应 用 随 机 过 程

李育强 姚强

2020年 9 月

内容简介

全书可分为三个部分. 第一个部分(第1、2章)介绍随机过程的预备知识, 第二个部分(第3、4章)介绍离散时间马氏链模型及其应用, 第三个部分(第5、6、7)分别介绍更新过程、布朗运动与离散时间鞅. 本书着重于对随机过程的基本知识、方法和思想的诠释, 并结果大量示例阐述随机过程在统计、管理、金融以及经济等方向的实际应用. 为了方便学习, 每节后还配有一定量的习题.

本书面向大学非数学专业的中、低年级学生, 也可供统计、管理等学科研究人员参考.

前　　言

"随机过程"以介绍若干研究复杂随机现象的理论、方法与应用为主要内容, 长期以来作为专业课程和研究类课程在高校若干专业的高年级或研究生阶段开设. 然而, 随着社会与科技的发展, 越来越多的学科意识到该课程的重要性, 开设该课程的专业越来越多, 开设的时间也被安排到大学二、三年级. 在这种背景下, 精选教学内容使其在保留课程特色与精华的同时降低学习门槛成为该课程教学中需要面对的最大挑战. 这本《应用随机过程》就是编者在多年的教学反思与实践基础上, 参考国内外有关著作编著而成的.

本书设定读者只具有数学分析(高等数学), 高等代数(线性代数)和初等概率论(概率论与数理统计)等大学一、二年级的数学基础. 内容着重于随机过程的基本理论和基本方法的介绍, 注重实际应用. 为了保证本书的独立性与理论体系的相对完整性，也为了便于读者自学使用, 本书在内容上做了如下的编排.

首先, 注重夯实与强化读者已有的数学基础. 本书第一章简要地回顾了初等概率论的基本内容, 并补充介绍了其中的条件数学期望, 分布函数的特征刻画以及随机变量的收敛等重、难点内容. 我们还通过第一章的练习帮助读者回顾一些学习随机过程所需的数学分析(高等数学)和高等代数知识.

其次, 注重与初等概率论课程的衔接. 我们在导入随机过程时从初等概率论涉及的随机变量序列出发, 以贝努利过程(独立同分布的贝努利随机变量序列), 随机游动(独立同分布随机变量的部分和序列), 泊松过程(独立同指数分布随机变量的部分和序列的逆)等具体模型作为本书的第二章内容, 帮助读者初步认识和了解随机过程所讨论的问题、所涉及的思想和方法以及相关的应用.

再次, 注重降低阅读本书的数学门槛. 读者只需在熟练掌握高等数学、线性代数以及初等概率论知识基础上都能顺利阅读与理解本书介绍的全部内容. 为了理论体系的相对完整, 我们对那些通常以测度论或实变函数等大学高年级数学知识为基础展开、但对理解随机过程理论和方法所必须的概念和结论都做了相应地改写和注释.

最后, 注重示例说明与学习效果检验. 本书配有各类示例超过120个, 帮助读者理解相关的概念、理论与方法. 书中许多例子具有一定的应用背景, 有助于读者理解如何应用随机过程解决实际问题. 同时, 为了便于读者检验自己的学习效果, 本书包含练习超过200 道并将练习分节设置. 根据练习的难易程度, 我们通过加*号对练习加以区分, 读者可以通过完成每章节后的练习加深对课程内容的理解(带*的内容与练习, 初学者可略过).

全书共七章, 除前面提到的第一、第二章外, 在第三章我们系统地介绍了马氏链的状态分类、平稳分布、转移概率的极限理论和遍历性定理. 在第四章我们结合可逆马氏链、隐马氏链、分枝过程等随机模型进一步展示了马氏链在各个领域方面的应用. 第五章介绍了更新过程的有关概念、方法、极限理论及其应用. 第六章围绕布朗运动, 我们主要介绍了它的定义与分布、极值与首达时的分布以及布朗运动的简单应用. 第七章简要地介绍了离散时间鞅的相关理论, 包括停时定理、鞅的不等式以及鞅的极限定理等内容. 考虑到平稳过程是时间序列的理论基础, 本书不包含这个专题. 本书第四、五、六、七章之间, 除个别例子与练习外, 内容相对独立.

依据编者的经验, 全部讲完本书内容需要54-72学时. 具体章节学时(含习题课)安排如下: 第一章6-10学时; 第二章8-10学时; 第三章10-14学时; 第四章6-8学时; 第五章8-10学时; 第六章8-10学时; 第七章8-10学时. 对于更短学时的教学, 编者建议在完整介绍前三章基础上, 根据培养目标, 有选择地介绍后面四章中的部分章节内容.

全书内容由李育强执笔, 姚强负责其中部分练习的补充、验算以及全书最后的统一与校、审等工作. 本书初稿在华东师范大学统计学院的统计学与金融工程专业反复试用过多次, 收到许多同学的反馈意见. 在此对这些专业的同学表示衷心的感谢.

本书是华东师范大学2019校精品教材建设专项的成果. 在教材编写过程中得到华东师范大学教务处、经济与管理学部以及统计学院等部门的大力支持, 在此也一并表示感谢. 由于编者水平有限, 错误之处在所难免, 恳请广大读者批评指正.

编　者

2020年9月20日

目录

第一章　预备知识

本章我们将简要回顾概率论有关知识并补充或深化一些必要内容.

1.1　随机变量及其分布

1.1.1　样本空间, 随机事件与概率

给定一个集合Ω, 若它包含了所研究随机现象的全部基本事件, 我们就称这个集合为样本空间. 称Ω的任意一个子集为事件; 称由Ω中若干子集构成的集类$\mathfrak{F}$为σ-代数, 若它满足

(1) $\Omega \in \mathfrak{F}$;

(2) $A \in \mathfrak{F} \Rightarrow \bar{A} \in \mathfrak{F}$;

(3) $A_k \in \mathfrak{F}, k = 1, 2, \cdots \Rightarrow \bigcup\limits_{k\geq 1} A_k \in \mathfrak{F}$.

P是$\mathfrak{F} \to [0,1]$上的函数, 用来度量$\mathfrak{F}$中事件发生的可能性. 我们称它为概率, 如果它满足以下条件

(1) 对任意$A \in \mathfrak{F}$, $P(A) \geq 0$;

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

(3) $A_k \in \mathfrak{F}, k = 1, 2, \cdots$互不相交$\Rightarrow P(\bigcup\limits_{k\geq 1} A_k) = \sum\limits_{k\geq 1} P(A_k)$.

称$P(A)$为事件A在概率P下的值, 简称为A的概率.

每个概率P都与一个σ-代数相关联. 当样本空间上给定一个概率P后, 关联σ-代数中的每一个事件A都有一个概率$P(A)$存在, 对那些不在关联σ-代数中的事件则没有明确的概率对应. 从直观上理解, 概率P对应于刻画某种随机规律的指标, 而相关联的σ-代数就是能够刻画出这种随机规律的事件集. 我们称$\mathfrak{F}$中每个元素为概率P下的随机事件, 简称为随机事件(在不引起混淆时也可简称为事件).

称n个随机事件$A_1, \cdots, A_n$是独立的, 若对任意$2 \leq k \leq n$个事件$A_{i_1}, \cdots, A_{i_k}$,

$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}).$$

任意给定两个随机事件A, B. 假定事件B发生的概率为正, 那么在事件B发生前提下事件A发生的条件概率(随机事件下的条件概率)为

$$P(A|B)=P(AB)/P(B).$$

本质上, 随机事件下的条件概率可以看作将样本空间Ω"缩小"为B后的概率.

随机事件概率有如下的基本计算公式.

(1) (全概率公式) 设随机事件$B_1,\cdots,B_n$为样本空间Ω的一个分割, 即$B_1,\cdots$, B_n 互不相交且它们的并集为全空间Ω, 那么对任意一个随机事件A,

$$P(A)=\sum_{k=1}^{n}P(AB_k).$$

(2) (乘法公式) 对任意两个随机事件A,B,

$$P(AB)=P(B)P(A|B).$$

(3) (贝叶斯(Bayes)公式) 设随机事件$B_1,\cdots,B_n$为样本空间Ω的一个分割, 那么

$$P(B_i|A)=\frac{P(AB_i)}{\sum_{k=1}^{n}P(A|B_k)P(B_k)}.$$

通常把样本空间Ω, 随机事件集$\mathfrak{F}$以及概率P三者看作一个整体, 称为概率空间, 记作$(\Omega,\mathfrak{F},P)$.

1.1.2 单个随机变量及相关刻画

称定义在概率空间$(\Omega,\mathfrak{F},P)$上的实值函数$X(\omega)$ 为随机变量, 若对任意$x\in\mathbb{R}$,

$$\{X\le x\}=\{\omega;X(\omega)\le x\}\in\mathfrak{F}.$$

通常简记随机变量为X, Y, Z. 称随机变量X,Y是几乎必然或几乎处处相等的(记作$X=Y$ a.s.), 如果$P(X=Y)=1$. 称函数

$$F(x)=P(X\le x)$$

为随机变量X的分布函数(简称为分布). F是分布函数当且仅当F右连续,单调不降而且$F(-\infty)=0$, $F(+\infty)=1$.

♠**注记1.1.1.** *随机变量的上述定义是我们在初等概率论课程学习随机变量时的一种常见定义. 值得强调的是, 单凭这个定义有时并不足以帮助我们准确辨识随机变量. 事实上, 随机变量除了具有该定义所展示的函数属性外还有分布属性(每个随机变量都对应到一个分布函数)以及与其他随机变量的关联属性(需要考虑联合分布)等等.*

(A) 离散型随机变量

称一个随机变量为离散型随机变量, 若该随机变量可能的不同取值只有有限多个或可用自然数的次序将所有取值一一排列出来(可列的). 设离散型随机变量X可能的不同取值为

$$x_0, x_1, \cdots, x_n, \cdots,$$

我们称

$$p_i = P(X = x_i), \qquad i = 0, 1, \cdots,$$

为X的概率分布列(简称为分布列). 称

$$E(X^k) = \sum_{i=0}^{\infty} x_i^k p_i, \qquad k \geq 1,$$

为X的k阶(原点)矩, 如果上式右边的求和有意义. 特别地, $k = 1$时称其为X的数学期望(均值). 称

$$Var(X) = \sum_{i=0}^{\infty} (x_i - E(X))^2 p_i$$

为X的方差.

下面我们列举几个常见离散型随机变量的分布.

(1) 两点分布$\mathrm{b}(1, p)$: $X \in \{0, 1\}$且$P(X = 1) = p,\ P(X = 0) = 1 - p,\ p \in (0, 1)$.

(2) 二项分布$\mathrm{b}(n, p)$: $X \in \{0, 1, \cdots, n\}$且

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}, \qquad k = 0, 1, \cdots, n, \ \ p \in (0, 1).$$

(3) 几何分布$\mathrm{Ge}(p)$: $X \in \{1, 2, \cdots, n, \cdots\}$且

$$P(X = k) = p(1 - p)^{k-1}, \qquad k \geq 1, \ \ p \in (0, 1).$$

(4) 泊松分布$\mathrm{P}(\lambda)$(称参数λ为强度): $X \in \{0, 1, 2, \cdots, n, \cdots\}$且

$$P(X = k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}, \qquad k \geq 0, \quad \lambda > 0.$$

通常我们简称离散型随机变量所刻画的分布为离散分布.

(B) 连续型随机变量

若存在非负函数$f(x)$使得随机变量X的分布函数

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

则称随机变量X是连续型的, 称函数$f(x)$为是X的概率密度函数.

一般地, 若$F(x)$在t的左导数$F'_{-}(t)$存在, 即

$$\lim_{s \to t-} \frac{F(s) - F(t)}{s - t} = \lim_{s \to t-} \frac{P(s < X \leq t)}{t - s}$$

存在且有限, 那么我们可以取概率密度函数f在t点的值为该极限. 特别, 若F可导, 那么概率密度函数f就可以取成F的一阶导数F'.

对任意$k \geq 1$, 称

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x)\mathrm{d}x$$

为X的k阶(原点)矩, 如果上式右边积分有意义. 特别地, 称$E(X)$为X的数学期望(均值). 称

$$Var(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)\mathrm{d}x$$

为X的方差.

一些典型连续型随机变量的分布如下.

(1) 均匀分布: X服从(a, b)上均匀分布$\mathrm{U}(a, b)$, 若概率密度函数为

$$f(x) = \begin{cases} 1/(b-a), & x \in (a, b), \\ 0, & \text{其他}. \end{cases}$$

(2) 指数分布: X服从$(0, \infty)$上参数$\lambda > 0$的指数分布$\mathrm{Exp}(\lambda)$, 若概率密度函数为

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{其他}. \end{cases}$$

(3) 伽马分布: X服从$(0, \infty)$上形状参数$a > 0$, 尺度参数$\lambda > 0$ 的伽马分布$\mathrm{Ga}(a, \lambda)$, 若概率密度函数为

$$f(x) = \begin{cases} \frac{\lambda^a x^{a-1}}{\Gamma(a)} \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{其他}. \end{cases}$$

(4) 正态分布: X服从$\mathbb{R}$上均值为μ, 方差为σ^2的正态分布$\mathrm{N}(\mu, \sigma^2)$, 若概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

通常我们简称连续型随机变量所刻画的分布为连续分布.

1.1.3 有限多个随机变量

同时考虑概率空间$(\Omega, \mathfrak{F}, P)$上的随机变量$X_1, \cdots, X_n$. 通常称

$$X = (X_1, \cdots, X_n)$$

为n元随机变量, 记作$X \in \mathbb{R}^n$. 当$n > 1$时, 统称为多元随机变量. 为表示方便, 我们也称$X = (X_1, \cdots, X_n)$ 为n维向量值随机变量(简称为向量值随机变量).

(A) 联合分布、协方差

设$X = (X_1, X_2, \cdots, X_n)$是$n$元随机变量, 对任意$(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, 令

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n).$$

称n元函数$F(x_1,x_2,\cdots,x_n)$为X的联合分布函数. 在多元随机变量背景下, 我们也称单个随机变量X_i的分布函数$F_i(x)=P(X_i\le x)$为X_i在该多元联合分布下的边际或边缘分布.

若$X_1,X_2,\cdots,X_n$都是离散型随机变量, 记所有可能的取值为

$$\{x_{i,j};\ i=1,\cdots,n,j=0,1,\cdots\}.$$

那么X的联合分布列为

$$p_{k_1,k_2,\cdots k_n}=P(X_1=x_{1,k_1},X_2=x_{2,k_2},\cdots,X_n=x_{n,k_n}).$$

相应地, 随机变量X_i的边际分布列

$$p_k^{(i)}=P(X_i=x_{i,k})=\sum_{\substack{k_r=0\\ r\ge 1,r\ne i}}^{\infty}P(X_1=x_{1,k_1},\cdots,X_i=x_{i,k},\cdots,X_n=x_{n,k_n}).$$

称X为n元连续型随机变量, 若存在非负函数$f(x_1,x_2,\cdots,x_n)$使得

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\mathrm{d}t_1\int_{-\infty}^{x_2}\mathrm{d}t_2\cdots\int_{-\infty}^{x_n}f(t_1,t_2,\cdots,t_n)\mathrm{d}t_n.$$

此时称该多元函数$f(x_1,x_2,\cdots,x_n)$为$(X_1,\cdots,X_n)$的联合概率密度函数. X_i的概率密度函数(也称为边际或边缘概率密度函数)$f_i(x)$可以用联合概率密度函数表示为

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(t_1,\cdots,t_{i-1},t,t_{i+1}\cdots t_n)\mathrm{d}t_1\cdots\mathrm{d}t_{i-1}\mathrm{d}t_{i+1}\cdots\mathrm{d}t_n.$$

若X是连续型随机变量, 那么在分布函数F的n阶混合偏导

$$\frac{\partial^n F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_n}$$

存在的点$(x_1,x_2,\cdots,x_n)$ 处, 可令

$$f(x_1,\cdots,x_n)=\frac{\partial^n F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_n}.$$

n元随机变量X的函数$g(X)=g(X_1,X_2,\cdots,X_n)$的数学期望定义为

$$E(g(X))=\begin{cases}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_n)f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n, & \text{连续型},\\ \sum_{k_i=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_n=0}^{\infty}g(x_{1,k_1},\cdots,x_{n,k_n})p_{k_1,\cdots,k_n}, & \text{离散型}.\end{cases}$$

任意两个随机变量之间的协方差为

$$\begin{aligned}Cov(X_i,X_j)&=E((X_i-E(X_i)(X_j-E(X_j)))\\ &=\begin{cases}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_i-E(X_i))(x_j-E(X_j))f_{i,j}(x_i,x_j)\mathrm{d}x_i\mathrm{d}x_j, & \text{连续型},\\ \sum_{k_i=0}^{\infty}\sum_{k_j=0}^{\infty}(x_{i,k_i}-E(X_i))(x_{j,k_j}-E(X_j))p_{k_i,k_j}^{(i,j)}, & \text{离散型},\end{cases}\end{aligned}$$

这里$f_{i,j}(x_i,x_j)$, $p_{k_i,k_j}^{(i,j)}$表示由X_i,X_j这两个随机变量所得到的联合概率密度函数或联合分布列. 称n阶方阵$\mathbf{D}=(Cov(X_i,X_j))$ 为n 个随机变量$X_1,X_2,\cdots,X_n$的协方差矩阵.

n元正态分布的是一种非常重要的连续型多元联合分布. 当$n=2$时我们

说 (X_1,X_2) 服从二元正态分布$\mathrm{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 若它们的联合概率密度函数$f(x_1,x_2)$ 为

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{2(1-\rho^2)}\Big(\frac{(x_1-\mu_1)^2}{\sigma_1^2}-2\rho\frac{x_1-\mu_1}{\sigma_1}\frac{x_2-\mu_2}{\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\Big)\Big\},$$

其中$\mu_1=E(X_1),\mu_2=E(X_2)$, $\sigma_1^2=VarX_1,\sigma_2^2=VarX_2$, 相关系数

$$\rho=\frac{Cov(X_1,X_2)}{\sigma_1\sigma_2}.$$

引进协方差矩阵

$$\mathbf{D}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}.$$

那么$|\mathbf{D}|=\sigma_1^2\sigma_2^2(1-\rho^2)$,

$$\mathbf{D}^{-1}=\frac{1}{|\mathbf{D}|}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}=\frac{1}{1-\rho^2}\begin{pmatrix}\sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1}\\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2}\end{pmatrix}.$$

记$x-\mu=(x_1-\mu_1,x_2-\mu_2)^T$. 利用矩阵表示, 二元正态概率密度函数可写成

$$f(x_1,x_2)=\frac{1}{2\pi\sqrt{|\mathbf{D}|}}\mathrm{e}^{-\frac{1}{2}(x-\mu)^TD^{-1}(x-\mu)}.$$

一般地, n元正态分布的概率密度函数为

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{D}|}}\mathrm{e}^{-\frac{1}{2}(x-\mu)^T\mathbf{D}^{-1}(x-\mu)}.$$

这里$\mathbf{D}$为正定矩阵, $x=(x_1,\cdots,x_n)^T,\mu=(\mu_1,\cdots,\mu_n)^T$, 其中$\mu_1,\cdots,\mu_n$ 为常数. 记该n元正态分布为$N(\mu,\mathbf{D})$. 若随机变量$(X_1,\cdots,X_n)$服从分布$N(\mu,\mathbf{D})$, 那么$\mathbf{D}$ 就是$X_1,\cdots,X_n$之间的协方差矩阵, 而且对任意$1\le k\le n$, $E(X_k)=\mu_k$.

下面这个例子介绍了参数为$(p_1,p_2,\cdots,p_m)$的多项分布.

▶**例1.1.2.** *向m个罐子里抛硬币, 硬币落入第i个罐子的概率设为p_i. 连续独立地抛n次. 以X_i表示落入第i个罐子的硬币数. 求$(X_1,X_2,\cdots,X_m)$ 的联合分布. 其中$p_1+\cdots+p_m=1$.*

解 以n_i表示观测到的落入第i个罐子的硬币数, 那么$0\le n_i\le n$, 而且$n_1+\cdots+n_m=n$. 随机事件

$$\{X_i=n_i,i=1,2,\cdots,m\}$$

发生的组合数为

$$C_n^{n_1}C_{n-n_1}^{n_2}\cdots C_{n-n_1-\cdots-n_{m-2}}^{n_{m-1}}=\frac{n!}{n_1!n_2!\cdots n_m!},$$

由实验的独立同分布性可知每一组合发生的概率为$p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m}$, 因此

$$P(X_i=n_i,i=1,2,\cdots,n)=\begin{cases}\frac{n!}{n_1!n_2!\cdots n_m!}p_1^{n_1}p_2^{n_2}\cdots p_m^{n_m}, & n_i\ge 0,\sum_{i=1}^m n_i=n,\\ 0, & \text{其它}.\end{cases}$$

容易看出任意前k项$(k<m)$的分布列为

$$P(X_i=n_i,i=1,2,\cdots,k)=\begin{cases}\frac{n!}{n_1!\cdots n_k!m_k!}p_1^{n_1}\cdots p_k^{n_k}q_k^{m_k}, & n_i\geq 0,\sum_{i=1}^k n_i\leq n,\\ 0, & \text{其它},\end{cases}$$

其中$q_k=1-\sum_{i=1}^k p_i$, $m_k=n-\sum_{i=1}^k n_i$. □

(B) 随机变量的独立性

称n个随机变量$X_1,X_2,\cdots,X_n$是独立的, 如果对任意$x_1,x_2,\cdots,x_n\in\mathbb{R}$,

$$P(X_1\leq x_1,X_2\leq x_2,\cdots,X_n\leq x_n)=P(X_1\leq x_1)P(X_2\leq x_2)\cdots P(X_n\leq x_n).$$

称无穷多个随机变量$X_1,X_2,\cdots,X_n,\cdots$是独立的, 若对任意$k\geq 2$, X_1, X_2, $\cdots$, X_k 是独立的.

若$X_1,X_2,\cdots,X_n$为连续型随机变量, 那么$X_1,X_2,\cdots,X_n$独立当且仅当它们的联合概率密度函数等于各自概率密度函数的乘积, 即

$$f(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n).$$

随机变量相互独立的概念和性质还可推广到多元随机变量的情形.

♣**定义1.1.3.** 设$\mathbf{X}_i=(X_{i,1},\cdots,X_{i,k_i})$, $i=1,\cdots,n$, 分别是$k_1,\cdots,k_n$维的随机变量, 其中每个$k_i\geq 1$. 称多元随机变量$\mathbf{X}_1,\cdots,\mathbf{X}_n$ 是独立的, 若任意的$x_{i,j}\in(-\infty,+\infty)$, 其中$1\leq j\leq k_i$, $1\leq i\leq n$,

$$P(X_{i,j}\leq x_{i,j},1\leq j\leq k_i,1\leq i\leq n)=\prod_{i=1}^n P(X_{i,j}\leq x_{i,j},1\leq j\leq k_i).$$

由多元随机变量独立性的定义可知, 若随机变量$X_1,\cdots,X_n$是独立的, 将其分成不相交的任意k组, 将每组看成一个多元随机变量, 记作$\mathbf{X}_1,\cdots,\mathbf{X}_k$, 那么这$k$个多元随机变量也是独立的. 若随机变量$Y_i$只是$\mathbf{X}_i$的函数, 那么$Y_1,\cdots,Y_k$也相互独立.

(C) 随机变量函数的分布

设n元随机变量$X=(X_1,\cdots,X_n)$的取值区域为$\mathcal{D}$, Φ是从$\mathcal{D}$ 到$\mathcal{A}\in\mathbb{R}^m$的函数, 令$m$元随机变量$Y=(Y_1,\cdots,Y_m)=\Phi(X)$, 那么对任意集合$B\subset\mathcal{A}$,

$$P(Y\in B)=P(X\in\Phi^{-1}(B)),$$

其中$\Phi^{-1}(B)$表示集合B在$\mathcal{D}$内的原像集.

特别地, 若$m=n$且Φ为可逆函数, 那么

(1) 若X是离散的, 对任意$y=(y_1,\cdots,y_n)$, 设$\Phi^{-1}(y)=(x_1,\cdots,x_n)$, 那么

$$P(Y_1=y_1,\cdots,Y_n=y_n)=P(X_1=x_1,\cdots,X_n=x_n).$$

(2) 若X是连续型随机变量, Φ连续、可微且雅可比(Jacobi)行列式

$$J_\Phi=\left|\frac{\partial y_i}{\partial x_j}\right|_{i,j=1,\cdots,n}$$

处处不为0, 其中y_i表示函数Φ的第i 个分量. 那么对任意$(u_1,\cdots,u_n)\in\mathcal{A}$,

$$f_Y(u_1,\cdots,u_n)=\frac{f_X(\Phi^{-1}(u_1,\cdots,u_n))}{|J_\Phi|_{\Phi^{-1}(u_1,\cdots,u_n)}},$$

其中f_Y, f_X分别表示Y,X的联合概率密度函数, $|J_\Phi|_{\Phi^{-1}(u_1,\cdots,u_n)}$表示雅可比行列式$J_\Phi$在$(x_1,\cdots,x_n)=\Phi^{-1}(u_1,\cdots,u_n)$点的绝对值.

▶**例1.1.4.** 设$\{T_k;k\geq 1\}$是非负随机变量序列$\{W_i;i\geq 1\}$的部分和序列, 即

$$T_k=\sum_{i=1}^k W_i,\quad k\geq 1.$$

若$T=(T_1,\cdots,T_n)$ 的联合概率密度函数为

$$f_T(t_1,\cdots,t_n)=\begin{cases}\lambda^n \mathrm{e}^{-\lambda t_n}, & 0<t_1<t_2<\cdots<t_n\\ 0, & \text{其它}.\end{cases}$$

那么$W_i, i=1,\cdots,n$, 服从参数为λ的指数分布且相互独立.

证明 定义一个从$\mathcal{A}=\{(t_1,\cdots,t_n);0\leq t_1\leq t_2\leq\cdots\leq t_n\}$到$\bar{\mathbb{R}}_+^n$的映射$\Phi$, 使得

$$\Phi(t_1,t_2,\cdots,t_n)=(t_1,t_2-t_1,\cdots,t_n-t_{n-1}).$$

显然Φ是连续、可微的可逆函数, 雅可比行列式$J_\Phi\equiv 1$. 令

$$W=(W_1,\cdots,W_n)=\Phi(T_1,\cdots,T_n).$$

那么$(W_1,\cdots,W_n)$在$\bar{\mathbb{R}}_+^n$ 上的联合概率密度函数

$$\begin{aligned}f_W(x_1,x_2,\cdots,x_n)&=\frac{f_T(\Phi^{-1}(x_1,\cdots,x_n))}{|J_\Phi|_{\Phi^{-1}(x_1,\cdots,x_n)}}\\&=f_T(\Phi^{-1}(x_1,\cdots,x_n))=\lambda^n\mathrm{e}^{-\lambda\sum_{i=1}^n x_i}.\end{aligned}$$

由此可知W_1的概率密度函数为

$$f_{W_1}(x_1)=\int_0^\infty \mathrm{d}x_2\cdots\int_0^\infty\lambda^n\mathrm{e}^{-\lambda\sum_{i=1}^n x_i}\mathrm{d}x_n=\lambda\mathrm{e}^{-\lambda x_1}.$$

类似计算可得W_i, $i\geq 2$, 的概率密度函数为$f_{W_i}(x_i)=\lambda\mathrm{e}^{-\lambda x_i}$. 因此

$$f_W(x_1,x_2,\cdots,x_n)=\prod_{i=1}^n f_{W_i}(x_i).$$

所以$W_i, i=1,\cdots,n$, 是服从参数为λ的指数分布的独立同分布随机变量. □

(D) 独立随机变量的和与极值的分布

若X,Y独立且分布函数分别为$F_X(x)$和$F_Y(y)$, 那么$Z=X+Y$的分布函数

$$G(z)=P(Z\leq z)=P(X+Y\leq z)$$

可用卷积表示, $G = F_X * F_Y$, 即对任意$z \in \mathbb{R}$,

$$G(z) = F_X * F_Y(z) = \begin{cases} \sum_{x_k} p_X(x_k) F_Y(z - x_k), & X\text{为离散型随机变量}, \\ \int_{-\infty}^{\infty} f_X(x) F_Y(z - x)\mathrm{d}x, & X\text{为连续型随机变量}. \end{cases}$$

特别地, 若X, Y均是连续型随机变量, 那么Z也是连续型随机变量, 它的概率密度函数

$$g(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\mathrm{d}x.$$

若$X_1, X_2, \cdots, X_n$独立同分布, 设共同分布为F, 那么对任意$2 \le k \le n$, $Z_k = X_1 + \cdots + X_k$的分布函数

$$F_k(z) = P(X_1 + X_2 + \cdots + X_k \le z) = F^{*k}(z) = F^{*(k-1)} * F,$$

其中$F^{*1} = F$. 若F有概率密度函数f, 那么Z_k的概率密度函数

$$f_k(z) = \int_{-\infty}^{\infty} f_{k-1}(x) f(z - x)\mathrm{d}x,$$

其中$f_1 = f$. 此时它们中最大值和最小值的分布函数分别为

$$F_{\max}(z) = [F(z)]^n, \qquad F_{\min}(z) = 1 - (1 - F(z))^n,$$

概率密度函数分别是$f_{\max}(z) = n[F(z)]^{n-1} f(z)$和$f_{\min}(z) = n(1 - F(z))^{n-1} f(z)$.

►**例1.1.5.** (1)若X_1, $X_2, \cdots$, X_n是分别服从$\mathrm{Ga}(a_1, \lambda)$, $\mathrm{Ga}(a_2, \lambda)$, $\cdots$, $\mathrm{Ga}(a_n, \lambda)$分布的独立随机变量, 那么$X_1 + \cdots + X_n$服从$\mathrm{Ga}(a_1 + \cdots + a_n, \lambda)$分布.

(2)若X_1, $X_2, \cdots, X_n$是分别服从$\mathrm{N}(\mu_1, \sigma_1^2)$, $\mathrm{N}(\mu_2, \sigma_2^2), \cdots, \mathrm{N}(\mu_n, \sigma_n^2)$分布的独立随机变量, 那么$X_1 + \cdots + X_n$服从$\mathrm{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$分布.

作为本节的结束, 最后我们指出随机变量的两个基本不等式.

(1) 若随机变量X的p阶矩存在, 那么对任意$x > 0$,

$$P(|X| \ge x) \le \frac{E(|X|^p)}{x^p}, \qquad p > 0,$$

其中$p = 1$时称不等式为马尔可夫(Markov)不等式, $p = 2$时称其为切比雪夫(Chebyshev)不等式.

(2) 若随机变量X, Y的二阶矩都存在, 那么

$$[E(XY)]^2 \le E(X^2)E(Y^2),$$

即柯西-施瓦茨(Cauchy-Schwarz)不等式成立.

练习题1.1

1 假定下面出现的条件概率都有意义. 证明

$$P(AB|C) = P(A|C)P(B|AC) = P(B|C)P(A|BC).$$

2 对任意随机事件B, 记条件概率$P(B|A)$为$P_A(B)$, 若$P(AC)\neq 0$, 证明

$$P_A(B|C)=P(B|AC).$$

3 假定抛游戏币只有正面或反面两种结果. 现在抛两次游戏币, 分别以X,Y表示相应的结果: 正面取1, 反面取0. 试写出下面三种情形对应的概率空间, 分析X与Y是否相同. (1) 抛两次不同材质的游戏币(材质影响正反面出现的概率); (2)将两次抛游戏币看作是重复的独立试验; (3)若第一次抛出正面, 则第二次抛一个两面都是正面的游戏币, 反之则抛两面都是反面的游戏币.

4 设X_1,X_2,X_3独立且服从均匀分布$\mathrm{U}(0,1)$, 分别用$X_{(1)},X_{(2)},X_{(3)}$表示其中最小, 其次和最大的随机变量. 求$X_{(1)},X_{(2)},X_{(3)}$的联合概率密度函数.

5 若$k\geq 0$时非负数列$\{a_n(k);n\geq 1\}$单调不降且$\lim\limits_{n\to\infty}a_n(k)=a(k)$. 证明

$$\sum_{k=0}^{\infty}a(k)=\lim_{n\to\infty}\sum_{k=0}^{\infty}a_n(k).$$

6 利用习题5证明

(1) 若随机变量X的取值为$0,1,2,\cdots$, 那么$E(X)=\sum_{k=1}^{\infty}P(X\geq k)$.

(2) 若X是非负连续型随机变量, 那么$E(X)=\int_0^{\infty}P(X>x)\mathrm{d}x$.

7 任取整数集的子集S,U, 若对任意$i\in S$, $j\in U$, $b_{i,j}$非负, 试用习题5证明

$$\sum_{i\in S}\sum_{j\in U}b_{i,j}=\sum_{j\in U}\sum_{i\in S}b_{i,j}.$$

8* 利用习题5证明(无穷求和或积分的交换次序公式)

(1) 若对任意n,k, $b_{n,k}$是非负实数, 证明$\sum\limits_{n=1}^{\infty}\sum\limits_{k=1}^{n}b_{n,k}=\sum\limits_{k=1}^{\infty}\sum\limits_{n=k}^{\infty}b_{n,k}$.

(2) 若$f(x,y)$非负可积, 那么$\int_0^{\infty}\mathrm{d}x\int_0^{x}f(x,y)\mathrm{d}y=\int_0^{\infty}\mathrm{d}y\int_y^{\infty}f(x,y)\mathrm{d}x$.

9 证明若X的二阶矩存在, 那么$P(|X|>0)\geq[E(X)]^2/E(X^2)$.

10* 设正整数$k_1,\cdots,k_m$的最大公因数为d. 证明存在正整数N, 当$n>N$时存在非负整数$c_1,\cdots,c_m$使得$nd=c_1k_1+c_2k_2+\cdots+c_mk_m$.

1.2 条件数学期望

条件数学期望是概率计算时的重要工具. 这一节我们将在回顾条件数学期望的定义与性质基础上进一步补充与介绍条件数学期望的一些内容.

1.2.1 随机事件发生条件下的分布与期望

任意给定的两个随机变量X, Y, 利用初等概率中随机事件发生条件下的条件概率定义可知, 在事件$\{Y \in A\}$发生的条件下X的分布函数为

$$F_X(x|A) = P(X \leq x|Y \in A) = \frac{P(X \leq x, Y \in A)}{P(Y \in A)}.$$

注意到随机事件发生条件下的条件概率实质上是某个新的样本空间上的概率, 因此我们在初等概率论学习的许多问题可以在条件概率下讨论. 比如, 在$Y \in A$的条件下我们可以定义X的函数$g(X)$的数学期望为

$$E(g(X)|Y \in A) = \begin{cases} \frac{\sum_x g(x)P(X=x, Y \in A)}{P(Y \in A)}, & X\text{离散}, \\ \int_{-\infty}^{\infty} g(x)f(x|A)\mathrm{d}x, & X\text{连续}, \end{cases}$$

其中$f(x|A)$为分布函数$F_X(x|A)$的概率密度函数.

▶**例1.2.1.** 设(X, Y)服从$(0,1) \times (0,2)$上的均匀分布, 求$E(X|Y > 1)$.

解 令$A = (1, \infty)$. 当(X, Y)为连续型随机变量时,

$$f(x|A) = \frac{\int_A f(x,y)\mathrm{d}y}{\int_{-\infty}^{\infty}\mathrm{d}x \int_A f(x,y)\mathrm{d}y},$$

其中$f(x,y)$为(X, Y)的联合概率密度函数. 由于(X, Y)服从$(0,1) \times (0,2)$上均匀分布,

$$f(x,y) = 1/2, \quad 0 < x < 1, 0 < y < 2.$$

代入计算得

$$f(x|A) = 1, \quad 0 < x < 1.$$

从而

$$E(X|Y > 1) = \int_0^1 x f(x|A)\mathrm{d}x = 1/2. \qquad \square$$

特别地, 对二元离散型随机变量(X, Y), 设y是Y的一个可能值, 在$\{Y = y\}$的条件下, X的条件分布列为

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)},$$

其中x表示X的可能取值. 在$\{Y=y\}$的条件下, X的条件数学期望为

$$E(X|Y=y)=\frac{\sum_x xP(X=x,Y=y)}{P(Y=y)}=\sum_x xP(X=x|Y=y). \quad (1.2.1)$$

二元函数$\Phi(X,Y)$在$\{Y=y\}$的条件下变成$\Phi(X,y)$, 从而条件数学期望为

$$\begin{aligned}E(\Phi(X,Y)|Y=y)&=\frac{\sum_x \Phi(x,y)P(X=x,Y=y)}{P(Y=y)}\\&=\sum_x \Phi(x,y)P(X=x|Y=y).\end{aligned}$$

▶**例1.2.2.** 设X, X_1分别服从二项分布$b(n,p)$和$b(m,p)$且独立, $m<n$. 求

$$E(X|X+X_1=n).$$

解 令$Y=X+X_1$, 则$Y\sim b(n+m,p)$. 因此

$$\begin{aligned}E(X|X+X_1=n)&=\frac{\sum_k kP(X=k,Y=n)}{P(Y=n)}\\&=\frac{\sum_{k=n-m}^{n} kP(X=k,X_1=n-k)}{P(Y=n)}\\&=\frac{\sum_{k=n-m}^{n} kC_n^kC_m^{n-k}}{C_{n+m}^n}=\frac{\sum_{k=0}^{m}(n-k)C_n^{n-k}C_m^k}{C_{n+m}^n}.\end{aligned}$$

注意到

$$\sum_{k=0}^{m}(n-k)C_n^{n-k}C_m^k=\sum_{k=0}^{m}nC_{n-1}^{n-1-k}C_m^k=nC_{n+m-1}^{n-1},$$

我们可得

$$E(X|X+X_1=n)=\frac{n\sum_{k=0}^{m}C_{n-1}^kC_m^{m-k}}{C_{n+m}^m}=n\frac{C_{n+m-1}^{n-1}}{C_{n+m}^n}=\frac{n^2}{n+m}. \qquad \square$$

当(X,Y)是二元连续型随机变量时,虽然随机事件$\{Y=y\}$发生的概率为0, 但由初等概率论可知, 此时我们可以引进X的条件概率密度函数

$$f(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{f(x,y)}{\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x},$$

其中$f(x,y)$为(X,Y)的联合概率密度函数. 从而在$\{Y=y\}$的条件下, X的条件数学期望定义为

$$E(X|Y=y)=\int_{-\infty}^{\infty}xf(x|y)\mathrm{d}x=\frac{\int_{-\infty}^{\infty}xf(x,y)\mathrm{d}x}{f_Y(y)}. \quad (1.2.2)$$

类似, 对二元函数$\Phi(X,Y)$在$\{Y=y\}$的条件下的条件数学期望为

$$E(\Phi(X,Y)|Y=y)=\int_{-\infty}^{\infty}\Phi(x,y)f(x|y)\mathrm{d}x.$$

▶**例1.2.3.** 对任意$y>0$, 求$E(X|Y=y)$, 其中(X,Y)的联合概率密度函数为

$$f(x,y)=6/(1+x+y)^4, \qquad 0<x,y<\infty.$$

解 对任意$x>0$,

$$f(x|y)=\frac{f(x,y)}{\int_0^\infty f(x,y)\mathrm{d}x}=\frac{3(1+y)^3}{(1+x+y)^4}.$$

因此

$$E(X|Y=y)=\int_0^\infty xf(x|y)\mathrm{d}x=(1+y)^3\int_0^\infty \frac{3x}{(1+x+y)^4}\mathrm{d}x=\frac{1+y}{2}. \qquad \square$$

1.2.2 随机变量下的条件数学期望

一般而言, 当条件事件$\{Y=y\}$变化时, 对应的条件数学期望也会发生改变, 从而我们可以设计一个从$\Omega\longrightarrow\mathbb{R}$的映射, 使得对任意$\omega\in\Omega$,

$$\omega\in\Omega\longrightarrow E(X|Y=y),\qquad \text{其中 } y=Y(\omega). \tag{1.2.3}$$

然而这个映射可能有缺陷, 因为若(1.2.1)或(1.2.2)的分母为0, 表达式$E(X|Y=y)$没有意义. 但此时若令

$$\mathcal{Y}=\begin{cases}\{y;\ P(Y=y)>0\}, & Y\text{离散},\\ \{y;\ f_Y(y)\neq 0\}, & Y\text{连续},\end{cases}$$

以及$\Omega_Y=\{\omega;\ Y(\omega)\in\mathcal{Y}\}$, 那么映射在集合$\Omega_Y$上是明确、可行的.

♣**定义1.2.4.** *设(X,Y)是连续型或离散型二元随机变量, X的数学期望存在. 若存在随机变量Z使得*

(1) *Z为随机变量Y的函数, 即存在函数g使得$Z=g(Y)$;*

(2) *对任意$\omega\in\Omega_Y$,*

$$Z(\omega)=g(Y(\omega))=E(X|Y=y),\qquad \text{其中 } y=Y(\omega). \tag{1.2.4}$$

那么我们称Z为X关于随机变量Y的条件数学期望, 简称为条件数学期望.

从条件数学期望定义可看出函数g满足: 对任意$y\in\mathcal{Y}$, $g(y)=E(X|Y=y)$. 进一步, 条件数学期望Z只能保证在Ω_Y上是明确、唯一的. 对Ω_Y外的ω, 定义本身并没有明确要求; 因此条件数学期望可能是不唯一的. 但这样定义出来的任意两个可能的条件数学期望Z_1,Z_2必满足$\Omega_Y\subset\{Z_1=Z_2\}$. 注意到

$$P(\Omega_Y)=\begin{cases}\sum_{y\in\mathcal{Y}}P(Y=y)=1, & Y\text{离散},\\ \int_{\mathcal{Y}}f_Y(y)\mathrm{d}y=1, & Y\text{连续}.\end{cases}$$

这表明由定义1.2.4给出的任意两个条件数学期望在一个概率为1的事件上是完全一致的. 因此在几乎处处相等意义下, 定义1.2.4中的条件数学期望是唯一的.

♠**注记1.2.5.** *通常把该"唯一"的条件数学期望记作$E(X|Y)$. 此后凡涉及条件数学期望及其性质时, 若有必要都在"几乎处处成立"这一前提下理解, 即除了*

一个零概率事件, 对应的性质成立.

▶**例1.2.6.** 设$\Omega = (0,1)$且对任意$(a,b) \subset \Omega$, $P((a,b)) = b-a$. X, Y是Ω上的随机变量, 满足

$$X(\omega) = \begin{cases} -1, & \omega \in (0,1/2), \\ 1, & \omega \in [1/2,1), \end{cases} \qquad Y(\omega) = \begin{cases} 0, & \omega \in (0,1/4), \\ 1, & \omega = 1/4, \\ 2, & \omega \in (1/4,1). \end{cases}$$

求$E(X|Y)$.

解 由题设可知

$$P(Y=0) = 1/4,\ P(Y=1) = 0 \text{ 且 } P(Y=2) = 3/4.$$

因此$\mathcal{Y} = \{0,2\}$. 从而

$$g(0) = E(X|Y=0) = \frac{-P(X=-1,Y=0)+P(X=1,Y=0)}{P(Y=0)} = -1,$$
$$g(2) = E(X|Y=2) = \frac{-P(X=-1,Y=2)+P(X=1,Y=2)}{P(Y=2)} = 1/3.$$

因此

$$E(X|Y)(\omega) = \begin{cases} g(0) = -1, & \omega \in (0,1/4), \\ g(2) = 1/3, & \omega \in [1/4,1), \end{cases}$$

这里我们补充(自由)定义$E(X|Y)$在$\omega = 1/4$的值为$1/3$. □

▶**例1.2.7.** 设(X,Y)为二元正态随机变量, 联合概率密度函数如下, 求$E(X|Y)$.

$$f(x,y) = \frac{1}{\sqrt{3}\pi} \mathrm{e}^{-\frac{2}{3}(x^2-xy+y^2)}, \qquad (x,y) \in \mathbb{R}^2.$$

解 注意到

$$f(x,y) = \frac{1}{\sqrt{3}\pi} \mathrm{e}^{-\frac{2}{3}(x-y/2)^2} \mathrm{e}^{-\frac{y^2}{2}},$$

直接计算可知

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}}.$$

因此

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{\sqrt{3\pi/2}} \mathrm{e}^{-\frac{(x-y/2)^2}{3/2}},$$

由此可得, 对任意$y \in \mathbb{R}$,

$$g(y) = E(X|Y=y) = \int_{-\infty}^{\infty} x f(x|y)\mathrm{d}x = \frac{y}{2}.$$

所以$E(X|Y) = g(Y) = Y/2$. □

▶**例1.2.8.** (X,Y)的联合概率密度函数为$f(x,y)$. 设$a < b$ 且$P(a \le Y \le b) >$

0. 证明
$$\int_a^b E(X|Y=y)f_Y(y)\mathrm{d}y = E(X|a \le Y \le b)P(a \le Y \le b),$$
其中$f_Y(y)$为Y的概率密度函数.

证明 随机事件$A=\{a \le Y \le b\}$发生条件下, 随机变量X的概率密度函数为
$$f_{X|A}(x) = \frac{\int_a^b f(x,y)\mathrm{d}y}{\int_{-\infty}^{\infty}\int_a^b f(x,y)\mathrm{d}x\mathrm{d}y},$$
因此
$$\begin{aligned} E(X|a \le Y \le b) &= \int_{-\infty}^{\infty} x f_{X|A}(x)\mathrm{d}x = \frac{\int_{-\infty}^{\infty}\mathrm{d}x\int_a^b xf(x,y)\mathrm{d}y}{\int_{-\infty}^{\infty}\int_a^b f(x,y)\mathrm{d}x\mathrm{d}y} \\ &= \frac{\int_{-\infty}^{\infty}\mathrm{d}x\int_a^b xf(x,y)\mathrm{d}y}{P(A)}. \end{aligned}$$
另一方面, 记$E(X|Y)$为$g(Y)$, 那么
$$\begin{aligned} \int_a^b E(X|Y=y)f_Y(y)\mathrm{d}y &= \int_a^b \mathrm{d}y\int_{-\infty}^{\infty} xf(x|y)f_Y(y)\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\mathrm{d}x\int_a^b xf(x,y)\mathrm{d}y. \end{aligned}$$
因此
$$P(a \le Y \le b)E(X|a \le Y \le b) = \int_a^b E(X|Y=y)f_Y(y)\mathrm{d}y. \qquad \square$$

随机变量下的条件数学期望有如下重要性质.

■**性质1.2.9.** 假设以下数学期望与条件数学期望总有意义.

(1) 若X与Y独立, 则$E(X|Y)=E(X)$.

(2) 若$a \le \Phi(X,Y) \le b$, 则
$$a \le E(\Phi(X,Y)|Y) \le b,$$
从而$E(a|Y)=a$, 其中a,b为常数.

(3) $E(\sum\limits_{i=1}^n a_iX_i|Y) = \sum\limits_{i=1}^n a_iE(X_i|Y)$, 其中$a_i$非随机.

(4) $|E(X|Y)| \le E(|X||Y)$.

(5) $E(X)=E(E(X|Y))$. (重期望公式)

(6) 对几乎处处的$\omega \in \Omega$, 记$y=Y(\omega)$, 那么
$$E(\Phi(X,Y)|Y)(\omega) = E(\Phi(X,y)|Y=y).$$
特别地,
$$E(f(X)g(Y)|Y) = g(Y)E(f(X)|Y) \text{ 且 } E(g(Y)|Y)=g(Y).$$

证明 我们仅以连续型随机变量为例证明(1),(5)和(6).

(1) 按此前关于条件数学期望的解释, 我们只需要证明对任意$\omega \in \Omega_Y$, 等式成立即可. 任取$\omega \in \Omega_Y$, 记$y = Y(\omega)$,

$$\begin{aligned} E(X|Y)(\omega) &= E(X|Y=y) = \int_{-\infty}^{\infty} x f(x|y) dx \\ &= \int_{-\infty}^{\infty} x f_X(x) dx = E(X), \end{aligned}$$

其中$f_X(x)$表示随机变量X的概率密度函数.

(5) 直接计算得

$$\begin{aligned} E(E(X|Y)) &= \int_{-\infty}^{\infty} E(X|Y=y) f_Y(y) \mathrm{d}y \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x|y) \mathrm{d}x f_Y(y) \mathrm{d}y \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y) \mathrm{d}x \mathrm{d}y = E(X). \end{aligned}$$

(6) 为叙述简单, 仅考虑对任意给定的y函数$\phi_y(x) = \Phi(x,y)$关于x单调的情形. 此时记$Z = \Phi(X,Y)$, 那么(Z,Y)的联合概率密度函数为

$$g(z,y) = \frac{f(x,y)}{|\phi_y'(x)|}\Big|_{x=\phi_y^{-1}(z)}.$$

此时由随机变量下条件期望定义

$$E(\Phi(X,Y)|Y)(\omega) = E(Z|Y)(\omega) = E(Z|Y=y) = \frac{\int_{-\infty}^{\infty} z g(z,y) \mathrm{d}z}{\int_{-\infty}^{\infty} g(z,y) \mathrm{d}z},$$

其中$y = Y(\omega)$. 作变量代换$x = \phi_y^{-1}(z)$, 那么

$$\begin{aligned} E(\Phi(X,Y)|Y)(\omega) &= \frac{\int_{-\infty}^{\infty} \phi_y(x) f(x,y) \mathrm{d}x}{\int_{-\infty}^{\infty} f(x,y) \mathrm{d}x} \\ &= \frac{\int_{-\infty}^{\infty} \Phi(x,y) f(x,y) \mathrm{d}x}{\int_{-\infty}^{\infty} f(x,y) \mathrm{d}x} = E(\Phi(X,y)|Y=y). \end{aligned}$$

其他性质请读者自己完成; 参见习题2.6. □

▶**例1.2.10.** *某矿工身陷在有三个门的矿井中, 选择第1个门的通道行进2小时后他将到达安全地, 选择第2个门的通道后将在3个小时后回到原地, 选择第3 个门的通道后将在5个小时后回到原地. 假定这个矿工每次都随机地选取一个门, 问他到达安全地的平均时间是多少?*

解 记X为矿工到达安全地的时间, N为矿工第一次选择的通道. 由于矿工回到原地后仍然是随机选择通道, 因此当矿工回到原地后面对的问题与开始时相同. 由问题假设我们可得

$$E(X|N=1) = 2, \quad E(X|N=2) = 3 + E(X), \quad E(X|N=3) = 5 + E(X).$$

由重期望公式可得

$$
\begin{aligned}
&E(X)=E(E(X|N))\\
&=E(X|N=1)P(N=1)+E(X|N=2)P(N=2)+E(X|N=3)P(N=3)\\
&=\frac{1}{3}(2+3+5+2E(X)).
\end{aligned}
$$

由此可知$E(X)=10$. □

1.2.3 条件数学期望的推广与一般化

条件数学期望还可推广到多个条件随机变量的情形. 下面的定义以多元连续型随机变量为例. 对离散情形, 可类似定义.

♣**定义1.2.11.** 设$(X,Y_1,\cdots,Y_n)$是连续型随机变量且X的数学期望存在. 令

$$
\mathcal{Y}=\{(y_1,\cdots,y_n);\ f_Y(y_1,\cdots,y_n)\neq 0\},
$$

$$
\Omega_Y=\{\omega;\ (Y_1(\omega),\cdots,Y_n(\omega))\in\mathcal{Y}\},
$$

其中$f_Y(y_1,\cdots,y_n)$为$Y=(Y_1,\cdots,Y_n)$的联合概率密度函数. 若随机变量

$$
Z=g(Y_1,Y_2,\cdots,Y_n)
$$

满足: 对任意$\omega\in\Omega_Y$,

$$
\begin{aligned}
Z(\omega)&=g(Y_1(\omega),\cdots,Y_n(\omega))=E(X|Y_1=y_1,\cdots,Y_n=y_n)\\
&=\int_{-\infty}^{\infty}\frac{xf(x,y_1,\cdots,y_n)}{f_Y(y_1,\cdots,y_n)}\mathrm{d}x,
\end{aligned}
$$

其中$y_k=Y_k(\omega)$, $f(x,y_1,\cdots,y_n)$为$(X,Y_1,\cdots,Y_n)$的联合概率密度函数, 则称Z为X关于随机变量$Y_1,\cdots,Y_n$的条件数学期望, 记作$E(X|Y_1,\cdots,Y_n)$.

此时条件数学期望仍遵循定义1.2.4后对条件数学期望概念的解释和说明, 并且对应的性质1.2.9仍然成立. 比如

(1) 若X与$Y_1,\cdots,Y_n$独立, 则$E(X|Y_1,\cdots,Y_n)=E(X)$.

(2) $E(\sum\limits_{i=1}^{n}a_iX_i|Y_1,\cdots,Y_n)=\sum\limits_{i=1}^{n}a_iE(X_i|Y_1,\cdots,Y_n)$, 其中对任意$1\leq i\leq n$, a_i为常数.

(3) $E(\Phi(X,Y_1,\cdots,Y_n)|Y_1=y_1,\cdots,Y_n=y_n)$

$$
=E(\Phi(X,y_1,\cdots,y_n)|Y_1=y_1,\cdots,Y_n=y_n).
$$

(4) $E\big(E(\Phi(X,Y_1,\cdots,Y_n)|Y_1,\cdots,Y_n)\big)=E(\Phi(X,Y_1,\cdots,Y_n))$.

作为重期望公式的拓展, 我们还有如下的所谓平滑性质.

■**性质1.2.12.** 对任意$n>m$,

$$
E\big(E(h(X)|Y_1,\cdots,Y_n)\big|Y_1,\cdots,Y_m\big)=E(h(X)|Y_1,\cdots,Y_m).
$$

证明* 只考虑连续型随机变量的情形.

设$(X,Y_1,\cdots,Y_n)$的联合概率密度函数为$f(x,y_1,\cdots,y_n)$, $(Y_1,\cdots,Y_n)$的联合概率密度函数

$$f_{Y_{1:n}}(y_1,\cdots,y_n)=\int_{-\infty}^{\infty}f(x,y_1,\cdots,y_n)\mathrm{d}x,$$

而对任意$1\le m\le n$, $(Y_1,\cdots,Y_m)$的联合概率密度函数

$$f_{Y_{1:m}}(y_1,\cdots,y_m)=\int_{\mathbb{R}^{n-m+1}}f(x,y_1,\cdots,y_m,y_{m+1},\cdots,y_n)\mathrm{d}x\mathrm{d}y_{m+1}\cdots\mathrm{d}y_n.$$

记$g(Y_1,\cdots,Y_n)=E(h(X)|Y_1,\cdots,Y_n)$, 对任意$k\le n$, 令

$$\mathcal{Y}_{1:k}=\{(y_1,\cdots,y_k):f_{Y_{1:k}}(y_1,\cdots,y_k)\neq 0\}.$$

任取$(y_1,\cdots,y_m)\in\mathcal{Y}_{1:m}$,

$$\begin{aligned}
&E\big(E(h(X)|Y_1,\cdots,Y_n)\big|Y_1=y_1,\cdots,Y_m=y_m\big)\\
&\quad=E(g(Y_1,\cdots,Y_n)|Y_1=y_1,\cdots,Y_m=y_m)\\
&\quad=\frac{\int_{-\infty}^{\infty}\mathrm{d}y_{m+1}\cdots\int_{-\infty}^{\infty}(gf_{Y_{1:n}})(y_1,\cdots,y_m,y_{m+1},\cdots,y_n)\mathrm{d}y_n}{f_{Y_{1:m}}(y_1,\cdots,y_m)}.
\end{aligned}$$

由条件数学期望定义, 当$(y_1,\cdots,y_n)\in\mathcal{Y}_{1:n}$时,

$$g(y_1,\cdots,y_{m+1},\cdots,y_n)=\frac{\int_{-\infty}^{\infty}h(x)f(x,y_1,\cdots,y_m,y_{m+1},\cdots,y_n)\mathrm{d}x}{f_{Y_{1:n}}(y_1,\cdots,y_{m+1},\cdots,y_n)},$$

而当$(y_1,\cdots,y_n)\notin\mathcal{Y}_{1:n}$时,可以令$g(y_1,\cdots,y_m,y_{m+1},\cdots,y_n)=0$. 因此

$$\begin{aligned}
&E\big(E(h(X)|Y_1,\cdots,Y_n)\big|Y_1=y_1,\cdots,Y_m=y_m\big)\\
&=\frac{\int_{-\infty}^{\infty}\mathrm{d}y_{m+1}\cdots\int_{-\infty}^{\infty}\mathrm{d}y_n\int_{-\infty}^{\infty}h(x)f(x,y_1,\cdots,y_m,y_{m+1},\cdots,y_n)\mathrm{d}x}{f_{Y_{1:m}}(y_1,\cdots,y_m)}\\
&=E(h(X)|Y_1=y_1,\cdots,Y_m=y_m).
\end{aligned}$$

由条件数学期望定义及其说明可知结论成立. □

♠注记1.2.13. *条件数学期望还可推广到更一般的情形, 具体的细节不再展开. 要指出的是此后凡涉及随机变量下的条件期望时, 我们总认为它已有恰当定义, 而且相应的性质1.2.9 和性质1.2.12也成立.*

▶例1.2.14. *设X,Y为独立同分布随机变量, 求$E(X|X+Y)$.*

解 由X,Y为独立同分布随机变量可知

$$E(Y|X+Y)=E(Y|Y+X)=E(X|X+Y).$$

因此

$$\begin{aligned}
2E(X|X+Y)&=E(X|X+Y)+E(Y|X+Y)\\
&=E(X+Y|X+Y)=X+Y.
\end{aligned}$$

由此可得$E(X|X+Y)=(X+Y)/2$. □

▶例1.2.15. 设$\{X_n; n \geq 1\}$是一族独立随机变量, 对任意$n \geq 1$, $S_n = X_1 + \cdots + X_n$, 证明对任意$n \geq 1$, $E(S_{n+1}|S_1, \cdots, S_n) = E(S_{n+1}|S_n)$.

证明 由独立性假设可知$S_{n+1} - S_n = X_{n+1}$与$X_1, \cdots, X_n$独立, 进而与$S_1, \cdots, S_n$独立, 因此由条件数学期望性质可得

$$\begin{aligned}
E(S_{n+1}|S_1, \cdots, S_n) &= E(S_n|S_1, \cdots, S_n) + E(X_{n+1}|S_1, \cdots, S_n) \\
&= S_n + E(X_{n+1}) \\
&= E(S_n|S_n) + E(X_{n+1}|S_n) = E(S_{n+1}|S_n).
\end{aligned}$$

□

▶例1.2.16. 设某仪器寿命服从参数为λ的指数分布, 假定仪器损坏后能及时更换新仪器, 而且仪器之间的寿命相互没有影响, 问从仪器新装上开始的t时间内平均使用的仪器数量.

[**分析**] 记$N(t)$为t时间内使用的仪器数量, 问题所求为$E(N(t))$. 记第一个仪器的寿命为W. 显然, 当$W = w > t$时$N(t) = 1$. 当$W = w \leq t$时, 记从w 到t这段时间用的仪器个数为$N(w, t)$, 那么$N(t) = 1 + N(w, t)$. 注意到$N(w, t)$ 只与w时刻之后用的新仪器有关, 根据假设, 作为一个随机变量与第一台仪器的寿命W无关而且与从0开始$t - W$时间内的仪器使用数量$N(t - w)$ 分布相同, 即

$$E(N(t)|W = w) = \begin{cases} 1, & w > t, \\ 1 + E(N(t - w)), & w \leq t. \end{cases}$$

解 记$N(t)$为$[0, t]$内使用仪器数, 第一台仪器寿命为W, 那么$W \sim \mathrm{Exp}(\lambda)$. 由条件数学期望性质可知

$$\begin{aligned}
E(N(t)) = E(E(N(t)|W)) &= \int_0^{\infty} E(N(t)|W = w)\lambda \mathrm{e}^{-\lambda w} \mathrm{d}w \\
&= \int_t^{\infty} \lambda \mathrm{e}^{-\lambda w} \mathrm{d}w + \int_0^t (1 + E(N(t - w)))\lambda \mathrm{e}^{-\lambda w} \mathrm{d}w.
\end{aligned}$$

令$u(t) = E(N(t))$, 则

$$u(t) = 1 + \int_0^t \lambda \mathrm{e}^{-\lambda(t-s)} u(s) \mathrm{d}s = 1 + \lambda \mathrm{e}^{-\lambda t} \int_0^t \mathrm{e}^{\lambda s} u(s) \mathrm{d}s.$$

两边关于t求导得,

$$u'(t) = -\lambda(u(t) - 1) + \lambda u(t) = \lambda.$$

故$u(t) = \lambda t + c$. 再由$u(0) = E(N(0)) = 1$得$c = 1$. 因此

$$E(N(t)) = u(t) = 1 + \lambda t.$$

□

利用其示性函数$1_A(\omega) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A, \end{cases}$ 我们也可以定义随机事件A在给定随机变量下的条件概率.

♣**定义1.2.17.** 对任意随机事件A, 称

$$P(A|Y_1,\cdots,Y_n)=E(1_A|Y_1,\cdots,Y_n)$$

为随机事件A在随机变量$Y_1,\cdots,Y_n$条件下的条件概率.

显然随机变量条件下的条件概率是个$[0,1]$ 区间取值的随机变量.

▶**例1.2.18.** 设二维连续型随机变量的联合概率密度函数

$$f(x,y)=\lambda^2 y\mathrm{e}^{-\lambda y(1+x)},\quad x,y\geq 0,$$

求条件概率$P(X<Y|Y)$.

解 对任意$y\geq 0$, 由定义可知

$$\begin{aligned}P(X<Y|Y=y)=E(1_{\{X<y\}}|Y=y)&=\frac{\int_0^y\lambda^2 y\mathrm{e}^{\lambda y(1+x)}\mathrm{d}x}{\int_0^\infty\lambda^2 y\mathrm{e}^{\lambda y(1+x)}\mathrm{d}x}\\&=\frac{\lambda(\mathrm{e}^{-\lambda y}-\mathrm{e}^{-\lambda y(1+y)})}{\lambda\mathrm{e}^{-\lambda y}}=1-\mathrm{e}^{-\lambda y^2},\end{aligned}$$

因此$P(X<Y|Y)=1-\mathrm{e}^{-\lambda Y^2}$. □

与数学期望和方差的关系一样, 利用条件数学期望也可定义条件方差

$$Var(X|Y)=E\big((X-E(X|Y))^2|Y\big)=E(X^2|Y)-\big(E(X|Y)\big)^2.$$

显然, 条件方差仍是随机变量. 对几乎处处$\omega\in\Omega$, 记$y=Y(\omega)$,

$$Var(X|Y)(\omega)=E\big((X-E(X|Y))^2|Y=y\big)=E\big((X-E(X|Y=y))^2|Y=y\big).$$

▶**例1.2.19.** 接例*1.2.3*, 求$Var(X|Y)$.

解 由例1.2.3易知$E(X|Y)=(1+Y)/2$, 而且对任意$y>0$,

$$E(X^2|Y=y)=(1+y)^3\int_0^\infty\frac{3x^2}{(1+x+y)^4}\mathrm{d}x=(1+y)^2.$$

因此$E(X^2|Y)=(1+Y)^2$, 进而

$$Var(X|Y)=(1+Y)^2-(1+Y)^2/4=3(1+Y)^2/4.$$ □

▲**命题1.2.20.** 若X的二阶矩存在, 那么

$$Var(X)=E(Var(X|Y))+Var(E(X|Y)).$$

证明 由重期望公式可得

$$\begin{aligned}Var(X)&=E\big((X-E(X))^2\big)=E\big(E\big((X-E(X))^2|Y\big)\big)\\&=E\big(E\big((X-E(X|Y)+E(X|Y)-E(X))^2|Y\big)\big)\\&=E\big(E\big[(X-E(X|Y))^2|Y\big]\big)+E\big(E\big[(E(X|Y)-E(X))^2|Y\big]\big)\\&=E(Var(X|Y))+E\big((E(X|Y)-E(X))^2\big)\\&=E(Var(X|Y))+Var(E(X|Y)).\end{aligned}$$ □

▶**例1.2.21.** 给定一簇随机变量$\{X_t;\ t\in T\}$. 设N是取值在T上的随机变量. 由此可定义复合随机变量X_N使得对任意ω, $X_N(\omega)=X_{N(\omega)}(\omega)$. 现假设$Y_k$, $k=1,2,\cdots$,是独立同分布的随机变量, 二阶矩存在. N 是非负整数值的随机变量, 与$\{Y_k;k\geq 1\}$ 独立且二阶矩也存在. 令$X_N=\sum_{k=1}^N Y_k$. 求复合随机变量X_N的均值与方差.

解 对任意$\omega\in\varOmega$, 记$n=N(\omega)$. 由独立性和条件期望性质可得

$$\begin{aligned}
E(X_N|N)(\omega)&=E\Big(\sum_{k=1}^N Y_k\Big|N=n\Big)=E\Big(\sum_{k=1}^n Y_k\Big|N=n\Big)=nE(Y_1),\\
Var(X_N|N)(\omega)&=E\Big(\big(\sum_{k=1}^N Y_k-nE(Y_1)\big)^2\Big|N=n\Big)\\
&=E\Big(\big(\sum_{k=1}^n\big(Y_k-E(Y_k)\big)\big)^2\Big|N=n\Big)\\
&=E\Big(\big(\sum_{k=1}^n\big(Y_k-E(Y_k)\big)\big)^2\Big)=nVar(Y_1).
\end{aligned}$$

因此$E(X_N|N)=NE(Y_1)$, $Var(X_N|N)=NVar(Y_1)$. 从而

$$\begin{aligned}
E(X_N)&=E(E(X_N|N))=E(N)E(Y_1);\\
Var(X_N)&=E(Var(X_N|N))+Var(E(X_N|N))\\
&=E(N)Var(Y_1)+[E(Y_1)]^2Var(N).
\end{aligned}$$

□

练习题1.2

1 已知$X_m,m=1,\cdots,n$, 是独立同分布的随机变量, 而且

$$P(X_1=1)=p,\qquad P(X_1=0)=1-p.$$

任取整数$1\leq k\leq n$, 对任意$1\leq m\leq n$, 求$P(X_m=1|X_1+X_2+\cdots+X_n=k)$.

2 若连续型随机变量X,Y独立, 概率密度函数分别为$f(x)$和$g(x)$, 证明

$$E(\Phi(X,Y)|X)=\int_{\mathbb{R}}\Phi(X,y)g(y)\mathrm{d}y.$$

3 设二元随机变量(X,Y)的联合概率密度函数为

$$f(x,y)=6x^2y,\ \ 0<x,y<1.$$

求$E(X|Y)$, $E(X|X+Y)$.

4 设抛游戏币出现正反面的概率分别是p, $1-p$. 每次抛游戏币都是独立的. 假定直至连续出现k次正面抛游戏币才结束, 证明抛掷游戏币次数N的平均

值

$$E(N) = \frac{1-p^k}{p^k(1-p)}.$$

5 一对夫妇希望至少有一个男孩和一个女孩而且只要男孩女孩都有了就不再生育. 假设男孩的出生率为p, $0 < p < 1$. 求他们孩子中男孩比例的预期.

6 在连续型随机变量情形下证明性质1.2.9中(2),(3),(4).

7 设X, Y是两有界随机变量且$E(X|Y) = Y$, $E(Y|X) = X$, 证明

$$P(X = Y) = 1.$$

1.3 特征函数

这一节我们回顾与介绍刻画随机变量分布的三个主要工具: 特征函数, 拉普拉斯(Laplace)变换与概率母函数.

1.3.1 特征函数

♣**定义1.3.1.** *设随机变量X的分布函数为F. 我们称复值函数*

$$\begin{aligned}\psi_X(t) &= E(\mathrm{e}^{itX}) = \int_{-\infty}^{\infty} \mathrm{e}^{itx}\mathrm{d}F(x)\\ &= \int_{-\infty}^{\infty} \cos(tx)\mathrm{d}F(x) + i\int_{-\infty}^{\infty} \sin(tx)\mathrm{d}F(x)\\ &= E(\cos(tX)) + iE(\sin(tX))\end{aligned}$$

为随机变量X(或分布F)的特征函数, 其中$t \in (-\infty, \infty)$, i为虚数单位.

由定义可知ψ_X作为变量t的函数在$\mathbb{R}$上一致连续, 而且对任意$t \in \mathbb{R}$,

$$|\psi_X(t)| \le \psi_X(0) = 1.$$

特征函数与分布函数之间有如下关联.

★**定理1.3.2.** (唯一性定理) *特征函数与概率分布相互唯一确定. 即不同的分布函数有不同的特征函数; 特征函数不同则对应的分布函数也不同.*

★**定理1.3.3.** (连续性定理) *设随机变量X_n的特征函数为$\psi_{X_n}(t)$. 那么存在随机变量X使得X_n依分布收敛到X(记作$X_n \xrightarrow{d} X$, 其中依分布收敛的定义可参见本章第四节)的充要条件是对任意$t \in \mathbb{R}$, $\psi_{X_n}(t)$ 收敛而且*

$$\lim_{t\to 0}\lim_{n\to\infty} \psi_{X_n}(t) = 1.$$

■**性质1.3.4.** *特征函数具有如下性质:*

(1) $\overline{\psi_X(t)} = \psi_X(-t) = \psi_{-X}(t)$, 其中$\overline{\psi_X(t)}$表示复函数$\psi_X(t)$的共轭.

(2) $\psi_{aX+b}(t) = \psi_X(at)\mathrm{e}^{ibt}$.

(3) 若X, Y独立, 那么$\psi_{X+Y}(t) = \psi_X(t)\psi_Y(t)$, $t \in \mathbb{R}$.

利用特征函数我们还可以方便地求出X的各阶原点矩.

■**性质1.3.5.** *若X的n阶绝对矩$E(|X|^n) < \infty$, 那么对任意$k \le n$,*

$$E(X^k) = i^{-k}\psi_X^{(k)}(0).$$

几个常用的分布的特征函数如下:

(1) 指数分布$\mathrm{Exp}(\lambda)$的特征函数为$\psi(t) = \lambda/(\lambda - it)$.

(2) 伽马分布$\mathrm{Ga}(a,\lambda)$的特征函数为$\psi(t)=[\lambda/(\lambda-it)]^a$.

(3) 正态分布$\mathrm{N}(\mu,\sigma^2)$的特征函数是$\psi(t)=\exp\left\{i\mu t-\frac{\sigma^2t^2}{2}\right\}$.

特征函数也可推广到多元随机变量的情形. 设$X=(X_1,\cdots,X_n)$为n元随机变量, 称n元复值函数

$$\psi_{(X_1,X_2,\cdots,X_n)}(t_1,\cdots,t_n)=E(\mathrm{e}^{i(t_1X_1+\cdots+t_nX_n)})$$

为X的特征函数, 其中$t_1,\cdots,t_n\in\mathbb{R}$.

与一元随机变量的特征函数一样, 多元特征函数也具有唯一性.

★**定理1.3.6.** *n元特征函数与n维随机变量的分布函数相互唯一确定.*

由此我们可得如下利用特征函数判别随机变量独立性的方法.

▲**命题1.3.7.** *随机变量$X_1,\cdots,X_n$相互独立当且仅当对任意的$t_1,\cdots,t_n\in\mathbb{R}$,*

$$\psi_{(X_1,X_2,\cdots,X_n)}(t_1,\cdots,t_n)=\prod_{k=1}^n\psi_{X_k}(t_k).$$

►**例1.3.8.** *设$X=(X_1,\cdots,X_n)$服从n元正态分布$N(\mu,\mathbf{D})$, 其中μ为n维向量, $\mathbf{D}$为n阶正定矩阵, 求X的特征函数.*

解 对任意$t_1,t_2,\cdots,t_n\in\mathbb{R}$, X的特征函数

$$\phi(t_1,t_2,\cdots,t_n)=E(\mathrm{e}^{i\sum_{k=1}^n t_kX_k}).$$

为了表示方便, 令$t=(t_1,t_2,\cdots,t_n)^T$, $x=(x_1,x_2,\cdots,x_n)^T$, 那么

$$\phi(t_1,t_2,\cdots,t_n)=\int_{\mathbb{R}^n}\mathrm{e}^{it^Tx}\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{D}|}}\mathrm{e}^{-\frac{(x-\mu)^T\mathbf{D}^{-1}(x-\mu)}{2}}\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

由矩阵$\mathbf{D}$的对称性可得

$$\begin{aligned}&\frac{(x-\mu-i\mathbf{D}t)^T\mathbf{D}^{-1}(x-\mu-i\mathbf{D}t)}{2}\\&=\frac{(x-\mu)^T\mathbf{D}^{-1}(x-\mu)}{2}-\frac{t^T\mathbf{D}^T\mathbf{D}^{-1}\mathbf{D}t}{2}\\&\quad-\frac{i(x-\mu)^T\mathbf{D}^{-1}\mathbf{D}t+it^T\mathbf{D}^T\mathbf{D}^{-1}(x-\mu)}{2}\\&=\frac{(x-\mu)^T\mathbf{D}^{-1}(x-\mu)}{2}-\frac{t^T\mathbf{D}t}{2}+it^T\mu-it^Tx.\end{aligned}$$

因此

$$\begin{aligned}\phi(t_1,t_2,\cdots,t_n)&=\int_{\mathbb{R}^n}\frac{\mathrm{e}^{it^T\mu-\frac{t^T\mathbf{D}t}{2}}}{(2\pi)^{n/2}\sqrt{|\mathbf{D}|}}\mathrm{e}^{-\frac{(x-\mu-i\mathbf{D}t)^T\mathbf{D}^{-1}(x-\mu-i\mathbf{D}t)}{2}}\mathrm{d}x_1\cdots\mathrm{d}x_n\\&=\mathrm{e}^{it^T\mu-\frac{t^T\mathbf{D}t}{2}}.\end{aligned}$$

□

1.3.2 拉普拉斯变换

对非负随机变量分布的刻画, 用所谓的拉普拉斯(Laplace)变换会更方便.

♣**定义1.3.9.** *若X为非负随机变量, 称函数$L_X(\theta)=E(\mathrm{e}^{-\theta X})$ 为X的拉普拉斯变换, 其中变量$\theta\in[0,\ +\infty)$.*

记X的分布函数为F. 拉普拉斯变换$L_X(\theta)$可用积分形式表示成

$$L_X(\theta)=\int_0^\infty \mathrm{e}^{-\theta x}\mathrm{d}F(x).$$

此时也称$L_X(\theta)$是分布F的拉普拉斯变换, 也常记做$L_F(\theta)$.

当$F(x)$是具有密度为$f(x)$的连续分布时,

$$L_F(\theta)=\int_0^\infty \mathrm{e}^{-\theta x}f(x)\mathrm{d}x.$$

当$F(x)$是具有分布列$p_k=P(X=x_k)$, $k=1,2,\cdots,\infty$, 的离散分布时,

$$L_F(\theta)=\sum_{k=1}^{\infty}\mathrm{e}^{-\theta x_k}p_k.$$

特别地, 若x_k取非负整数, 令$p_k=P(X=k)$, 那么

$$L_X(\theta)=\sum_{k=0}^{\infty}\mathrm{e}^{-\theta k}p_k.$$

★**定理1.3.10.** *(拉普拉斯变换反演公式) 设F是$[0,\infty)$上的分布函数, $L_F(\theta)$是F的拉普拉斯变换, 那么对F的任意连续点x, 有*

$$F(x)=\lim_{n\to\infty}\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n),$$

其中$\lfloor\cdot\rfloor$表示不超过$\cdot$的最大整数.

证明* 只证明F为连续分布的情形. 设其概率密度函数为f.

对拉普拉斯变换$L_F(\theta)$在$(0,\infty)$内求其k阶导数得

$$L_F^{(k)}(\theta)=\int_0^\infty(-y)^k\mathrm{e}^{-\theta y}f(y)\mathrm{d}y.$$

对任意给定的$x\geq 0$, 取$\theta=n$, 则

$$\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n)=\int_0^\infty\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}f(y)\mathrm{d}y.$$

设X_i, $i\geq 1$, 独立且服从强度为y的泊松分布, 那么$E(X_1)=Var(X_1)=y$且

$$X_1+X_2+\cdots+X_n$$

服从强度为ny的泊松分布. 此时

$$P(X_1+X_2+\cdots+X_n\leq\lfloor nx\rfloor)=\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}.$$

对任意$\varepsilon>0$, 当$y-x\ge\varepsilon$时, 由切比雪夫不等式,

$$\begin{aligned}\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}&=P\Big(\frac{X_1+X_2+\cdots+X_n-ny}{n}\le\frac{\lfloor nx\rfloor-ny}{n}\Big)\\&\le P\Big(\frac{|X_1+X_2+\cdots+X_n-ny|}{n}\ge\frac{|\lfloor nx\rfloor-ny|}{n}\Big)\\&\le\frac{n^2}{(ny-\lfloor nx\rfloor)^2}\frac{Var(X_1)}{n}\le\frac{y}{n(y-x)^2}\le\frac{1}{n}(\frac{x}{\varepsilon^2}+\frac{1}{\varepsilon}).\end{aligned}$$

另一方面, 当$y-x\le-\varepsilon$时,

$$\begin{aligned}1-\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}&=P(X_1+X_2+\cdots+X_n>\lfloor nx\rfloor)\\&\le P(X_1+X_2+\cdots+X_n\ge nx)\\&=P\Big(\frac{X_1+X_2+\cdots+X_n-ny}{n}\ge x-y\Big)\\&\le\frac{1}{(x-y)^2}\frac{y}{n}\le\frac{x}{n\varepsilon^2}.\end{aligned}$$

因此

$$\begin{aligned}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n)-F(x)\Big|&\le\int_0^{(x-\varepsilon)\vee 0}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}-1\Big|f(y)\mathrm{d}y\\&\quad+\int_{x+\varepsilon}^{\infty}\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}f(y)\mathrm{d}y+2\int_{(x-\varepsilon)\vee 0}^{x+\varepsilon}f(y)\mathrm{d}y\\&\le\frac{1}{n}(\frac{x}{\varepsilon^2}+\frac{1}{\varepsilon})+2(F(x+\varepsilon)-F(x-\varepsilon)),\end{aligned}$$

其中对任意$a,b\in\mathbb{R}, a\vee b=\max\{a,b\}$. 令$n\to\infty$, 我们可得

$$\limsup_{n\to\infty}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n)-F(x)\Big|\le 2(F(x+\varepsilon)-F(x-\varepsilon)).$$

由于ε的任意性以及

$$\lim_{\varepsilon\to 0}F(x+\varepsilon)-F(x-\varepsilon)=0,$$

我们可得

$$\limsup_{n\to\infty}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n)-F(x)\Big|=0.$$

因此

$$\lim_{n\to\infty}\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL_F^{(k)}(n)=F(x).$$

由定理1.3.10 我们可以立即得到如下的拉普拉斯变换的唯一性定理.

★定理1.3.11. $[0,\infty)$上不同的概率分布有不同的拉普拉斯变换.

对于拉普拉斯变换也有所谓的连续性定理. 定理的证明超出了本课程范畴, 省略. 感兴趣的同学可以参考 [1]和 [12]中有关内容.

★**定理1.3.12.** 设非负随机变量X_n的拉普拉斯变换为$L_{X_n}(\theta)$. 那么存在随机变量X使得$X_n \xrightarrow{d} X$的充要条件是对任意$\theta > 0$, $L_{X_n}(\theta)$收敛而且

$$\lim_{\theta\to 0}\lim_{n\to\infty} L_{X_n}(\theta) = 1.$$

♠**注记1.3.13.** *定理1.3.12可以推广为如下的所谓广义连续性定理: 非负随机变量$X_n \xrightarrow{d} X$当且仅当存在$\lambda > 0$使得对任意$\theta > \lambda$, $L_{X_n}(\theta) \to L_X(\theta)$.*

此外, 由拉普拉斯变换定义容易验证如下性质成立.

■**性质1.3.14.** (1) 对任意$\theta > 0$, $0 \le L_X(\theta) \le L_X(0) = 1$.

(2) 对任意$a > 0, b > 0$, $L_{aX+b}(\theta) = L_X(a\theta)\mathrm{e}^{-b\theta}$.

(3) 若X, Y独立非负随机变量, 那么$L_{X+Y}(\theta) = L_X(\theta)L_Y(\theta)$, $\theta \ge 0$.

(4) 对任意正整数k, $E(X^k) = (-1)^k L_X^{(k)}(0)$, 其中$L_X^{(k)}$表示$L_X$的$k$阶导数.

▶**例1.3.15.** 若X服从均匀分布$\mathrm{U}(0,1)$, 那么$L_X(\theta) = (1-\mathrm{e}^{-\theta})/\theta$.

▶**例1.3.16.** 若X服从指数分布$\mathrm{Exp}(\lambda)$, 那么

$$L_X(\theta) = \int_0^\infty \mathrm{e}^{-\theta x}\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x = \lambda/(\theta+\lambda).$$

▶**例1.3.17.** 若X服从伽马分布$\mathrm{Ga}(a,\lambda)$, 那么

$$L_X(\theta) = \int_0^\infty \mathrm{e}^{-\theta x}\frac{\lambda^a x^{a-1}}{\Gamma(a)}\mathrm{e}^{-\lambda x}\mathrm{d}x = \int_0^\infty \frac{\lambda^a x^{a-1}}{\Gamma(a)}\mathrm{e}^{-(\theta+\lambda)x}\mathrm{d}x = \left(\frac{\lambda}{\theta+\lambda}\right)^a.$$

▶**例1.3.18.** 若X_n服从二项分布$b(n,\frac{1}{n})$, 那么对任意$\theta > 0$, 当$n\to\infty$时

$$\begin{aligned} L_{X_n}(\theta) &= \sum_{k=0}^n C_n^k(1-\frac{1}{n})^{n-k}\frac{1}{n^k}\mathrm{e}^{-\theta k} \\ &= (1-\frac{1-\mathrm{e}^{-\theta}}{n})^n \to \mathrm{e}^{-(1-\mathrm{e}^{-\theta})}. \end{aligned}$$

注意$\mathrm{e}^{-(1-\mathrm{e}^{-\theta})}$是强度为1的泊松分布的拉普拉斯变换. 因此当$n\to\infty$时, X_n依分布收敛到一个强度为1的泊松分布的随机变量.

1.3.3 概率母函数

在研究只取有限或无限多个非负整数值的随机变量时, 用所谓概率母函数来代替特征函数或拉普拉斯变换较为方便.

♣**定义1.3.19.** 设X为非负整数值的随机变量, 称函数$\phi_X(s) = E(s^X)$ 为X的概率母函数, 其中变量$s \in [-1,1]$.

设X的分布列为

$$p_k = P(X=k), \qquad k \ge 0,$$

那么

$$\phi_X(s) = \sum_{k=0}^{\infty} p_k s^k.$$

对照拉普拉斯变换定义, 容易看出$s \in (0,1]$时, 概率母函数

$$\phi_X(s) = L_X(-\ln s).$$

因此我们容易将拉普拉斯变换的性质平移到概率母函数. 对此我们不再赘述, 请读者自己补充. 下面我们主要介绍和解释概率母函数性质的一些变化.

■**性质1.3.20.** 若非负整数值随机变量X的n阶矩存在, 那么对任意$k \le n$以及$s \in [-1,1]$, 概率母函数ϕ_X的k阶导数$\phi_X^{(k)}(s)$存在, 且X的n阶矩可由$\phi_X(s)$前n阶导数在$s=1$的值表示. 特别地,

$$E(X) = \phi_X'(1), \qquad E(X^2) = \phi_X''(1) + \phi_X'(1).$$

证明 由微积分中幂级数一致收敛性质可知结果成立, 细节请读者补充. □

★**定理1.3.21.** (反演公式) 若非负整数值随机变量X的概率母函数为$\phi_X(s)$. 那么对任意$k \ge 0$, $P(X=k) = \phi_X^{(k)}(0)/k!$.

证明 由$\phi_X(s) = \sum_{n=0}^{\infty} p_n s^n$可知

$$\phi_X^{(k)}(s) = \sum_{n=k}^{\infty} p_n s^{n-k} n(n-1)\cdots(n-k+1).$$

因此$\phi_X^{(k)}(0) = k! p_k$. □

▶**例1.3.22.** 设X服从几何分布$\mathrm{Ge}(p)$, 那么

$$\phi_X(s) = \sum_{k=1}^{\infty} s^k p(1-p)^{k-1} = ps\sum_{k=0}^{\infty} s^k (1-p)^k = \frac{ps}{1-(1-p)s}.$$

▶**例1.3.23.** 设X_1, X_2, $\cdots$ 是一列服从均匀分布$\mathrm{U}(0,1)$的独立随机变量. Y服从几何分布$\mathrm{Ge}(p)$且与$\{X_m; m \ge 1\}$ 独立, 令$S = \sum\limits_{m=1}^{Y} X_m$, 求$S$的拉普拉斯变换$L_S(\theta)$.

解 由定义和条件数学期望的性质可得

$$\begin{aligned} L_S(\theta) &= E(\mathrm{e}^{-\theta S}) = E(E(\mathrm{e}^{-\theta S}|Y)) = E(E(\mathrm{e}^{-\theta\sum_{m=1}^{Y} X_m}|Y)) \\ &= \sum_{k=1}^{\infty} E(\mathrm{e}^{-\theta\sum_{m=1}^{Y} X_m}|Y=k)p(1-p)^{k-1} \\ &= \sum_{k=1}^{\infty} E(\mathrm{e}^{-\theta\sum_{m=1}^{k} X_m}|Y=k)p(1-p)^{k-1}. \end{aligned}$$

由于

$$E(\mathrm{e}^{-\theta\sum_{m=1}^{k}X_m}|Y=k)=E(\mathrm{e}^{-\theta\sum_{m=1}^{k}X_m})=[E(\mathrm{e}^{-\theta X_1})]^k=\Big(\frac{1-\mathrm{e}^{-\theta}}{\theta}\Big)^k,$$

因此

$$L_S(\theta)=\sum_{k=1}^{\infty}\Big(\frac{1-\mathrm{e}^{-\theta}}{\theta}\Big)^k p(1-p)^{k-1}=\frac{p(1-\mathrm{e}^{-\theta})}{\theta-(1-p)(1-\mathrm{e}^{-\theta})}.$$ □

容易将概率母函数的定义推广到多元非负整数值的随机变量, 得到联合分布的概率母函数. 为简单, 下面仅介绍二元随机变量情形, 三元及以上情形有类似的结果.

♣**定义1.3.24.** 设X_1,X_2为非负整数值的随机变量, 联合分布列为

$$p_{n,m}=P(X_1=n,X_2=m),\qquad n,m\geq 0,$$

称$[-1,1]\times[-1,1]$上的二元函数

$$\phi_{(X_1,X_2)}(s,t)=E(s^{X_1}t^{X_2})=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}s^n t^m p_{n,m}$$

为二元随机变量(X_1,X_2)的概率母函数.

与定理1.3.21一样, 此时我们有反演公式

$$p_{n,m}=\frac{\partial^{n+m}\phi_{(X_1,X_2)}(s,t)}{n!m!\partial s^n\partial t^m}\Big|_{\substack{s=0\\t=0}}.$$

▲**命题1.3.25.** 非负整数值随机变量X_1,X_2相互独立的充要条件是对任意的$s,t\in[-1,1]$,

$$\phi_{(X_1,X_2)}(s,t)=\phi_{X_1}(s)\phi_{X_2}(t)=\phi_{(X_1,X_2)}(s,1)\phi_{(X_1,X_2)}(1,t).$$

♠**注记1.3.26.** 拉普拉斯变换和概率母函数可以不做任何改变地适用于以正概率取$+\infty$的非负随机变量. 此时, 独立随机变量和的拉普拉斯变换(概率母函数)仍为各自拉普拉斯变换(概率母函数)的乘积.

练习题1.3

1 设X是连续型非负随机变量, 证明$L_X(t)$在$[0,+\infty)$内是一致连续的.

2 设X是连续型非负随机变量, 分布函数为$F(x)$, 证明对任意$t>0$, 其拉普拉斯变换$L_X(t)$满足

$$\frac{1-L_X(t)}{t}=\int_0^{\infty}\mathrm{e}^{-tx}(1-F(x))\mathrm{d}x.$$

3 $\{X_k;k\geq 1\}$是一列独立同分布的随机变量, 服从指数分布$\mathrm{Exp}(\lambda)$. 假定$\{X_k;k\geq 1\}$与服从泊松分布$\mathrm{P}(\lambda)$的随机变量N独立. 令$S=\sum_{m=1}^{N}X_m$. 求S的拉普拉斯变换$L_S(\theta)$.

4 X是连续型非负随机变量, $E(X^{2n}) < +\infty$. 对任意$k \le 2n$, 令$\mu_k = E(X^k)$. 证明对任意$\theta > 0$,

$$\sum_{k=0}^{2n-1} \frac{(-\theta)^k \mu_k}{k!} \le L_X(\theta) \le \sum_{k=0}^{2n} \frac{(-\theta)^k \mu_k}{k!}.$$

5 设独立同分布的标准正态分布随机变量簇$W_m, m \ge 1$与$\{N(n), n \ge 1\}$ 独立, 其中$N(n)$服从泊松分布$\mathrm{P}(n)$. 令$Y_n = \sum_{k=1}^{N(n)} W_k$. 求$Y_n$的特征函数, 并证明当$n \to +\infty$时, $Y_n/\sqrt{n}$依分布收敛到一个服从标准正态分布的随机变量.

6 记非负整数值随机变量ξ的概率母函数为$A(s)$. 证明

(1) 若$E(\xi) > 1$(包括$E(\xi) = +\infty$), 那么在区间$[0,1)$ 内$A(s) = s$有唯一解;

(2) 若$E(\xi) \le 1$且$P(\xi = 1) < 1$, 那么$A(s) = s$在$[0,1)$ 内无解.

7 若随机变量X的分布列为$P(X = n) = n/2^{n+1}, n \ge 1$, 求$X$的概率母函数$\phi(s)$.

8 已知概率母函数$\phi(s) = s + \frac{1}{1+\gamma}(1-s)^{1+\gamma}$, $\gamma \in (0,1]$, 求对应的概率分布列.

9 证明性质1.3.20.

1.4 收敛性与极限定理

在初等概率论中我们还简单地研究过随机变量序列的收敛问题. 主要介绍了三类收敛, 并得到了一些大数定律与中心极限定理. 为了后续方便, 这一节我们将对这些内容做一些展开和补充.

1.4.1 三类收敛性

在概率论中我们已经了解随机变量的依分布收敛, 依概率收敛和几乎必然收敛这三种收敛方式. 回顾它们的定义如下: 设$\{X_n; n \geq 1\}$为概率空间$(\Omega, \mathfrak{F}, P)$上一列随机变量.

(1) 依分布收敛: 对任意n, 记X_n的分布函数为$F_n(x)$, 称X_n依分布收敛到随机变量X(记作$X_n \xrightarrow{d} X$), 若对X分布函数$F(x)$的每个连续点x 都有

$$\lim_{n\to\infty} F_n(x) = F(x).$$

(2) 依概率收敛: 称X_n依概率收敛到随机变量X(记作$X_n \xrightarrow{P} X$), 若对任意$\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0,$$

即对任意$\varepsilon > 0$以及任意$\delta > 0$, 存在$N(\varepsilon, \delta)$, 当$n \geq N$时,

$$P(|X_n - X| > \varepsilon) \leq \delta.$$

(3) 几乎必然(处处)收敛: 称随机变量X_n几乎必然(几乎处处, 以概率1)收敛到随机变量X(记作$X_n \xrightarrow{a.s.} X$), 若

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

♠**注记1.4.1.** *依概率收敛和几乎必然收敛意义下的极限, 在几乎处处相等意义下是唯一的. 此外*

(1) $X_n \xrightarrow{d} X$当且仅当X_n的特征函数$\psi_n(t)$收敛到X 的特征函数$\psi(t)$. 若X_n是非负随机变量, 还可将特征函数用拉普拉斯变换代替.

(2) X_n几乎必然收敛到X还可以等价地表述成: 存在一个集合$A \subset \Omega$使得$P(A) = 0$且对任意$\omega \in \Omega \setminus A$, 数列$X_n(\omega) \to X(\omega)$. 用$\varepsilon - N$ 语言表示如下. $X_n \xrightarrow{a.s.} X$ 当且仅当存在一个集合$A \in \mathfrak{F}$ 满足$P(A) = 0$, 对任意$\omega \in \Omega \setminus A$ 以及任意$\varepsilon > 0$, 存在一个$N = N(\varepsilon, \omega)$ 使得当$n > N$时$|X_n(\omega) - X(\omega)| < \varepsilon$.

(3) 几乎必然收敛蕴含了依概率收敛, 依概率收敛蕴含了依分布收敛; 但是, 反过来都不成立.

下面这个例子主要帮助大家进一步认识几乎必然收敛概念.

▶**例1.4.2.** 给定三个随机变量序列$\{X_n; n\geq 1\}, \{Y_n; n\geq 1\}, \{Z_n; n\geq 1\}$. 已知对任意$n\geq 1$, $P(X_n\leq Y_n\leq Z_n)=1$而且当$n\to\infty$时, X_n和Z_n几乎必然收敛到同一个极限X. 证明$n\to\infty$时Y_n也几乎必然收敛到X.

证明 记$A=\{\omega; X_n(\omega)\not\to X(\omega)\}$, $B=\{\omega; Z_n(\omega)\not\to X(\omega)\}$以及

$$C_n=\{\omega; X_n(\omega)\leq Y_n(\omega)\leq Z_n(\omega)\text{不成立}\}.$$

那么由题设可知$P(A)=P(B)=P(C_n)=0$. 令

$$D=A\cup B\cup\Big(\bigcup_{n=1}^{\infty}C_n\Big).$$

那么$P(D)=0$, 而且对任意$\omega\in\Omega\setminus D$,

$$X_n(\omega)\to X(\omega),\ Z_n(\omega)\to X(\omega),\ X_n(\omega)\leq Y_n(\omega)\leq Z_n(\omega), n=1,2,\cdots$$

同时成立. 此时由数列收敛的夹逼准则可知$Y_n(\omega)\to X(\omega)$. 这表明$Y_n\stackrel{a.s.}{\to}X$. □

关于这三类收敛的更多例子, 性质和应用, 感兴趣的同学可以回顾概率论中有关内容. 比如: 利用这三种收敛, 初等概率论已经介绍了如下的大数定律和中心极限定理.

★**定理1.4.3.** (弱大数定律) 设$\{X_n; n\geq 1\}$为一列独立同分布的随机变量. 若$E(|X_n|)<\infty$, 那么当$n\to\infty$时$\frac{1}{n}\sum_{k=1}^n X_k$依概率收敛到$E(X_1)$.

★**定理1.4.4.** (中心极限定理) 设$\{X_n; n\geq 1\}$ 为一列独立同分布的随机变量, 二阶矩存在, 记其均值和标准差分别为μ和σ, 那么对任意$x\in(-\infty,+\infty)$,

$$P\Big(\frac{\sum_{i=1}^n(X_i-\mu)}{\sigma\sqrt{n}}\leq x\Big)\to\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u,$$

即$n\to\infty$时, $\frac{\sum_{i=1}^n(X_i-\mu)}{\sigma\sqrt{n}}$依分布收敛到一个服从标准正态分布的随机变量.

极限与收敛性是研究和了解随机过程性质的常用手段. 在本节剩余部分, 我们将对几乎必然收敛和依概率收敛做些必要的补充.

1.4.2 Borel-Cantelli 引理

▼**引理1.4.5.** 已知随机变量X以及一列随机变量$\{X_n; n\geq 1\}$, 令

$$A_{n,m}=\{|X_n-X|<\frac{1}{m}\},$$

那么

$$\{\lim_{n\to\infty}X_n=X\}=\bigcap_{m=1}^{\infty}\bigcup_{N=1}^{\infty}\bigcap_{n\geq N}A_{n,m}.$$

证明 对任意$\omega\in\{\lim_{n\to\infty}X_n=X\}$以及任意$m>0$, 存在$N(m,\omega)$, 当$n\geq N$时,

$$|X_n(\omega)-X(\omega)|<1/m,$$

因此
$$\omega \in \bigcap_{n\geq N(m,\omega)} \{|X_n - X| < \frac{1}{m}\} \subset \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \{|X_n - X| < \frac{1}{m}\}.$$
再由m的任意性可得
$$\omega \in \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \{|X_n - X| < \frac{1}{m}\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} A_{n,m}.$$
反之, 任取
$$\omega \in \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} A_{n,m},$$
对任意$m > 1$, 存在N, 当$n > N$时$\omega \in A_{n,m}$. 这表明
$$|X_n(\omega) - X(\omega)| \leq \frac{1}{m}.$$
所以$\lim\limits_{n\to\infty} X_n(\omega) = X(\omega)$, 即$\omega \in \{\lim\limits_{n\to\infty} X_n = X\}$. 因此引理结论成立. □

★**定理1.4.6.** 已知随机变量X以及一列随机变量$\{X_n; n \geq 1\}$. 下列结论等价:

(1) $X_n \overset{a.s.}{\to} X$.

(2) $P\big(\bigcup_{m=1}^{\infty}\bigcap_{N=1}^{\infty}\bigcup_{n\geq N}\{|X_n - X| \geq \frac{1}{m}\}\big) = 0$.

(3) 对任意$\varepsilon > 0$, $P\big(\bigcap_{N=1}^{\infty}\bigcup_{n\geq N}\{|X_n - X| \geq \varepsilon\}\big) = 0$.

(4) 对任意$\varepsilon > 0$, $\lim\limits_{N\to\infty} P\big(\bigcup_{n\geq N}\{|X_n - X| \geq \varepsilon\}\big) = 0$.

证明 (1) $\Leftrightarrow$ (2): 引理1.4.5的自然推论.

(2) $\Rightarrow$ (3): 显然成立.

(3) $\Rightarrow$ (2): 注意到对任意$m > 0$, 存在$\varepsilon > 0$使得$1/m > \varepsilon$, 从而由(3)可知
$$P(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|X_n - X| \geq \frac{1}{m}\}) = 0.$$
因此
$$P(\bigcup_{m=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|X_n - X| \geq \frac{1}{m}\}) \leq \sum_{m=1}^{\infty} P(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|X_n - X| \geq \frac{1}{m}\}) = 0.$$
即(2)成立.

(3) $\Leftrightarrow$ (4): 此时注意到
$$\bigcup_{n\geq N+1} \{|X_n - X| \geq \varepsilon\} \subset \bigcup_{n\geq N} \{|X_n - X| \geq \varepsilon\}.$$
由概率的连续性可得
$$P(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|X_n - X| \geq \varepsilon\}) = \lim_{N\to\infty} P(\bigcup_{n\geq N} \{|X_n - X| \geq \varepsilon\}) = 0,$$
从而(3)与(4)等价. □

♠**注记1.4.7.** 注意到若$\omega \in \bigcap_{n=1}^{\infty}\bigcup_{k\geq n} A_k$, 则意味着对任意$n > 1$, 至少存在

一个A_k, $k>n$, 使得$\omega\in A_k$, 反之也成立. 这意味着, 事件$\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k$发生当且仅当有无穷多个$A_k$事件发生. 即

$$\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k=\{无穷多个A_k事件发生\}\triangleq\{A_k,\mathrm{i.o.}\},$$

其中i.o是*infinitely often*的缩写.

★**定理1.4.8.** (*Borel-Cantelli*引理) 设$\{A_n;n\geq 1\}$是$(\Omega,\mathfrak{F},P)$中一列随机事件, $p_k=P(A_k)$,

1) 若$\sum_{k=1}^{\infty}p_k<\infty$, 那么$P(A_k,\ \mathrm{i.o.})=0$.

2) 若$\sum_{k=1}^{\infty}p_k=\infty$且$A_k$相互独立, 那么$P(A_k,\ \mathrm{i.o.})=1$.

证明 1)当$\sum_{k=1}^{\infty}p_k<\infty$时, 对任意$\varepsilon>0$, 存在正整数$N$, 使得

$$0\leq P(\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k)\leq P(\bigcup_{k\geq N}A_k)\leq\sum_{k\geq N}P(A_k)=\sum_{k\geq N}p_k<\varepsilon.$$

令$\varepsilon\to 0$, 可得$P(A_k,\ \mathrm{i.o.})=0$.

2) 由不等式$1-x\leq e^{-x}$及随机事件独立性可得, 对任意$n\geq 1$,

$$\begin{aligned}0\leq P(\bigcap_{k\geq n}\bar{A}_k)&=\prod_{k\geq n}P(\bar{A}_k)=\prod_{k\geq n}(1-P(A_k))\\&\leq\prod_{k\geq n}e^{-p_k}=\mathrm{e}^{-\sum_{k\geq n}p_k}=0.\end{aligned}$$

因此$P(\bigcup_{n\geq 1}\bigcap_{k\geq n}\bar{A}_k)=0$, 即$P(A_k,\mathrm{i.o})=1$. □

定理结论(1)告诉我们, 当(1)中条件成立时,

$$P(A_k,\ \mathrm{i.o.})=P\Big(\bigcap_{n=1}^{\infty}\bigcup_{k\geq n}A_k\Big)=0.$$

那么

$$P\Big(\bigcup_{n=1}^{\infty}\bigcap_{k\geq n}\bar{A}_k\Big)=1.$$

这表明几乎所有的基本事件ω都落在集合$\cup_{n=1}^{\infty}\cap_{k\geq n}\bar{A}_k$内. 从而存在一个$N(\omega)$, 使得$\omega\in\cap_{k\geq N}\bar{A}_k$, 也就是说当$k>N$后, 事件$A_k$不再发生. 因此该定理的第一个结论告诉我们, 若条件成立, 那么对几乎所有的ω, 事件A_k只发生有限次.

利用定理1.4.6和定理1.4.8我们还可以得到如下判断几乎必然收敛的充分条件. 我们用一个注记的形式给出.

♠**注记1.4.9.** 若对任意$\varepsilon>0$,

$$\sum_{n=1}^{\infty}P(|X_n-X|>\varepsilon)<\infty,\tag{1.4.1}$$

那么$X_n\stackrel{a.s.}{\to}X$.

▲**命题1.4.10.** 若$X_n \xrightarrow{P} X$, 那么存在$\{X_n\}$的一个子列$\{X_{n_k}\}$使得

$$X_{n_k} \xrightarrow{a.s.} X.$$

证明 因为$X_n \xrightarrow{P} X$, 对任意$k \geq 1$, $P(|X_n - X| > 1/2^k) \to 0$. 因此, 存在$n_k$使得

$$P(|X_{n_k} - X| > 1/2^k) < 1/2^k,$$

而且n_k 单调不降. 由此, 对任意$\varepsilon > 0$, 令M为正常数使得$\varepsilon > 1/2^M$,

$$\begin{aligned}\sum_{k=1}^{\infty} P(|X_{n_k} - X| > \varepsilon) &\leq \sum_{k=1}^{M} P(|X_{n_k} - X| > \varepsilon) + \sum_{k=M+1}^{\infty} P(|X_{n_k} - X| > 1/2^k) \\ &\leq M + \sum_{k=M+1}^{\infty} \frac{1}{2^k} \leq M + 1 < \infty,\end{aligned}$$

即(1.4.1)对序列$\{X_{n_k}\}$成立, 从而由注记1.4.9可知本命题成立. □

▶**例1.4.11.** 设$\{X_k; k \geq 1\}$为一列独立同分布的随机变量, 且X_k服从二项分布$b(1, p)$, 试用*Borel-Cantelli*引理证明

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = p, \quad \text{a.s.} \tag{1.4.2}$$

证明 对任意$\varepsilon > 0$, 令

$$A_k = \Big\{ \Big|\frac{1}{k} \sum_{i=1}^{k} X_i - p\Big| > \varepsilon \Big\}.$$

那么

$$P(A_k) = P\Big(\Big|\sum_{i=1}^{k} (X_i - p)\Big| > k\varepsilon \Big) \leq \frac{1}{k^4 \varepsilon^4} E\Big(\big(\sum_{i=1}^{k} (X_i - p)\big)^4 \Big).$$

直接计算可知

$$E\Big(\big(\sum_{i=1}^{k} (X_i - p)\big)^4 \Big) = kpq(p^3 + q^3) + 3k(k-1)p^2 q^2. \tag{1.4.3}$$

因此$\sum\limits_{k=1}^{\infty} P(A_k) < \infty$. 由Borel-Cantelli引理可知(1.4.2)成立. □

事实上例1.4.11的结论对一阶矩存在的独立同分布的随机变量序列都成立. 我们不加证明地给出如下的所谓强大数定律.

★**定理1.4.12.** (强大数定律) 设$\{X_n; n \geq 1\}$为一列独立同分布的随机变量. 若$E(|X_n|) < \infty$, 那么

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = E(X_1), \quad a.s.$$

1.4.3 柯西基本列*

♣**定义1.4.13.** 称随机变量序列$\{X_n; n \geq 1\}$是几乎必然意义下的基本列或柯西列, 若存在一个零概率集A, 使得对任意$\omega \in \Omega \setminus A$ 以及任意$\varepsilon > 0$, 存在一个$N = N(\varepsilon, \omega)$, 当$n, m > N$时, $|X_n(\omega) - X_m(\omega)| < \varepsilon$.

由于数列收敛当且仅当它是基本列, 我们可得如下结论.

▲**命题1.4.14.** $\{X_n; n \geq 1\}$几乎必然收敛当且仅当$\{X_n; n \geq 1\}$是几乎必然意义下的基本列.

与定理1.4.6类似, 我们有如下结果. 证明请读者自己补充.

★**定理1.4.15.** 设$\{X_n; n \geq 1\}$为一列随机变量, 下列结论等价:

(1) $\{X_n; n \geq 1\}$ 是几乎必然意义下的基本列.

(2) 对任意$\varepsilon > 0$,

$$P\Big(\bigcap_{n=1}^{\infty} \bigcup_{v=1}^{\infty} \{|X_{n+v} - X_n| \geq \varepsilon\}\Big) = 0.$$

(3) 对任意$\varepsilon > 0$, 当$n \to \infty$时,

$$P\Big(\bigcup_{v=1}^{\infty} \{|X_{n+v} - X_n| \geq \varepsilon\}\Big) \to 0.$$

由定理1.4.15, 我们易得下面的推论.

♦**推论1.4.16.** $\{X_n; n \geq 1\}$为几乎必然意义下的基本列, 若存在非负数列$\{\delta_n; n \geq 1\}$ 使得

$$\sum_{n=1}^{\infty} \delta_n < \infty \text{ 而且 } \sum_{n=1}^{\infty} P(|X_{n+1} - X_n| \geq \delta_n) < \infty.$$

证明 注意到对任意$\varepsilon > 0$, 存在k 使得$\sum\limits_{n=k}^{\infty} \delta_n < \varepsilon$, 进而对任意$m > k$,

$$\begin{aligned}\bigcup_{v=1}^{\infty} \{|X_{m+v} - X_m| \geq \varepsilon\} &\subset \bigcup_{v=1}^{\infty} \{|X_{m+v} - X_m| \geq \sum_{n=m}^{m+v-1} \delta_n\} \\ &\subset \bigcup_{v=1}^{\infty} \bigcup_{n=m}^{m+v-1} \{|X_{n+1} - X_n| \geq \delta_n\} = \bigcup_{n=m}^{\infty} \{|X_{n+1} - X_n| \geq \delta_n\}.\end{aligned}$$

因此由假设条件以及定理1.4.15可知本推论成立. □

♣**定义1.4.17.** 称随机变量序列$\{X_n; n \geq 1\}$是依概率收敛意义下的基本列或柯西列, 若对任意$\varepsilon > 0$, $\delta > 0$ 存在一个$N = N(\varepsilon, \delta)$ 使得当$n, m > N$ 时,

$$P(|X_n(\omega) - X_m(\omega)| \geq \varepsilon) \leq \delta.$$

由关系式$\{|X_n-X_m|\ge\varepsilon\}\subset\{|X_n-X|\ge\varepsilon/2\}\cup\{|X_m-X|\ge\varepsilon/2\}$ 可知依概率收敛的序列一定是依概率收敛的基本列, 进而几乎必然收敛的基本列一定是依概率收敛的基本列.

▲**命题1.4.18.** 依概率收敛的基本列存在子列使其为几乎必然收敛的基本列.

证明 设$\{X_n;n\ge 1\}$是任一依概率收敛的基本列. 取$\varepsilon=1/2^k$, $\delta=1/2^k$. 存在单调增加的n_k, 使得对任意$m>n_k$,

$$P(|X_m(\omega)-X_{n_k}(\omega)|\ge 1/2^k)\le 1/2^k.$$

取子列$\{X_{n_k};k\ge 1\}$, 那么

$$\sum_{k=1}^{\infty}P(|X_{n_{k+1}}-X_{n_k}|\ge 1/2^k)<\infty.$$

由推论1.4.16可知$\{X_{n_k};k\ge 1\}$就是几乎必然意义下的基本列. □

记依概率收敛意义下的基本列$\{X_n;n\ge 1\}$的极限为X, 由

$$\{|X_n-X|\ge\varepsilon\}\subset\{|X_{n_k}-X|\ge\varepsilon/2\}\cup\{|X_n-X_{n_k}|\ge\varepsilon/2\},$$

可知$\{X_n;n\ge 1\}$依概率收敛. 因此$\{X_n;n\ge 1\}$依概率收敛当且仅当它是依概率收敛的基本列.

1.4.4 收敛的简单性质

■**性质1.4.19.** 对依概率收敛和几乎必然收敛, 以下运算性质成立.

(1) 若对任意k,$X_n^{(k)}\xrightarrow{a.s.}X^{(k)}$ (或$X_n^{(k)}\xrightarrow{P}X^{(k)}$), c_k为常数, 那么

$$\sum_{k=1}^{m}c_kX_n^{(k)}\xrightarrow{a.s.}\sum_{k=1}^{m}c_kX^{(k)}\ \left(\text{相应地}\sum_{k=1}^{m}c_kX_n^{(k)}\xrightarrow{P}\sum_{k=1}^{m}c_kX^{(k)}\right).$$

(2) 若$X_n\xrightarrow{a.s.}X$(或$X_n\xrightarrow{P}X$), $Y_n\xrightarrow{a.s.}Y$(相应地$Y_n\xrightarrow{P}Y$), 那么

$$X_nY_n\xrightarrow{a.s.}XY\ (\text{相应地}X_nY_n\xrightarrow{P}XY).$$

(3) 若$Y_n\xrightarrow{a.s.}Y$(或$Y_n\xrightarrow{P}Y$)且$P(Y=0)=0$, 那么

$$1/Y_n\xrightarrow{a.s.}1/Y\ \ (\text{相应地}1/Y_n\xrightarrow{P}1/Y),$$

其中当$Y_n=0$时$1/Y_n$可以恰当地补充定义.

证明 我们只证(2), 其他作为习题请读者自己完成(参见本节习题8).

若$X_n\xrightarrow{a.s.}X$ 且$Y_n\xrightarrow{a.s.}Y$, 那么存在零概率集A,B使得对任意$\omega\in\Omega\setminus A$和任意$\omega_1\in\Omega\setminus B$有

$$X_n(\omega)\to X(\omega),\quad Y_n(\omega_1)\to Y(\omega_1).$$

令$C=A\cup B$, 那么C也是零概率集, 而且对任意$\omega\in\Omega\setminus C$,

$$X_n(\omega)Y_n(\omega)\to X(\omega)Y(\omega).$$

这表明$X_nY_n \xrightarrow{a.s.} XY$.

若$X_n \xrightarrow{P} X$ 且$Y_n \xrightarrow{P} Y$. 对任意$\delta > 0$, 存在$M > 0$ 使得

$$P(|Y| \geq M) < \delta/4, \quad P(|X| \geq M-1) < \delta/8.$$

由于$X_n \xrightarrow{P} X$, 因此存在N_1, 当$n > N_1$时, $P(|X_n - X| > 1) < \delta/8$, 进而

$$\begin{aligned} P(|X_n| < M) &\geq P(|X_n - X| \leq 1, |X| < M-1) \\ &\geq 1 - P(|X_n - X| > 1) - P(|X| \geq M-1) = 1 - \delta/4. \end{aligned}$$

由此可得$P(|X_n| \geq M) \leq \delta/4$. 进一步地, 由依概率收敛条件, 对任意$\varepsilon > 0$, 存在$N > N_1$, 使得当$n \geq N$时,

$$P(|Y_n - Y| \geq \varepsilon/(2M)) \leq \delta/4, \quad P(|X_n - X| \geq \varepsilon/(2M)) \leq \delta/4.$$

因此当$n > N$时,

$$\begin{aligned} P(|X_nY_n - XY| \geq \varepsilon) &= P(|X_nY_n - X_nY + X_nY - XY| \geq \varepsilon) \\ &\leq P(|X_nY_n - X_nY| + |X_nY - XY| \geq \varepsilon) \\ &\leq P(|X_n||Y_n - Y| \geq \varepsilon/2) + P(|Y||X_n - X| \geq \varepsilon/2) \\ &\leq P(|X_n| \geq M) + P(|Y_n - Y| \geq \varepsilon/2M) \\ &\qquad + P(|Y| \geq M) + P(|X_n - X| \geq \varepsilon/2M) \\ &\leq \delta/4 + \delta/4 + \delta/4 + \delta/4 = \delta. \end{aligned}$$

这表明$X_nY_n \xrightarrow{P} XY$. □

♠**注记1.4.20.** 上面我们介绍的几乎必然收敛和依概率收敛隐含地假定了极限X几乎必然有限. 在几乎必然收敛意义下, 利用微积分中数列收敛到无穷的定义, 我们也可以考虑极限X以正概率取到正无穷和/或负无穷的情形, 比如当收敛到正无穷时,

$$\{X_n(\omega) \to +\infty\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{X_n \geq m\}, \tag{1.4.4}$$

而$X_n \to +\infty$ $a.s.$ 则意味着除去一个零概率的例外集A, 对任意$\omega \in \Omega \setminus A$都有$X_n(\omega) \to +\infty$.

对于依分布收敛而言, 类似于性质1.4.19 的运算性质一般不成立, 但我们不加证明地给出如下的斯卢茨基(Slutsky)引理.

▼**引理1.4.21.** 若$X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$, 其中c为常数. 那么

$$X_n + Y_n \xrightarrow{d} X + c, \qquad X_nY_n \xrightarrow{d} cX.$$

为了后面引用方便, 在本节最后, 我们介绍一个与随机变量数学期望有关的极限定理. 为此, 我们先给出数学期望的全期望公式.

▼**引理1.4.22.** 设事件集$\{A_i; i\geq 1\}$两两不交且$\cup_{i=1}^{\infty}A_i=\Omega$, 即$\{A_i; i\geq 1\}$是样本空间$\Omega$的一个分割. 若$X$是非负随机变量或$X$的数学期望存在, 那么

$$E(X)=\sum_{i=1}^{\infty}E(X1_{A_i}). \qquad \text{(全期望公式)}$$

证明 只证X是非负连续的情形, 对其他情形感兴趣的读者可参阅[9]. 此时,

$$E(X)=\int_0^{\infty}xf_X(x)\mathrm{d}x=\lim_{M\to\infty}\int_0^{M}xf_X(x)\mathrm{d}x=\lim_{M\to\infty}E(X1_{\{X\leq M\}}).$$

由于$\{A_i; i\geq 1\}$是样本空间Ω的一个分割, 对任意$n\geq 1$,

$$1_{\Omega}=\sum_{i=1}^{n}1_{A_i}+1_{\cup_{i=n+1}^{\infty}A_i}.$$

因此

$$\begin{aligned}E(X1_{\{X\leq M\}})&=E\Big(X1_{\{X\leq M\}}\Big(\sum_{i=1}^{n}1_{A_i}+1_{\cup_{i=n+1}^{\infty}A_i}\Big)\Big)\\&=\sum_{i=1}^{n}E(X1_{\{X\leq M\}}1_{A_i})+E(X1_{\{X\leq M\}}1_{\cup_{i=n+1}^{\infty}A_i}).\end{aligned}$$

再由$1=\sum_{i=1}^{\infty}P(A_i)$可知$n\to\infty$时,

$$E(X1_{\{X\leq M\}}1_{\cup_{i=n+1}^{\infty}A_i})\leq MP(\cup_{i=n+1}^{\infty}A_i)\to 0.$$

所以

$$\begin{aligned}E(X1_{\{X\leq M\}})&=\lim_{n\to\infty}\sum_{i=1}^{n}E(X1_{\{X\leq M\}}1_{A_i})\\&=\sum_{i=1}^{\infty}E(X1_{\{X\leq M\}}1_{A_i})\leq\sum_{i=1}^{\infty}E(X1_{A_i}).\end{aligned}$$

由此可得

$$E(X)\leq\sum_{i=1}^{\infty}E(X1_{A_i}).$$

另一方面, 对任意$n\geq 1$,

$$E(X)\geq E(X1_{\cup_{i=1}^{n}A_i})=\sum_{i=1}^{n}E(X1_{A_i}).$$

令$n\to\infty$得

$$E(X)\geq\sum_{i=1}^{\infty}E(X1_{A_i}).$$

综合两方面, 可知本引理成立. □

★**定理1.4.23.** (单调收敛定理) 设X_n为非负随机变量且$X_n\leq X_{n+1}$对任意$n\geq 1$几乎处处成立. 记$X=\lim\limits_{n\to\infty}X_n$, *a.s.* 那么

$$E(\lim_{n\to\infty}X_n)=E(X)=\lim_{n\to\infty}E(X_n).$$

证明 由条件可知, 存在一个概率为0的事件A, 使得在$\Omega\backslash A$上, $\{X_n,\ n\geq 1\}$为单调不降数列, 从而极限存在(可能为∞), 记其为X. 因此$X_n \stackrel{a.s.}{\to} X$.

(1) 若$P(X=\infty)=p>0$, 由(1.4.4)可知对任意$M>0$,

$$\begin{aligned}p = P(X_n\to+\infty) &= P(\bigcap_{m\geq 1}\bigcup_{N\geq 1}\bigcap_{n>N}\{X_n>m\})\\ &\leq P(\bigcup_{N\geq 1}\bigcap_{n>N}\{X_n>M\}) = \lim_{N\to\infty}P(\bigcap_{n>N}\{X_n>M\}).\end{aligned}$$

这表明存在$m(M)$, 当$n>m$ 时,

$$P(X_n>M)>p/2,$$

从而$E(X_n)>Mp/2$. 先令$n\to\infty$, 再令$M\to\infty$, 可得

$$\lim_{n\to\infty}E(X_n)=\infty=E(X).$$

(2) 若$P(X=\infty)=0$. 此时由全期望公式可得

$$E(X_n)=\sum_{k=1}^{\infty}E(X_n 1_{\{k-1\leq X<k\}}),\quad E(X)=\sum_{k=1}^{\infty}E(X1_{\{k-1\leq X<k\}}).$$

对任意$n,k\geq 1$, 记

$$a_n(k)=E(X_n1_{\{k-1\leq X<k\}}),\ \ a(k)=E(X1_{\{k-1\leq X<k\}}).$$

显然$a_n(k)$, $a(k)$非负, 而且$a_n(k)$随n 增加单调不降. 注意到

$$\begin{aligned}|a_n(k)-a(k)| &\leq E(|X_n-X|1_{\{k-1\leq X<k\}})\\ &= E(|X_n-X|1_{\{k-1\leq X<k\}\cap\{|X_n-X|>\varepsilon\}})\\ &\quad +E(|X_n-X|1_{\{k-1\leq X<k\}\cap\{|X_n-X|\leq\varepsilon\}})\\ &\leq kP(|X_n-X|>\varepsilon)+\varepsilon.\end{aligned}$$

由于$X_n\to X, a.s.$ 因此随$n\to\infty$, $kP(|X_n-X|>\varepsilon)\to 0$. 再由$\varepsilon$的任意小性可知, $n\to\infty$时$a_n(k)\to a(k)$. 由数列级数收敛知识(见本章第一节习题5)可得

$$E(X)=\sum_{k=1}^{\infty}a(k)=\lim_{n\to\infty}\sum_{k=1}^{\infty}a_n(k)=\lim_{n\to\infty}E(X_n).$$ □

练习题1.4

1 若存在随机变量X使得随机序列$S_n/n\to X$, a.s. 成立, 而且非负整数值随机序列$N_n\to\infty$, a.s. 成立. 证明$S_{N_n}/N_n\to X,\ \ a.s.$

2 证明: 若对任意$\varepsilon>0$, 存在N使得当$n>N$时$P(|X_n-X|<\varepsilon)=1$, 那么$X_n\stackrel{a.s.}{\to}X$. 举例说明逆命题不成立.

3 试构造两个随机变量序列$\{X_n;n\geq 1\}$,$\{Y_n;n\geq 1\}$使得$X_n\stackrel{d}{\to}X$, $Y_n\stackrel{P}{\to}Y$但$X_n+Y_n\stackrel{d}{\not\to}X+Y$.

4 设$f(x)$在$[0,+\infty)$上连续有界, 单调上升且$f(0)=0$, 证明随机变量X_n依概率收敛于0当且仅当$\lim\limits_{n\to\infty} E(f(|X_n|))=0$.

5 设$\{X_k;k\geq 1\}$为一列相互独立随机变量, 而且对任意$k\geq 1$, $E(X_k)=0$, $Var(X_k)=k$. 对任意$n\geq 1$, 令

$$Y_n=\sum_{k=1}^{n} X_k/k.$$

证明对任意$r>1/2$, 当$n\to\infty$时Y_n/n^r几乎必然收敛到0.

6 对任意给定的随机变量序列$\{X_n;n\geq 1\}$, 令$Y_n=\sup\limits_{k\geq 1}|X_{n+k}-X_n|$. 证明$X_n$几乎必然收敛当且仅当对任意$\varepsilon>0$, $\lim\limits_{n\to\infty} P(Y_n>\varepsilon)\to 0$.

7 验证(1.4.3)成立.

8* 试给出性质1.4.19中(1)和(3)的证明.

9* 试给出定理1.4.15的证明.

第二章　简单随机模型

本章我们将简要地引入随机过程的基本概念并介绍两类简单随机模型: 直线上简单随机游动与泊松过程. 这两类随机过程在一定意义上可以看作是独立同分布随机变量部分和序列以及部分和序列的“逆”, 在应用领域被广泛使用.

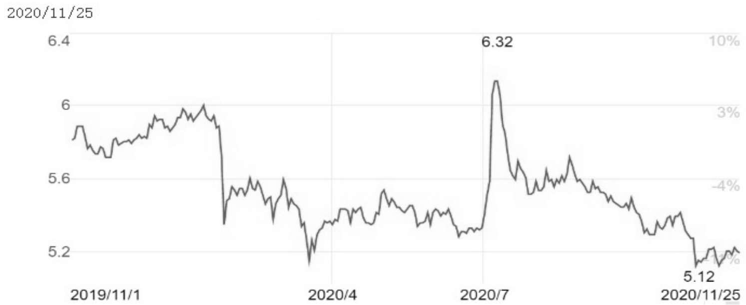
2.1　随机过程简介

直观而言, 随机过程是人们为认识复杂随机现象动态变化规律而建立的一种数学模型, 通常以一族(通常是无穷多的)随机变量的形式出现.

2.1.1　随机过程的基本概念

先看下面这个例子.

▶**例2.1.1.** 观察一只股票的价格波动, 是否存在“稳赢”的投资策略? 一般地, 股票价格变化具有一定的随机性(如下图所示).



用随机变量$X(t)$ 表示股票在t 时刻的价格, 如果想通过买卖该股票获利, 那么我们应该掌握随机变量族$\{X(t);\ t\geq 0\}$的统计规律. 当我们想通过某种策略获利, 比如5元买入6元卖出, 我们首先要解决的问题就是对应的买入时机和卖出时机是否存在? 在什么时候?

以上例子表明, 许多问题会涉及动态变化的随机现象. 为了解决问题, 我们可以用一族(一般是无穷多)随机变量刻画(描述)这些现象. 通常我们就称这族随机变量是一个随机过程. 介绍随机过程的各种统计规律和研究方法就是随机过程课程的主要内容.

♣**定义2.1.2.** *称$X = \{X(t);\ t \in T\}$为概率空间$(\Omega, \mathfrak{F}, P)$ 上的随机过程, 若它是该概率空间上一族随机变量(可以是向量值的). 称$X(t)$, $t \in T$, 的所有可能取值域S为随机过程X的状态空间, 称T为指标集或参数集, 特别地, 若指标集具有时间属性时, 也称之为时间参数.*

为表示方便, 我们有时也把随机过程$X = \{X(t);\ t \in T\}$记成$X = (X_t)_{t\in T}$或$X = \{X_t; t \in T\}$, 简记成$(X_t)_{t\in T}$或$\{X_t; t \in T\}$. 无需强调参数集时还可简记成(X_t)或$\{X_t\}$. 特别地, 无需强调随机过程中的变量时可简写为随机过程X.

从映射(函数)角度看, 随机过程也是$T \times \Omega \mapsto S$的二元映射. 当$t$和$\omega$都给定时, $X(t,\omega)$是一个实数或向量; 当只给定t时, $X(t,\cdot)$就是随机变量; 当只给定$\omega \in \Omega$时, $X(\cdot,\omega)$就是一个指标集T 上的函数:

$$X(\cdot,\omega) : T \to S \ \ 对任意t \in T, \ \ X(\cdot,\omega)(t) = X(t,\omega).$$

通常我们称函数$X(\cdot,\omega)$ 为随机过程X的轨道(函数)或样本(函数).

按照指标集和状态空间的不同, 通常我们把随机过程分成四类: 连续参数离散状态、连续参数连续状态、离散参数离散状态、离散参数连续状态随机过程.

♠**注记2.1.3.** *对随机过程还可从其他角度进行分类. 比如, 若对所有$t \in T$, $X(t) \in \mathbb{R}^d$, 其中$d > 1$, 那么我们称X为d-维随机过程或d-维向量值随机过程; 若$T \subset \mathbb{R}^n$, 其中$n > 1$(此时T往往对应空间或时空指标), 我们又称X 为随机场. 若无特别说明, 本书此后只讨论实值随机过程, 即S, T 都是$\mathbb{R}$或$\mathbb{R}$的子集, 比如自然数集$\mathbb{N}$, 整数集$\mathbb{Z}$, 正整数集$\mathbb{Z}_+$, 非负实数集$\bar{\mathbb{R}}_+$, 正实数集$\mathbb{R}_+$或者某个实数区间等等.*

2.1.2 随机过程的刻画

对有限个随机变量$X_1, \cdots, X_m$, 我们一般用联合分布函数

$$F(x_1, x_2, \cdots, x_m) = P(X_1 \le x_1, X_2 \le x_2, \cdots, X_m \le x_m),$$

刻画它们的统计规律. 对于随机过程, 我们可以借鉴这种思想.

♣**定义2.1.4.** *任取$X = \{X_t; t \in T\}$的k个参数, $t_1, t_2, \cdots, t_k$, 称$X_{t_1}, \cdots, X_{t_k}$的联合分布函数*

$$F_{t_1,t_2,\cdots,t_k}(x_1, x_2, \cdots, x_k) = P(X_{t_1} \le x_1, X_{t_2} \le x_2, \cdots, X_{t_k} \le x_k)$$

为随机过程X的一个k阶有限维分布, 称函数族

$$\{F_{t_1,t_2,\cdots,t_k}(x_1,x_2,\cdots,x_k);\ t_1,\cdots,t_k\in T, k\geq 1\}$$

为随机过程X的有限维分布族.

随机过程X在$t\in T$的一阶有限维分布, 即X_t的分布, 有时也被称为X在t的边际(边缘)分布. 若指标集T含最小元素, 记作t_0, 则人们常称X_{t_0}的分布为X的初始分布, 称X_{t_0}的值为X的初值.

♠注记2.1.5. 当S只有至多可数个元素时, 有限维分布族可用如下有限维分布列代替,

$$P_{t_1,\cdots,t_k}(x_1,\cdots,x_k)=P(X_{t_1}=x_1,\cdots,X_{t_k}=x_k),\ \ x_1,\cdots,x_k\in S.$$

■性质2.1.6. 随机过程的有限维分布族满足

(1) 对称性, 即对任意k以及$(1,2,\cdots,k)$的一个置换$(i_1,i_2,\cdots,i_k)$,

$$F_{t_1,t_2,\cdots,t_k}(x_1,x_2,\cdots,x_k)=F_{t_{i_1},t_{i_2},\cdots,t_{i_k}}(x_{i_1},x_{i_2},\cdots,x_{i_k}).$$

(2) 相容性, 即对任意$n>m$,

$$F_{t_1,\cdots,t_m,\cdots,t_n}(x_1,\cdots,x_m,+\infty,\cdots,+\infty)=F_{t_1,\cdots,t_m}(x_1,\cdots,x_m).$$

证明 性质由有限维分布定义直接可得. □

下面的柯尔莫哥洛夫存在性定理(Kolmogorov Existence Theorem)表明任何给定的具有相容性的有限维分布族都存在随机过程与之对应. 证明省略.

★定理2.1.7. 设$\{F_{t_1,t_2,\cdots,t_k}(x_1,x_2,\cdots,x_k);\ t_1,\cdots,t_k\in T,k\geq 1\}$是满足对称性和相容性的概率分布函数族, 则存在概率空间$(\Omega,\mathfrak{F},P)$及其上的一个随机过程$X=\{X(t);t\in T\}$使得X的有限维分布族恰好是给定的概率分布函数族.

直观地看, 若能就某个随机现象给出分布函数或分布函数族, 我们就在一定意义上刻画了该随机现象的统计规律. 柯尔莫哥洛夫存在性定理为人们提供了在保证不改变统计规律的前提下用随机过程模型研究随机现象的理论支撑.

♣定义2.1.8. 设$X=\{X_t;t\in T\}$, $Y=\{Y_t;t\in T\}$都是概率空间$(\Omega,\mathfrak{F},P)$上的随机过程, 称Y是X 的一个版本(*Version*), 若X,Y的有限维分布族相同; 称Y是X的一个修正(*Modification*), 若对任意$t\in T$, $P(X_t=Y_t)=1$; 称X与Y不可区分的(*Indistinguishable*), 若存在一个集合$A\in\mathfrak{F}$, 使得$P(A)=1$而且对任意$\omega\in A$, $X_t(\omega)=Y_t(\omega)$对所有$t\in T$ 都成立.

由定义容易看出, 在上述三个衡量两个过程关系的概念中“不可区分”要求最苛刻. 两个随机过程是不可区分的, 那么他们一定互为修正; 两个过程互为修正

则一定互为版本. 一般而言, 这两个结论的逆都不真. 但若X,Y 为离散参数随机过程, 那么"不可区分"与" 修正"等价(参见本节习题1).

♣**定义2.1.9.** 设$X=\{X_t;t\in T\}$, $Y=\{Y_u;u\in S\}$都是概率空间$(\Omega,\mathfrak{F},P)$上的随机过程, 若对任意$k,m\geq 1$, $t_1,\cdots,t_k\in T$, $u_1,\cdots,u_m\in S$,

$$(X_{t_1},\cdots,X_{t_k}) \text{ 与 } (Y_{u_1},\cdots,Y_{u_m})$$

独立, 则称X和Y是独立的.

►**例2.1.10.** 设$\{W_i;\ i\geq 1\}$为一列独立随机变量, $\{S_n;n\geq 1\}$为$\{W_i;i\geq 1\}$的部分和序列, N为一个固定常数. 对任意$k\geq 1$, 令$T_k=S_{N+k}-S_N$. 那么随机过程$\{T_k;k\geq 1\}$与$\{S_n;1\leq n\leq N\}$独立.

事实上, 对任意$m\geq 1$, 任取$k_1,\cdots,k_m\geq 1$, 由W_i, $i\geq 1$独立同分布可知

$$(T_{k_1},T_{k_2},\cdots,T_{k_m})=\Big(\sum_{i=N+1}^{N+k_1}W_i,\sum_{i=N+1}^{N+k_2}W_i,\cdots,\sum_{i=N+1}^{N+k_m}W_i\Big)$$

与

$$(S_1,S_2,\cdots,S_N)=(W_1,W_1+W_2,\cdots,\sum_{i=1}^{N}W_i)$$

是独立的. □

我们也可以通过随机变量的数值特征在一定程度上了解随机过程.

♣**定义2.1.11.** 设$X=\{X(t);t\in T\}$是定义在概率空间$(\Omega,\mathfrak{F},P)$上的随机过程, 分别称如下的函数

$$m_X(t)=E(X(t)),\qquad D_X(t)=Var(X(t)),$$
$$R_X(s,t)=E(X(s)X(t)),\qquad \mathrm{Cov}_X(s,t)=Cov(X(s),X(t)),$$

为随机过程X的均值函数, 方差函数, 自相关函数与协方差函数.

2.1.3 典型随机过程

♣**定义2.1.12.** 称随机过程$X=\{X(t);t\in T\}$是独立增量过程, 若对任意n以及$t_1<\cdots<t_n\in T$,

$$X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1})$$

相互独立(若指标集T有最小元素t_0, 那么还要求

$$X(t_0),X(t_1)-X(t_0),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1})$$

相互独立). 进一步, 若对任意$t,t+h,s,s+h\in T$,

$$X(t+h)-X(t)\ \text{ 与 }\ X(s+h)-X(s)$$

的分布相同, 则称X为平稳独立增量过程.

如果$X=\{X(t);t\in\bar{\mathbb{R}}_+\}$为平稳独立增量过程, 均值函数$m(t)$与方差函数$D(t)$都存在且连续. 那么对任意$t\geq 0$,

$$m_X(t)=Ct+E(X(0)),\quad D_X(t)=C_1t+Var(X(0)),\tag{2.1.1}$$

其中$C=E(X(1))-E(X(0))$, $C_1=Var(X(1)-X(0))$.

当$T=\mathbb{N}$时独立增量过程就是一列相互独立随机变量的部分和, 而初值为0的平稳独立增量过程就是一列独立同分布随机变量的部分和.

平稳独立增量过程有很多, 常见的例子包括本章后面要介绍的随机游动, 泊松过程以及第六章介绍的布朗运动.

数学期望有限的独立增量过程还具有如下两个重要性质: (为了方便理解, 假设$T=\mathbb{N}$)

(1) 对任意$s,t\in T$且$s<t$, 以及对任意$x\in\mathbb{R}$,

$$P(X(t)<x|X(0),\cdots,X(s))=P(X(t)<x|X(s)).$$

(2) 若对任意$s,t\in T$, $E(X(t)-X(s))=0$, 那么对任意$s<t\in T$,

$$E(X(t)|X(0),\cdots,X(s))=X(s).$$

下面我们检验性质(1),对性质(2)的检验留请读者自己完成.

首先注意到

$$P(X(t)<x|X(0),\cdots,X(s))=P(X(t)-X(s)<x-X(s)|X(0),\cdots,X(s)).$$

记$Y=X(t)-X(s)$, 由于X是独立增量过程, 因此Y与

$$X(0),X(1)-X(0),\cdots,X(s)-X(s-1)$$

独立, 进而与$X(0),\cdots,X(s)$独立. 注意上面条件概率中$X(0),\cdots,X(s)$是已知条件, 记$z=X(s)$, 那么

$$\begin{aligned}P(X(t)<x|X(0),\cdots,X(s))&=P(Y<x-z|X(0),\cdots,X(s-1),X(s)=z)\\&=P(Y<x-z)|_{z=X(s)}.\end{aligned}$$

另一方面, 类似推理可知

$$\begin{aligned}P(X(t)<x|X(s))&=P(X(t)-X(s)<x-X(s)|X(s))\\&=P(Y<x-X(s)|X(s))=P(Y<x-z)|_{z=X(s)}.\end{aligned}$$

综合这两方面可得性质(1).

通常称满足性质(1)的随机过程为马氏过程, 满足性质(2)的过程为鞅(过程). 我们将分别在第三,四章和第七章进一步介绍这两类过程.

除了独立增量过程, 马氏过程和鞅外, 在随机过程领域讨论比较多的还有所谓的严平稳过程和平稳过程. 我们以指标集$T=\bar{\mathbb{R}}_+$或$\mathbb{N}$为例给出定义如下.

♣**定义2.1.13.** 若对任意的$t_1<\cdots<t_n\in T$以及$h\in T$,
$$(X(t_1),\cdots,X(t_n)) \text{ 与 } (X(t_1+h),\cdots,X(t_n+h))$$
同分布, 则称对随机过程$X=\{X(t);t\in T\}$为严平稳过程.

严平稳过程的一个具体例子就是独立同分布随机变量序列. 在应用时严平稳过程要求较苛刻, 人们更多考虑如下的所谓平稳过程.

♣**定义2.1.14.** 称随机过程X为平稳过程, 若对任意$t\in T$, $E(X(t))$ 与t无关, 且对任意$t,h\in T$, 协方差$Cov(X(t),X(t+h))$与t无关.

显然若X是平稳过程, 那么对任意$t\in T$, $E(X^2(t))<\infty$, 即$X(t)$的二阶矩存在(简称这样的过程为二阶矩过程); 二阶矩存在的严平稳过程一定是平稳过程.

平稳过程的理论是时间序列分析的基础, 本书对平稳过程的介绍从略.

在本节最后, 我们介绍一个具体的简单随机过程并举例说明它的应用.

♣**定义2.1.15.** 称随机过程$X=\{X_n;n\in\mathbb{Z}_+\}$为伯努利过程, 若$\{X_n;n\in\mathbb{Z}_+\}$为独立同分布($i.i.d$)的随机变量序列, 而且
$$P(X_n=1)=p,\qquad P(X_n=0)=q=1-p,\quad 0<p<1.$$

伯努利过程常用来描述一列独立重复实验, 通常将$P(X_n=1)=p$看作第n次实验成功概率, $P(X_n=0)=q$ 看作第n次实验失败的概率. 显然, 伯努利过程是严平稳的.

令$T_1=\inf\{n>0;X_n=1\}$, 其中约定$\inf\emptyset=+\infty$. 直观地看T_1表示第一次实验成功的时刻. 但是这个时刻是随机的, 而且可以取$+\infty$. 进一步分析可知, 判定$\{T_1=n\}$这个事件是否发生, 我们可以通过观察n及n之前实验的结果: 若n之前实验都没成功而在n时刻实验恰好成功, 那么事件$\{T_1=n\}$发生; 否则该事件没有发生, 即
$$\{T_1=n\}=\{X_n=1,X_1,\cdots,X_{n-1}\text{均不为}1\}. \tag{2.1.2}$$
在随机过程研究中, 具有像T_1这种性质的(广义)随机变量是非常重要的工具; 一般地, 我们称这种随机变量为**停时**, 具体定义如下.

♣**定义2.1.16.** 称非负整数值的随机变量(可取$+\infty$) T是关于随机过程 $X=\{X_n;n\geq 0\}$ 的停时(*Stopping time*), 若对任意非负整数n, 随机事件$\{T=n\}$可以通过随机变量 $X_0,\cdots,X_n$ 表示, 也即随机变量$1_{\{T=n\}}$ 可表示成随机变量$X_0,\cdots,X_n$的函数. 当无需强调"参考"过程X时, 我们简称T为停时.

♠**注记2.1.17.** 定义2.1.16只是给出了关于离散时间随机过程的停时定义; 对这类停时性质的讨论, 我们将在第七章第一节展开. 对于更一般的停时定义, 感兴趣的同学可以参阅 [9].

回到伯努利过程. 对任意$k>1$, 同样可定义$T_k=\inf\{n>T_{k-1};X_n=1\}$. 那么$T_k$ 表示第k 次伯努利实验成功的时刻, 而且也是停时.

任取数列$\{f(n);n\geq 1\}$. 由

$$f(T_k)=f(T_{k-1}+1)X_{T_{k-1}+1}+\cdots+f(T_k-1)X_{T_k-1}+f(T_k)X_{T_k}$$

可知, 若$\sum\limits_{n\geq 1}|f(n)|<\infty$, 那么

$$\sum_{k=1}^{\infty}f(T_k)=\sum_{n=1}^{\infty}f(n)X_n.$$

此时, 对任意$\varepsilon>0$, 存在$m>0$使得

$$\sum_{n=1}^{m}f(n)X_n-\varepsilon\leq\sum_{k=1}^{\infty}f(T_k)\leq\sum_{n=1}^{m}f(n)X_n+\varepsilon,$$

进而我们可得

$$E\Big(\sum_{k=1}^{\infty}f(T_k)\Big)=\lim_{m\to\infty}E(\sum_{n=1}^{m}f(n)X_n)=p\sum_{n=1}^{\infty}f(n). \tag{2.1.3}$$

▶**例2.1.18.** 设在某种仪器在单位时间内损坏事件构成一个伯努利过程. 假定仪器损坏后更换的费用为c元. 为使仪器损坏时用户能免费更换, 厂方在卖出产品时应合理定价, 假设投资1元到下一个单位时间可增值到$1/\alpha$元, 其中$0<\alpha<1$. 问厂家预期应至少多定价多少?

解: 以T_k表示第k次损坏一个仪器的时间, 此时更换仪器要c元, 为保证此时能免费更换, 厂家需额外收取$c\alpha^{T_k}$ 元费用. 因此, 要保证长期免费更换厂家应多收取的费用为$\sum_{k=1}^{\infty}c\alpha^{T_k}$. 由(2.1.3)可知, 预期应增收费用为

$$E\Big(\sum_{k=1}^{\infty}c\alpha^{T_k}\Big)=cp\sum_{k=1}^{\infty}\alpha^k=\frac{cp\alpha}{1-\alpha}.$$ □

练习题2.1

1 试解释对离散参数随机过程而言, “不可区分”与“修正”等价.

2 验证(2.1.1)中$m(t)$与$D(t)$的表达式对任意$t\geq 0$成立.

3 在参数T为非负整数的条件下验证独立增量过程的性质(2).

4 证明随机过程$\{X_t;\ t\in T\}$与$\{Y_t;\ t\in T\}$独立当且仅当对任意$m\geq 1$ 以及任意选取的$t_1,\cdots,t_m\in T$, $(X_{t_1},\cdots,X_{t_m})$与$(Y_{t_1},\cdots,Y_{t_m})$独立.

5 已知$\{X(t);\ t\in[0,+\infty)\}$为平稳独立增量过程. 对任意$s>0$, 令$Y(t)=X(s+t)-X(s)$, 证明$\{Y(t);t\in[0,+\infty)\}$也是平稳独立增量过程.

6 证明伯努利过程不是独立增量过程.

2.2　直线上简单随机游动

随机游动是一类简单、直观的随机模型, 有着广泛的应用. 本节将简要地介绍直线上简单随机游动的理论和应用.

2.2.1　模型及其刻画

♣**定义2.2.1.** *已知$\{X_n; n \geq 0\}$为相互独立的随机变量序列而且对任意$n \geq 1$, X_n的分布相同, 令*

$$W_n = \sum_{k=0}^{n} X_k, \quad n \geq 0.$$

称随机过程$W = \{W_n;\ n \geq 0\}$为(一维或直线上)随机游动(Random Walk).

♠**注记2.2.2.** *若$X_n \in \mathbb{R}^d, n \geq 0$, 但定义2.2.1的其他条件不变, 那么我们称$W$为$d$-维随机游动.*

为讨论方便, 若无特别声明, 本节我们介绍的随机游动都是**整数值的随机游动**, 即$X_n, n \geq 0$, 都是整数值随机变量. 称满足

$$P(X_n = 1) = p, \qquad P(X_n = -1) = q = 1 - p, \quad n \geq 1,$$

的整数值随机游动为简单随机游动或(q, p)-简单随机游动, 其中$p = q = 1/2$时简称W为简单对称随机游动.

从初等概率论角度看, 随机游动W实质就是独立同分布随机变量的部分和序列. 从而随机游动是平稳独立增量过程. 当我们把X_n形象地理解成沿直线随机走出的一步或一个单位时间的位移时, W就描述了沿直线上行走时随机走过的路线. 这也是我们称W 为随机游动的一种直观解释, 相应地,我们有时也称X_n为(第n步的)步长.

对于描述随机现象动态变化的模型, 人们除了因为关心随机现象的趋势变化而研究它的长期、整体性质外, 还会因为关心随机现象发展过程中的随机事件而研究模型的局部, 微观性质. 对随机游动而言, 其长期趋势我们可以通过概率论里学过的大数定律和中心极限定理有所了解. 在对随机游动局部性质研究中一个基本问题就是: 给定随机游动的初始位置(即初始值), 过了一段时间后它会到哪里？用概率的语言来说, 就是对任意$x, y \in \mathbb{Z}$, 讨论条件概率

$$p_n(x, y) = P(W_n = y | W_0 = x), \qquad n \geq 0.$$

通常我们称该条件概率为n步转移概率. 其中0步转移概率(即$n=0$)定义如下:

$$p_0(x,y)=\delta_x(y)=\begin{cases}1, & y=x;\\ 0, & y\neq x.\end{cases}$$

为表示方便, 我们常简记$p_n(0,y)$为$p_n(y)$, 并将一步转移概率$p_1(x,y)$和$p_1(y)$分别简记为$p(x,y)$和$p(y)$.

对n步转移概率$p_n(x,y)$, 利用随机事件下条件概率的定义以及X_i, $i\geq 0$, 的独立性可知

$$\begin{aligned}p_n(x,y) &= \frac{P(W_n=y,W_0=x)}{P(W_0=x)}\\ &= \frac{P(\sum_{i=1}^n X_i=y-x,X_0=x)}{P(X_0=x)}\\ &= P(\sum_{i=1}^n X_i=y-x). \qquad (2.2.1)\end{aligned}$$

由此可知

$$p_n(y-x)=p_n(0,y-x)=P(\sum_{i=1}^n X_i=y-x)=p_n(x,y). \qquad (2.2.2)$$

进一步地, 对任意$k\geq 0$, $n\geq 1$, 利用X_i, $i\geq 1$, 的独立同分布性质可知, 条件概率

$$\begin{aligned}P(W_{k+n}=y|W_k=x) &= \frac{P(W_{n+k}=y,W_k=x)}{P(W_k=x)}\\ &= \frac{P(\sum_{i=k+1}^{n+k} X_i=y-x,W_k=x)}{P(W_k=x)}\\ &= P\Big(\sum_{i=k+1}^{n+k} X_i=y-x\Big) \qquad (2.2.3)\\ &= P(\sum_{i=1}^n X_i=y-x).\end{aligned}$$

将这个结果与(2.2.1),(2.2.2)联系后可知

$$P(W_{k+n}=y|W_k=x)=p_n(x,y)=p_n(y-x)=P(\sum_{i=1}^n X_i=y-x). \qquad (2.2.4)$$

因此$p_n(x,y)$不仅表示随机游动从时刻0由x出发经n步到达位置y的概率, 也表示从任一时刻所在位置x出发经过n步到位置y的概率, 他们都等于从位置0出发经n步到达位置$y-x$ 的概率. 若已知步长X_n的分布为F, 那么

$$p_n(x,y)=F^{*n}(\{y-x\}), \qquad (2.2.5)$$

这里$F^{*n}(\{y-x\})$表示分布为F的n重卷积的随机变量落在$y-x$点的概率, 即

$$F^{*n}(\{y-x\})=F^{*n}(y-x)-F^{*n}((y-x)-),$$

其中$F^{*n}((y-x)-)$为函数$F^{*n}(\cdot)$在$y-x$点的左极限.

▶**例2.2.3.** 设随机游动W的步长分布为

$$P(X_n=-2)=P(X_n=-1)=P(X_n=1)=1/3.$$

求转移概率$p_4(3,1)$.

[分析] 由$p_4(3,1)=p_4(0,-2)$可知所求概率等价于从0出发, 经过4步最终位移为-2的概率. 由于每步只有三种可能$+1,-1,-2$, 问题相当于在三个分别放有标号为$+1$,-1和-2的球的坛子里总共随机地选取四个球, 问最后四个球的标号之和为-2 的概率.

解 分别以m,n,k表示选择-2, -1和1的次数, 那么

$$\begin{cases} m+n+k=4, \\ -2m-n+k=-2. \end{cases}$$

由此可得$m=2k-2$, $n=6-3k$. 由于m,n,k均为非负整数, 因此

$$(m,n,k)=(0,3,1)\text{或}(2,0,2).$$

所以

$$p_4(3,1)=C_4^1\frac{1}{3^4}+C_4^2\frac{1}{3^4}=10/81.$$

□

对简单随机游动, 我们可以非常方便地计算$\{p_n(x,y);x,y\in\mathbb{Z},n\in\mathbb{N}\}$. 事实上根据上面的分析, 对于简单随机游动而言每次步长只能是$+1$(向前一步)或-1(向后一步), 要从位置x出发经n步到达y, 意味着其中有$(n+y-x)/2$ 步向前, $(n+x-y)/2$ 步向后. 所以

$$p_n(x,y)=\begin{cases} C_n^{(n+y-x)/2}p^{(n+y-x)/2}q^{(n+x-y)/2}, & n-|y-x|\text{为非负偶数}, \\ 0, & \text{其他}. \end{cases} \tag{2.2.6}$$

利用转移概率, 我们还能方便地写出随机游动的任意有限维分布. 任取

$$0\le k_0<k_1<\cdots<k_n,\quad r_0,r_1,\cdots,r_n\in\mathbb{Z}.$$

由(2.2.3)和(2.2.4)可得

$$\begin{aligned} &P(W_{k_0}=r_0,W_{k_1}=r_1,\cdots,W_{k_n}=r_n) \\ &=P\Big(W_{k_0}=r_0,\sum_{i=k_0+1}^{k_1}X_i=r_1-r_0,\cdots,\sum_{i=k_{n-1}+1}^{k_n}X_i=r_n-r_{n-1}\Big) \\ &=P(W_{k_0}=r_0)P\Big(\sum_{i=k_0+1}^{k_1}X_i=r_1-r_0\Big)\cdots P\Big(\sum_{i=k_{n-1}+1}^{k_n}X_i=r_n-r_{n-1}\Big) \\ &=P(W_{k_0}=r_0)p_{k_1-k_0}(r_0,r_1)p_{k_2-k_1}(r_1,r_2)\cdots p_{k_n-k_{n-1}}(r_{n-1},r_n). \end{aligned} \tag{2.2.7}$$

进一步, 若我们已知W的初始分布, 记作μ, 那么由全概率公式可得

$$P(W_{k_0}=r_0)=\sum_{i\in\mathbb{Z}}P(W_0=i,W_{k_0}=r_0)$$

$$= \sum_{i\in\mathbb{Z}} P(W_0 = i)P(W_{k_0} = r_0|W_0 = i)$$
$$= \sum_{i\in\mathbb{Z}} \mu(i)p_{k_0}(i, r_0). \tag{2.2.8}$$

可见初始分布以及转移概率族$\{p_n(x,y); x, y \in \mathbb{Z}, n \geq 1\}$是我们认识随机游动的随机规律的基础.

▶**例2.2.4.** *已知初值为0的随机游动W的步长分布为*
$$P(X_n = k) = \frac{1}{2^{k+2}}, \quad k \geq -1.$$
求概率$P(W_1 = 1, W_2 = 0, W_7 = 1)$.

解 注意到W的初值为0. 由有限维分布计算公式(2.2.7)和(2.2.8)可知
$$P(W_1 = 1, W_2 = 0, W_7 = 1) = P(W_1 = 1)p(1,0)p_5(0,1)$$
$$= P(W_0 = 0)p(0,1)p(1,0)p_5(0,1) = p(0,1)p(1,0)p_5(0,1).$$
记W的步长分布为F, 那么由(2.2.5)可知
$$p(0,1) = F(\{1\}) = 1/8; \quad p(1,0) = F(\{-1\}) = 1/2;$$
$$p_5(0,1) = F^{*5}(\{1\}) = P\Big(\sum_{i=1}^{5} X_i = 1\Big) = P\Big(\sum_{i=1}^{5}(X_i + 1) = 6\Big).$$
注意到$X_i + 1$是独立同分布的非负随机变量, 对应的概率母函数
$$g(s) = \sum_{k=0}^{\infty} P(X_i + 1 = k)s^k = \sum_{k=0}^{\infty} \frac{s^k}{2^{k+1}} = \frac{1}{2-s},$$
因此$\sum_{i=1}^{5}(X_i + 1)$的概率母函数为
$$G(s) = [g(s)]^5 = \frac{1}{(2-s)^5}.$$
由概率母函数的反演公式可知
$$P\Big(\sum_{i=1}^{5}(X_i + 1) = 6\Big) = \frac{1}{6!}G^{(6)}(0) = \frac{10!}{6!4!}\frac{1}{2^{11}} = C_{10}^4/2^{11}.$$
综上可得$P(W_1 = 1, W_2 = 0, W_7 = 1) = C_{10}^4/2^{15}$. □

2.2.2 基本性质及其应用

随机游动也是平稳独立增量过程. 利用这一特性我们可以得到随机游动的许多重要性质, 我们用定理形式介绍其中两条.

★**定理2.2.5.** *设W是整数值随机游动. 对任意$m, k \in \mathbb{N}$以及$i_0, i_1, \cdots, i_{m+1} \in \mathbb{Z}$,*
$$P(W_{m+k} = i_{m+1}|W_m = i_m, \cdots, W_0 = i_0) = P(W_{m+k} = i_{m+1}|W_m = i_m).$$

证明 对任意$m \geq 1$, 由(2.2.7)可知

$$\begin{aligned}&P(W_{m+k}=i_{m+1},\cdots,W_1=i_1,W_0=i_0)\\&=P(X_0=i_0)\prod_{r=0}^{m-1}p(i_r,i_{r+1})p_k(i_m,i_{m+1}),\end{aligned}$$

并且

$$P(W_m=i_m,\cdots,W_0=i_0)=P(X_0=i_0)\prod_{r=0}^{m-1}p(i_r,i_{r+1}).$$

两式做商后可得

$$P(W_{m+1}=i_{m+1}|W_m=i_m,\cdots,W_0=i_0)=p_k(i_m,i_{m+1}).$$

由此及(2.2.4)可知定理2.2.5成立. □

本质上定理2.2.5是上一节介绍平稳独立增量过程时提到的马氏性在随机游动上的具体展现. 这里我们结合随机游动的特性, 表达更具体、证明更简洁.

★**定理2.2.6.** (对称原理) 对任意$n \geq 1$, 区间$I \subset \mathbb{R}$以及任意随机游动,

$$\begin{aligned}&P(W_1>0,\cdots,W_{n-1}>0,W_n\in I|W_0=0)\\&=P(W_1<W_n,\cdots,W_{n-1}<W_n,W_n\in I|W_0=0).\end{aligned}$$

证明 令$Y_1=X_n$, $Y_2=X_{n-1},\cdots,Y_n=X_1$, $Y_0=X_0$. 对任意$k=0,1,\cdots,n$, 定义新的随机游动$S=\{S_k;k\geq 0\}$ 使得

$$S_k=\sum_{i=0}^{k}Y_i.$$

那么由Y_k的设定, 对任意$k=1,\cdots,n$,

$$S_k=\sum_{i=n}^{n-k+1}X_i+X_0=W_n-\sum_{i=1}^{n-k}X_i=W_n-W_{n-k}+W_0.$$

由于$X_1,\cdots,X_n$是独立同分布的随机变量, 因此$(X_1,\cdots,X_n)$ 与 $(Y_1,\cdots,Y_n)$有相同分布. 从而

$$(W_0,W_1,\cdots,W_n)\text{ 与 }(S_0,S_1,\cdots,S_n)$$

分布相同. 因此

$$\begin{aligned}&P(W_1>0,\cdots,W_{n-1}>0,W_n\in I|W_0=0)\\&=P(S_1>0,\cdots,S_{n-1}>0,S_n\in I|S_0=0)\\&=P(W_n-W_{n-1}+W_0>0,\cdots,W_n-W_1+W_0>0,W_n\in I|W_0=0)\\&=P(W_1<W_n,\cdots,W_{n-1}<W_n,W_n\in I|W_0=0).\end{aligned}$$ □

♠**注记2.2.7.** 从上述对称原理的证明可以看出, 对称原理只依赖于$X_n,n\geq 0$, 的相互独立性以及X_n, $n\geq 1$, 的同分布性质, 与X_n的值域形态无关.

为了进一步说明对称原理的应用, 我们引入一类特殊的随机游动:

♣**定义2.2.8.** *称整数值随机游动W是右连续的(或不带右跳的), 若其步长满足:*

$$P(X_n=1)>0 \text{ 而且 } \sum_{k=-\infty}^{1} P(X_n=k)=1.$$

显然简单随机游动是右连续的随机游动.

对任意整数i, 令$\tau_i=\inf\{n\geq 1,\ W_n=i\}$. 那么$\tau_i$是一个停时. 事实上

$$\{\tau_i=n\}=\{W_1\neq i,\cdots,W_{n-1}\neq i,W_n=i\}. \tag{2.2.9}$$

通常我们称这样定义的τ_i为状态i的首回时. 直观地看, 若$W_0=i$, τ_i表示随机游动W首次回到位置i的时间. 当$W_0\neq i$时τ_i也是随机游动W首次到达i的时间. 注意, 当整数值随机游动W右连续且初值为0时对任意$i>0$,

$$\{\tau_i=n\}=\{W_1<W_n,\cdots,W_{n-1}<W_n,W_n=i\}.$$

★**定理2.2.9.** *若整数值随机游动W右连续, 那么对任意正整数m和n,*

$$P(\tau_m=n|W_0=0)=\frac{m}{n}P(W_n=m|W_0=0).$$

我们先证明一个排列组合方面的结论.

▼**引理2.2.10.** *对任意正整数m, 在满足$x_1+\cdots+x_n=m$, 其中$x_i\leq 1$, $i=1,\cdots,n$, 的任意一个整数解(可以是负数)$(u_1,u_2,\cdots,u_n)$的n个循环排列中恰有m个循环排列使得其前$n-1$ 项部分和都严格小于m.*

例如, 取$n=5$, $m=2$, 那么$(-2,1,1,1,1)$ 是方程$x_1+\cdots+x_n=m$满足条件的一组解, 这一组的循环排列有5种, 分别是

$$(-2,1,1,1,1),(1,-2,1,1,1),(1,1,-2,1,1),(1,1,1,-2,1),(1,1,1,1,-2),$$

其中只有第一, 第二两种排列满足前$n-1=4$项部分和严格小于$m=2$. 事实上排列$(-2,1,1,1,1)$ 和$(1,-2,1,1,1)$的前4项部分和分别是

$$(-2,-1,0,1) \text{ 和 } (1,-1,0,1).$$

证明* 设$(u_1,\cdots,u_n)$是方程$x_1+\cdots+x_n=m$的整数解. 由于$(u_1,\cdots,u_n)$的循环都是整数解, 因此我们可以选择一个好的循环作为分析的对象. 记$\{u_i\}$的部分和序列为$S_1,\cdots,S_n$. 设该部分和序列的最大值为M并令$l=\min\{k,S_k=M\}$. 那么$M\geq m$, $u_l=1$且对任意$l<i\leq n$, $u_{l+1}+\cdots+u_i\leq 0$. 选择循环排列$(u_{l+1},\cdots,u_n,u_1,\cdots,u_l)$ 作为我们分析对象, 并将其重新记为$(v_1,v_2,\cdots,v_n)$. 可得:

(1) 对任意$k<n$, 部分和$v_1+\cdots+v_k<v_1+\cdots+v_n=m$.

(2) 由于部分和$v_1+\cdots+v_{n-l}=u_{l+1}+\cdots+u_n\le 0$而且$v_i$最大只能取到$+1$, 因此对任意$1\le j\le m$, 一定存在一个指标$n_k$ 使得部分和

$$v_1+\cdots+v_{n_k}=j$$

并记这些指标中最小的指标为k_j, 其中k_0规定为0. 由(1)知$k_m=n$,并且由于部分和序列$v_1+\cdots+v_i$是从0值开始而且v_i最大只能取到$+1$可得k_j 随着j的增加而增加.

(3) 对任一$j<m$, $(v_{k_j+1},\cdots,v_{k_{j+1}})$一定满足

$$v_{k_j+1}+\cdots+v_{k_{j+1}}=1,\ \ v_{k_{j+1}}=1.$$

若$k_j+1\le i<k_{j+1}$, 则$v_{k_j+1}+\cdots+v_i\le 0$ (否则由$i\ge k_{j+1}$导出矛盾).

由此我们可以验证: $(v_1,v_2,\cdots,v_n)$的循环排列中所有的

$$(v_{k_j+1},\cdots v_n,v_1,\cdots,v_{k_j}),\quad j=0,1,\cdots,m-1,$$

都满足引理要求, 而其他的任何一个循环

$$(v_i,\cdots,v_n,v_1,\cdots,v_{i-1})$$

都不满足要求, 因为此时存在一个j使得$k_j+1<i\le k_{j+1}$, 该循环的前$n-i+k_j$项的部分和为

$$\begin{aligned}v_i+\cdots+v_n+v_1+\cdots+v_{k_j}&=v_1+\cdots+v_n-(v_{k_j+1}+\cdots+v_{i-1})\\&=m-(v_{k_j+1}+\cdots+v_{i-1})\ge m.\end{aligned}$$

所以在$(v_1,v_2,\cdots,v_n)$的循环排列中有且仅有m个满足引理要求. □

利用该引理, 我们很容易证明定理2.2.9. 证明如下.

证明 记$A=\{W_n=m,W_0=0\}=\{X_1+\cdots+X_n=m\}$,

$$\begin{aligned}B&=\{\tau_m=n,W_0=0\}\\&=\{W_1<W_n,\cdots,W_{n-1}<W_n,W_n=m,W_0=0\}\\&=\{X_1+\cdots+X_k<m,k=1,\cdots,n-1,X_1+\cdots+X_n=m\}.\end{aligned}$$

任取A中一个基本事件$u=(u_1,\cdots,u_n)$, 记由该基本事件通过循环所得到的n个(包括可能相同的)排列构成的集合为T_u, 剔除其中重复的排列后所得到的集合为Γ_u. 显然Γ_u中的每个排列都是A中的基本事件. 此时

$$A=\bigcup_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}\Gamma_u,$$

并且由于A中至多有可数个基本事件u, 上式只涉及至多可数个事件Γ_u的并.

(1) 若$T_u=\Gamma_u$, 由引理2.2.10, 对任意$u\in A$, Γ_u中有且仅有m 个循环排列事件属于B. 注意到X_i是独立同分布的随机变量, 每个循环排列发生概率相同,

因此
$$\frac{P(B\cap\Gamma_u)}{P(A\cap\Gamma_u)}=\frac{m}{n}. \tag{2.2.10}$$

(2) 若$T_u\neq\Gamma_u$, 则必有Γ_u中一个基本事件对应的排列在T_u中出现了多次, 比如$k(>1)$次. 不妨假设这个基本事件就是u, 并且对u分别轮换了$0=i_1-1<i_2-1<\cdots<i_k-1$次后回到$u$, 这表明

$$\begin{aligned}&(u_1,\cdots,u_{i_2-1},u_{i_2},\cdots,u_{i_k},\cdots,u_n)\\&\quad=(u_{i_2},\cdots,u_{i_k},\cdots,u_n,u_1,\cdots,u_{i_2-1},)\\&\quad=\cdots=(u_{i_k},\cdots,u_n,u_1,\cdots,u_{i_2},\cdots,u_{i_k-1}).\end{aligned}$$

因此
$$(u_1,\cdots,u_{i_2-1})=(u_{i_2},\cdots,u_{i_3-1})=\cdots=(u_{i_k},\cdots,u_n).$$

由此可知Γ_u中所有基本事件都在T_u中恰好出现k次. 进而由引理2.2.10可知$B\cap\Gamma_u$与$A\cap\Gamma_u$中的基本事件数之比仍然是m/n, 从而(2.2.10)也成立.

因此, 由$B\subset A$可得
$$P(B\cap\Gamma_u)=P(B\cap A\cap\Gamma_u)=\frac{P(B\cap\Gamma_u)}{P(A\cap\Gamma_u)}P(A\cap\Gamma_u)=\frac{m}{n}P(A\cap\Gamma_u).$$
进而, 由全概率公式,
$$P(B)=P(B\cap\bigcup_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}\Gamma_u)=\sum_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}P(B\cap\Gamma_u)=\frac{m}{n}\sum_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}P(A\cap\Gamma_u)=\frac{m}{n}P(A).$$
所以定理2.2.9成立. □

▶**例2.2.11.** 设W是(q,p)简单随机游动. 对任意$i\geq 0$, 求$P(\tau_i=n|W_0=0)$.

解 先考虑$i>0$的情形. 注意到(q,p)简单随机游动是右连续的. 由定理2.2.9以及(2.2.6) 可知,
$$\begin{aligned}P(\tau_i=n|W_0=0)&=\frac{i}{n}P(W_n=i|W_0=0)\\&=\begin{cases}\frac{i}{n}C_n^{(n+i)/2}p^{(n+i)/2}q^{(n-i)/2}, & n-i\text{为非负偶数},\\ 0, & \text{其他}.\end{cases}\end{aligned}$$
下面考虑$i=0$的情形. 此时由于简单随机游动必须偶数步才可能从位置0回到0, 因此, 对任意$k\geq 0$,
$$P(\tau_0=2k+1|W_0=0)=0.$$
另一方面, 由全概率公式可得
$$\begin{aligned}P(\tau_0=2k|W_0=0)&=P(\tau_0=2k,W_1=1|W_0=0)\\&\quad+P(\tau_0=2k,W_1=-1|W_0=0).\end{aligned}$$

利用随机事件下条件概率公式
$$P(AB|C) = P(A|C)P(B|AC)$$
可得
$$\begin{aligned}&P(\tau_0 = 2k, W_1 = 1|W_0 = 0)\\&\quad = P(W_1 = 1|W_0 = 0)P(\tau_0 = 2k|W_1 = 1, W_0 = 0)\\&\quad = pP(W_{2k} = 0, W_i \neq 0, 1 < i < 2k|W_1 = 1, W_0 = 0),\end{aligned}$$
其中最后一个等号由(2.2.9)可得. 利用随机游动的平稳独立增量性质可知
$$\begin{aligned}&P(W_{2k} = 0, W_i \neq 0, 1 < i < 2k|W_1 = 1, W_0 = 0)\\&\quad = P\Big(W_{2k} - W_1 = -1, W_i - W_1 \neq -1, 1 < i < 2k\Big|W_1 = 1, W_0 = 0\Big)\\&\quad = P\Big(W_{2k} - W_1 = -1, W_i - W_1 \neq -1, 1 < i < 2k\Big)\\&\quad = P\Big(W_{2k-1} - W_0 = -1, W_i - W_0 \neq -1, 0 < i < 2k-1\Big)\\&\quad = P\Big(W_{2k-1} - W_0 = -1, W_i - W_0 \neq -1, 0 < i < 2k-1\Big|W_0 = 0\Big)\\&\quad = P(W_{2k-1} = -1, W_i \neq -1, 0 < i < 2k-1|W_0 = 0)\\&\quad = P(\tau_{_{-1}} = 2k-1|W_0 = 0),\end{aligned}$$
其中第二个等号用到增量独立性, 第三个等号用到增量平稳性, 第四个等号又用到增量独立性.

注意从0出发的(q,p)简单随机游动首达-1的时间与从0出发的(p,q)简单随机游动首达1的时间分布相同, 由前面$i \geq 1$部分的计算可知
$$P(\tau_{_{-1}} = 2k-1|W_0 = 0) = \frac{1}{2k-1}\frac{(2k-1)!}{k!(k-1)!}q^k p^{k-1} = \frac{(2k-2)!}{k!(k-1)!}q^k p^{k-1}.$$
从而
$$P(\tau_0 = 2k, W_1 = 1|W_0 = 0) = pP(\tau_{_{-1}} = 2k-1|W_0 = 0) = \frac{(2k-2)!}{k!(k-1)!}q^k p^k.$$
同理可得
$$P(\tau_0 = 2k, W_1 = -1|W_0 = 0) = qP(\tau_1 = 2k-1|W_0 = 0) = \frac{(2k-2)!}{k!(k-1)!}q^k p^k.$$
因此
$$P(\tau_0 = 2k|W_0 = 0) = 2\frac{(2k-2)!}{k!(k-1)!}q^k p^k = \begin{cases}2pq, & k = 1,\\ \frac{(2k-3)!!}{2^k}\frac{(4pq)^k}{k!}, & k \geq 2.\end{cases} \qquad \square$$

对简单随机游动还有如下的反射原理.

▼**引理2.2.12.** (反射原理) 设W为**简单随机游动**, 对任意正整数x, y, 从(n, x)到$(n+m, y)$并途中位置回到零点的轨道数与从$(n, -x)$到 $(n+m, y)$的轨道数相同. 这里所谓随机游动的轨道是指将随机游动W中各位置(n, W_n)在平面上绘点并

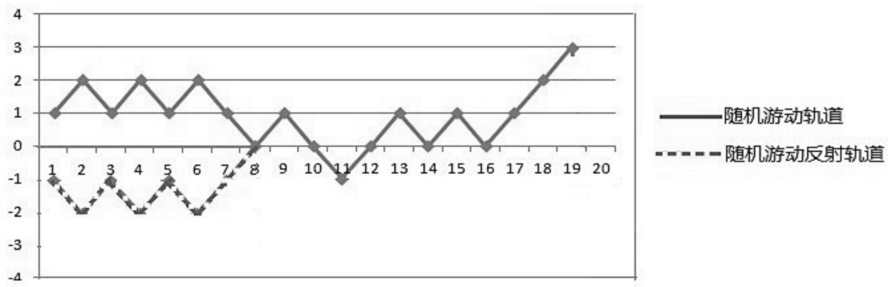
按n从小到大的次序用直线依次相连所得到的折线.

证明 令A表示从(n,x)到$(n+m,y)$并在中间某个时刻到达零点的所有轨道, B 为从$(n,-x)$到 $(n+m,y)$的所有轨道.

任取A中一个轨道

$$(n,x),(n+1,s_1),\cdots,(n+k,s_k),\cdots,(n+m,y).$$

设沿轨道从n开始首次到达位置零点的时刻为$n+k$, 即$s_k=0$, 但s_1到s_{k-1}都大于0(如下图蓝色折线所示). 将(n,x) 到 $(n+k,0)$的轨道沿时间轴翻转得到一条连接$(n,-x)$ 与$(n+k,0)$的轨道(如下图红线所示), 再与原来$(n+k,0)$ 到$(n+m,y)$轨道拼接, 得到一条从$(n,-x)$到$(n+m,y)$的轨道. 显然这种轨道产生方式是一对一的, 因此A中轨道数不超过B中轨道数.



反过来, 任取B中一轨道

$$(n,-x),(n+1,s_1),\cdots,(n+k,s_k),\cdots,(n+m,y).$$

由于简单随机游动每次只能移动一步, 而$-x<0<y$, 因此由$(n,-x)$到$(n+m,y)$的过程中必有某个时刻到达位置0. 设$n+k$是首个这样的时刻. 同样可将$(n,-x)$到$(n+k,0)$的轨道沿时间轴翻转得到一条由(n,x)到$(n+k,0)$的轨道, 进而得到一条从(n,x) 到$(n+m,y)$的轨道. 显然这种轨道产生方式也是一对一的, 因此B中轨道数不超过A中轨道数.

综合两方面可知A,B中轨道数相同. □

▲**命题2.2.13.** *若W是***简单对称随机游动***, 那么对任意$x,y>0$,*

$$P(W_{n+m}=y,\min_{n<k<n+m}W_k\le 0|W_n=x)=P(W_{n+m}=y|W_n=-x).$$

证明 注意到简单对称随机游动每一步向上或向下走的机会是相等的. 将W的任意两个可能点(n,x), $(n+m,y)$连接起来的任一轨道都是等可能的, 而且与从(n,x)出发走过m步的任一轨道发生的概率一样. 注意到从(n,x) 出发走过m 步的轨道数为2^m, 由引理2.2.12得

$$P(W_{n+m}=y,\min_{n<k<n+m}W_k\le 0|W_n=x)$$

$$= \frac{A\text{中轨道数}}{2^m} = \frac{B\text{中轨道数}}{2^m} = P(W_{n+m} = y | W_n = -x). \qquad \square$$

下面我们举两个应用随机游动基本性质解决问题的例子.

▶**例2.2.14.** *娱乐公司为了吸引顾客推出一种特别活动. 活动规则如下: 顾客可以参与一种输赢概率各50%的下注游戏, 获胜者赢得一个筹码, 反之则输掉一个筹码. 公司为顾客第一局买单, 即获胜的顾客赢得一个筹码而失败的顾客则无需付出. 若顾客继续游戏, 则按正常游戏规则执行. 假设某个顾客参与了该活动但没有购买任何筹码, 包括第一局, 问他至少能玩该游戏10局的机会是多少?*

解 以W_n表示该顾客第n局结束后手上有的筹码数, 那么W是初值为0的简单对称随机游动. $W_1 > 0$意味着$W_1 = 1$, 而且$W_9 > 0$保证了$W_{10} \geq 0$. 因此所求概率为

$$\begin{aligned}
&P(W_1 > 0, \cdots, W_{10} \geq 0 | W_0 = 0) \\
&\quad = P(W_1 > 0, W_2 > 0, \cdots, W_9 > 0 | W_0 = 0) \\
&\quad = \sum_{k=1}^{\infty} P(W_1 = 1, \min_{1<m<9} W_m > 0, W_9 = 2k-1 | W_0 = 0).
\end{aligned}$$

利用$\{\min_{1<m<9} W_m > 0\}$的对立事件$\{\min_{1<m<9} W_m \leq 0\}$得

$$\begin{aligned}
&P(W_1 > 0, \cdots, W_{10} \geq 0 | W_0 = 0) \\
&\quad = \sum_{k=1}^{\infty} \big[P(W_1 = 1, W_9 = 2k-1 | W_0 = 0) \\
&\qquad\qquad - P(W_1 = 1, \min_{1<m<9} W_m \leq 0, W_9 = 2k-1 | W_0 = 0)\big].
\end{aligned}$$

由随机事件下条件概率性质(参见习题1.1)可得

$$\begin{aligned}
&P(W_1 > 0, \cdots, W_{10} \geq 0 | W_0 = 0) \\
&\quad = \sum_{k=1}^{\infty} P(W_1 = 1 | W_0 = 0)\big[P(W_9 = 2k-1 | W_1 = 1, W_0 = 0) \\
&\qquad\qquad - P(\min_{1<m<9} W_m \leq 0, W_9 = 2k-1 | W_1 = 1, W_0 = 0)\big].
\end{aligned}$$

定理2.2.5表明

$$P(W_9 = 2k-1 | W_1 = 1, W_0 = 0) = P(W_9 = 2k-1 | W_1 = 1) = p_8(1, 2k-1).$$

此外, 由随机游动的独立增量性可以证明(细节请读者补充)

$$\begin{aligned}
&P(\min_{1<m<9} W_m \leq 0, W_9 = 2k-1 | W_1 = 1, W_0 = 0) \\
&\quad = P(\min_{1<m<9} W_m \leq 0, W_9 = 2k-1 | W_1 = 1).
\end{aligned}$$

因此, 结合命题2.2.13可得

$$P(W_1 > 0, \cdots, W_{10} \geq 0)$$

$$= P(W_1 = 1|W_0 = 0)\sum_{k=1}^{\infty}\left[p_8(1, 2k-1) - p_8(-1, 2k-1)\right]$$

$$= \frac{1}{2}\sum_{k=1}^{\infty}\left[p_8(2k-2) - p_8(2k)\right] = \frac{p_8(0)}{2} = \frac{35}{256}. \qquad \square$$

▶**例2.2.15.** 在一次选举中候选人A得到了n张票, 候选人B得到m张票, $n > m$. 问计票过程中A始终领先B的可能性有多大?

解 令$X_0 = 0$. 对$i \geq 1$, 以$X_i = 1$表示第i个人投票给了A, $X_i = -1$表示第i个人投票给了B. 假定X_i, $i \geq 1$, 是独立同分布的, 投票给A, B的概率分别为$p, q = 1-p$. 以随机游动$W_k = X_0 + X_1 + \cdots + X_k$表示计票过程中得票变化情况. 结果为$W_{n+m} = n - m > 0$. 所求问题转为求$(q, p)$随机游动的如下条件概率

$$P(W_1 > 0, W_2 > 0, \cdots, W_{n+m-1} > 0|W_{n+m} = n-m, W_0 = 0).$$

由对称原理及定理2.2.9可得

$$\begin{aligned}
&P(W_1 > 0, W_2 > 0, \cdots, W_{n+m-1} > 0, W_{n+m} = n-m|W_0 = 0)\\
&\quad = P(W_1 < W_{n+m}, \cdots, W_{n+m-1} < W_{n+m}, W_{n+m} = n-m|W_0 = 0)\\
&\quad = \frac{n-m}{n+m}P(W_{n+m} = n-m|W_0 = 0).
\end{aligned}$$

从而

$$P(W_1 > 0, W_2 > 0, \cdots, W_{n+m-1} > 0|W_{n+m} = n-m, W_0 = 0) = \frac{n-m}{n+m}. \qquad \square$$

练习题2.2

1 已知W是初值为0的随机游动, 步长分布为

$$P(X_n = -2) = P(X_n = -1) = P(X_n = 1) = 1/3.$$

(1) 求概率$P(W_3 = -2)$, (2) 求概率$P(W_3 = -2, W_7 = 0)$.

(3) 求概率$P(W_1 > 0, W_2 > 0, \cdots, W_5 > 0, W_6 = 2)$.

2 设整数值随机游动W的步长分布为

$$P(X_n = k) = \frac{1}{(k+1)!\mathrm{e}}, \quad k \geq -1.$$

(1) 对任意$x \in \mathbb{Z}$, 求$p_n(x)$.

(2) 求概率$P(W_1 < 2, W_2 < 2, \cdots, W_5 < 2, W_6 = 0|W_0 = 2)$.

3 (接例2.2.14) 若此人玩了至少10 局, 试预测他在第10 局结束时的筹码数.

4 考虑一个简单的股票价格变化模型. 假设股票只有涨跌两种状态, 且机会均等. 当股票上涨时涨幅为6%, 当股票下跌时跌幅以同等概率取−4%, −8%. 若股票每日的波动是独立的. 自某个设定的观察日开始,

(1) 求在首次下跌后, 股票平均总涨跌幅.

(2) 求在10天内上涨天数总比下跌天数多的概率.

(3) 在首次观察到上涨比下跌多一天时, 计算股票平均总涨跌幅.

5 分别记甲乙两种药物对某种疾病的治愈率为p_1, p_2. 为了比较它们的治愈率大小, 安排了系列如下的临床对比试验: 每次试验同时治疗两个病人, 一个接受甲药治疗, 另一个接受乙药治疗. 观察每次治疗效果. 假设每次试验是独立的; 病人对药物的疗效无本质影响. 试用恰当的整数值随机游动模型对该系列试验中治愈病人数差异的变化过程建模. 对任意$x \in \mathbb{Z}$, 试写该模型的一步转移概率$p(x)$.

6 对初值为0的简单对称随机游动W,

(1) 对任意$n \geq 1$, 证明$P(W_1 \neq 0, W_2 \neq 0, \cdots, W_{2n} \neq 0) = P(W_{2n} = 0)$.

(2) 令$F_1(s) = \sum_{k=1}^{\infty} P(\tau_1 = k | W_0 = 0) s^k$ 其中$|s| < 1$, 求$F_1(s)$.

7 对初值为0的简单随机游动W, 证明: 对任意正整数N以及与N同奇偶的整数$a \in [-N, N]$,

$$P(W_k < k, k = 1, \cdots, N-1 | W_N = a) = \frac{N-a}{2N}.$$

2.3 泊松过程

泊松过程是一种应用广泛的简单随机过程, 它与指数分布密切相关, 因而具有良好的性质.

2.3.1 计数过程

用数学模型刻画随机现象时, 最基本的工作就是统计随机现象发生的次数. 由此所得到的数学模型常被称为计数过程, 其准确定义如下:

♣**定义2.3.1.** 称非负整数值的随机过程$N=\{N(t);t\geq 0\}$为计数过程或点过程, 若其样本函数$N(\cdot,\omega)$为右连续的单调不降函数. 若还存在一个集合A使得$P(A)=1$而且

$$A\subset\{\omega;N(t,\omega)-N(t-,\omega)\leq 1\text{对任意}t\geq 0\},$$

则称N 是简单计数过程, 其中$N(t-,\omega)$ 表示函数$N(\cdot,\omega)$在t 点的左极限.

本质上, 简单计数过程要求依概率1保证在任一时刻至多发生一次随机事件.

对任意给定的$\omega\in\Omega$, 由于计数过程的样本函数$N(t,\omega)$关于t单调不降且右连续. 因此对任意给定的整数$k\geq 0$,

$$T_k(\omega)=\inf\{t\geq 0;\ N(t,\omega)\geq k\}$$

有意义. 由定义可知,

(1) T_k是第k次事件发生的时刻, 通常称之为第k次(随机事件)到达时刻. 由$N(t)$关于t的单调不降性可知T_k随k增加也是单调不降的而且$T_0=0$.

(2) 对任意$u\geq 0$以及$\omega\in\Omega$,

$$u\in\{t\geq 0;N(t,\omega)\geq N(u,\omega)\},$$

因此$T_{N(u,\omega)}(\omega)\leq u$. 这表明对任意$t\geq 0$, $T_{N(t)}\leq t$. 因此由T_n单调不降性可知

$$N(t)\leq\sup\{k;T_k\leq t\}.$$

(3) 对任意$\omega\in\Omega$, 若$T_k(\omega)\leq t$, 那么由T_k定义, 存在一列t_n单调下降地收敛到$T_k(\omega)$并且$N(t_n,\omega)\geq k$. 由N 的右连续性,

$$N(T_k(\omega),\omega)=\lim_{n\to\infty}N(t_n,\omega)\geq k.$$

即对任意ω, $N(T_k(\omega),\omega)\geq k$, (若$N$为简单计数过程, 则对几乎所有的$\omega$都有$N(T_k)=k$.) 因此由$N(t)$关于$t$的单调不降性可知

$$N(t)\geq\sup\{N(T_k);T_k\leq t\}\geq\sup\{k;T_k\leq t\}.$$

综上所述可知, 计数过程N可表示成

$$N(t)=\sup\{k\geq 0; T_k\leq t\}, \tag{2.3.1}$$

即对任意ω, $N(t,\omega)=\sup\{k; T_k(\omega)\leq t\}$, 其中约定$\sup\emptyset=0$. 由此可知

$$\{N(t)=k\}=\{T_k\leq t<T_{k+1}\}=\{T_{k+1}>t\}\setminus\{T_k>t\}, \tag{2.3.2}$$

即$N(t)$表示时刻t之前(包含t)发生的随机事件总数.

若记$W_k=T_k-T_{k-1}$, 那么W_k表示第$k-1$个事件与第k个事件之间的时间间隔. 一般地我们称W_k为N的第k个间隔时间, 称$W=\{W_k; k\geq 1\}$为N 的时间间隔序列. 此时事件发生的时间$\{T_k; k\geq 0\}$就是时间间隔序列的部分和, 即对任意$k\in\mathbb{N}$,

$$T_k=\sum_{i=1}^{k}W_i,$$

其中约定$\sum\limits_{n=1}^{0}=0$. 进而

$$N(t)=\sup\{n\geq 0; \sum_{k=1}^{n}W_k\leq t\}. \tag{2.3.3}$$

♠**注记2.3.2.** 设$S=\{S_k; k\geq 0\}$是随机变量X_i, $i\geq 1$, 的部分和序列. 对任意给定的ω, 当X_i恒正时, S_k是严格增加的. 映射$k\to S_k$ 可逆. 逆映射

$$S^{-1}: D=\{x_k;\ x_k=S_k, k\geq 0\}\mapsto\mathbb{N} \text{ 使得 } S^{-1}(x_k)=k.$$

容易看出这个逆映射还可推广到非负实数集$\bar{\mathbb{R}}_+$上, 使得对任意$x\in\bar{\mathbb{R}}_+$, 当$x_k\leq x<x_{k+1}$时$S^{-1}(x)=k$, 即

$$S^{-1}(x)=\sup\{k\geq 0;\ S_k\leq x\}. \tag{2.3.4}$$

如果X_i可以为0, 那么S_k作为k的函数不一定可逆, 但(*2.3.4*)定义的函数$S^{-1}(x)$仍有意义, 而且$S^{-1}(x)$就是使部分和$S_k\leq x$成立的最大求和项数. 通常, 我们称过程

$$S^{-1}=\{S^{-1}(t);\ t\geq 0\}$$

为部分和序列$\{S_k;\ k\geq 0\}$的逆过程.

对照(*2.3.3*)和(*2.3.4*) 可以发现, 计数过程N可以看作是部分和序列$\{T_k; k\geq 0\}$ 的逆过程.

2.3.2 泊松过程及其刻画

许多时候人们会把碰到的计数过程简化成一种特殊的计数模型, 使得事件发生的时间间隔构成一列相互独立且服从同一个指数分布的随机变量. 例如

▶**例2.3.3.** 假设某房间只有一盏电灯, 而且电灯泡坏了能得到及时更换. 灯泡的寿命通常假定在一段时间内服从相同的指数分布, 而且不同灯泡的寿命是独立的, 若以$N(t)$表示在$[0,t]$ 时间内更换下来的灯泡数, 此时N 对应的时间间隔序列就可以看成是一列服从相同指数分布的独立随机变量.

我们称这样一类特殊的计数随机模型为泊松(Poisson)过程.

♣**定义2.3.4.** 称计数过程$N=\{N(t);\ t\geq 0\}$是强度为λ的泊松($Poisson$)过程, 若N 具有独立同分布的时间间隔序列$\{W_k;\ k\geq 1\}$且间隔时间服从参数为λ的指数分布.

若无特别说明, 在本节剩下部分我们总以$\{W_k;\ k\geq 1\}$表示泊松过程的时间间隔序列, 以$T_k=\sum_{i=1}^k W_i$表示第k个事件达到的时刻.

由(2.3.3)可知

$$N(t)=\sup\{k\geq 0; T_k\leq t\}=\sup\{k\geq 0;\sum_{n=1}^k W_n\leq t\},\quad t\geq 0.$$

进一步, 由$\{N(0)>0\}\subset\{W_1=0\}$以及

$$\{存在t\geq 0使得N(t)-N(t-)\geq 2\}=\{存在某个k使得W_k=0\},$$

可得

$$P(N(0)=0)=1-P(N(0)\geq 1)\geq 1-P(W_1=0)=1,$$

而且

$$P(存在t\geq 0使得N(t)-N(t-)\geq 2)\leq\sum_{k=1}^\infty P(W_k=0)=0.$$

因此泊松过程是初值为0的简单计数过程.

下面的命题在一定程度上解释了为什么我们称$N=\{N(t)\}$为泊松过程, 同时也告诉我们, 泊松过程的所谓“强度”其实就是就是泊松过程记录的随机事件在单位时间内发生次数的平均数.

▲**命题2.3.5.** 记强度为λ的泊松过程为N, 那么对任意$t>0$, $N(t)$服从强度为λt的泊松分布. 因此$E(N(t))=Var(N(t))=\lambda t$.

证明 注意到泊松过程的时间间隔W_k是独立同分布的指数随机变量. 由概率论知识可知任意n个相互独立且服从参数为λ的指数分布的随机变量之和服从伽马分布$\mathrm{Ga}(n,\lambda)$，其概率密度函数

$$\gamma_n(x)=\begin{cases}\frac{\lambda^n x^{n-1}\mathrm{e}^{-\lambda x}}{(n-1)!}, & x>0,\\ 0, & x\leq 0.\end{cases}$$

因此对任意$k \geq 1$, $T_k = \sum_{i=1}^{k} W_i$的密度函数为$\gamma_k(x)$. 由(2.3.2)可知

$$\begin{aligned} P\big(N(t)=k\big) &= P(T_k \leq t < T_{k+1}) = P(T_k \leq t) - P(T_{k+1} \leq t) \\ &= \int_0^t [\gamma_k(x) - \gamma_{k+1}(x)]\mathrm{d}x = \int_0^t \left[\frac{\lambda^k x^{k-1}}{(k-1)!} - \frac{\lambda^{k+1}x^k}{k!}\right]\mathrm{e}^{-\lambda x}\mathrm{d}x. \end{aligned}$$

由分部积分可得

$$P\big(N(t)=k\big) = \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}.$$

这表明$N(t)$服从强度为λt的泊松分布. □

我们知道泊松过程是独立同分布随机变量部分和序列的“逆过程”, 由于独立同分布随机变量部分和序列是平稳独立增量过程, 因此也可以说泊松过程是一类平稳独立增量过程的逆过程. 不仅如此, 下面定理还告诉我们, 泊松过程自己也是平稳独立增量过程.

★定理2.3.6. 强度为λ的泊松过程是一个平稳独立增量过程且对任意$t, s \geq 0$, $N(t+s) - N(s)$服从强度为λt的泊松分布, 即对任意整数$k \geq 0$,

$$P(N(s+t) - N(s) = k) = \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}. \tag{2.3.5}$$

证明* 只需证明对任意$n \geq 1$以及$0 = t_0 < t_1 < \cdots < t_n$和非负整数$k_1, \cdots, k_n$, 总有

$$P\Big(N(t_r) = \sum_{i=1}^{r} k_i, r = 1, 2, \cdots, n\Big) = \mathrm{e}^{-\lambda t_n}\prod_{i=1}^{n}\frac{[\lambda(t_i - t_{i-1})]^{k_i}}{k_i!}, \tag{2.3.6}$$

(参见本节习题11). 对任意$r \geq 1$, 记$s_r = \sum_{i=1}^{r} k_i$. 由等式(2.3.2)可得

$$\Big\{N(t_r) = \sum_{i=1}^{r} k_i, r = 1, 2, \cdots, n\Big\} = \bigcap_{r=1}^{n}\{T_{s_r} \leq t_r, T_{s_r+1} > t_r\}.$$

为表示方便, 对任意$n \geq 1$, 令

$$\begin{aligned} A_n &= \bigcap_{r=1}^{n}\{T_{s_r} \leq t_r, T_{s_r+1} > t_r\} \\ &= \{T_{s_1} \leq t_1, T_{s_1+1} > t_1, \cdots, T_{s_n} \leq t_n, T_{s_n+1} > t_n\}, \end{aligned}$$

以及

$$\begin{aligned} B_n &= \{T_{s_n} \leq t_n, T_{s_n+1} > t_n\}, \\ C_n &= \{T_{s_1} \leq t_1, T_{s_1+1} > t_1, \cdots, T_{s_n} \leq t_n\}. \end{aligned}$$

此时(2.3.6)可改写为

$$P\big(A_n\big) = \mathrm{e}^{-\lambda t_n}\prod_{i=1}^{n}\frac{[\lambda(t_i - t_{i-1})]^{k_i}}{k_i!}. \tag{2.3.7}$$

当$n=1$时由命题2.3.5可知(2.3.7)显然成立. 下设$n=m-1$时成立, 即

$$P\big(A_{m-1}\big)=\mathrm{e}^{-\lambda t_{m-1}}\prod_{i=1}^{m-1}\frac{[\lambda(t_i-t_{i-1})]^{k_i}}{k_i!}.$$

当$n=m$时, 注意到$A_m=A_{m-1}B_m$. 我们分情况讨论如下:

(1)当$k_m=0$时, $s_m=s_{m-1}$, $B_m=\{T_{s_{m-1}}<t_m,T_{s_{m-1}+1}>t_m\}$. 由重期望公式可知

$$\begin{aligned}P(A_m)&=P(A_{m-1}B_m)=P(C_{m-1}\cap\{T_{s_{m-1}+1}>t_m\})=E(1_{C_{m-1}}1_{\{T_{s_{m-1}+1}>t_m\}})\\&=E(E(1_{C_{m-1}}1_{\{T_{s_{m-1}+1}>t_m\}}|T_1,\cdots,T_{s_{m-1}})).\end{aligned}$$

注意到C_{m-1}是一个由$T_1,\cdots,T_{s_{m-1}}$表示的随机事件, 而T_k是服从相同指数分布的独立随机变量的部分和序列, $T_{s_{m-1}+1}-T_{s_{m-1}}=W_{s_{m-1}+1}$与$T_1,\cdots,T_{s_{m-1}}$独立, 利用条件期望性质可得

$$\begin{aligned}P(A_m)&=E\big(1_{C_{m-1}}P(T_{s_{m-1}+1}>t_m|T_1,\cdots,T_{s_{m-1}})\big)\\&=E\big(1_{C_{m-1}}P(W_{s_{m-1}+1}>t_m-T_{s_{m-1}}|T_1,\cdots,T_{s_{m-1}})\big)\\&=E\big(1_{C_{m-1}}P(W_{s_{m-1}+1}>t_m-T_{s_{m-1}})\big)\\&=E\big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\big)\\&=\mathrm{e}^{-\lambda(t_m-t_{m-1})}E\big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_{m-1}-T_{s_{m-1}})}\big).\end{aligned}\tag{2.3.8}$$

(2)当$k_m=1$时, $s_m=s_{m-1}+1$, $B_m=\{T_{s_m}<t_m,T_{s_m+1}>t_m\}$, 因此

$$\begin{aligned}P(A_m)&=P(A_{m-1}B_m)=P(C_{m-1},t_{m-1}<T_{s_m}<t_m,T_{s_m+1}>t_m)\\&=E\Big(1_{C_{m-1}}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\lambda\mathrm{e}^{-\lambda s}\mathrm{d}s\int_{t_m-T_{s_{m-1}}-s}^{\infty}\lambda\mathrm{e}^{-\lambda u}\mathrm{d}u\Big)\\&=E\big(1_{C_{m-1}}\lambda(t_m-t_{m-1})\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\big)\\&=\lambda(t_m-t_{m-1})\mathrm{e}^{\lambda(t_m-t_{m-1})}E\big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_{m-1}-T_{s_{m-1}})}\big),\end{aligned}\tag{2.3.9}$$

其中在第三个等号与上面推导一样用了条件期望和W_k独立同指数分布等性质.

(3)当$k_m>1$时, $s_m\geq s_{m-1}+2$, $B_m=\{T_{s_m}<t_m,T_{s_m+1}>t_m\}$, 因此

$$\begin{aligned}P(A_m)&=P(A_{m-1}B_m)=P(C_{m-1}\cap\{T_{s_{m-1}+1}>t_{m-1}\}\cap B_m)\\&=P(C_{m-1},t_m-T_{s_{m-1}}>T_{s_{m-1}+1}-T_{s_{m-1}}>t_{m-1}-T_{s_{m-1}},\\&\qquad T_{s_m}-T_{s_{m-1}+1}<t_m-T_{s_{m-1}+1},T_{s_m+1}-T_{s_m}>t_m-T_{s_m}).\end{aligned}\tag{2.3.10}$$

因为对任意$k>l$, $T_k-T_l=\sum_{i=l+1}^{k}W_i$的密度函数为$\gamma_{k-l}(x)$, 与(2.3.8)的同样推导可知, (2.3.10)等于

$$E\Big(1_{C_{m-1}}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\gamma_1(s)\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\gamma_{k_m-1}(u)\mathrm{d}u\int_{t_m-s-u-T_{s_{m-1}}}^{\infty}\gamma_1(v)\mathrm{d}v\Big).$$

直接积分计算可得

$$\begin{aligned}
&\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\gamma_1(s)\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\gamma_{k_m-1}(u)\mathrm{d}u\int_{t_m-s-u-T_{s_{m-1}}}^{\infty}\gamma_1(v)\mathrm{d}v\\
&=\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\lambda\mathrm{e}^{-\lambda s}\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\gamma_{k_m-1}(u)\mathrm{d}u\int_{t_m-T_{s_{m-1}}-s-u}^{\infty}\lambda\mathrm{e}^{-\lambda v}\mathrm{d}v\\
&=\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\frac{\lambda^{k_m}u^{k_m-2}}{(k_m-2)!}\mathrm{d}u\\
&=\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\frac{\lambda^{k_m}(t_m-t_{m-1})^{k_m}}{k_m!}.
\end{aligned}$$

因此

$$\begin{aligned}
P(A_m)&=E\Big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\frac{\lambda^{k_m}(t_m-t_{m-1})^{k_m}}{k_m!}\Big)\\
&=\frac{\lambda^{k_m}(t_m-t_{m-1})^{k_m}}{k_m!}\mathrm{e}^{-\lambda(t_m-t_{m-1})}E\Big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_{m-1}-T_{s_{m-1}})}\Big).
\end{aligned}\tag{2.3.11}$$

另一方面重, 复(2.3.8)的推导过程可知(事实上只需在(2.3.8)中令$t_m=t_{m-1}$),

$$P\big(A_{m-1}\big)=E\big(1_{C_{m-1}}\mathrm{e}^{-\lambda(t_{m-1}-T_{s_{m-1}})}\big).\tag{2.3.12}$$

由归纳假设以及(2.3.8), (2.3.9),(2.3.11) 和(2.3.12)可知$n=m$时(2.3.7)也成立. 进而由归纳原理, (2.3.7)对一切$n\geq 1$ 都成立. □

更有意思的是, 下一个定理在一定程度上告诉我们具有平稳独立增量性质的简单计数过程是且只能是泊松过程.

★定理2.3.7. *N为泊松过程当且仅当N是初值为0的非零简单计数过程且具有平稳独立增量.*

为证明定理2.3.7, 我们需要如下的引理.

▼引理2.3.8. *$\{W_k;\ k\geq 1\}$ 是服从参数为λ的指数分布的独立随机变量序列当且仅当对任意$n\geq 1$及任意$0<t_1\leq t_2\leq\cdots\leq t_n$,*

$$P(T_k\leq t_k,k=1,2,\cdots,n)=\int_0^{t_1}\mathrm{d}s_1\int_{s_1}^{t_2}\mathrm{d}s_2\cdots\int_{s_{n-1}}^{t_n}\lambda^n\mathrm{e}^{-\lambda s_n}\mathrm{d}s_n,\tag{2.3.13}$$

其中对任意$k\geq 1$, $T_k=\sum_{i=1}^k W_i$.

证明 由数学归纳法可得必要性, 留作习题. 充分性由例1.1.4可得, 此略. □

▼引理2.3.9. *若计数过程N初值为零且具有平稳独立增量, 那么对任意$0\leq t_0<t_1<\cdots<t_n<t_{n+1}$以及非负整数$k_0<k_1<\cdots<k_{n+1}$,*

$$\begin{aligned}
&P\big(N(t_{n+1})\geq k_{n+1}|N(t_m)=k_m,N(t_m-)=k_m-1,0\leq m\leq n\big)\\
&=P\big(N(t_{n+1}-t_n)\geq k_{n+1}-k_n\big).
\end{aligned}$$

证明 由于计数过程单调不降, 对任意$t \ge 0$, $N(t-) = \lim\limits_{s \to t-} N(s)$. 因此,

$$P\big(N(t_{n+1}) \ge k_{n+1}, N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \le m \le n\big)$$
$$= \lim_{\varepsilon \to 0} P\big(N(t_{n+1}) \ge k_{n+1}, N(t_m) = k_m, N(t_m - \varepsilon) = k_m - 1, 0 \le m \le n\big).$$

不妨设$0 < \varepsilon < \min_{1 \le k \le n}\{t_k - t_{k-1}\}$, 由平稳独立增量性可得

$$\begin{aligned} &P\big(N(t_{n+1}) \ge k_{n+1}, N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \le m \le n\big) \\ &= P\big(N(t_{n+1}) - N(t_n) \ge k_{n+1} - k_n\big) \\ &\quad \times \lim_{\varepsilon \to 0} P\big(N(t_m) = k_m, N(t_m - \varepsilon) = k_m - 1, 0 \le m \le n\big) \\ &= P\big(N(t_{n+1} - t_n) \ge k_{n+1} - k_n\big) \\ &\quad \times P\big(N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \le m \le n\big). \end{aligned}$$

由此可知引理2.3.9成立. □

定理2.3.7的证明* 必要性由定理2.3.6和泊松过程简单性质可得. 下证充分性. 为此, 我们需证明W_k, $k \ge 1$, 为独立同分布随机变量且服从参数为λ 的指数分布. 由引理2.3.8, 我们只需证明(2.3.13) 对任意$n \ge 1$ 成立.

当$n = 1$时, 对任意$t \ge 0$, 注意到

$$P(T_1 \le t) = P(N(t) \ge 1) = 1 - P(N(t) = 0).$$

令$f(t) = P(N(t) = 0)$. 对任意$s, t > 0$, 由平稳独立增量性以及$N(0) = 0$可得

$$f(s+t) = P(N(s+t) = 0) = P(N(s+t) = 0, N(t) = 0) = f(s)f(t).$$

由于$N(t)$右连续且不恒为0, $0 \le f(t) < 1$且右连续. 因此

$$f(t) \equiv 0 \ \text{ 或存在}\lambda > 0\text{使得}f(t) = \mathrm{e}^{-\lambda t}.$$

若$f(t) \equiv 0$, 则$P(N(t+s) - N(s) \ge 1) = 1 - f(t) = 1$, 从而对任意$1 > s > 0$, 由

$$\{N(1) - N(s) \ge 2\} \supset \{N(1) - N(\frac{1+s}{2}) \ge 1\} \cap \{N(\frac{1+s}{2}) - N(s) \ge 1\},$$

可知

$$P(N(1) - N(s) \ge 2) = 1.$$

令$s \to 1-$, 得

$$P(N(1) - N(1-) \ge 2) = 1.$$

这与N为简单计数过程矛盾, 因此

$$P(T_1 \le t) = P(N(t) \ge 1) = \int_0^t \lambda \mathrm{e}^{-\lambda s} \mathrm{d}s, \tag{2.3.14}$$

即$n = 1$时(2.3.13)成立.

下设$n = m - 1$时(2.3.13)也成立. 那么$n = m$时, 由归纳假设得

$$P(T_k \le t_k, 1 \le k \le m) = \int_0^{t_1} \mathrm{d}s_1 \cdots \int_{s_{m-2}}^{t_{m-1}} \Big[\lambda^{m-1} \mathrm{e}^{-\lambda s_{m-1}}$$

$$\times P(T_m \le t_m | T_k = s_k, 1 \le k < m)\Big] \mathrm{d}s_{m-1}. \quad (2.3.15)$$

由于N是初值为0的简单计数过程且有平稳独立增量, 由引理2.3.9和(2.3.14)可知

$$\begin{aligned} &P(T_m \le t_m | T_k = s_k, 1 \le k < m) \\ &\quad = P(N(t_m) \ge m | N(s_k) = k, N(s_k-) = k-1, 1 \le k < m) \\ &\quad = P(N(t_m - s_{m-1}) \ge 1) \\ &\quad = \int_0^{t_m - s_{m-1}} \lambda \mathrm{e}^{-\lambda s_m} \mathrm{d}s_m = \int_{s_{m-1}}^{t_m} \lambda \mathrm{e}^{-\lambda(s_m - s_{m-1})} \mathrm{d}s_m. \end{aligned}$$

将其代入(2.3.15)整理后即得(2.3.13). □

由定理2.3.6和2.3.7容易推出, 强度为λ的泊松过程的增量$\tilde{N}(t) = N(t+s) - N(s)$服从泊松分布$\mathrm{P}(\lambda t)$并且随机过程$\{\tilde{N}(t); t \ge 0\}$是$N$的一个版本, 即$\tilde{N}(t)$是与$N$有相同有限维分布族的泊松过程.

▶**例2.3.10.** 假设某银行顾客按强度为λ的泊松过程N到达, 已知到s时刻为止, 到达银行的顾客数是10. 问(1)再经过s单位时间, 新增顾客的平均数c. (2) 再等待5位顾客的平均时间t.

解 (1) 所求新增顾客平均数$c = E(N(2s) - N(s))$. 由平稳独立增量性可知

$$c = E(N(s)) = \lambda s.$$

(2)对任意t, 令$\tilde{N}(t) = N(t+s) - N(s)$. 那么由$N$的平稳独立增量性可知$\tilde{N}$ 仍然具有平稳增量性, 从而还是强度为λ的泊松过程. 因此要再等待5位顾客对$\tilde{N}$ 而言意味着还要计数到5. 所以所求平均时间$t = E(T_5)$. 由T_5服从$\mathrm{Ga}(5, \lambda)$分布可知$t = 5/\lambda$. □

2.3.3 到达时间的条件分布

由引理2.3.8可知强度为λ的泊松过程的前n个事件发生时间$(T_1, \cdots, T_n)$的联合概率密度函数

$$f(t_1, \cdots, t_n) = \lambda^n \mathrm{e}^{-\lambda t_n}, \qquad 0 \le t_1 < \cdots < t_n.$$

下面这个结果刻画了在给定$N(t) = n$条件下它们联合分布的变化.

★**定理2.3.11.** 设$\{T_n;\ n \ge 1\}$是强度为λ的泊松过程$N = \{N(t);\ t \ge 0\}$的到达时刻序列, 那么对任意$0 \le s_1 < t_1 \le s_2 < \cdots \le s_n < t_n \le t$,

$$P\big(T_j \in (s_j, t_j], 1 \le j \le n | N(t) = n\big) = \frac{n!}{t^n} \prod_{j=1}^{n} (t_j - s_j).$$

证明 任取$0 = t_0 \le s_1 < t_1 \le s_2 < \cdots \le s_n < t_n \le t$. 我们有

$$
\begin{aligned}
&P\big(T_j \in (s_j, t_j], 1 \le j \le n \big| N(t) = n\big) \\
&= P\Big(N(t_j) - N(s_j) = 1, N(s_j) - N(t_{j-1}) = 0, 1 \le j \le n \Big| N(t) = n\Big) \\
&= \frac{P\big(N(t_j) - N(s_j) = 1, N(s_j) - N(t_{j-1}) = 0, 1 \le j \le n, N(t) - N(t_n) = 0\big)}{P(N(t) = n)}.
\end{aligned}
$$

由于泊松过程是平稳独立增量过程,

$$
\begin{aligned}
&P\big(N(t_j) - N(s_j) = 1, N(s_j) - N(t_{j-1}) = 0, 1 \le j \le n, N(t) - N(t_n) = 0\big) \\
&= P(N(t) - N(t_n) = 0) \prod_{j=1}^{n} \Big[P(N(t_j) - N(s_j) = 1) P(N(s_j) - N(t_{j-1}) = 0)\Big] \\
&= P(N(t - t_n) = 0) \prod_{j=1}^{n} \Big[P(N(t_j - s_j) = 1) P(N(s_j - t_{j-1}) = 0)\Big].
\end{aligned}
$$

再由$N(t)$服从强度为λt的泊松分布可知

$$
\begin{aligned}
P(N(t_j - s_j) = 1) &= \lambda(t_j - s_j)\mathrm{e}^{-\lambda(t_j - s_j)}, \\
P(N(s_j - t_{j-1}) = 0) &= \mathrm{e}^{-\lambda(s_j - t_{j-1})}, \\
P(N(t - t_n) = 0) &= \mathrm{e}^{-\lambda(t - t_n)}, \\
P(N(t) = n) &= \frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}.
\end{aligned}
$$

综合上述结果可得

$$
\begin{aligned}
&P\big(T_j \in (s_j, t_j], 1 \le j \le n | N(t) = n\big) \\
&= \mathrm{e}^{-\lambda(t - t_n)} \prod_{j=1}^{n} \Big[\lambda(t_j - s_j)\mathrm{e}^{-\lambda(t_j - t_{j-1})}\Big] \Big/ \Big[\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}\Big] \\
&= \frac{n!}{t^n} \prod_{j=1}^{n} (t_j - s_j).
\end{aligned}
$$

定理得证. □

已知n个独立$(0, t]$上均匀分布随机变量$X_1, X_2, \cdots, X_n$的次序统计量

$$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$$

的概率密度函数为

$$p(x_1, \cdots, x_n) = n! t^{-n}, \quad 0 < x_1 < x_2 < \cdots < x_n \le t,$$

从而对任意$0 \le s_1 < t_1 < \cdots \le s_n < t_n \le t$, 概率

$$P(X_{(j)} \in (s_j, t_j), 1 \le j \le n) = \frac{n!}{t^n} \prod_{j=1}^{n} (t_j - s_j).$$

比较上式与定理2.3.11的结论, 我们可知$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$的联合分布与给定$N(t) = n$的条件下的$(T_1, T_2, \cdots, T_n)$的联合分布相同. 即给定条件$N(t) = n$后,

从随机规律性上看, $T_1,\cdots,T_n$可看作n个独立的$(0,t]$上均匀分布随机变量的次序统计量.

▶例2.3.12. (接例*2.3.10*) 求第*10*位顾客到达的时间T_{10}的分布和均值.

解 由定理2.3.11, 在$N(s)=10$条件下, T_{10}与10个独立的$(0,s]$上均匀分布随机变量的最大次序统计量分布相同. 因此T_{10}的概率密度函数为

$$f(x)=10\frac{x^9}{s^{10}},\quad x\in(0,\ s].$$

平均时间

$$E(T_{10})=\int_0^s xf(x)\mathrm{d}x=10s\int_0^1 x^{10}\mathrm{d}x=\frac{10s}{11}.\qquad\square$$

▶例2.3.13. 已知$\{N(t);t\geq 0\}$是强度为λ的泊松过程, (1) 求前n个事件中最小到达时间间隔的分布. (2) 若已知$N(t)=n$, 求前n个事件中最小到达时间间隔的分布.

解 (1) 以W_k表示第k次事件发生的时间间隔. 那么W_k服从指数分布$\mathrm{Exp}(\lambda)$而且相互独立. 记前n个事件中最小时间间隔为T, 那么

$$T=\min\{W_1,W_2,\cdots,W_n\}.$$

进而对任意$x>0$,

$$P(T>x)=P(W_k>x,k=1,\cdots,n)=\mathrm{e}^{-n\lambda x}.$$

因此T服从指数分布$\mathrm{Exp}(n\lambda)$.

(2) 记T_k为第k个事件的到达时间, 由定理2.3.11可知, $(T_1,\cdots,T_n)$的联合概率密度函数在$N(t)=n$条件下为

$$f(t_1,\cdots,t_n)=\frac{n!}{t^n},\qquad 0<t_1<t_2<\cdots<t_n\leq t.$$

此时, 对$0<x<t/n$, 由

$$\{T>x\}=\{T_k-T_{k-1}>x,k=1,2,\cdots,n\},$$

其中$T_0=0$, 以及在$N(T)=n$条件下

$$\begin{aligned}&\{T_k-T_{k-1}>x,k=1,2,\cdots,n\},\\&=\{x<T_1<T_2-x<T_2<\cdots<T_n-x<T_n\leq t\}\\&=\{x<T_1<t-(n-1)x,T_1+x<T_2<t-(n-2)x,\cdots,\\&\qquad T_{n-2}+x<T_{n-1}<t-x,T_{n-1}+x<T_n\leq t\},\end{aligned}$$

可知

$$\begin{aligned}P(T>x|N(t)=n)&=P(T_k-T_{k-1}>x,k=1,\cdots,n|N(t)=n)\\&=\int_x^{t-(n-1)x}\mathrm{d}t_1\int_{t_1+x}^{t-(n-2)x}\mathrm{d}t_2\cdots\int_{t_{n-2}+x}^{t-x}\mathrm{d}t_{n-1}\int_{t_{n-1}+x}^{t}\frac{n!}{t^n}\mathrm{d}t_n\end{aligned}$$

$$= \left(1-\frac{nx}{t}\right)^n.$$

因此

$$P(T \le x|N(t)=n) = \begin{cases} 0, & x \le 0, \\ 1-\left(1-\frac{nx}{t}\right)^n, & 0 < x < t/n, \\ 1, & x \ge t/n. \end{cases}$$ □

作为定理2.3.11的应用, 我们有如下推论.

♦**推论2.3.14.** 设n元函数$f(x_1, x_2 \cdots, x_n)$对于任意n级排列$(t_1, t_2, \cdots, t_n)$都有

$$f(x_1, x_2, \cdots, x_n) = f(x_{t_1}, x_{t_2}, \cdots, x_{t_n}),$$

那么

$$E(f(T_1, T_2, \cdots, T_n)|N(t)=n) = E(f(X_1, X_2, \cdots, X_n)),$$

其中$X_1, \cdots, X_n$服从$(0, t]$上均匀分布的独立随机变量.

证明 由定理2.3.11可知

$$E(f(T_1, T_2, \cdots, T_n)|N(t)=n) = E(f(X_{(1)}, X_{(2)}, \cdots, X_{(n)})).$$

由于次序统计量只是$X_1, \cdots, X_n$的一种排列, 由f的条件假设可知

$$E(f(X_{(1)}, X_{(2)}, \cdots, X_{(n)})) = E(f(X_1, \cdots, X_n)),$$

进而结论成立. □

▶**例2.3.15.** 旅客依强度为λ的泊松过程到达车站, 若火车在时刻t离站, 求在$(0, t)$ 区间内到达的旅客的平均总等待时间.

解 记第k位旅客到达时刻为T_k, t之前到达总旅客数位$N(t)$, 则总等待时间为

$$T = \sum_{i=1}^{N(t)} (t - T_i).$$

所求平均总等待时间为

$$E(T) = E\Big(\sum_{i=1}^{N(t)} (t-T_i)\Big) = E\Big(E\Big(\sum_{i=1}^{N(t)} (t-T_i)|N(t)\Big)\Big)$$
$$= \sum_{n=1}^{\infty} E\Big(\sum_{i=1}^{n} (t-T_i)|N(t)=n\Big) P(N(t)=n).$$

显然改变T_i的顺序不会影响$\sum_{i=1}^{n}(t-T_i)$的结果, 由推论2.3.14,

$$E(T) = \sum_{n=1}^{\infty} E\Big(\sum_{i=1}^{n} (t-X_i)\Big) P(N(t)=n)$$
$$= \sum_{n=1}^{\infty} \Big(nt - E\Big(\sum_{i=1}^{n} X_i\Big)\Big) P(N(t)=n)$$

$$= \sum_{n=1}^{\infty} \frac{nt}{2} P(N(t)=n) = \frac{t}{2} E(N(t)) = \frac{\lambda t^2}{2}. \qquad \square$$

▶**例2.3.16.** 对任意$k \geq 1$, 设T_k是强度为λ的泊松过程N的第k个到达时刻, 证明: 若$\int_0^{\infty} |f(t)| \mathrm{d}t < \infty$, 那么

$$E\Big(\sum_{n=1}^{\infty} f(T_n)\Big) = \lambda \int_0^{\infty} f(s) \mathrm{d}s.$$

证明* 先设f非负. 注意到$t \to \infty$时$N(t) \to \infty$ a.s. (参见本节习题9). 由单调收敛定理(定理1.4.23)以及全期望公式,

$$\begin{aligned} E\Big(\sum_{n=1}^{\infty} f(T_n)\Big) &= \lim_{t\to\infty} E\Big(\sum_{n=1}^{N(t)} f(T_n)\Big) \\ &= \lim_{t\to\infty} \sum_{k=1}^{\infty} E\Big(\sum_{n=1}^{k} f(T_n) \Big| N(t)=k\Big) P(N(t)=k). \end{aligned}$$

由推论2.3.14可知

$$E\Big(\sum_{n=1}^{k} f(T_n) \Big| N(t)=k\Big) = k \int_0^t \frac{f(s)}{t} \mathrm{d}s = \frac{k}{t} \int_0^t f(s) \mathrm{d}s.$$

因此

$$\begin{aligned} E\Big(\sum_{n=1}^{\infty} f(T_n)\Big) &= \lim_{t\to\infty} \Big[\frac{1}{t} \int_0^t f(s) \mathrm{d}s \sum_{k=1}^{\infty} k P(N(t)=k)\Big] \\ &= \lambda \lim_{t\to\infty} \int_0^t f(s) \mathrm{d}s = \lambda \int_0^{\infty} f(s) \mathrm{d}s. \end{aligned}$$

对一般的f, 注意到$f = f^+ - f^- = f \vee 0 - (-f) \vee 0$. 由$f$非负情形的讨论可知

$$E\Big(\sum_{n=1}^{\infty} |f(T_n)|\Big) < \infty, \quad E\Big(\sum_{n=1}^{\infty} f^+(T_n)\Big) < \infty, \quad E\Big(\sum_{n=1}^{\infty} f^-(T_n)\Big) < \infty.$$

这表明级数$\sum\limits_{n=1}^{\infty} f(T_n)$, $\sum\limits_{n=1}^{\infty} f^+(T_n)$ 和$\sum\limits_{n=1}^{\infty} f^-(T_n)$ 都几乎必然收敛. 因此

$$\sum_{n=1}^{\infty} f(T_n) = \sum_{n=1}^{\infty} (f^+(T_n) - f^-(T_n)) = \sum_{n=1}^{\infty} f^+(T_n) - \sum_{n=1}^{\infty} f^-(T_n)$$

几乎处处成立. 从而

$$\begin{aligned} E\Big(\sum_{n=1}^{\infty} f(T_n)\Big) &= E\Big(\sum_{n=1}^{\infty} f^+(T_n)\Big) - E\Big(\sum_{n=1}^{\infty} f^-(T_n)\Big) \\ &= \lambda \int_0^{\infty} [f^+(t) - f^-(t)] \mathrm{d}t = \lambda \int_0^{\infty} f(t) \mathrm{d}t. \end{aligned} \qquad \square$$

2.3.4 稀疏过程

有时我们需要对计数过程$N = \{N(t)\}$记录的随机事件进行分类. 比如我们记录到达银行的顾客, 可以将顾客按性别分成两类, 对应地计数过程就会被分拆. 特别地, 如果我们要求每个事件**独立于其他事件**以概率p_i归为第i 类事件, 其中 $i = 1, \cdots, m$且 $\sum_{i=1}^m p_i = 1$, 那么我们称记录第i类事件发生次数的计数过程$N_i = \{N_i(t)\}$为N的稀疏(Thinning)过程.

显然, 对任意t,

$$N(t) = \sum_{i=1}^m N_i(t).$$

而且根据稀疏过程的构造要求, 我们容易知道: 对任意$s < t$, 在给定$(s, t]$时段内总的随机事件数$N(t) - N(s)$条件下, 该区间内各类事件发生次数

$$(N_1(t) - N_1(s), N_2(t) - N_2(s), \cdots, N_m(t) - N_m(s))$$

的联合分布是以$p_1, p_2, \cdots, p_m$为发生概率的多项分布(参见例1.1.2), 而且与其它不相交区间发生的事件独立. 即对任意非负整数$k_i, 1 \le i \le m$, 令$k = \sum_{i=1}^m k_i$, 那么

$$\begin{aligned} &P(N_i(t) - N_i(s) = k_i, 1 \le i \le m | N(t) - N(s) = k) \\ &= \frac{k!}{k_1! k_2! \cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}, \end{aligned} \tag{2.3.16}$$

而且对任意$0 = t_0 < t_1 < \cdots < t_n$, 以及非负整数$k_{i,j}, 1 \le i \le m, 1 \le j \le n$,

$$\begin{aligned} &P\big(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \le i \le m, 1 \le j \le n | N(t_l) - N(t_{l-1}), l = 1, \cdots, n\big) \\ &= \prod_{j=1}^n P\big(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \le i \le m | N(t_j) - N(t_{j-1})\big). \end{aligned} \tag{2.3.17}$$

★**定理2.3.17.** *若N是强度为λ的泊松过程, 那么它的稀疏过程$N_1, N_2, \cdots, N_m$是独立的泊松过程, 强度分别$\lambda p_1, \cdots, \lambda p_m$.*

证明* 我们先证明对任意$i \ge 1$, N_i是强度为λp_i的泊松过程.

对任意$n \ge 1$, 任取$0 = t_0 < t_1 < \cdots < t_n$以及非负整数$k_1, \cdots, k_n$. 由$N$ 的独立增量性和(2.3.17) 可知

$$\begin{aligned} &P(N_i(t_j) - N_i(t_{j-1}) = k_j, j = 1, \cdots, n) \\ &= E\Big[P\big(N_i(t_j) - N_i(t_{j-1}) = k_j, j = 1, \cdots, n | N(t_l) - N(t_{l-1}), l = 1, \cdots, n\big)\Big] \\ &= E\Big[\prod_{j=1}^n P\big(N_i(t_j) - N_i(t_{j-1}) = k_j | N(t_j) - N(t_{j-1})\big)\Big] \end{aligned}$$

$$= \prod_{j=1}^{n} E\Big[P\big(N_i(t_j) - N_i(t_{j-1}) = k_j | N(t_j) - N(t_{j-1})\big)\Big]$$
$$= \prod_{j=1}^{n} P(N_i(t_j) - N_i(t_{j-1}) = k_j).$$

因此N_i具有独立增量. 注意到对任意$t > s \geq 0$以及非负整数k,

$$P(N_i(t) - N_i(s) = k) = \sum_{r=0}^{\infty} P(N_i(t) - N_i(s) = k, N(t) - N(s) = k + r)$$
$$= \sum_{r=0}^{\infty} P(N_i(t) - N_i(s) = k | N(t) - N(s) = k + r) P(N(t) - N(s) = k + r).$$

由(2.3.16)得

$$P(N_i(t) - N_i(s) = k) = \sum_{r=0}^{\infty} C_{k+r}^{k} p_i^k (1 - p_i)^r \frac{[\lambda(t-s)]^{k+r}}{(k+r)!} \mathrm{e}^{-\lambda(t-s)}$$
$$= \frac{[p_i \lambda(t-s)]^k \mathrm{e}^{-\lambda(t-s)}}{k!} \sum_{r=0}^{\infty} \frac{[\lambda(1-p_i)(t-s)]^r}{r!}$$
$$= \frac{[p_i \lambda(t-s)]^k \mathrm{e}^{-\lambda(t-s)}}{k!} \mathrm{e}^{\lambda(1-p_i)(t-s)} = \frac{[p_i \lambda(t-s)]^k}{k!} \mathrm{e}^{-\lambda p_i (t-s)}.$$

因此N_i有平稳增量且增量$N_i(t) - N_i(s)$服从强度为$p_i\lambda(t-s)$的泊松分布. 由定理2.3.7 可知N_i是强度为λp_i的泊松过程.

对任意$n \geq 1$以及任意$0 = t_0 < t_1 < \cdots < t_n$, 任取非负整数$k_{i,j}$, $1 \leq i \leq m, 1 \leq j \leq n$, 并令$k_j = \sum_{i=1}^{m} k_{i,j}$以及

$$A = \{N(t_j) - N(t_{j-1}) = k_j,\ 1 \leq j \leq n\}.$$

由条件概率公式和(2.3.17)可得

$$P(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n)$$
$$= P\big(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n \big| A\big) P(A)$$
$$= P(A) \prod_{j=1}^{n} P\big(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \leq i \leq m \big| N(t_j) - N(t_{j-1}) = k_j\big).$$

再由(2.3.16)以及N的增量独立且服从泊松分布可知

$$P(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n)$$
$$= \prod_{j=1}^{n} \Big[\frac{k_j!}{k_{1,j}! k_{2,j}! \cdots k_{m,j}!} p_1^{k_{1,j}} \cdots p_m^{k_{m,j}} \frac{[\lambda(t_j - t_{j-1})]^{k_j}}{k_j!} \mathrm{e}^{-\lambda(t_j - t_{j-1})}\Big]$$
$$= \prod_{i=1}^{m} \Big[\prod_{j=1}^{n} \frac{[\lambda p_i (t_j - t_{j-1})]^{k_{i,j}}}{k_{i,j}!} \mathrm{e}^{-\lambda p_i (t_j - t_{j-1})}\Big]$$
$$= \prod_{i=1} P\big(N_i(t_j) - N_i(t_{j-1}) = k_{i,j}, 1 \leq j \leq n\big).$$

因此$(N_i(t_1),\cdots,N_i(t_n)), i = 1,\cdots,m$ 相互独立. 由随机过程独立性判别条件(参见本章第一节习题4)可知稀疏过程$N_i, i=1,\cdots,m$, 独立. $\square$

▶**例2.3.18.** *某商场顾客以一个强度为每小时20人的泊松过程到达, 其中每个到达的顾客独立地以10%的可能为男性, 以90%的可能为女性. (1) 问半个小时内至少有1个男顾客到达的概率p. (2) 在一小时内恰有4个男顾客到达条件下, 求这一小时到达顾客总数的期望n.*

解 以$N_1(t), N_2(t)$分别记录t之前男, 女顾客到达的人数. 那么N_1, N_2分别是强度为2和18的泊松过程且独立. 所以

(1)所求概率$p = P\big(N_1(0.5)\geq 1\big) = 1 - P\big(N_1(0.5)=0\big) = 1-\mathrm{e}^{-1}$.

(2)所求期望为

$$\begin{aligned} n &= E(N(1)|N_1(1)=4) = E(N_1(1)+N_2(1)|N_1(1)=4)\\ &= 4+E(N_2(1)|N_1(1)=4) = 4+E\big(N_2(1)\big) = 4+18 = 22. \end{aligned}$$ $\square$

2.3.5 非齐次泊松过程

泊松过程要求强度为常数, 这个条件使泊松过程简单易处理, 但与现实情况会有较大差异. 所谓非齐次泊松过程是泊松过程的一种推广, 它的定义如下:

♣**定义2.3.19.** *称计数过程$N=\{N(t);t\geq 0\}$是一个强度为$\lambda(t)$的非齐次泊松过程, 若*

(1) $N(0)=0$;

(2) N具有独立增量;

(3) 对任意$0\leq s<t$, $N(t)-N(s)$服从强度为$\int_s^t \lambda(u)\mathrm{d}u$的泊松分布.

由定义可知非齐次的泊松过程N在任意时刻t服从强度为$\int_0^t \lambda(u)\mathrm{d}u$ 的泊松分布. 因此此计数过程首个事件发生时间T_1的分布满足

$$P(T_1>t) = P(N(t)=0) = \mathrm{e}^{-\int_0^t \lambda(u)\mathrm{d}u},$$

从而T_1的概率密度函数为

$$f(t) = \lambda(t)\mathrm{e}^{-\int_0^t \lambda(u)\mathrm{d}u}, \quad t\geq 0.$$

▶**例2.3.20.** *若某网店在不同时段顾客访问事件是独立的, 在晚上11:00到11:30, 顾客按10人/小时的速度访问, 而11:30-12:00按6人/小时的速度访问. 求在晚上11:00-12:00没有顾客访问该网店的概率.*

解 由于假定不同时间段顾客访问事件是独立的, 而且不同时间访问速度不同, 因此可以用非齐次泊松过程刻画$N(t)$. 以晚11:00为起始时间, 在$0 < t \le 0.5$时间段内$N(t)$的强度为10, 在$0.5 < t \le 1$时间段内强度为6. 因此所求概率

$$p = P(N(1) = 0) = \mathrm{e}^{-\int_0^1 \lambda(u)\mathrm{d}u} = \mathrm{e}^{-\int_0^{0.5} 10\mathrm{d}u + \int_{0.5}^1 6\mathrm{d}u} = \mathrm{e}^{-8}. \qquad \square$$

非齐次泊松过程也会自然出现在泊松过程的稀疏过程中, 推广定理2.3.17得

★定理2.3.21. *若强度为λ的泊松过程在任意时刻t记录的事件以概率$p_i(t)$ 归为第i类事件, 而且与其他事件独立, 其中 $i = 1, \cdots, m$ 且 $\sum_{i=1}^m p_i(t) = 1$. 那么记录第i类事件发生次数的计数过程$N_i = \{N_i(t)\}$就是一个强度为$\lambda p_i(t)$的非齐次泊松过程, 而且$N_1, N_2, \cdots, N_m$相互独立.*

证明* 首先, 从定理条件假设容易看出, (2.3.17)仍然成立.

对任给的$0 \le s < t$和非负整数n, 任取非负整数$n_1, \cdots, n_m$使得

$$n_1 + \cdots + n_m = n.$$

将指标$1, 2, \cdots, n$随机地分成m组使得每组中指标个数分别是$n_1, n_2, \cdots, n_m$. 将所有这种分法的集合记成A. 由组合数计算可知A中共有$K = \frac{n!}{n_1!n_2!\cdots n_m!}$ 种分法, 并用记号$b_i, 1 \le i \le K$ 表示. 以$B_i(j), 1 \le j \le m$ 表示在b_i这种分法下落入第j组的n_j 个指标构成的集合. 对任意$(x_1, x_2, \cdots, x_n) \in (s, t]^n$, 定义

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^K \prod_{r \in B_i(1)} p_1(x_r) \prod_{r \in B_i(2)} p_2(x_r) \cdots \prod_{r \in B_i(m)} p_m(x_r).$$

(1) 任意调换$(x_1, x_2, \cdots, x_n)$的次序$\big($记调换后的结果为$(y_1, y_2, \cdots, y_n)\big)$, 由组合的无次序性以及乘法的可交换性可知

$$\begin{aligned} f(y_1, y_2, \cdots, y_n) &= \sum_{i=1}^K \prod_{r \in B_i(1)} p_1(y_r) \prod_{r \in B_i(2)} p_2(y_r) \cdots \prod_{r \in B_i(m)} p_m(y_r) \\ &= \sum_{i=1}^K \prod_{r \in B_i(1)} p_1(x_r) \prod_{r \in B_i(2)} p_2(x_r) \cdots \prod_{r \in B_i(m)} p_m(x_r) \\ &= f(x_1, x_2, \cdots, x_n). \end{aligned}$$

(2) 考察条件概率

$$P\big(N_i(t) - N_i(s) = n_i, 1 \le i \le m | T_1, \cdots, T_n, N(t) - N(s) = n\big),$$

其中$T_1, \cdots, T_n$表示N在(s, t)内n次事件发生的时刻. 此时事件

$$\{N_i(t) - N_i(s) = n_i, 1 \le i \le m\}$$

可看作在n个具有m种可能结果的独立随机实验中(每次实验结果出现的概率可以不同), 各种结果出现次数分别为$n_1, n_2, \cdots, n_m$的随机事件. 因此构成该随机

事件的所有可能的组合方法恰为集合A, 而A每种组合b_i出现的概率为

$$\prod_{r\in B_i(1)} p_1(T_r) \prod_{r\in B_i(2)} p_2(T_r)\cdots \prod_{r\in B_i(m)} p_m(T_r).$$

因此

$$\begin{aligned}
&P(N_i(t)-N_i(s)=n_i, 1\le i\le m|T_1,\cdots,T_n,N(t)-N(s)=n)\\
&=\sum_{i=1}^{K}\prod_{r\in B_i(1)} p_1(T_r) \prod_{r\in B_i(2)} p_2(T_r)\cdots \prod_{r\in B_i(m)} p_m(T_r)\\
&=f(T_1,T_2,\cdots,T_n).
\end{aligned}$$

由泊松过程的平稳独立增量性质, 推论2.3.14以及条件数学期望的性质可知

$$\begin{aligned}
&P(N_i(t)-N_i(s)=n_i, 1\le i\le m|N(t)-N(s)=n)\\
&=E(f_n(T_1,\cdots,T_n)|N(t)-N(s)=n)=E(f(X_1,\cdots,X_n))\\
&=\sum_{i=1}^{K}E\Big(\prod_{r\in B_i(1)} p_1(X_r) \prod_{r\in B_i(2)} p_2(X_r)\cdots \prod_{r\in B_i(m)} p_m(X_r)\Big)\\
&=\frac{n!}{n_1!n_2!\cdots n_m!}\Big[\frac{1}{t-s}\int_s^t p_1(u)\mathrm{d}u\Big]^{n_1}\cdots\Big[\frac{1}{t-s}\int_s^t p_m(u)\mathrm{d}u\Big]^{n_m}, \quad (2.3.18)
\end{aligned}$$

其中$X_1,\cdots,X_n$为n个独立的服从(s,t)上均匀分布的随机变量.

将(2.3.18)替换(2.3.16), 重复定理2.3.17的证明可知定理成立. □

▶**例2.3.22.** 假设商场顾客按强度为K人/小时的泊松过程到达, 每个顾客在商场停留的时间服从参数为λ的指数分布. 商场监控系统在c时开始工作、$c+d$时停止. 假设顾客被监控系统捕捉到的概率p 与商场监控系统工作时顾客在商场停留的时间u满足关系式$p=1-\mathrm{e}^{-u}$, 求商场监控系统工作d 小时捕捉到的平均顾客数m (假定顾客行为是相互独立的).

解 记任意时刻t进入商场的顾客停留时间为U_t. 那么由题设可知, 当$t\le c$时, 顾客被监控到的概率

$$p_t=E(1_{\{c-t\le U_t\}}(1-\mathrm{e}^{-(U_t-c+t)\wedge d}))=\frac{1-\mathrm{e}^{-d(1+\lambda)}}{1+\lambda}\mathrm{e}^{\lambda(t-c)},$$

其中对任意$x,y\in\mathbb{R}$, $x\wedge y=\min\{x,y\}$. 当$d+c>t>c$时, 顾客被监控到的概率

$$p_t=E(1-\mathrm{e}^{-U_t\wedge(d+c-t)})=\frac{1-\mathrm{e}^{(1+\lambda)(t-c-d)}}{1+\lambda}.$$

因此在$[0,c+d]$时段内, 被监控到的顾客按强度为Kp_t的非齐次泊松过程到达. 从而被监控的平均顾客数为

$$m=K\int_0^{d+c}p_t\mathrm{d}t=K\Big[\frac{(1-\mathrm{e}^{-d(1+\lambda)})(1-\mathrm{e}^{-\lambda c})}{\lambda(1+\lambda)}+\frac{d}{1+\lambda}-\frac{1-\mathrm{e}^{-(1+\lambda)d}}{(1+\lambda)^2}\Big]. \quad □$$

2.3.6 复合泊松过程

♣**定义2.3.23.** *若$N=\{N(t);t\geq 0\}$为泊松过程, $\{Y_n;\ n\geq 1\}$为一列独立同分布的随机变量且与N独立. 对任意$t\geq 0$, 令*

$$X(t)=\sum_{n=1}^{N(t)}Y_n,$$

称$X=\{X(t);\ t\geq 0\}$为复合泊松过程.

复合泊松过程$X(t)$的特征函数

$$E(\mathrm{e}^{iuX(t)})=E(\mathrm{e}^{iu\sum_{k=1}^{N(t)}Y_k})=E\big(E\big(\mathrm{e}^{iu\sum_{k=1}^{N(t)}Y_k}|N(t)\big)\big).$$

假定$E(\mathrm{e}^{itY_n})=\psi(t)$, 由于$\{Y_n\}$独立同分布, 而且与泊松过程$N$独立,

$$E(\mathrm{e}^{iu\sum_{k=1}^{N(t)}Y_k}|N(t))=\psi^{N(t)}(u),$$

进而

$$E(\mathrm{e}^{iuX(t)})=E(\psi^{N(t)}(u)).$$

若N是强度为λ的泊松过程, 那么由$N(t)\sim \mathrm{P}(\lambda t)$可得,

$$E(\mathrm{e}^{iuX(t)})=\sum_{k=0}^{\infty}\psi^k(u)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}=\mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda t\psi(u)}=\mathrm{e}^{\lambda t(\psi(u)-1)}.$$

由此及特征函数与矩的关系(性质1.3.5)可得(细节请读者自己补充)

$$E(X(t))=\lambda tE(Y_1),$$

以及

$$E(X^2(t))=[\lambda tE(Y_1)]^2+\lambda tE(Y_1^2).$$

这也是例1.2.21中的一般公式在复合泊松情形下的结果.

复合泊松过程在包括管理, 金融, 保险等许多领域有着广泛的应用.

▶**例2.3.24.** *假设汽车以强度为10辆/分钟的泊松过程到达某高速公路收费站, 并且每辆汽车交的通行费是独立同分布的, 平均每辆车的通行费为10元. 问该高速收费站1小时能平均收费多少?*

解 以$N=\{N(t)\}$表示强度为10辆/分钟的泊松过程, Y_k表示第k辆车交的通行费. 那么复合泊松过程

$$X(t)=\sum_{n=1}^{N(t)}Y_n,$$

就表示到t时刻, 收费站收取的汽车通行费用. 依题设, 所求平均收费为

$$E(X(60))=\lambda tE(Y_1)\Big|_{\substack{\lambda=10\\ t=60}}=6000\text{元}.$$ □

对于泊松过程的更多推广我们将在第五章进一步介绍, 本节不再赘述.

练习题2.3 以下总设$N=\{N(t);t\geq 0\}$是强度为λ的泊松过程.

1 已知T服从参数为μ的指数分布且与N独立, 对任意$k\in\mathbb{N}$, 求$P(N(T)=k)$.

2 对任意$0<s<t$, 求条件概率$P(N(s)=k|N(t)=n)$, 其中$0\leq k\leq n$为整数.

3 以T_k表示N的第k次随机事件发生的时间. 对任意$t>x>0$, 求
(1) $P(t-T_k>x|N(t)=k)$, (2) $P(t-T_{N(t)}>x)$.

4 以W_k表示N的第k次事件的间隔时间, 对任意$x>0$, 求$P(W_{N(t)+1}>x)$.

5 电影院观众到达服从强度为50人每小时的泊松过程, 其中40%是男性, 60%是女性. 假定观众到达是完全独立的. (1)求前三个到达的观众是女性的概率. (2)已知最后两名观众离放映不到5分钟到达, 问他们到达时距离放映不到2分钟的概率. (3) 假定观众50% 可能不买爆米花, 30% 买1袋爆米花, 20%买两袋爆米花. 以N_0,N_1,N_2分别表示一小时内没买, 买一袋, 买两袋爆米花人数, 求(N_0,N_1,N_2)的联合分布.

6 考虑一个简单的高速公路入口模型. 假设汽车以速率为每分钟5辆的泊松过程到达高速公路甲入口, 进入高速后每辆汽车立即独立地以40%的可能去A 地, 60%的可能去往方向不同的B地, 并且以均匀分布选择$(80,120)$ 间的某一速度匀速行驶. 某车从甲入口进入高速后以100公里每小时速度前往B地, 求20分钟内超过该车或被该车超过的汽车数量平均数.

7 保险公司要为承保的意外事件理赔. 假设意外事件以强度为λ的泊松过程到达, 每个意外事件的理赔金额服从参数为μ的指数分布且相互独立, 试用复合泊松过程模型计算一个单位时间内总的理赔金额的方差.

8 设$T_0=0$, $\{T_k;k\geq 1\}$ 是单调不降的非负随机变量序列. 对任意$t\geq 0$, 令$N(t)=\sup\{k\geq 0;\ T_k\leq t\}$. 证明
(1) 对任意$k\geq 1$, $T_k=\inf\{t;\ N(t)\geq k\}$. (2) $N(t)+1=\inf\{k;\ T_k>t\}$.

9 证明, $t\to\infty$时$N(t)$几乎必然收敛到∞.

10* 证明引理2.3.8的必要性.

11* 若$N(0)=0$且对任意$n\geq 1$以及$0=t_0<t_1<\cdots<t_n$和非负整数$k_1,\cdots,k_n$, (2.3.6)总成立, 证明$N=\{N(t);\ t\geq 0\}$是平稳独立增量过程且对任意$t,s\geq 0$, $N(t+s)-N(s)$服从强度为λt的泊松分布.

12* 设N_1, N_2分别是强度为λ_1, λ_2的泊松过程且相互独立. 证明$N=N_1+N_2$是强度为$\lambda_1+\lambda_2$的泊松过程.

13* 证明复合泊松过程仍然是平稳独立增量过程.

14* 若例2.3.10中N的强度λ未知, 试给出λ的极大似然估计, 并由此计算c和t.

第三章　离散时间马尔可夫链

上一章介绍的两类随机过程都具有平稳独立增量性质. 这种性质给我们讨论带来很大的方便, 但在现实问题中这样的性质并不常有. 人们更容易观察或更能近似观察到所谓的马尔可夫性质. 对这种现象建模可以使用所谓的马尔可夫过程. 本章介绍其中一类简单的模型——离散时间离散状态的马尔可夫过程.

3.1　马尔可夫链与转移概率矩阵

3.1.1　条件独立与马尔科夫链

♣**定义3.1.1.** 称随机事件A, B在事件C(要求$P(C)>0$)发生条件下独立, 若

$$P(AB|C)=P(A|C)P(B|C).$$

简称A, B条件独立.

对条件独立, 我们还有下面的性质.

■**性质3.1.2.** 设C, BC,以及AC均为概率非0事件, 那么以下三条等价:

(1) A, B在事件C下条件独立.

(2) $P(A|BC)=P(A|C)$.

(3) $P(B|AC)=P(B|C)$.

证明 (1) $\Rightarrow$ (2), (3): 由$P(AB|C)=P(A|C)P(B|C)$可得

$$P(ABC)=P(A|C)P(B|C)P(C)=P(AC)P(B|C)=P(A|C)P(BC),$$

从而

$$P(A|BC)=P(A|C),\qquad P(B|AC)=P(B|C).$$

(2) $\Rightarrow$ (1): 由$P(A|BC)=P(A|C)$可知

$$P(ABC)=P(A|C)P(BC).$$

两边除以$P(C)$得$P(AB|C)=P(A|C)P(B|C)$.

类似可得(3) $\Rightarrow$ (1). 证毕. □

♠**注记3.1.3.** 由条件独立定义可以看出, 在$P(C) > 0$条件下, 若$P(BC) = 0$或$P(AC) = 0$, 那么A, B关于C必然条件独立. 因此根据命题3.1.2, 我们约定: $P(BC) = 0$而$P(C) > 0$时,

$$P(A|BC) = P(A|C),$$

$P(AC) = 0$而$P(C) > 0$时,

$$P(B|AC) = P(B|C).$$

将条件独立的假设用到随机过程上就可定义所谓的马尔科夫性质(Markov Property). 直观而言就是要求在随机过程已知任一时刻t的确切状态后, t之后发生的随机事件与t 之前的随机事件独立. 比如

▶**例3.1.4.** 观察例2.2.14中某个顾客手中的筹码。若已知他在某个时刻t(现在)手中筹码为x, 经验告诉我们, 他在此后(未来)的筹码数的变化取决于此刻的筹码数而与t 之前(过去) 的筹码多少无关.

▶**例3.1.5.** 假定一个罐子里有红白两种颜色共$2N$个球, 做如下的抽球游戏并观察其中红球的动态变化: 从罐中无放回地随机抽出一个球, 如果抽出的是红球就放回一个白球, 如果抽出是白球就放入一个红球. 以X_n 表示第n次抽取后罐里红球数. 若已知第k(现在)次抽取后罐里有m个红球, 那么k 之后(未来)从罐里取出红/白球的情况与k之前(过去)取出红/白球情况无关.

称具有这种马尔可夫性质的随机过程为马尔可夫过程. 本课程只介绍其中具有离散参数和离散状态的情形, 由于可数状态空间中元素总可以通过适当标号区分, 因此本章内容对状态离散的向量值随机过程也成立. 但为叙述方便, 以下我们总设状态空间S 为$\mathbb{Z}$ 或它的子集. 离散时间离散状态条件下的马尔可夫性质定义如下:

♣**定义3.1.6.** 设$X = \{X_n;\ n \geq 0\}$是概率空间$(\Omega, \mathfrak{F}, P)$上的随机过程, 状态空间$S \subset \mathbb{Z}$, 称$X$具有马尔可夫性质(简称为马氏性), 若对任意$n \geq 1$, $i_0, \cdots, i_{n-1}, i, j \in S$, 只要$P(X_n = i) > 0$, 就有

$$P(X_{n+1} = j|X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j|X_n = i). \quad (3.1.1)$$

此时称X为离散时间的马尔可夫链, 简称为马氏链.

注意, 虽然在马氏性定义中我们要求对任意$n \geq 1$, $i_0, \cdots, i_{n-1}, i, j \in S$, 只要$P(X_n = i) > 0$, 就有(3.1.1)成立, 但结合注记3.1.3 可知, 其实我们只需在

$$P(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) > 0 \quad (3.1.2)$$

条件下检验(3.1.1)是否成立即可. 此后, 在检验(3.1.1)是否成立时, 我们总是不加说明地假定(3.1.2)成立.

需要指出的是, 在定义3.1.6中, 我们只考虑了参数集$T=\mathbb{N}$的情形. 对更一般的参数集$T=\{k,k+1,\cdots,k+N\}$(其中$k\in\mathbb{Z}$, N是一个有限或无穷的正整数), 我们可以非常容易地平移马氏链的定义——将定义3.1.6中$X_0,X_1,\cdots,X_n,X_{n+1}$替换为$X_k,X_{k+1},\cdots,X_{k+n},X_{k+n+1}$, 其中$n\in\{k+1,k+2,\cdots,k+N-1\}$, 其它保持不变. 尽管如此, 为了讨论方便, 若无特别申明, 此后我们都默认马氏链的参数集$T=\mathbb{N}$.

下面我们看一个利用定义验证马氏链的例子.

▶**例3.1.7.** 设ξ_1, ξ_2, $\cdots$为一列独立同分布的非负整数值随机变量. 令

$$X_0=K,\qquad X_n=\begin{cases}0\vee(X_{n-1}-\xi_n), & k<X_{n-1}\le K,\\ 0\vee(K-\xi_n), & X_{n-1}\le k,\end{cases}\quad n\ge 1,$$

其中$k<K$为给定的两个正整数. 那么$X=\{X_n;n\ge 0\}$为离散时间马氏链.

证明 显然X的状态空间$S=\{0,1\cdots,K\}$. 对任意$n\ge 1$ 以及$i_1,\cdots,i_{n+1}\in S$, 当$P(X_n=i_n)>0$时,

$$\begin{aligned}&P(X_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_1=i_1,X_0=K)\\ &\quad=\begin{cases}P((K-\xi_{n+1})\vee 0=i_{n+1}|X_n=i_n,\cdots,X_0=K), & i_n\le k,\\ P((i_n-\xi_{n+1})\vee 0=i_{n+1}|X_n=i_n,\cdots,X_0=K), & k<i_n\le K.\end{cases}\end{aligned}$$

对任意$i\ge 1$, 由X_i的定义可知X_i为X_{i-1}与ξ_i的函数, 如此递归代入后可知X_i是随机变量$\xi_1,\cdots,\xi_i$的函数. 这表明事件

$$\{X_n=i_n,\cdots,X_1=i_1,X_0=K\}$$

由随机变量$\xi_1,\cdots,\xi_n$确定. 再由$\{\xi_i;i\ge 1\}$的独立性可知

$$\begin{aligned}&P(X_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_1=i_1,X_0=K)\\ &\quad=\begin{cases}P((K-\xi_{n+1})\vee 0=i_{n+1}), & i_n\le k,\\ P((i_n-\xi_{n+1})\vee 0=i_{n+1}), & k<i_n\le K,\end{cases}\\ &\quad=\begin{cases}P((K-\xi_{n+1})\vee 0=i_{n+1}|X_n=i_n), & i_n\le k,\\ P((i_n-\xi_{n+1})\vee 0=i_{n+1}|X_n=i_n), & k<i_n\le K,\end{cases}\\ &\quad=P(X_{n+1}=i_{n+1}|X_n=i_n).\end{aligned}$$

所以X为离散时间马氏链. □

需要强调的是, 在马氏链的定义中, 随机过程X在时刻n的状态必须是完全已知的. 若n时刻的状态不明确, 那么即使是马氏链, 相应的(3.1.1)也可能不成立. 例如, 由马氏链定义和第二章第二节定理2.2.5可知(q,p)-随机游动W是一个马氏

链, 现在设$P(W_0=0)=P(W_0=2)=1/2$ 且$p\neq q$, 那么

$$p=P(W_2=2|W_1\in\{1,3\},W_0=0),$$

但是

$$\begin{aligned}P(W_2=2|W_1\in\{1,3\})&=\frac{P(W_2=2,W_1\in\{1,3\})}{P(W_1\in\{1,3\})}\\&=\frac{P(W_2=2,W_1\in\{1,3\},W_0=0)+P(W_2=2,W_1\in\{1,3\},W_0=2)}{P(W_1\in\{1,3\},W_0=0)+P(W_1\in\{1,3\},W_0=2)}.\end{aligned}$$

注意到

$$\begin{aligned}&P(W_2=2,W_1\in\{1,3\},W_0=0)=P(W_2=2,W_1=1,W_0=0)=\frac{1}{2}p^2,\\&P(W_2=2,W_1\in\{1,3\},W_0=2)\\&\quad=P(W_2=2,W_1=1,W_0=2)+P(W_2=2,W_1=3,W_0=2)\\&\quad=\frac{1}{2}qp+\frac{1}{2}pq,\\&P(W_1\in\{1,3\},W_0=0)+P(W_1\in\{1,3\},W_0=2)\\&\quad=P(W_1=1,W_0=0)+P(W_1=1,W_0=2)+P(W_1=3,W_0=2)\\&\quad=\frac{1}{2}(p+1).\end{aligned}$$

因此

$$P(W_2=2|W_1\in\{1,3\})=\frac{p(1+q)}{1+p}\neq p=P(W_2=2|W_1\in\{1,3\},W_0=0).$$

3.1.2 马氏链的等价刻画

在马氏性中我们还可以对代表过去与未来的随机事件用更宽泛的形式表示.

★**定理3.1.8.** X是马氏链当且仅当对任意$n\geq 1$, 非负整数$0\leq t_1<t_2<\cdots<t_n<t$, 以及整数$i_1,\cdots,i_{n+1}\in S$, 若$P(X_{t_n}=i_n)>0$, 则有

$$P(X_t=i_{n+1}|X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)=P(X_t=i_{n+1}|X_{t_n}=i_n).\quad(3.1.3)$$

证明* 只需证明必要性. 对任意$n\geq k\geq 0$, 由全概率公式和(3.1.1)可得

$$\begin{aligned}&P(X_{n+1}=i_{n+1}|X_k=i_k,\cdots,X_{n-1}=i_{n-1},X_n=i_n)\\&\quad=\frac{P(X_{n+1}=i_{n+1},X_n=i_n,\cdots,X_k=i_k)}{P(X_n=i_n,\cdots,X_k=i_k)}.\\&\quad=\sum_{j_0,\cdots,j_{k-1}\in D^k}\frac{P(X_{n+1}=i_{n+1},\cdots,X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{P(X_n=i_n,\cdots,X_k=i_k)}\\&\quad=\sum_{j_0,\cdots,j_{k-1}\in D^k}\Big[\frac{P(X_{n+1}=i_{n+1},\cdots,X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{P(X_n=i_n,\cdots,X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}\end{aligned}$$

$$\times\frac{P(X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{P(X_n=i_n,\cdots,X_k=i_k)}\Big]$$
$$=P(X_{n+1}=i_{n+1}|X_n=i_n), \tag{3.1.4}$$

其中

$$D^k=\{(j_0,\cdots,j_k);\ P(X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)>0\}.$$

下面我们证明对任意$m\geq 1$, $n\geq 0$,

$$\begin{aligned}&P(X_{n+1}=i_{n+1},\cdots,X_{n+m}=i_{n+m}|X_k=i_k,X_{k+1}=i_{k+1}\cdots,X_n=i_n)\\&\qquad=P(X_{n+1}=i_{n+1},\cdots,X_{n+m}=i_{n+m}|X_n=i_n),\ \ 0\leq k\leq n.\end{aligned} \tag{3.1.5}$$

当$m=1$时(3.1.5)就是我们已证明的(3.1.4), 因而对任意$n\geq 0$以及$0\leq k\leq n$成立. 设$m\leq l-1$时(3.1.5)对任意$n\geq 0$以及$0\leq k\leq n$成立, 那么当$m=l$ 时, 对任意$n\geq 0$ 以及$0\leq k\leq n$, 记

$$A=\{X_k=i_k,\cdots,X_{n-1}=i_{n-1}\},\ B=\{X_n=i_n\},\ C=\{X_{n+1}=i_{n+1}\},$$
$$D=\{X_{n+2}=i_{n+2},\cdots,X_{n+l}=i_{n+l}\}.$$

此时由归纳假设可得

$$P(D|ABC)=P(D|C)=P(D|BC),\quad P(C|AB)=P(C|B).$$

因此

$$\begin{aligned}(3.1.5)\text{的左边}&=P(CD|AB)=P(ABCD)/P(AB)\\&=P(D|ABC)P(C|BA)=P(D|C)P(C|B)=P(D|CB)P(C|B)\\&=P(DC|B)=(3.1.5)\text{的右边}.\end{aligned}$$

由数学归纳法可知(3.1.5)显然对任意$m\geq 1$, $n\geq 0$以及$0\leq k\leq n$成立.

最后我们注意到, 由全概率公式可知

$$\begin{aligned}&P(X_t=i_{n+1}|X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)\\&\quad=\frac{P(X_t=i_{n+1},X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)}{P(X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)}\\&\quad=\sum_{\substack{0\leq s<t,\ j_s\in S\\ s\neq t_1,\cdots,t_n}}\frac{P(X_t=i_{n+1},X_{t_k}=i_k,X_s=j_s,t_1\leq s<t,s\neq t_k,1\leq k\leq n)}{P(X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)}.\end{aligned}$$

再由(3.1.5)并结合注记3.1.3得

$$\begin{aligned}&P(X_t=i_{n+1}|X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)\\&\quad=\sum_{j_s\in S,\ t_n<s<t}P(X_t=i_{n+1},X_s=j_s,t_n<s<t|X_{t_n}=i_n)\\&\qquad\times\sum_{\substack{0\leq s<t_n\ j_s\in S\\ s\neq t_1,\cdots,t_n}}\frac{P(X_{t_k}=i_k,X_s=j_s,t_1\leq s<t_n,s\neq t_k,1\leq k\leq n)}{P(X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)}\\&\quad=P(X_t=i_{n+1}|X_{t_n}=i_n).\end{aligned}$$

因此X是马氏链. $\square$

进一步, 由(3.1.5)易证对任意一个时刻n之后的随机事件

$$A=\{X_{n+k_1}=i_{n+1},\cdots,X_{n+k_l}=i_{n+l}\},$$

和任意一个n之前的随机事件

$$B=\{X_{t_1}=i_1,\cdots,X_{t_m}=i_m\},$$

其中$0\le t_1<\cdots<t_m<n$, 都有

$$P(A|X_n=i_n,B)=P(A|X_n=i_n). \tag{3.1.6}$$

具体证明请读者完成(参见本节习题10).

3.1.3 转移概率矩阵与C-K方程

由(3.1.3)和(3.1.1)可知, 对离散时间马氏链, 下面的条件概率是基本的

$$P(X_{n+m}=j|X_n=i).$$

♣**定义3.1.9.** 称条件概率$P(X_{n+m}=j|X_n=i)$为马氏链X由时刻n状态i经m步转移到状态j的转移概率, 记作$p_{i,j}^{(n,n+m)}$或$p(n,i;n+m,j)$. 若对任意i,j,m, $p_{i,j}^{(n,n+m)}$与n无关, 即

$$P(X_{n+m}=j|X_n=i)=P(X_m=j|X_0=i)$$

对任意i,j,n,m成立, 则称转移概率是时齐的(或平稳的). 此时常将$p_{i,j}^{(n,n+m)}$简记成$p_{i,j}^{(m)}$或$p_m(i,j)$, 称之为从状态i到状态j 的m步转移概率. 特别地, 简记$p_{i,j}^{(1)}$为$p_{i,j}$或$p(i,j)$, 并简称为从状态i 到状态j 的转移概率. 约定

$$p_{i,j}^{(0)}=\delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

♣**定义3.1.10.** 称马氏链X是时齐的, 如果X的转移概率是时齐的.

★**定理3.1.11.** 设$X=\{X_n;\ n\ge 0\}$是时齐马氏链, 则对任意$n,m\ge 1$及$i,j\in S$,

$$p_{i,j}^{(n+m)}=\sum_{k\in S}p_{i,k}^{(n)}p_{k,j}^{(m)}. \tag{3.1.7}$$

由此可知$p_{i,j}^{(n)}=\sum\limits_{j_1\in\mathbf{S}}\cdots\sum\limits_{j_{n-1}\in\mathbf{S}}p_{i,j_1}p_{j_1,j_2}\cdots p_{j_{n-1},j}$.

证明 只需证明(3.1.7). 由全概率公式可知

$$p_{i,j}^{(n+m)}=P(X_{n+m}=j|X_0=i)=\sum_{k\in S}P(X_{n+m}=j,X_n=k|X_0=i).$$

再由条件概率公式与马氏性(3.1.3)可得

$$
\begin{aligned}
P(X_{n+m}=j,X_n=k|X_0=i) &= \frac{P(X_{n+m}=j,X_n=k,X_0=i)}{P(X_0=i)} \\
&= \frac{P(X_{n+m}=j,X_n=k,X_0=i)}{P(X_n=k,X_0=i)}\frac{P(X_n=k,X_0=i)}{P(X_0=i)} \\
&= P(X_{n+m}=j|X_n=k)P(X_n=k|X_0=i)=p_{k,j}^{(m)}p_{i,k}^{(n)}.
\end{aligned}
$$

由此可知(3.1.7)成立. □

通常称(3.1.7)为Chapmann-Kolmogorov(C-K)方程. 证明C-K方程中所用全概率公式、条件概率公式以及马氏性的组合方法是研究马氏链的基本技巧.

若无特别说明, 本书此后所涉及马氏链X都假定是时齐的.

♣**定义3.1.12.** *对任意给定的$n\in\mathbb{N}$, 将马氏链$X=\{X_k;\ k\geq 0\}$所有的n步转移概率排成矩阵形式$(p_{i,j}^{(n)})_{i,j\in S}$, 其中所有$p_{i,i}^{(n)}$的连线构成矩阵主对角线. 我们称此矩阵为$X$的$n$步转移概率矩阵, 记作$\mathbf{P}^{(n)}$. 特别地, 当$n=1$ 时简称为转移概率矩阵, 记作$\mathbf{P}$; 当$n=0$时$\mathbf{P}^{(0)}=\mathbf{I}$为与恒同变换对应的单位矩阵.*

当n步转移概率矩阵需要被显式地表达出来时, 我们会面临一个状态作为出发状态该安排在哪一行或作为到达状态该安排在哪一列这样的问题. 也就是确定状态空间与行、列的一一对应问题. 这样的一一对应往往是不唯一的. 通常会根据$p_{i,j}^{(n)}$的规律选择恰当的对应法则使得n步转移概率矩阵易于书写和便于理解. **若无特别声明, 本书此后总按状态从小到大的排序从上到下与矩阵的行对应**. 无论怎样, 我们强调, 按照n 步转移概率矩阵的表示, 状态i 对应的行刻画了马氏链从状态i 出发, 经过n 步后的分布; 而与状态j对应的列, 则由从不同状态出发, 经过n步到达状态j的概率构成.

当S只有有限个状态时, 若我们还要求所有的n步转移矩阵中状态空间与矩阵行、列之间的一一对应关系是不变的, 那么由C-K方程以及矩阵乘法可知, 对任意$n,m>0$,

$$
\begin{aligned}
\mathbf{P}^{(n+m)} &= \big(p_{i,j}^{(n+m)}\big)_{i,j\in S}=\Big(\sum_{k\in S}p_{i,k}^{(n)}p_{k,j}^{(m)}\Big)_{i,j\in S} \\
&= \big(p_{i,j}^{(n)}\big)_{i,j\in S}\big(p_{i,j}^{(m)}\big)_{i,j\in S}=\mathbf{P}^{(n)}\mathbf{P}^{(m)}=\mathbf{P}^{n+m}.
\end{aligned}
$$

因此, 对状态有限的马氏链, 若转移概率矩阵已知, 那么我们可以通过矩阵运算获得它的n步转移概率矩阵, 而且状态与行、列之间的对应保持不变.

▶**例3.1.13.** *如果把社会划分为若干个阶层, 比如上,中,下3种, 分别以1, 2, 3表示. 为了研究社会阶层的流动情况, 任意选取某个家族. X_n 表示该家族第n 代所处的社会阶层. 一个简单的社会学模型认为第$n+1$代的社会地位只取决于其*

父代的社会地位. 若假定

$$P(X_{n+1}=1|X_n=1)=0.8, \quad P(X_{n+1}=2|X_n=1)=0.1,$$
$$P(X_{n+1}=3|X_n=1)=0.1, \quad P(X_{n+1}=1|X_n=2)=0.3,$$
$$P(X_{n+1}=2|X_n=2)=0.4, \quad P(X_{n+1}=3|X_n=2)=0.3,$$
$$P(X_{n+1}=1|X_n=3)=0.05, \quad P(X_{n+1}=2|X_n=3)=0.1,$$
$$P(X_{n+1}=3|X_n=3)=0.85.$$

那么$\{X_n\}$构成一个时齐马氏链, 状态空间$S=\{1,2,3\}$, 转移概率是

$$p_{1,1}=0.8, \quad p_{1,2}=0.1, \quad p_{1,3}=0.1;$$
$$p_{2,1}=0.3, \quad p_{2,2}=0.4, \quad p_{2,3}=0.3;$$
$$p_{3,1}=0.05, \quad p_{3,2}=0.1, \quad p_{3,3}=0.85.$$

显然转移概率矩阵$\mathbf{P}=(p_{i,j})_{i,j\in S}$是一个3阶方阵. 我们将矩阵的第$1,2,3$行分别标记为出发状态$1,2,3$, 相应地, 矩阵的第$1,2,3$列就要标记为到达状态$1,2,3$. 这样, 得到转移概率矩阵为

$$\mathbf{P}=\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.05 & 0.1 & 0.85 \end{pmatrix}.$$

为了利用转移概率矩阵$\mathbf{P}$计算一般的n步转移概率矩阵$\mathbf{P}^{(n)}$, 我们首先利用代数知识对矩阵$\mathbf{P}$做些分析: 容易算出$\mathbf{P}$有三个不同特征值1, 0.3和0.75, 对应特征向量可取为

$$\mathbf{a}_1=(1,1,1)^T, \ \ \mathbf{a}_2=(1,-6,1)^T \text{ 和 } \mathbf{a}_3=(-26,-6,19)^T.$$

令矩阵$\mathbf{A}=(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3)$. 直接计算可知

$$\mathbf{A}^{-1}=\frac{1}{315}\begin{pmatrix} 108 & 45 & 162 \\ 25 & -45 & 20 \\ -7 & 0 & 7 \end{pmatrix}.$$

从而由矩阵的相似对角化理论可得

$$\mathbf{P}=\frac{1}{315}\begin{pmatrix} 1 & 1 & -26 \\ 1 & -6 & -6 \\ 1 & 1 & 19 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.75 \end{pmatrix}\begin{pmatrix} 108 & 45 & 162 \\ 25 & -45 & 20 \\ -7 & 0 & 7 \end{pmatrix}.$$

因此, 对任意$n\ge 1$,

$$\mathbf{P}^{(n)}=\mathbf{P}^n=\frac{1}{315}\begin{pmatrix} 1 & 1 & -26 \\ 1 & -6 & -6 \\ 1 & 1 & 19 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.3^n & 0 \\ 0 & 0 & 0.75^n \end{pmatrix}\begin{pmatrix} 108 & 45 & 162 \\ 25 & -45 & 20 \\ -7 & 0 & 7 \end{pmatrix}.$$

计算后可知n步转移概率矩阵$\mathbf{P}^{(n)}$为

$$\frac{1}{315}\begin{pmatrix} 108+25\cdot 0.3^n+182\cdot 0.75^n & 45-45\cdot 0.3^n & 162+20\cdot 0.3^n-182\cdot 0.75^n \\ 108-150\cdot 0.3^n+42\cdot 0.75^n & 45+270\cdot 0.3^n & 162-120\cdot 0.3^n-42\cdot 0.75^n \\ 108+25\cdot 0.3^n-133\cdot 0.75^n & 45-45\cdot 0.3^n & 162+20\cdot 0.3^n+133\cdot 0.75^n \end{pmatrix}.$$

比如$n=5$时可算得

$$\mathbf{P}^{(5)}=\begin{pmatrix} 0.480 & 0.143 & 0.377 \\ 0.373 & 0.145 & 0.482 \\ 0.243 & 0.143 & 0.614 \end{pmatrix}.$$

这表明, 在这个模型下, 现在处于上层的家庭, 5代后仍以48%的概率保持在上层, 以14.3%的概率进入中层, 而以37.7%的概率坠入下层; 而现在为中层和下层的家庭则在5代后分别有37.3%和24.3%的概率进入上层.

如果我们令$n\to\infty$, 那么

$$\lim_{n\to\infty}\mathbf{P}^{(n)}=\frac{1}{315}\begin{pmatrix} 108 & 45 & 162 \\ 108 & 45 & 162 \\ 108 & 45 & 162 \end{pmatrix}.$$

由此可以看出在该模型下, 无论家庭现在所处阶层如何, 经过足够多代的繁衍生息后, 有相同的机会步入上、中、下层社会. □

对状态无穷的马氏链我们注意到, 对给定了行标和列标的无穷阶矩阵$A=(a_{ij})_{i,j\in S}$, $B=(b_{i,j})_{i,j\in S}$, 我们仍然可以利用计算

$$c_{i,j}=\sum_{k\in S}a_{i,k}b_{k,j}$$

定义一个新的无穷阶矩阵$C=(c_{i,j})_{i,j\in S}$并将其仍记作AB. 借助于这种运算与记号, 由C-K方程容易看出, 对无穷状态的转移概率矩阵同样有下面的关系成立.

$$\mathbf{P}^{(n+m)}=\mathbf{P}^{(n)}\mathbf{P}^{(m)}=\mathbf{P}^{n+m}.$$

▶**例3.1.14.** 由(*2.2.4*)易知直线上的(q,p)随机游动是一个时齐马氏链. 对任意$i\in\mathbb{Z}$, 转移概率$p_{i,i+1}=p$, $p_{i,i-1}=q$, 当$j\neq i+1,i-1$时$p_{i,j}=0$. 显然转移该率矩阵$\mathbf{P}=(p_{i,j})_{i,j\in\mathbb{Z}}$是一个行列数均无穷的矩阵. 为了写出该转移矩阵, 我们先任意指定一行以及一列对应到状态0, 然后将指定行上下的行以及指定列左右的列按由近及远的次序依次与状态$-1,1,-2,2,\cdots$对应(如果没有对应的行或列, 则在相应位置添补); 最后对任意i,j, 令状态i对应的行与状态j对应的列交叉位

置上的元素为$p_{i,j}$, 由此我们得到(q,p) 随机游动的转移概率矩阵如下

$$\mathbf{P}=(p_{i,j})_{i,j\in\mathbb{Z}}=\begin{array}{cc}\begin{pmatrix}\cdots & \cdot & \cdot & \cdot & \cdot & \cdots\\ \cdots & 0 & p & 0 & 0 & \cdots\\ \cdots & q & 0 & p & 0 & \cdots\\ \cdots & 0 & q & 0 & p & \cdots\\ \cdots & 0 & 0 & q & 0 & \cdots\\ \cdots & \cdot & \cdot & \cdot & \cdot & \cdots\end{pmatrix} & \begin{array}{c}\cdot\\ -1\\ 0\\ 1\\ 2\\ \cdot\end{array}\end{array}$$

可以检验当$n\geq 2$时, $\mathbf{P}^{(n)}=(p_{i,j}^{(n)})_{i,j\in\mathbb{Z}}=\mathbf{P}^n$, 其中$p_{i,j}^{(n)}$为公式(*2.2.6*)中$p_n(i,j)$. 具体的检验过程请读者自己完成.

♠**注记3.1.15.** 对任意的$n\geq 1$, 马氏链的n步转移概率矩阵都满足:

(1) $p_{i,j}^{(n)}\geq 0,\ i,j\in S$, (2) $\sum_{j\in S}p_{i,j}^{(n)}=1,\ i\in S$.

一般称满足这两条件的矩阵为随机矩阵.

3.1.4 有限维分布

下面我们用转移概率给出马氏链的有限维分布.

★**定理3.1.16.** 设马氏链X的转移概率矩阵为$(p_{i,j})_{i,j\in S}$, 初始分布为$(\mu_i)_{i\in S}$. 那么对任意$0=t_0\leq t_1<t_2<\cdots<t_n$以及$i_1,\cdots,i_n\in S$,

$$P(X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)=\sum_{i\in S}\mu_i p_{i,i_1}^{(t_1-t_0)}p_{i_1,i_2}^{(t_2-t_1)}\cdots p_{i_{n-1},i_n}^{(t_n-t_{n-1})}, \quad (3.1.8)$$

其中对任意$k=1,2,\cdots,n$, 记$n_k=t_k-t_{k-1}$, 由*C-K*方程可知

$$p_{i_{k-1},i_k}^{(t_k-t_{k-1})}=p_{i_{k-1},i_k}^{(n_k)}=\sum_{j_1\in\mathbf{S}}\cdots\sum_{j_{n_k-1}\in\mathbf{S}}p_{i_{k-1},j_1}p_{j_1,j_2}\cdots p_{j_{n_k-1},i_k}. \quad (3.1.9)$$

证明 只需证明(3.1.8). 由条件概率公式与马氏性, $P(X_{t_1}=i_1,\cdots,X_{t_n}=i_n)$等于

$$\begin{aligned}&P(X_{t_n}=i_n|X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)P(X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)\\ &\quad=p_{i_{n-1},i_n}^{(t_n-t_{n-1})}P(X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)=\cdots\\ &\quad=p_{i_{n-1},i_n}^{(t_n-t_{n-1})}\cdots p_{i_1,i_2}^{(t_2-t_1)}\sum_{i\in S}P(X_{t_1}=i_1,X_0=i)\\ &\quad=\sum_{i\in S}\mu_i p_{i,i_1}^{(t_1)}p_{i_1,i_2}^{(t_2-t_1)}\cdots p_{i_{n-1},i_n}^{(t_n-t_{n-1})}.\end{aligned}$$

□

注意到(3.1.8)中多步转移概率通过(3.1.9)转化为一步转移概率. 定理3.1.16表明只要知道了马氏链的初始分布和转移概率矩阵就可以算出它的任意有限维分布函数, 继而计算出由该马氏链确定的随机事件的概率.

▶**例3.1.17.** 设$\{\xi_n; n\geq 1\}$为一列独立同分布的非负整数值随机变量且与非负整数值随机变量X_0独立. 对$n\geq 1$, 令

$$X_n=\begin{cases}X_{n-1}-1+\xi_n, & X_{n-1}>0,\\ \xi_n, & X_{n-1}=0.\end{cases}$$

证明$X=\{X_n;\ n\geq 0\}$为时齐马氏链. 若记$p_k=P(\xi_1=k)$, 求X的转移概率矩阵. 进一步若设$\mu_k=P(X_0=k)$, 求$P(X_0=1, X_1=2, X_3=4)$.

证明 显然X的状态空间$S=\mathbb{N}$. 对任意$n\geq 1$以及非负整数$i_0,i_1,\cdots,i_n,i_{n+1}$, 当$P(X_n=i_n)>0$时,

$$\begin{aligned}&P(X_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_1=i_1,X_0=i_0)\\ &=\begin{cases}P(i_n-1+\xi_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_0=i_0), & i_n>0,\\ P(\xi_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_0=i_0), & i_n=0.\end{cases}\end{aligned}$$

与例3.1.7类似分析可知$\{X_n=i_n,\cdots,X_1=i_1,X_0=i_0\}$由随机变量$X_0,\xi_1,\cdots,\xi_n$确定. 由独立性假设可得

$$\begin{aligned}&P(X_{n+1}=i_{n+1}|X_n=i_n,\cdots,X_1=i_1,X_0=i_0)\\ &=\begin{cases}P(i_n-1+\xi_{n+1}=i_{n+1}), & i_n>0,\\ P(\xi_{n+1}=i_{n+1}), & i_n=0,\end{cases}\\ &=\begin{cases}P(i_n-1+\xi_{n+1}=i_{n+1}|X_n=i_n), & i_n>0,\\ P(\xi_{n+1}=i_{n+1}|X_n=i_n), & i_n=0,\end{cases}\\ &=P(X_{n+1}=i_{n+1}|X_n=i_n).\end{aligned}$$

所以X为离散时间马氏链. 直接计算可知对任意$n\in\mathbb{N}$,

$$\begin{aligned}P(X_{n+1}=j|X_n=i)&=\begin{cases}P(\xi_{n+1}=j+1-i), & i\neq 0,\\ P(\xi_n=j), & i=0,\end{cases}\\ &=\begin{cases}p_{j+1-i}, & i\neq 0\text{且}j+1-i\geq 0,\\ 0, & i\neq 0\text{且}j+1-i<0,\\ p_j, & i=0.\end{cases}\end{aligned}$$

因此X是时齐的马氏链, 而且当$i=0$时, $p_{0,j}=p_j$, $j\in\mathbb{N}$; 当$i>0$时,

$$p_{i,j}=p_{j+1-i},\ j\geq i-1;\quad p_{i,j}=0,\ j<i-1.$$

进而X的转移概率矩阵可写成

$$\mathbf{P}=(p_{i,j})_{i,j\in\mathbb{N}}=\begin{array}{cc}\begin{pmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ 0 & p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & 0 & p_0 & p_1 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \end{pmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \cdot \end{array}\end{array}$$

由定理3.1.16以及C-K方程可知, 所求概率

$$\begin{aligned}P(X_0=1,X_1=2,X_3=4)&=\mu_1 p_{1,2}p_{2,4}^{(2)}=\mu_1 p_{1,2}\sum_{k=1}^{5}p_{2,k}p_{k,4}\\&=\mu_1 p_2(p_0p_4+p_1p_3+p_2^2+p_3p_1+p_4p_0)\\&=\mu_1 p_2(2p_0p_4+2p_1p_3+p_2^2).\end{aligned}$$

□

练习题3.1

1 设$X=\{X_n;n\geq 0\}$是定义在概率空间$(\Omega,\mathfrak{F},P)$上的时齐马氏链. 以P_i表示事件$\{X_0=i\}$发生情形下的条件概率, 即对任意$A\in\mathfrak{F}$, $P_i(A)=P(A|X_0=i)$. 证明, 对任意时间$m>n>0$以及任意状态i,j,k,

$$P_i(X_m=k|X_n=j)=P(X_{m-n}=k|X_0=j).$$

2 设$\{N(t);t\geq 0\}$为泊松过程. 任取$t_0>0$. 证明$\{N(nt_0);n\geq 0\}$是马氏链.

3 以$X=\{X_n;n\geq 0\}$表示成功概率为p的伯努利过程, 对任意$n\geq 1$, 令$Y_n=X_n+X_{n-1}$. 试判断过程$Y=\{Y_n;n\geq 1\}$是否是马氏链, 并写明理由.

4 已知$X=\{X_n;n\geq 0\}$为马氏链, 对任意$N\geq 1$以及$0\leq n\leq N$, 令$Y_n=X_{N-n}$. 证明$Y=\{Y_n;0\leq n\leq N\}$也是马氏链.

5 在例3.1.7中设ξ_1取0, 1, 2, 3, 4的概率分别为0.1, 0.2, 0.3, 0.2, 0.2. 再设$k=2$, $K=4$. 试写出该马氏链的转移概率矩阵.

6 (Wright-Fisher 模型) 考虑一个有N个基因的固定群体, 基因是A或a这两种类型. 假定这个群体在$n+1$时基因的状态是通过时刻n 的状态变异得到. 进一步假设每个基因发生变异的机会是等可能的, 由A变异为a的概率为p, 由a 变异为A 的概率为q. 以X_n表示n时刻基因中A 的个数, 那么此时$\{X_n\}$就是个时齐马氏链. 试写出该模型的转移概率矩阵.

7 证明一般马氏链的C-K方程: 设马氏链$X=\{X_n;n\geq 0\}$的状态空间为S, 证明对任意$n,m,k\geq 0$, $i,j\in S$, 都有

$$p_{i,j}^{(n,n+m+k)}=\sum_{l\in S}p_{i,l}^{(n,n+m)}p_{l,j}^{(n+m,m+m+k)}.$$

8 设$X=\{X_n;n\geq 0\}$是状态空间为S的一般马氏链, 证明X是时齐的当且仅当对任意$n,m\geq 0$, $i,j\in S$, 都有$p_{i,j}^{(n,n+1)}=p_{i,j}^{(m,m+1)}$.

9 已知马氏链X的转移概率矩阵$\mathbf{P}=\begin{pmatrix}1 & 0 & 0\\ 1/4 & 1/2 & 1/4\\ 0 & 0 & 1\end{pmatrix}$, 求$\mathbf{P}^{(n)}$.

10* 试证明(3.1.6).

11* 设X是状态空间为S的时齐马氏链, 证明对任意$n\geq 1$, $i\in S$以及$S_m\subset S$, $m\geq 0$,

$$P\left(\bigcap_{m=n+1}^{\infty}\{X_m\in S_m\}\Big|X_n=i,\bigcap_{m=0}^{n-1}\{X_m\in S_m\}\right)$$
$$=P\left(\bigcap_{m=1}^{\infty}\{X_m\in S_{n+m}\}\Big|X_0=i\right).$$

3.2 状态分类

对马氏链的状态分类有助于我们在一定程度上掌握各种状态动态出现的关联与特征, 并引导我们进一步了解马氏链的一般规律. 为了叙述简洁, 我们强调, 此后若无特别说明, 总设马氏链.$X = \{X_n;\ n \geq 0\}$的转移概率矩阵$\mathbf{P} = (p_{i,j})_{i,j\in S}$. 为书写方便, 我们还分别简记$E(\cdot|X_0 = i)$和$P(\cdot|X_0 = i)$为$E_i(\cdot)$ 和$P_i(\cdot)$.

3.2.1 互通、本质与不可约

在研究动态变化的随机现象时, 人们常常会问两个给定的随机现象是否会先后出现? 对于用马氏链刻画的动态随机现象而言, 这样的问题可以借助于n步转移概率来刻画. 我们给出如下刻画状态之间关联的概念.

♣**定义3.2.1.** *如果存在$n \geq 0$使得$p_{i,j}^{(n)} > 0$, 则称状态i可达状态j, 记作$i \to j$. 反之, 则对任意$n \geq 0$, $p_{i,j}^{(n)} = 0$, 称状态i不可达状态j, 记作$i \not\to j$. 若$i \to j$且$j \to i$, 则称i, j互通, 记作$i \leftrightarrow j$.*

根据定义, 由$p_{i,j}^{(0)} = \delta(i,j)$可知$i \leftrightarrow i$, 即每一个状态都与自己是互通的. 对于马氏链模型的两个不同状态i, j, 可达与互通的直观地解释就是: $i \to j$当且仅当在观察到状态i之后我们有机会观察到状态j ; $i \not\to j$当且仅当在观察到状态i之后我们几乎不可能观察到状态j; $i \leftrightarrow j$当且仅当先观察到i之后可能观察到j, 先观察到j之后也可能观察到i.

▲**命题3.2.2.** *互通是一种等价关系. 即互通关系满足*
(1) 反身性 $i \leftrightarrow i$; (2)对称性;$i \leftrightarrow j \Rightarrow j \leftrightarrow i$; (3)传递性 $i \leftrightarrow j, j \leftrightarrow k \Rightarrow i \leftrightarrow k$.

证明 只需证明(3)传递性. 由$i \leftrightarrow j, j \leftrightarrow k$可知, 存在非负整数$n, m, r, l$使得

$$p_{i,j}^{(n)} > 0, \quad p_{j,i}^{(m)} > 0, \quad p_{j,k}^{(r)} > 0, \quad p_{k,j}^{(l)} > 0.$$

由C-K方程可知,

$$p_{i,k}^{(n+r)} = \sum_{l\in S} p_{i,l}^{(n)} p_{l,k}^{(r)} \geq p_{i,j}^{(n)} p_{j,k}^{(r)} > 0.$$

类似可知$p_{k,i}^{(l+m)} \geq p_{k,j}^{(l)} p_{j,i}^{(m)} > 0$. 因此$i \leftrightarrow k$. □

对于任意状态i, 利用互通关系我们可以构造一个集合使其包含所有与i互通的状态并将其记作$C(i)$, 即

$$C(i) = \{k \in S;\ k \leftrightarrow i\}.$$

通常我们称这个集合为包含状态i的互通类. 由互通等价性可知$C(i)$中任意两状态互通而且对不同的状态i,j

$$C(i)\cap C(j)\neq\emptyset\Leftrightarrow C(i)=C(j).$$

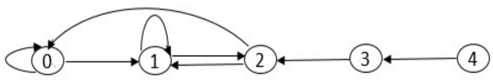
这表明若$C(i)\neq C(j)$, 则$C(i)\cap C(j)=\emptyset$.

▶**例3.2.3.** 若马氏链X的状态空间为$S=\{0,1,2,3,4\}$, 转移概率矩阵为

$$\mathbf{P}=\begin{pmatrix}1/3 & 2/3 & 0 & 0 & 0\\ 0 & 1/2 & 1/2 & 0 & 0\\ 3/4 & 1/4 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\end{pmatrix}.$$

由$0\to 1$, $1\to 2$, $2\to 0$可知$\{0,1,2\}\subset C(0)$. 再由$0\not\to 3$, $0\not\to 4$ 可知$C(0)=\{0,1,2\}$. 同样可知$C(3)=\{3\}$, $C(4)=\{4\}$. □

对于有限状态的马氏链, 我们可以通过观察所谓状态一步转移图来写出包含各个状态的互通类. 一步转移图构造如下: 先将状态空间中各元素用点分开表示, 并将各点用对应的状态标记, 然后对任意两点i,j, 若$p_{i,j}>0$ 则从i引一条方向指向j 的有向连线, 否则则不连线. 例如, 根据例3.2.3的转移概率矩阵我们可以画出马氏链X的一步转移图如下.



通常我们称这些与状态对应的点为顶点. 若顶点i,j之间的存在从i指向j的连线, 则称该连线为从i到j的边, 记作$\overrightarrow{ij}$. 若边$\overrightarrow{ij}$ 存在或者若存在若干顶点$j_1,\cdots,j_k$使得边$\overrightarrow{ij_1},\overrightarrow{j_1j_2},\cdots,\overrightarrow{j_{k-1}j_k},\overrightarrow{j_kj}$ 都存在, 则称顶点i,j是有向连通的, 仍记作$i\to j$. 若$i\to j$且$j\to i$, 则称顶点i,j是双向连通的, 也记作$i\leftrightarrow j$. 比如从例3.2.3 的一步转移图可知: 作为顶点, $0\leftrightarrow 1$, $0\leftrightarrow 2$, $1\leftrightarrow 2$, 但没有顶点与3或4双向连通.

容易证明$i\neq j$时, 状态$i\to j$当且仅当在一步转移图中对应顶点$i\to j$(参见本节习题1). 因此

$$C(i)=\{i\}\cup\{j;\text{顶点}i,j\text{双向连通}\}.$$

利用这个等式, 结合上面关于例3.2.3 的一步转移图的分析, 我们可以直接写出例3.2.3 的各个互通类.

利用互通关系, 我们还可以考虑一类特殊的状态如下.

♣**定义3.2.4.** 若对任意$j\in S$, 由$i\to j$可得$j\to i$, 则称状态i是本质的.

由定义可知, 由任意本质状态i出发的所有可达状态必然落在互通类$C(i)$中. 由此我们容易得到下面的结果.

■**性质3.2.5.** 若i是本质的, 那么$C(i)$中所有状态都是本质的. 此时我们称$C(i)$为本质类.

证明 任取$j \in C(i)$, 若$j \to k$, 则由$i \to j$可知$i \to k$, 再由i本质可知, $k \to i$, 从而$k \in C(i)$, $k \to j$. 即j是本质的. □

一种特别的本质状态是所谓的吸收态.

♣**定义3.2.6.** 若$p_{i,i} = 1$, 则称i为吸收的.

显然若i是吸收的, 那么$C(i) = \{i\}$, 从而是本质的; 但反之不成立.

利用本质互通类可以帮助我们在一定条件下简化状态空间. 事实上由性质3.2.5可知, 若i是本质的, 那么对任意$j \in C(i)$, j可达的状态只能在$C(i)$中. 换句话说, 若$X_0 \in C(i)$且$C(i)$是本质的, 那么对所有$n \geq 1$, $X_n \in C(i)$; 此时我们可以将X 的状态空间简化为$C(i)$.

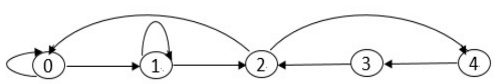
♣**定义3.2.7.** 若马氏链X的所有状态都是互通的, 即存在$i \in S$, $S = C(i)$, 则称X是不可约的($irreducible$).

由互通类性质可知, 若X是不可约的, 那么对任意$i \in S$都有$C(i) = S$.

▶**例3.2.8.** 若马氏链X的状态空间$S = \{0, 1, 2, 3, 4\}$, 转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} 1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 3/4 & 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

那么X的状态一步转移图为

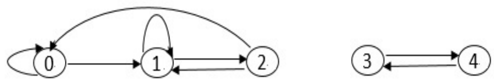


由图容易看出$C(0) = \{0, 1, 2, 3, 4\}$, 从而X是不可约的. □

▶**例3.2.9.** 若马氏链X的状态空间仍为$\{0, 1, 2, 3, 4\}$, 但转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} 1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 3/4 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

那么它的状态一步转移图为



由图容易看出$C(0)=\{0,1,2\}$, $C(3)=\{3,4\}$. 从而X的状态空间包含2个互通类$C(0)$, $C(3)$. □

对不可约马氏链而言所有状态都是本质的. 从而无论现在观察到什么状态, 面对不确定的未来, 我们可以断言所有的状态都有可能出现.

3.2.2 周期性

下面我们讨论研究动态变化随机现象时常会问的另外一个问题, 如果某个随机现象会反复出现, 那么它可能出现的时间有什么规律呢？下面这个概念可以帮助我们对这个问题做些初步的分析.

♣**定义3.2.10.** 设i是马氏链X的一个状态, 若存在一个正整数m使得$p_{i,i}^{(m)}>0$, 则称使得$p_{i,i}^{(n)}>0$的所有正整数$n(n\geq 1)$的最大公约数为状态i的周期, 记作d_i. 若对所有$n\geq 1$, $p_{i,i}^{(n)}=0$, 则约定i 的周期是∞. 若i 的周期为1, 即$d_i=1$, 则称i是非周期的.

直观地看, $\{n\geq 1;\ p_{i,i}^{(n)}>0\}$就是马氏链模型先观察到状态$i$之后能重复观察到状态$i$的所有可能的间隔时间, 因此周期就是这些间隔时间的最大公约数, 这表明, 若n不是d_i的倍数, 那么在间隔为n的时间状态i不可能出现. 严格地说就是若$d_i\nmid n$, 则$p_{i,i}^{(n)}=0$.

▲**命题3.2.11.** 同一个互通类中的状态周期相同, 即对任意$j\in C(i)$, $d_i=d_j$.

证明 由$i\leftrightarrow j$可知, 存在正整数$n,m>0$使得$p_{i,j}^{(n)}>0$, $p_{j,i}^{(m)}>0$. 由C-K方程可得

$$p_{i,i}^{(n+m)}\geq p_{i,j}^{(n)}p_{j,i}^{(m)}>0,\quad p_{j,j}^{(n+m)}\geq p_{j,i}^{(m)}p_{i,j}^{(n)}>0.$$

因此$d_i|n+m$, $d_j|n+m$. 另外, 对任意$l>0$, 若$p_{i,i}^{(l)}>0$, 则

$$p_{j,j}^{(n+m+l)}\geq p_{j,i}^{(m)}p_{i,i}^{(l)}p_{i,j}^{(n)}>0.$$

这表明$d_j|l+m+n$, 从而$d_j|l$. 由此可得$d_j|d_i$. 类似可得$d_i|d_j$. 因此$d_i=d_j$. □

♣**定义3.2.12.** 若马氏链X的每一个状态都有相同的周期d, 则称X是d周期马氏链. 特别地, 若$d=1$, 则称X是非周期的.

▶**例3.2.13.** (q,p)-简单随机游动W是个周期$d=2$的不可约马氏链. 事实上, 对任意$i,j\in\mathbb{Z}$ (不妨设$j>i$),

$$p_{i,j}^{(j-i)}=p^{j-i}>0,\quad p_{j,i}^{(j-i)}=q^{j-i}>0.$$

所以$i\leftrightarrow j$, 即$\mathbb{Z}=C(0)$, 从而W是不可约的. 对任意$i\in\mathbb{Z}$, 由

$$p_{i,i}^{(n)}=p_{0,0}^{(n)}=\begin{cases}0, & n=2k-1,\\ C_{2k}^k p^k q^k, & n=2k,\end{cases}$$

可知$d_i=2$, 因此W的周期$d=2$. □

▶**例3.2.14.** 考虑例3.1.17中定义的马氏链X, 若$0<p_0$且$p_0+p_1<1$, 那么X是不可约的非周期马氏链.

证明 由$0<p_0$以及$0<p_0+p_1<1$可知, 必存在$k\geq 2$使得$p_k>0$. 对任意$i,j\geq 0$, 不妨设$j>i$,

(1) 由$p_{j,i}^{(j-i)}=p_0^{j-i}>0$可知$j\to i$.

(2) 当$i\neq 0$时, 令$n=\lceil\frac{j-i}{k-1}\rceil$, 其中$\lceil\cdot\rceil$表示不小于$\cdot$的最小整数. 对任意$m\geq 1$, 记$l_m=i+(k-1)m$. 由

$$p_{i,l_n}^{(n)}\geq p_{i,l_1}p_{l_1,l_2}\cdots p_{l_{n-1},l_n}=p_k^n>0,$$

可知$i\to l_n$. 注意到$l_n\geq j$, 而且, 若$l_n>j$, 则由(1)可知$l_n\to j$. 因此$i\to j$.

(3) 当$i=0$时, 由$p_k>0$可知$0\to k$, 而(1)和(2)表明$k\to j$, 因此$0\to j$.

综上可知对任意$i,j\in S$, $i\leftrightarrow j$. 因此X不可约. 注意到$p_{0,0}=p_0>0$, 状态0的周期为$d_0=1$. 由命题3.2.11可知X的所有状态周期均为1, 因此X是非周期的. □

★**定理3.2.15.** 设状态i的周期为d, 则存在正整数N使得对任意$n\geq N$, $p_{i,i}^{(nd)}>0$.

证明 由周期定义, 存在$1\leq n_1,\cdots,n_s$使得$p_{i,i}^{(n_k)}>0$, $k=1,\cdots,s$, 而且$n_1,\cdots,n_s$的最大公约数为d. 由第一章第一节习题10, 存在正整数N, 对任意$n\geq N$, 存在正整数$r_1,\cdots,r_s$使得

$$nd=r_1n_1+r_2n_2+\cdots+r_sn_s.$$

从而

$$p_{i,i}^{(nd)}\geq\prod_{k=1}^s(p_{i,i}^{(n_k)})^{r_k}>0.$$

这表明结论成立. □

由此易得如下两个推论.

♦**推论3.2.16.** 记状态i的周期为d_i, 若$p_{j,i}^{(m)} > 0$, 则存在$N > 0$使得$p_{j,i}^{(m+nd_i)} > 0$对所有的$n \geq N$都成立.

♦**推论3.2.17.** 若状态i是非周期的, 那么存在$N > 0$使得对任意$n > N$, $p_{i,i}^{(n)} > 0$.

需要指出的是, 许多周期马氏链的问题都可以借助适当的转化变成非周期马氏链的相关问题. 对于不可约马氏链, 本节习题13给了一种将周期马氏链转化为非周期马氏链的方法.

3.2.3 常返与非常返

前面我们介绍了刻画马氏模型两个状态是否可能先后出现的指标和一个状态可能重复出现的时间规律性的指标, 现在我们来讨论一个状态是否必然可重复观察到, 更准确地说是否几乎必然可重复观察到. 为此, 我们定义状态被首次重复观察到的时间如下.

对任意状态$i \in S$, 令$\tau_i = \inf\{n \geq 1;\ X_n = i\}$, 其中按约定$\inf \emptyset = +\infty$. 通常我们称$\tau_i$为$X$首次回到状态$i$的时刻, 简称为首回时.

当$X_0 = i$时, τ_i就是状态i首次重复出现的时间. 这样的τ_i在我们介绍随机游动时也定义过. 与(2.2.9)一样, 我们可知τ_i是一个停时, 而且

$$\{\tau_i \leq n\} = \bigcup_{k=1}^{n}\{\tau_i = k\} = \bigcup_{k=1}^{n}\{X_k = i, X_{k-1} \neq i, \cdots, X_1 \neq i\}.$$

进一步, 对任意$i, j \in S$, 令

$$f_{i,j}^{(n)} = P_i(\tau_j = n) = P(X_n = j, X_{n-1} \neq j, \cdots, X_1 \neq j | X_0 = i), \quad (3.2.1)$$

$$f_{i,j} = P_i(\tau_j < \infty) = \sum_{n=1}^{\infty} f_{i,j}^{(n)}. \quad (3.2.2)$$

由此可知,

$$f_{i,j} = \sum_{n=1}^{\infty} P(X_n = j, X_{n-1} \neq j, \cdots, X_1 \neq j | X_0 = i) = P_i(\bigcup_{k=1}^{\infty}\{X_k = j\}),$$

进而我们可以得到如下不等式

$$0 \leq f_{i,j}^{(n)} \leq p_{i,j}^{(n)} \leq f_{i,j} \leq 1. \quad (3.2.3)$$

特别, 若$i \not\to j$, 则$f_{i,j} = 0$,

直观地看, 当$j \neq i$时, $f_{i,j}^{(n)}$表示从i出发后在第n步首次到达j的概率, 而$f_{i,j}$表示从状态i出发, 在有限时间内达到状态j的概率. 当$i = j$时$f_{i,i}^{(n)}$表示从i出发后第n步首次回到i的概率, 而$f_{i,i}$表示从i 出发有限时间内回到状态i的概率. 因此若$f_{i,i} = 1$, 则表示在已经观察到状态i的条件下, 几乎必然可以重复观察到状态i.

♣**定义3.2.18.** 称状态i是常返的(*recurrent*), 如果$f_{i,i}=1$; 否则称为非常返的、暂留的或瞬过的(*transient*).

下面我们分析常返状态的性质并给出一个等价判别准则.

■**性质3.2.19.** 若i是常返的而且$i\to j$, 那么$f_{j,i}=1$.

证明 只需考虑$j\neq i$的情形, 对任意$n\geq 1$, 定义

$$e_{i,j}^{(n)}=P_i(X_n=j,X_v\neq i,v=1,\cdots,n-1). \tag{3.2.4}$$

它表示从i出发, 中途不仅过i而在第n步到达j的概率(也被称为禁忌概率). 因为$i\to j$, 我们可以观察到一列状态$i=i_0,i_1,i_2,\cdots,i_m=j$使得

$$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_{m-1},j}>0.$$

令$k=m-\max\{u;i_u=i,0\leq u\leq m-1\}$, 那么$k>0$而且

$$e_{i,j}^{(k)}\geq p_{i,i_{m-k+1}}p_{i_{m-k+1},i_{m-k+2}}\cdots p_{i_{m-1},j}>0.$$

若$f_{j,i}<1$, 那么

$$\begin{aligned}
&P_i(\tau_i=\infty)\geq P_i(X_k=j,X_n\neq i,1\leq n\neq k)\\
&\quad=P_i(X_k=j,X_n\neq i,1\leq n<k)P(X_v\neq i,v>k|X_k=j,X_n\neq i,1\leq n<k)\\
&\quad=e_{i,j}^{(k)}P(X_n\neq i,n\geq 1|X_0=j)=e_{i,j}^{(k)}(1-f_{j,i})>0.
\end{aligned}$$

此与i常返矛盾, 因此$f_{j,i}=1$. □

由性质3.2.19以及(3.2.3)可知, 若i常返, 那么i是本质的.

★**定理3.2.20.** 状态i常返当且仅当$\sum\limits_{n=0}^{\infty}p_{i,i}^{(n)}=\infty$. 若$i$非常返, 则

$$\sum_{n=0}^{\infty}p_{i,i}^{(n)}=\frac{1}{1-f_{i,i}}<\infty.$$

证明 对任意$i,j\in S$, 利用首次到达j的时间, 可将事件分拆为不相交事件的并集

$$\{X_n=j\}=\bigcup_{k=1}^{n}\{\tau_j=k,X_n=j\}.$$

因此由全概率公式和概率乘法公式可得

$$\begin{aligned}
p_{i,j}^{(n)}&=\sum_{k=1}^{n}P_i(\tau_j=k,X_n=j)=\sum_{k=1}^{n}P_i(X_n=j|\tau_j=k)P_i(\tau_j=k)\\
&=\sum_{k=1}^{n}P_i(X_n=j|\tau_j=k)f_{i,j}^{(k)}.
\end{aligned}$$

由马氏性和时齐性可知

$$P_i(X_n=j|\tau_j=k)=P(X_n=j|X_k=j,X_v\neq j,1\leq v<k,X_0=i)$$

$$= P(X_{n-k} = j | X_0 = j) = p_{j,j}^{(n-k)}. \tag{3.2.5}$$

因此

$$p_{i,j}^{(n)} = \sum_{k=1}^{n} p_{j,j}^{(n-k)} f_{i,j}^{(k)}. \tag{3.2.6}$$

由此可得, 对任意$s \in [0,1)$,

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} s^n = 1 + \sum_{n=1}^{\infty} p_{i,i}^{(n)} s^n = 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{n} s^k f_{i,i}^{(k)} s^{n-k} p_{i,i}^{(n-k)}.$$

由第一章第一节习题8可知, 上式右边我们可以交换求和次序, 从而

$$\begin{aligned}\sum_{n=0}^{\infty} p_{i,i}^{(n)} s^n &= 1 + \sum_{k=1}^{\infty} f_{i,i}^{(k)} s^k \sum_{n=k}^{\infty} p_{i,i}^{(n-k)} s^{n-k} \\ &= 1 + \sum_{k=1}^{\infty} f_{i,i}^{(k)} s^k \sum_{n=0}^{\infty} p_{i,i}^{(n)} s^n.\end{aligned}$$

注意到对所有的$s \in [0,1)$,

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} s^n < \infty, \qquad \sum_{k=1}^{\infty} f_{i,i}^{(k)} s^k < 1.$$

因此我们有

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} s^n = \frac{1}{1 - \sum_{k=1}^{\infty} f_{i,i}^{(k)} s^k}, \quad s \in [0,1).$$

再注意到$f_{i,i} = \lim\limits_{s \to 1-} \sum_{k=1}^{\infty} f_{i,i}^{(k)} s^k$, 上式两边令$s \to 1-$后可得

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} = \frac{1}{1 - f_{i,i}}.$$

由此可知若$f_{i,i} = 1$, 则$\sum\limits_{n=0}^{\infty} p_{i,i}^{(n)} = \infty$, 而且

$$f_{i,i} < 1 \Rightarrow \sum_{n=0}^{\infty} p_{i,i}^{(n)} = \frac{1}{1 - f_{i,i}},$$

即定理结论成立. □

♠**注记3.2.21.** 由定理证明中(*3.2.6*)还可知

$$\sum_{n=1}^{\infty} p_{i,j}^{(n)} = \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)} = \sum_{k=1}^{\infty} f_{i,j}^{(k)} \sum_{r=0}^{\infty} p_{j,j}^{(r)} = f_{i,j} \sum_{n=0}^{\infty} p_{j,j}^{(n)}.$$

注意到$f_{i,j} \le 1$, 我们总有

$$\sum_{n=1}^{\infty} p_{i,j}^{(n)} \le \sum_{n=0}^{\infty} p_{j,j}^{(n)}.$$

由定理3.2.20还容易推出

♦**推论3.2.22.** 若状态i常返, 那么$C(i)$中所有状态都常返, 此时称$C(i)$为常返类.

证明 对任意$j\in C(i)$, $i\leftrightarrow j$. 存在$m,n>0$使得$p_{i,j}^{(m)}>0, p_{j,i}^{(n)}>0$. 因此

$$\sum_{k=0}^{\infty}p_{j,j}^{(k)}\geq\sum_{k=m+n}^{\infty}p_{j,i}^{(n)}p_{i,i}^{(k-n-m)}p_{i,j}^{(m)}=p_{j,i}^{(n)}p_{i,j}^{(m)}\sum_{k=0}^{\infty}p_{i,i}^{(k)}=+\infty,$$

这表明j是常返的. □

由推论3.2.22 可知, 一个互通类中状态或全是常返的或全是非常返的, 若全是常返的称该互通类为常返类, 否则称为非常返类. 若不可约链的状态空间为常返类, 则称该马氏链为不可约常返链, 否则称为不可约非常返链.

▶**例3.2.23.** (q,p)-简单随机游动的任意状态i是常返的当且仅当$p=q=1/2$.

证明 由于(q,p)-随机游动是不可约的, 我们只需考虑状态0. 由$p_{0,0}^{(2n+1)}=0$以及

$$p_{0,0}^{(2n)}=C_{2n}^{n}(pq)^n=\frac{(2n)!}{n!n!}(pq)^n=\frac{(2n-1)!!}{2^n n!}(4pq)^n,$$

可得

$$\sum_{k=0}^{\infty}p_{0,0}^{(k)}s^k=1+\sum_{n=1}^{\infty}p_{0,0}^{(2n)}s^{2n}=1+\sum_{k=1}^{\infty}\frac{(2n-1)!!}{2^n n!}(4pqs^2)^n.$$

由幂函数展式

$$(1-x)^{-\frac{1}{2}}=1+\frac{1}{2}x+\frac{1\cdot 3}{2^2 2!}x^2+\cdots+\frac{(2n-1)!!}{2^n n!}x^n+\cdots,\ |x|<1,$$

可知

$$\sum_{k=0}^{\infty}p_{0,0}^{(k)}s^k=(1-4pqs^2)^{-1/2},\qquad |s|<1.$$

因此

$$\sum_{k=0}^{\infty}p_{0,0}^{(k)}=\lim_{s\to 1-}\sum_{k=0}^{\infty}p_{0,0}^{(k)}s^k=\lim_{s\to 1-}\frac{1}{\sqrt{1-4pqs^2}}=\begin{cases}+\infty, & p=q=1/2,\\ 1/|p-q|, & p\neq q.\end{cases}$$

可见(q,p)-简单随机游动常返当且仅当$p=q=1/2$, . □

▶**例3.2.24.** (高维简单对称随机游动) 直线上简单随机游动也可推广到高维空间. 以对称随机游动为例:

(1) 若$X=\{X_n;n\geq 0\}$是取值在$\mathbb{Z}^2=\{(i,j);i,j\in\mathbb{Z}\}$的时齐马氏链使得对任意$(i,j)\in\mathbb{Z}^2$,

$$p_{(i,j),(i+1,j)}=p_{(i,j),(i-1,j)}=p_{(i,j),(i,j+1)}=p_{(i,j),(i,j-1)}=1/4,$$

则称X是平面上简单对称随机游动(或平面整数格子点上对称随机游动).

(2) 若$X=\{X_n;n\geq 0\}$是取值在$\mathbb{Z}^3=\{(i,j,k);i,j,k\in\mathbb{Z}\}$的时齐马氏链使得对任意$(i,j,k)\in\mathbb{Z}^3$,

$$\begin{aligned}p_{(i,j,k),(i+1,j,k)}&=p_{(i,j,k),(i-1,j,k)}=p_{(i,j,k),(i,j+1,k)}=p_{(i,j,k),(i,j-1,k)}\\&=p_{(i,j,k),(i,j,k+1)}=p_{(i,j,k),(i,j,k-1)}=1/6,\end{aligned}$$

则称X是空间简单对称随机游动(或空间整数格子点上对称随机游动).

容易证明平面和空间的简单对称随机游动都是周期为2、不可约的马氏链. 此外我们还容易证明平面上简单对称随机游动是常返的, 但空间简单对称随机游动是非常返的. 事实上, 对空间简单对称随机游动, 其n步转移概率

$$p_{(0,0,0),(0,0,0)}^{(n)} = \sum_{r+l+s=k} \frac{(2k)!}{r!r!l!l!s!s!} \frac{1}{6^{2k}}, \quad n = 2k\text{为偶数},$$

而当$n = 2k-1$为奇数时$p_{(0,0,0),(0,0,0)}^{(n)} = 0$. 因此

$$\begin{aligned}
\sum_{n=0}^{\infty} p_{(0,0,0),(0,0,0)}^{(n)} &= \sum_{k=0}^{\infty} \sum_{r+l+s=k} \frac{(2k)!}{r!r!l!l!s!s!} \frac{1}{6^{2k}} \\
&= \sum_{k=0}^{\infty} \frac{1}{6^{2k}} \frac{(2k)!}{k!k!} \sum_{r=0}^{k} \sum_{l=0}^{k-r} \left[\frac{k!}{r!l!(k-r-l)!}\right]^2 \\
&\le \sum_{k=0}^{\infty} \frac{1}{6^{2k}} \frac{(2k)!}{k!k!} \max_{0\le r\le m\le k} \left\{\frac{k!}{r!(m-r)!(k-m)!}\right\} \sum_{r=0}^{k} \sum_{l=0}^{k-r} \frac{k!}{r!l!(k-r-l)!}.
\end{aligned}$$

注意到$\frac{k!}{r!l!(k-r-l)!}$是多项式$(x+y+z)^k$展开后的系数, 因此

$$\sum_{r=0}^{k} \sum_{l=0}^{k-r} \frac{k!}{r!l!(k-r-l)!} = 3^k.$$

另一方面, 对任意满足条件$0 \le r \le m \le k$的非负整数r, m,

$$\frac{k!}{r!(m-r)!(k-m)!} = C_k^m C_m^r.$$

当正整数$m \ge r$时, 组合数$C_m^r \le C_m^{\lfloor m/2 \rfloor}$. 因此

$$\frac{k!}{r!(m-r)!(k-m)!} \le C_k^m C_m^{\lfloor m/2 \rfloor}.$$

又由

$$\frac{C_k^m C_m^{\lfloor m/2 \rfloor}}{C_k^{m-1} C_{m-1}^{\lfloor (m-1)/2 \rfloor}} = \frac{k+1-m}{m - \lfloor m/2 \rfloor}, \quad 1 \le m \le k,$$

可知$C_k^m C_m^{\lfloor m/2 \rfloor}$在$m = m_k = k - \lfloor k/3 \rfloor$处取最大值. 由此可得

$$\sum_{n=0}^{\infty} p_{(0,0,0),(0,0,0)}^{(n)} \le \sum_{k=0}^{\infty} \frac{3^k}{6^{2k}} \frac{(2k)!}{k!(k-m_k)!\lfloor m_k/2 \rfloor!(m_k - \lfloor m_k/2 \rfloor)!}.$$

利用Stirling公式

$$k! \asymp \sqrt{2\pi k}(k/\mathrm{e})^k,$$

其中符号“$\asymp$”表示它两边的值相比随$k \to \infty$而收敛到1, 并注意到$m_k \asymp 2k/3$, 当k充分大时由近似计算可得

$$\frac{(2k)!}{k!(k-m_k)!\lfloor m_k/2 \rfloor!(m_k - \lfloor m_k/2 \rfloor)!} \asymp \frac{3\sqrt{3}}{2\pi\sqrt{\pi}} \frac{2^{2k}3^k}{k^{3/2}}.$$

因此存在常数$M > 0$使得

$$\sum_{n=0}^{\infty} p_{(0,0,0),(0,0,0)}^{(n)} \leq M \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} < \infty.$$

这表明$(0,0,0)$是非常返的, 从而由不可约性可知空间简单对称随机游动是非常返的. 平面上简单对称随机游动的常返性检验更简单, 参见本节习题9. □

▲命题3.2.25. 对任意$m \geq 1$, 记$g_{i,i}(m) = P_i$(至少有m个$n \geq 1$使得$X_n = i$), 并记$g_{i,i} = P_i(\{X_n = i\}, \text{ i.o.})$, 那么$g_{i,i}(m) = f_{i,i}^m$. 从而若$i$常返, 则$g_{i,i} = 1$; 若$i$非常返, 则$g_{i,i} = 0$.

证明 由全概率公式以及条件概率公式可得

$$\begin{aligned} g_{i,i}(m+1) &= P_i(\text{至少有}m+1\text{个}n \geq 1\text{使得}X_n = i) \\ &= \sum_{k=1}^{\infty} P_i(\tau_i = k, \text{至少有}m\text{个}n \geq k+1\text{使得}X_n = i) \\ &= \sum_{k=1}^{\infty} P_i(\text{至少有}m\text{个}n \geq k+1\text{使得}X_n = i|\tau_i = k)P_i(\tau_i = k). \end{aligned}$$

由本章第一节习题11不难证明

$$\begin{aligned} &P_i(\text{至少有}m\text{个}n \geq k+1\text{使得}X_n = i|\tau_i = k) \\ &\quad = P(\text{至少有}m\text{个}n \geq 1\text{使得}X_n = i|X_0 = i) = g_{i,i}(m). \end{aligned}$$

因此

$$g_{i,i}(m+1) = \sum_{k=1}^{\infty} g_{i,i}(m)P_i(\tau_i = k) = g_{i,i}(m)\sum_{k=1}^{\infty} f_{i,i}^{(k)} = g_{i,i}(m)f_{i,i}.$$

由此递归关系并注意到$g_{i,i}(1) = f_{i,i}$可得, 对任意$m \geq 1$,

$$g_{i,i}(m) = f_{i,i}^m.$$

令$m \to \infty$, 由概率连续性可知$g_{i,i} = \lim_{m\to\infty} f_{i,i}^m$, 因此后一结论自然成立. □

概念“常返”的定义中只要求状态几乎必然重复出现, 由命题3.2.25 可知, 这样的状态几乎必然可以无穷次地出现; 反之, 若一个状态是非常返的, 那么对几乎所有的ω, 都存在一个时刻$N(\omega)$ 使得在$N(\omega)$ 后该状态不再出现, 对后一论断的解释、证明, 我们留给同学自己完成(参见本节习题14).

★定理3.2.26. 记马氏链X首次回到状态i的时间为τ_i. 设$\tau_i < \infty$. 令$Y_n = X_{\tau_i+n}$, 即对任意ω, $Y_n(\omega) = X_{\tau_i(\omega)+n}(\omega)$. 那么$Y = \{Y_n;\ n \geq 0\}$ 是与X 有相同转移概率矩阵的马氏链且与$(\tau_i, X_0, \cdots, X_{\tau_i})$ 相互独立.

证明 首先注意到$Y_0 = X_{\tau_i} = i$是常量, 对任意$n \geq 1$, $k_0 \in S, k_1, \cdots, k_{n-1} \in$

$S\setminus\{i\}$, $m\ge 1$, $0\le n_1<\cdots<n_m$, $j_1,\cdots,j_m\in S$. 令$k_n=i$,

$$
\begin{aligned}
&P\Big(\bigcap_{l=0}^{n-1}\{X_l=k_l\}\cap\{\tau_i=n\}\cap\bigcap_{r=1}^{m}\{Y_{n_r}=j_r\}\Big)\\
&\quad=P\Big(\bigcap_{l=0}^{n}\{X_l=k_l\}\cap\bigcap_{r=1}^{m}\{X_{n+n_r}=j_r\}\Big)\\
&\quad=P\Big(\bigcap_{l=0}^{n-1}\{X_l=k_l\}\cap\{X_n=i\}\Big)P\Big(\bigcap_{r=1}^{m}\{X_{n+n_r}=j_r\}|X_n=i\Big)\\
&\quad=P\Big(\bigcap_{l=0}^{n-1}\{X_l=k_l\}\cap\{X_n=i\}\Big)P_i\Big(\bigcap_{r=1}^{m}\{X_{n_r}=j_r\}\Big).
\end{aligned}
$$

两边对全部的$n\ge 1$, $k_0\in S$以及$k_1,\cdots,k_{n-1}\in S\setminus\{i\}$求和得

$$P\Big(\bigcap_{r=1}^{m}\{Y_{n_r}=j_r\}\Big)=P_i\Big(\bigcap_{r=1}^{m}\{X_{n_r}=j_r\}\Big).$$

进而

$$
\begin{aligned}
&P\Big(\bigcap_{l=0}^{n-1}\{X_l=k_l\}\cap\{\tau_i=n\}\cap\bigcap_{r=1}^{m}\{Y_{n_r}=j_r\}\Big)\\
&\quad=P\Big(\bigcap_{l=0}^{n-1}\{X_l=k_l\}\cap\{X_n=i\}\Big)P_i\Big(\bigcap_{r=1}^{m}\{Y_{n_r}=j_r\}\Big).
\end{aligned}
$$

因此Y是与X有相同转移概率函数的马氏链且与$(\tau_i,X_0,\cdots,X_{\tau_i})$相互独立. □

定理3.2.26 表明将τ_i看作起点, 此后的随机过程Y本质就是0点从i出发的马氏链X, 而且τ_i之后的随机事件与τ_i之前的事件相互独立.

练习题3.2 (除特别说明外, 以下总设X为马氏链, 转移概率矩阵为$(p_{i,j})_{i,j\in S}$.)

1 若状态$i\neq j$, 证明$i\to j$当且仅当在对应的一步转移图中顶点i,j有向连通.

2 设X的状态空间$S=\{0,1,2,3,4\}$, 转移概率矩阵为

$$\begin{pmatrix}1/3&2/3&0&0&0\\0&1/3&1/3&1/3&0\\3/4&1/4&0&0&0\\0&0&0&0&1\\0&0&0&1/2&1/2\end{pmatrix}.$$

试写出状态空间S中所有互通类, 判断它们是否是本质类并说明理由.

3 若X的状态空间S可分为有限个不同的互通类, 那么这些互通类中至少存在一个本质互通类. 试以S中包含3个不同互通类为例给出详细证明.

4 若$i\leftrightarrow j$且记它们的周期为d, 若$p_{i,j}^{(n)}>0,p_{i,j}^{(m)}>0$, 证明$d|n-m$.

5 若状态i非常返, 那么对任意$j\in S$, $\lim\limits_{n\to\infty}p_{j,i}^{(n)}=0$.

6 若j为吸收状态, 那么$\lim\limits_{n\to\infty} p_{i,j}^{(n)} = f_{i,j}$.

7 若$C(i)$, $C(j)$是两不相同的本质互通类, 那么$f_{i,j}=0$.

8 元素有限的本质类必是常返类.

9 证明平面上简单对称随机游动是常返的.

10* 试构造马氏链X, 使得它的状态空间至少包含一个具有无穷多状态的非本质互通类(记作S_1), 而且当$X_0\in S_1$时$P(X_n\in S_1, n\geq 1)>0$.

11* 设S_1是X的一个具有有限状态的非本质互通类, $X_0\in S_1$. 证明对几乎所有的ω, 存在$N(\omega)$ 使得对所有的$n\geq N(\omega)$, $X_n(\omega)\notin S_1$.

12* 若马氏链X的转移概率矩阵能用分块矩阵表示为$\begin{pmatrix} A & B \\ O & C \end{pmatrix}$, 其中子块$A, C$都是方阵. 证明若$B$不是零子块, 那么$B$ 中非零行对应的状态都是非本质的.

13* 设X是不可约d周期马氏链, 状态空间为S.对任意$i\in S$, 令

$$S_k=\{j\in S;\ p_{i,j}^{(nd+k)}>0, n\geq 0\},\qquad k=0,\cdots,d-1.$$

证明(1) $S_k, 0\leq k\leq d-1$, 互不相交而且$S=\cup_{k=0}^{d-1}S_k$. (2)令$Y_n=X_{nd}$, 那么$Y=\{Y_n;\ n\geq 0\}$是以$\mathbf{P}^{(d)}$为一步转移概率矩阵的非周期马氏链. 特别地, 若$X_0\in S_k$, 那么Y是状态空间为S_k的非周期马氏链.

14* 若i是马氏链$X=\{X_n; n\geq 0\}$的非常返状态, 证明对几乎所有的ω, 都存在一个时刻$N(\omega)$ 使得对任意$n\geq N(\omega)$, $X_n(\omega)\neq i$.

15* 对任意$k\geq 1$, 令$\tau_i^{(k)}=\inf\{n>\tau_i^{(k-1)};\ X_n=i\}$以及$W_i^{(k)}=\tau_i^{(k)}-\tau_i^{(k-1)}$, 其中$\tau_i^{(0)}=0$. 证明: $W_i^{(k)}, k\geq 1$, 相互独立并且当$k\geq 2$时分布相同.

16* 若状态j的周期为d, 证明$\{n\geq 1;\ f_{j,j}^{(n)}>0\}$中元素的最大公约数也为$d$.

17* 令$\tau=\inf\{n\geq 0;\ \ X_n\geq 1\}$, 证明: 对任意$n\geq 0$,

$$P(X_{\tau+n}=j|X_\tau=i)=P(X_n=j|X_0=i),\quad i\geq 1.$$

18* 设不可约链X的状态空间$S=\{0,1,2,\cdots\}$, 若存在$\{y_i;\ i\in S\}$及某个状态j使得对任意$i\neq j$, $\sum\limits_{k\in S}p_{i,k}y_k\leq y_i$, 而且$\lim\limits_{i\to+\infty}y_i=+\infty$, 证明$X$常返.

3.3 首访概率与时间

在马氏链的研究和应用当中, 状态的首次到达时间或首次回访时间是非常重要的观察变量. 这一节我们将刻画这类观察变量的一些统计特征.

3.3.1 访问概率与分布

上一节我们定义了状态j的首回时τ_j, 并利用$f_{j,j} = P_j(\tau_j < \infty)$是否为1定义了状态的常返性. 稍多做一点分析后大家就会发现对任意的状态i 而言,

$$f_{i,j} = P_i(\tau_j < \infty) \tag{3.3.1}$$

给出了从i出发有限时间内能观察到状态j的概率, 从而一定程度上定量地给出了状态i,j先后出现的可能性.不仅如此,

$$f_{i,j}^{(n)} = P_i(\tau_j = n), \qquad n \geq 1,$$

还告诉我们在哪些时间上能观察到状态j并定量地给出了这种可能性大小. 这一小节我们就如何计算$f_{i,j}$和$f_{i,j}^{(n)}$做些初步的介绍.

对任意$i, j \in S$以及任意$u \in [0,1]$, 定义

$$F_{i,j}(u) = \sum_{n=1}^{\infty} f_{i,j}^{(n)} u^n.$$

显然当$f_{i,j} = 1$时, 序列$\{f_{i,j}^{(n)}\}$就是随机变量τ_j在初值为i条件下的分布列, $F_{i,j}(u)$就是该条件分布的概率母函数; 当$f_{i,j} < 1$ 时, 尽管$\{f_{i,j}^{(n)};\ n \geq 1\}$不再是$\tau_j$的条件分布列了, 但由幂级数与系数关系可知序列$\{f_{i,j}^{(n)}\}$ 与函数$F_{i,j}(u)$ 仍是一一对应的. 特别地, 若$i \not\to j$, 则$F_{i,j}(u) \equiv 0$. 因此, 无论怎样, $F_{i,j}(u)$决定了所有$f_{i,j}^{(n)}$的值, 并且总有$f_{i,j} = F_{i,j}(1)$.

关于$F_{i,j}(u)$ 我们有下面的结论.

★**定理3.3.1.** *对任意给定的$j \in S$以及$u \in [0,1]$, $\{F_{i,j}(u); i \in S\}$是线性方程组*

$$z_i = \sum_{k \neq j} u p_{i,k} z_k + u p_{i,j}, \qquad i \in S, \tag{3.3.2}$$

的最小非负解. 即$\{F_{i,j}(u); i \in S\}$本身是方程组(3.3.2)的非负解而且对该方程组的任意非负解 $\{g_i(u); i \in S\}$, $F_{i,j}(u) \leq g_i(u)$ 对一切$i \in S$ 都成立.

证明 对任意$m \geq 1$, 我们可知

$$\begin{aligned} f_{i,j}^{(m)} &= P_i(\tau_j = m) = \sum_{k \neq j} P_i(X_m = j, X_1 = k, X_v \neq j, 1 < v < m) \\ &= \sum_{k \neq j} P_i(X_1 = k) P(X_m = j, X_v \neq j, 2 < v < m | X_1 = k) = \sum_{k \neq j} p_{i,k} f_{k,j}^{(m-1)}. \end{aligned}$$

对任意$n \ge 1$, 记$A(i,n,u)=\sum\limits_{m=1}^{n} f_{i,j}^{(m)}u^m$, 那么

$$\begin{aligned}A(i,n+1,u) &= \sum_{m=1}^{n+1} f_{i,j}^{(m)}u^m = f_{i,j}^{(1)}u + \sum_{m=2}^{n+1} f_{i,j}^{(m)}u^m = p_{i,j}u + \sum_{m=2}^{n+1} f_{i,j}^{(m)}u^m \\ &= up_{i,j} + \sum_{k\neq j}\sum_{m=2}^{n+1} up_{i,k}f_{k,j}^{(m-1)}u^{m-1} = up_{i,j} + u\sum_{k\neq j} p_{i,k}A(k,n,u).\end{aligned}$$

注意到$\lim\limits_{n\to\infty} A(i,n,u) = F_{i,j}(u)$, 上式两边令$n\to\infty$得(参见第一章第一节习题5)

$$F_{i,j}(u) = up_{i,j} + u\sum_{k\neq j} p_{i,k}F_{k,j}(u).$$

因此$\{F_{i,j}(u); i\in S\}$确实为方程组(3.3.2)的非负解.

任取(3.3.2)的一组非负解$\{g_i(u); i\in S\}$, 由

$$g_i(u) = up_{i,j} + u\sum_{k\neq j} p_{i,k}g_k(u)$$

可知, 对任意$i\in S$,

$$g_i(u) \ge A(i,1,u) = up_{i,j}.$$

设对任意$i\in S$, $g_i(u)\ge A(i,k,u)$对 $k=n-1$成立, 那么由

$$\begin{aligned}A(i,n,u) &= up_{i,j} + \sum_{k\neq j} up_{i,k}A(k,n-1,u) \\ &\le up_{i,j} + \sum_{k\neq j} up_{i,k}g_k(u) = g_i(u)\end{aligned}$$

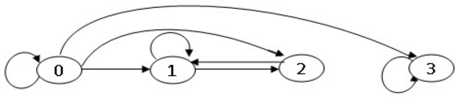
可知对任意$i\in S$, $g_i(u)\ge A(i,k,u)$对$k=n$也成立.

由数学归纳法可知, 对任意$i\in S$以及任意n都有$g_i(u)\ge A(i,n,u)$. 对任意$i\in S$, 令$n\to\infty$得$g_i(u)\ge F_{i,j}(u)$. 故$F_{i,j}(u)$为(3.3.2)的最小非负解. □

由定理3.3.1为通过求解代数方程找到$\{F_{i,j}(u);\ i\in S\}$分布提供了理论基础. 若X是有限状态马氏链, 我们就可以用代数方法求解线性方程组(3.3.2), 然后从中找到最小非负解即可. 特别地, 若线性方程组只有唯一解那么该唯一解就是$\{F_{i,j}(u);\ i\in S\}$.

▶**例3.3.2.** 设X的转移概率矩阵为$\begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, 状态空间为$\{0,1,2,3\}$. 求$P_0(\tau_1=k)$ 和 $F_{i,1}(u)$, 其中$i=0,1,2,3$, k为任意正整数.

解 由转移概率矩阵我们可以画出状态一步转移图如下.



由此可知$3 \not\to 1$, 因此对任意$u \in [0,1]$, $F_{3,1}(u) = 0$, 进而由定理3.3.1可得

$$\begin{cases} F_{0,1}(u) = up_{0,0}F_{0,1}(u) + up_{0,2}F_{2,1}(u) + up_{0,1}, \\ F_{2,1}(u) = up_{2,0}F_{0,1}(u) + up_{2,2}F_{2,1}(u) + up_{2,1}, \\ F_{1,1}(u) = up_{1,0}F_{0,1}(u) + up_{1,2}F_{2,1}(u) + up_{1,1}. \end{cases}$$

由前两个方程得

$$\begin{pmatrix} 1-up_{0,0} & -up_{0,2} \\ -up_{2,0} & 1-up_{2,2} \end{pmatrix} \begin{pmatrix} F_{0,1}(u) \\ F_{2,1}(u) \end{pmatrix} = \begin{pmatrix} up_{0,1} \\ up_{2,1} \end{pmatrix}.$$

将$p_{0,0} = p_{0,2} = p_{0,1} = 1/4,\ p_{2,0} = p_{2,2} = 0,\ p_{2,1} = 1$代入并计算得

$$F_{0,1}(u) = \frac{u+u^2}{4-u}, \quad F_{2,1}(u) = u.$$

将此结果代入第三个方程并注意到$p_{1,0} = 0, p_{1,1} = 1/3, p_{1,2} = 2/3$, 我们可得

$$F_{1,1}(u) = \frac{2}{3}u^2 + \frac{1}{3}u.$$

对任意$0 \le u \le 1$, 由幂级数展开式

$$\sum_{k=1}^{\infty} f_{0,1}^{(k)} u^k = F_{0,1}(u) = \left[\frac{u}{4} + \frac{u^2}{4}\right] \sum_{k=0}^{\infty} \frac{u^k}{4^k} = \frac{u}{4} + \sum_{k=2}^{\infty} \frac{5u^k}{4^k},$$

可知$f_{0,1}^{(1)} = 1/4, \quad f_{0,1}^{(k)} = 5/4^k, \quad k \ge 2.$ □

▶**例3.3.3.** 设W为直线上(q,p)-简单随机游动, 求$F_{0,1}(u)$, $u \in [0,1]$.

解 由定理3.3.1可知$\{F_{i,1}(u)\}_{i\in\mathbb{Z}}$是方程组

$$z_i = \sum_{k\ne 1} up_{i,k}z_k + up_{i,1}, \qquad i \in \mathbb{Z},$$

的最小非负解. 从而

$$F_{0,1}(u) = up + uqF_{-1,1}(u). \tag{3.3.3}$$

注意到对任意$n \ge 1$以及任意$i \in \mathbb{Z}$, $j \ge 2$, 由独立增量性可得

$$\begin{aligned} f_{i,i+j}^{(n)} &= P(\tau_{i+j} = n | W_0 = i) = P(W_n = i+j, W_k < i+j, 0<k<n | W_0 = i) \\ &= P(W_n - W_0 = j, W_k - W_0 < j, 0<k<n) \\ &= P(W_n = j, W_k < j, 0<k<n | W_0 = 0) = f_{0,j}^{(n)}. \end{aligned}$$

特别地,

$$\begin{aligned} f_{0,2}^{(n)} &= P(W_n = 2, W_k < 2, 0<k<n | W_0 = 0) \\ &= \sum_{j=1}^{n-1} P(W_n = 2, W_k < 2, j<k<n, W_j = 1, W_l < 1, 0<l<j | W_0 = 0). \end{aligned}$$

由马氏性可得

$$
\begin{aligned}
&P(W_n=2,W_k<2,j<k<n,W_j=1,W_l<1,0<l<j|W_0=0)\\
&\quad=P(W_j=1,W_l<1,0<l<j|W_0=0)\\
&\qquad\times P(W_n=2,W_k<2,j<k<n|W_j=1).
\end{aligned}
$$

注意到W还是时齐的,

$$
\begin{aligned}
&P(W_n=2,W_k<2,j<k<n|W_j=1)\\
&\quad=P(W_{n-j}=2,W_k<2,0<k<n-j|W_0=1)=f_{1,2}^{(n-j)}.
\end{aligned}
$$

因此

$$
f_{0,2}^{(n)}=\sum_{j=1}^{n-1}f_{0,1}^{(j)}f_{1,2}^{(n-j)}=\sum_{j=1}^{n-1}f_{0,1}^{(j)}f_{0,1}^{(n-j)}.
$$

由此可得

$$
\begin{aligned}
F_{-1,1}(u)=F_{0,2}(u)&=\sum_{n=2}^{\infty}f_{0,2}^{(n)}u^n=\sum_{n=2}^{\infty}\sum_{j=1}^{n-1}f_{0,1}^{(j)}u^j f_{0,1}^{(n-j)}u^{n-j}\\
&=\sum_{j=1}^{\infty}f_{0,1}^{(j)}u^j\sum_{k=1}^{\infty}f_{0,1}^{(k)}u^k=\big[F_{0,1}(u)\big]^2.
\end{aligned}
$$

将其代入(3.3.3), 此时方程两个可能的解分别为

$$
\frac{1+\sqrt{1-4pqu^2}}{2qu},\qquad \frac{1-\sqrt{1-4pqu^2}}{2qu}.
$$

注意到$F_{0,1}(u)$为最小非负解, 因此

$$
F_{0,1}(u)=\frac{1-\sqrt{1-4pqu^2}}{2qu}=\frac{2pu}{1+\sqrt{1-4pqu^2}},\qquad u\in[0,1]. \qquad \square
$$

注意到$f_{i,j}=F_{i,j}(1)$, 由定理3.3.1我们容易得到如下结论.

♦**推论3.3.4.** 对任意给定的$j\in S$, $\{f_{i,j};i\in S\}$是线性方程组

$$
z_i=\sum_{k\neq j}p_{i,k}z_k+p_{i,j},\qquad i\in S, \tag{3.3.4}
$$

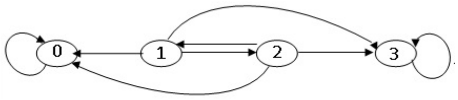
的最小非负解.

▶**例3.3.5.** 一家银行将贷款分类为全部付清(0), 信誉良好(1), 拖欠(2)或者呆账(3). 在银行每次清算时, 贷款按照如下转移概率在不同的类别之间转移:

$$
\begin{pmatrix}
1 & 0 & 0 & 0\\
0.1 & 0.8 & 0.1 & 0\\
0.1 & 0.4 & 0.4 & 0.1\\
0 & 0 & 0 & 1
\end{pmatrix}.
$$

求处于信誉良好级别的贷款最终全部付清的比例.

解 以$X=\{X_n;n\geq 0\}$表示贷款每次清算时状态的变化. 那么X是个时齐马氏链. 所求比例为概率$f_{1,0}$. 由转移概率矩阵我们可以画出状态一步转移图如下:



由于$3\not\to 0$, $f_{3,0}=0$. 由推论3.3.4可知

$$\begin{cases} f_{1,0}=p_{1,1}f_{1,0}+p_{1,2}f_{2,0}+p_{1,0}=0.8f_{1,0}+0.1f_{2,0}+0.1, \\ f_{2,0}=p_{2,1}f_{1,0}+p_{2,2}f_{2,0}+p_{2,0}=0.4f_{1,0}+0.4f_{2,0}+0.1. \end{cases}$$

由此解得$f_{1,0}=7/8$, $f_{2,0}=3/4$. 因此最终付清得比例为$f_{1,0}=87.5\%$. □

方程组(3.3.2)与(3.3.4)在理论上也很重要, 它给我们提供了一种通过判断方程组的解的性质确定马氏链常返性的方法.

★定理3.3.6. *设X不可约, 那么X常返当且仅当存在一个状态$j\in S$使得方程组(3.3.4)的任意一组非负解$\{u_i;\ i\in S\}$满足$u_i\geq 1$ 对一切$i\in S$成立.*

证明 "必要性": 若不可约链X常返, 则由推论3.2.22和性质3.2.19可知对任意$i,j\in S$, $f_{i,j}=1$. 从而方程组(3.3.2)的最小非负解$\{f_{i,j};\ i\in S\}$恒为1, 因此$u_i\geq 1$ 对一切$i\in S$成立.

"充分性": 由假设可知$\{u_i\equiv 1;i\in S\}$为(3.3.4)的最小非负解, 从而由推论3.3.4 可知$f_{j,j}=1$, 即j常返, 再由推论3.2.22 可知不可约链X为常返的. □

▶例3.3.7. *设马氏链X的状态空间为非负整数, 转移概率矩阵为*

$$\mathbf{P}=\begin{pmatrix} r_0 & p_0 & 0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

其中$p_0>0$, $r_0+p_0=1$, $q_i+r_i+p_i=1$, $p_i,q_i>0$, $i\geq 1$. 讨论X的常返性.

解 在条件假设下容易证明X是不可约马氏链. 考察$\{f_{n,0},n\geq 0\}$满足的方程组

$$\begin{cases} z_0=\sum_{k\neq 0}p_{0,k}z_k+p_{0,0}=p_0z_1+r_0 \\ z_1=\sum_{k\neq 0}p_{1,k}z_k+p_{1,0}=r_1z_1+p_1z_2+q_1 \\ z_n=\sum_{k\neq 0}p_{z,k}z_k+p_{n,0}=q_nz_{n-1}+r_nz_n+p_nz_{n+1}, \quad n\geq 2. \end{cases} \tag{3.3.5}$$

可得

$$z_{n+1}-z_n=\frac{q_n}{p_n}(z_n-z_{n-1})=\cdots=\frac{q_1q_2\cdots q_n}{p_1p_2\cdots p_n}(z_1-1).$$

由此可知(3.3.5)的非负解必满足$z_0 = p_0 z_1 + r_0$, 而且对任意$n \geq 1$,

$$z_{n+1} = z_1 + (z_1 - 1)\sum_{k=1}^{n} \frac{q_1 \cdots q_k}{p_1 \cdots p_k}.$$

(1)当$\sum_{k=1}^{\infty} \frac{q_1 \cdots q_k}{p_1 \cdots p_k} = \infty$时, 由$\{z_n, n \geq 0\}$非负可知$z_1 \geq 1$, 从而$z_n \geq 1$, $n \geq 0$. 由定理3.3.6, 此时X常返.

(2)当$A = \sum_{k=1}^{\infty} \frac{q_1 \cdots q_k}{p_1 \cdots p_k} < \infty$时, 可取$z_1 \in (A/(1+A), 1)$, 使得

$$0 < z_0 = p_0 z_1 + r_0 < 1 \quad \text{而且} \quad 0 < z_1 + A(z_1 - 1) < z_n \leq z_1 < 1$$

对所有$n \geq 1$都成立, 即(3.3.4)存在非负解$\{z_i, i \geq 0\}$使得$z_1 < 1$. 此时由定理3.3.6可知X是非常返的. □

3.3.2 平均访问时间

前面我们给出了函数$F_{i,j}(u)$满足的方程, 借助$F_{i,j}(u)$我们可以获得τ_j的分布信息. 自然地, 我们也可以考虑τ_j的一些数值特征. 比如当$f_{i,j} = 1$即$P_i(\tau_j < \infty) = 1$时,

$$\sum_{n=1}^{\infty} n f_{i,j}^{(n)} = \sum_{n=1}^{\infty} n P_i(\tau_j = n) = E_i(\tau_j)$$

表示了从i出发首次回到或到达状态j的平均时间. 虽然$f_{i,j} < 1$时, 上面这个等式不再成立, 但求和式

$$\sum_{n=1}^{\infty} n f_{i,j}^{(n)} = f_{i,j}\sum_{n=1}^{\infty} n \frac{f_{i,j}^{(n)}}{f_{i,j}} = f_{i,j}\sum_{n=1}^{\infty} n P_i(\tau_j = n | \tau_j < \infty) = f_{i,j} E_i(\tau_j | \tau_j < \infty)$$

仍能给我们提供一些重要信息. 在这一小节, 我们就来讨论这个求和式的计算与应用. 为此对任意$i \in S$和$j \in S$, 我们记

$$m_{i,j} = \sum_{n=1}^{\infty} n f_{i,j}^{(n)}. \tag{3.3.6}$$

显然, 若$i \not\to j$, 那么$m_{i,j} = 0$; 若i为吸收态, 那么$m_{i,i} = 1$; 若i是常返的, 那么$m_{i,i}$就是从i出发首次回到自己的平均时间间隔.

注意到

$$m_{i,j} = \frac{\mathrm{d}F_{i,j}(u)}{\mathrm{d}u}\Big|_{u=1}.$$

因此若解方程(3.3.2)得到函数$F_{i,j}(u)$, 我们便可以通过求导运算方便地算出$m_{i,j}$.

▶**例3.3.8.** *对例3.3.2中X, 求$m_{i,1}$, 其中$i = 0, 1, 2, 3$.*

解 对例3.3.2中结果$F_{i,1}(u)$求导, 并在导函数中令$u = 1$可得

$$m_{0,1} = 11/9, \quad m_{1,1} = 5/3, \quad m_{2,1} = 1, \quad m_{3,1} = 0.$$ □

通过求解方程组(3.3.2)得到函数$F_{i,j}(u)$有时候并不方便, 下面的定理给出了直接通过解方程组得到$m_{i,j}$的方法.

★**定理3.3.9.** 对任意给定的$j\in S$, $\{m_{i,j};i\in S\}$是线性方程组

$$z_i=\sum_{k\neq j}p_{i,k}z_k+f_{i,j},\qquad i\in S \tag{3.3.7}$$

的最小非负解.

证明 对任意$i,j\in S$以及$k\geq 1$, 记$b_{i,j}(n)=\sum_{m=1}^{n}mf_{i,j}^{(m)}$. 显然$n\to\infty$时$b_{i,j}(n)$单调不降且收敛到$m_{i,j}$. 此外, 我们有

$$\begin{aligned}b_{i,j}(n+1)&=\sum_{m=1}^{n+1}mf_{i,j}^{(m)}=\sum_{m=1}^{n}mf_{i,j}^{(m+1)}+\sum_{m=1}^{n+1}f_{i,j}^{(m)}\\&=\sum_{m=1}^{n+1}f_{i,j}^{(m)}+\sum_{m=1}^{n}\sum_{k\neq j}mP_i(X_{m+1}=j,X_1=k,X_v\neq j,2\leq v\leq m)\\&=\sum_{m=1}^{n+1}f_{i,j}^{(m)}+\sum_{k\neq j}p_{i,k}\sum_{m=1}^{n}mf_{k,j}^{(m)}=\sum_{k\neq j}p_{i,k}b_{k,j}(n)+\sum_{m=1}^{n+1}f_{i,j}^{(m)}.\end{aligned}$$

两边令$n\to\infty$得

$$m_{i,j}=f_{i,j}+\sum_{k\neq j}p_{i,k}m_{k,j}.$$

因此$\{m_{i,j}\}_{i\in S}$为方程组(3.3.7)的非负解. 任取(3.3.7)的一组非负解$\{u_i\}_{i\in S}$, 由

$$u_i=f_{i,j}+\sum_{k\neq j}p_{i,k}u_k$$

可知$u_i\geq b_{i,j}(1)=f_{i,j}^{(1)}$对任意$i\in S$成立.

设对任意$i\in S$, $u_i\geq b_{i,j}(k)$对$k=n-1$成立, 那么由

$$b_{i,j}(n)=\sum_{m=1}^{n+1}f_{i,j}^{(m)}+\sum_{k\neq j}p_{i,k}b_{k,j}(n-1)\leq f_{i,j}+\sum_{k\neq j}p_{i,k}u_k=u_i,$$

我们可知$u_i\geq b_{i,j}(k)$对$k=n$也成立.

由数学归纳法可知, 对任意$i\in S$以及任意$n\geq 1$都有

$$u_i\geq b_{i,j}(n).$$

令$n\to\infty$ 得$u_i\geq m_{i,j}$对所有$i\in S$成立, 因此$m_{i,j}$为(3.3.7)的最小非负解. □

▶**例3.3.10.** (接例*3.3.5*) 在假定信誉良好级别的贷款最终全部付清的条件下, 问平均要多少个周期才能让信誉良好级别的贷款最终全部付清(每清算一次算一个周期).

解 令$X_0=1$, $\tau_0=\min\{n\geq 1;\ X_n=0\}$. 那么$X=\{X_n;n\geq 0\}$表示信誉良好级别的贷款在偿付过程中的状态变化, $\{\tau_0<\infty\}$表示信誉良好级别的贷款最

终全部付清. 因此所求平均付清周期数

$$E(\tau_0|\tau_0<\infty,X_0=1)=\sum_{k=1}^{\infty}kP(\tau_0=k|\tau_0<\infty,X_0=1)$$
$$=\sum_{k=1}^{\infty}k\frac{P(\tau_0=k,\tau_0<\infty,X_0=1)}{P(\tau_0<\infty,X_0=1)}=\sum_{k=1}^{\infty}k\frac{P(\tau_0=k,X_0=1)}{P(\tau_0<\infty,X_0=1)}$$
$$=\sum_{k=1}^{\infty}k\frac{P(\tau_0=k|X_0=1)}{P(\tau_0<\infty|X_0=1)}=\sum_{k=1}^{\infty}k\frac{f_{1,0}^{(k)}}{f_{1,0}}=\frac{m_{1,0}}{f_{1,0}}.$$

由定理3.3.9可知, $m_{i,0},i=0,1,2,3$, 满足下列方程组.

$$\begin{cases}m_{0,0}=p_{0,1}m_{1,0}+p_{0,2}m_{2,0}+p_{0,3}m_{3,0}+f_{0,0},\\ m_{1,0}=p_{1,1}m_{1,0}+p_{1,2}m_{2,0}+p_{1,3}m_{3,0}+f_{1,0},\\ m_{2,0}=p_{2,1}m_{1,0}+p_{2,2}m_{2,0}+p_{2,3}m_{3,0}+f_{2,0},\\ m_{3,0}=p_{3,1}m_{1,0}+p_{3,2}m_{2,0}+p_{3,3}m_{3,0}+f_{3,0},\end{cases}$$

其中由例3.3.5可知$f_{1,0}=0.875$, $f_{2,0}=0.75$. 注意到$3\not\to 0$, 因此$m_{3,0}=0$. 将这些信息以及$p_{i,j}$的值代入上述方程组后可得

$$\begin{cases}m_{1,0}=0.8m_{1,0}+0.1m_{2,0}+0.875,\\ m_{2,0}=0.4m_{1,0}+0.4m_{2,0}+0.75.\end{cases}$$

由此可解得$m_{1,0}=7.5$. 因此所求平均付清周期数为$7.5/0.875=60/7$. □

▶**例3.3.11.** *设有$n+1$个选手依次两两进行比赛($n\geq 1$), 获胜者接连比赛下去直到有一名选手连续地击败其余n名选手才结束. 假设每场比赛中两名选手获胜的概率各为1/2, 求平均需比赛次数.*

解 以X_k表示第k次比赛结束时比赛胜者的连赢场次, 并令$X_0=0$. 那么$X=\{X_k;\ k\geq 0\}$为时齐马氏链, 转移概率矩阵

$$\mathbf{P}=\begin{array}{c}\begin{pmatrix}0&1&0&0&0&0&\cdots\\0&1/2&1/2&0&0&0&\cdots\\0&1/2&0&1/2&0&0&\cdots\\0&1/2&0&0&1/2&0&\cdots\\0&1/2&0&0&0&1/2&\cdots\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\end{pmatrix}\end{array}\begin{array}{c}0\\1\\2\\3\\4\\\vdots\end{array}$$

由此可知状态集$C(1)=\{1,2,\cdots\}$是X的一个互通类,

$$f_{1,1}=\sum_{m=1}^{\infty}f_{1,1}^{(m)}=\sum_{m=1}^{\infty}\frac{1}{2^m}=1.$$

所以$C(1)$是X的不可约常返类, 进而$f_{i,j}=1,i,j\geq 1$. 比赛结束的条件是$X_k=n$,

比赛结束的时间为$\tau_n = \inf\{k \geq 1, X_k = n\}$. 从而平均比赛次数

$$T = E_0(\tau_n) = m_{0,n}.$$

显然$m_{0,n} = m_{1,n} + 1$, 由定理3.3.9可知$m_{i,n}, 1 \leq i < n$,是方程组

$$\begin{cases} z_i = z_1/2 + z_{i+1}/2 + 1, & 1 \leq i < n-1, \\ z_{n-1} = z_1/2 + 1. \end{cases}$$

的最小非负解. 由此可得

$$\begin{cases} z_i = 2z_{i-1} - z_1 - 2, & 2 \leq i \leq n-1, \\ z_{n-1} = z_1/2 + 1. \end{cases}$$

因此对任意$2 \leq i \leq n-1$,

$$\begin{aligned} z_i &= 2^2 z_{i-2} - 2z_1 - 2^2 - z_1 - 2 \\ &= 2^{i-1} z_1 - (2^{i-2} + \cdots + 2 + 1) z_1 - (2^{i-1} + \cdots + 2) \\ &= z_1 - (2^i - 2). \end{aligned}$$

从而由$z_{n-1} = z_1 - 2^{n-1} + 2 = z_1/2 + 1$可知$z_1 = 2^n - 2$. 因此所求比赛次数

$$T = m_{0,n} = m_{1,n} + 1 = z_1 + 1 = 2^n - 1. \qquad \square$$

▶**例3.3.12.** *在例3.3.7中令$r_0 = 0$, $p_0 = 1$, 并且对一切$i \geq 1$, $r_i = 0$, $q_i = q$, $p_i = p$. 由此得到的马氏链X通常被称为带一个反射壁的(q,p)- 随机游动或非负整数值的反射(q,p)-随机游动(直观看就是在正半轴上运动的简单随机游动, 一旦到达位置0, 下一步就返回位置1). 对任意$n \geq 2$, 求$m_{0,n}$.*

解 我们将$q \leq p$的情形留给同学做练习, 这里我们只考虑$q > p$的情形. 此时由例3.3.7可知X常返, 对任意$i, j \geq 0$, $f_{i,j} = 1$. $\{m_{l,n}\}_{0 \leq l < n}$为下列方程组的解.

$$\begin{aligned} z_0 &= \sum_{k \neq n} p_{0,k} z_k + 1 = z_1 + 1, \\ z_l &= \sum_{k \neq n} p_{l,k} z_k + 1 = q z_{l-1} + p z_{l+1} + 1, \quad 1 \leq l \leq n-2, \qquad (3.3.8) \\ z_{n-1} &= \sum_{k \neq n} p_{n-1,k} z_k + 1 = q z_{n-2} + 1. \end{aligned}$$

记$r = q/p$. 由(3.3.8)可知, 对任意$2 \leq l \leq n-1$,

$$z_l - z_{l-1} = r(z_{l-1} - z_{l-2}) - (1+r),$$

以及

$$z_l - r z_{l-1} = z_{l-1} - r z_{l-2} - (1+r).$$

结合方程$z_0 - z_1 = 1$, 进一步可得

$$z_{n-1} - z_{n-2} = r^{n-2}(z_1 - z_0) - (1+r)(1 + r + \cdots + r^{n-3})$$

$$= -r^{n-2} - (1+r)(1+\cdots+r^{n-3}),$$

以及

$$\begin{aligned} z_{n-1} - rz_{n-2} &= z_1 - rz_0 - (n-2)(1+r) \\ &= (1-r)z_0 - 1 - (n-2)(1+r). \end{aligned}$$

再结合方程$z_{n-1} = qz_{n-2} + 1$, 即

$$(1+r)z_{n-1} = rz_{n-2} + 1 + r,$$

我们可得

$$\begin{cases} \frac{1+r-z_{n-1}}{r} = -r^{n-2} - (1+r)(1+\cdots+r^{n-3}), \\ 1+r-rz_{n-1} = (1-r)z_0 - 1 - (n-2)(1+r). \end{cases}$$

解方程得

$$z_{n-1} = \frac{2r^n - r - 1}{r-1}, \qquad z_0 = \frac{2r^{n+1} + n - nr^2 - 2r}{(r-1)^2}.$$

后者即为所求的$m_{0,n}$. □

♠**注记3.3.13.** 注意到当$i \neq j$时$f_{i,j}, m_{i,j}$只依赖于到达j之前各状态的转移概率, 与从j出发的转移概率$p_{j,\cdot}$无关. 从而改变$p_{j,\cdot}$不影响$f_{i,j}$, $m_{i,j}$的取值与判断.

▶**例3.3.14.** 考虑在整数点$\{0,1,2,\cdots,N\}$上的(q,p)-随机游动, 使得当X落在0位置时下一步就反弹为位置1, 当X落在位置N时下一步就反弹回$N-1$. 通常我们称这样的随机过程为带两个反射壁的随机游动. 它的转移概率矩阵为

$$\mathbf{P} = \begin{array}{c} \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ N-1 \\ N \end{array}$$

对这样的随机过程, 当我们考虑$f_{0,k}$和$m_{0,k}$时$(0 < k \leq N)$, 我们可以发现: 在到达k之前, 随机过程只是在状态集$\{0,1,\cdots,k-1\}$内变化, 而且这期间状态变化规律与例*3.3.12*中带单个反射壁的随机游动一致. 因此根据上面的注记, 这两种带反射壁的随机游动有相同的$f_{0,k}$和$m_{0,k}$.

练习题3.3

1 设X的状态空间为$\{1,2,3\}$, 转移概率矩阵为$\begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1/3 & 2/3 \end{pmatrix}$. 对任

意$n \geq 1$, 求$f_{1,2}^{(n)}$.

2 一个盗窃犯长期在$A=1, B=2, C=3$三地流传作案, 治安部门调查后发现他每年作案次数服从强度为3的泊松分布, 而连续两次作案的地点变化服从转移概率矩阵$\begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, 并且作案次数与作案地点无关. 已知此人刚刚在$A$ 地作案, 试估计一年内他在A地再作案的概率.

3 设X的状态空间为$\{1,2,3\}$, 转移概率矩阵为$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix}$. 求$f_{3,2}$ 和$m_{3,2}$.

4 若一篇文稿有n个错误, 每次校阅至少能发现一个, 但留下来的错误数在0到$n-1$之间等可能存在. 设原稿共有a个错误, 问为了改正全部错误平均需要校阅几次?

5 若马氏链X只有三个状态且$(m_{i,j})_{i,j=1,2,3} = \begin{pmatrix} 2 & 3 & 3 \\ 1 & 4 & 4 \\ 1 & 4 & 4 \end{pmatrix}$, 求转移概率矩阵$\mathbf{P}$.

6 (Ehrenfest模型) 设一个坛子内装有红白两色共N个球, 每次随机地从坛子中抽出一个球, 把它换成另一种颜色后放回. 以X_n表示经n次抽放后坛中的红球数, 那么$X=\{X_n; n \geq 0\}$ 为时齐马氏链. 若开始时坛内只有一个红球, 问平均要抽放多少次才能使坛内全是白球?

7 设X是(q,p)-随机游动. 对任意$k \in \mathbb{Z}$, 求$f_{0,k}$和$m_{0,k}$.

8 设X是非负整数值的反射(q,p)-随机游动. (1) 对任意$k \geq 0$, 求$f_{k,0}$, $m_{k,0}$, (2) 设$q \leq p$, 对任意$k \geq 0$; 求$f_{0,k}$, $m_{0,k}$.

9* 设X是非负整数值的带黏性壁的反射(q,p)-随机游动(就是在例3.3.12中令$1 > r_0 > 0, 0 < p_0 = 1 - r_0 < 1$, 其他设定不变), 对任意$k \geq 0$, 求$f_{0,k}$, $m_{0,k}$.

10* 设X是状态空间为$\{0,1,2,\cdots,N\}$的带两个反射壁的(q,p)-随机游动, 对任意$0 \leq i \leq j \leq N$, 求$f_{i,j}$, $m_{i,j}$.

11* 对任意给定的有限阶方阵$\mathbf{A}$以及向量$\mathbf{b}$, 证明: 若线性方程组$\mathbf{AX}=\mathbf{b}$存在恒正的最小非负解(即解的每个分量都是正的), 那么该最小非负解就是此方程组的唯一解.

3.4 正常返与平稳分布

转移概率极限与平稳分布是我们研究与描述马氏链长期行为的常用方法与工具. 借助于这些方法与工具, 这一节我们将揭示状态动态变化规律与随机现象长期规律之间的联系.

3.4.1 正常返与零常返

上一节我们介绍了计算首回时(回访周期)的分布及探讨了$m_{i,j}$的计算方法. 我们知道当i是常返状态时, $m_{i,i}$就是从i出发并再次回到i的平均时间. 利用这个回访周期的平均值, 我们可以把X的常返状态进一步区分为:

♣**定义3.4.1.** *设$i \in S$为X的常返状态. 称i是正常返的, 若$m_{i,i} < \infty$; 称i是零常返的, 若$m_{i,i} = +\infty$.*

由常返、非常返的定义以及停时$\tau_i = \inf\{n \geq 1; X_n = i\}$中$\inf \emptyset = +\infty$的约定, 我们可用随机变量与数学期望等基本概念将非常返、零常返、正常返这三者区别如下:

(A) i常返$\Leftrightarrow X_0 = i$条件下τ_i是正常随机变量, 即$P_i(\tau_i < +\infty) = 1$. 在此基础上

(A.1) i正常返$\Leftrightarrow$ τ_i的数学期望有限.

(A.2) i零常返$\Leftrightarrow$ τ_i的数学期望无穷.

(B) i非常返$\Leftrightarrow$ $X_0 = i$条件下τ_i是广义随机变量, 即$P_i(\tau_i = +\infty) > 0$.

直观地说, 状态i是正常返的意味着从i出发后, X能"较大概率"地在短时间内回访自己; 而零常返则意味着从i出发后, 虽然X一定会回访自己, 但那些需要非常长的时间后才能回到自己的事件发生概率较大, 以至于从平均角度看, 回访自己的时间间隔长得" 遥不可及". 显然正常返描述的仍然是状态随机变化现象. 有趣的是它还是区分马氏链转移概率平均极限是否为零的关键.

★**定理3.4.2.** *对一切$i \in S$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{i,i}^{(k)} = \begin{cases} 1/m_{i,i}, & i\text{正常返}, \\ 0, & \text{其他}. \end{cases}$$

证明定理之前我们先介绍一个关于数列的比极限的结论.

▼**引理3.4.3.** *设$\{a_n;\ n \geq 0\}$为一个不全为0的非负数列且满足条件*

$$\lim_{n\to\infty} \frac{a_n}{\sum_{k=0}^{n} a_k} = 0.$$

若$\{b_n;\ n\ge 0\}$为收敛数列, 那么

$$\lim_{n\to\infty}\frac{\sum_{k=0}^n a_k b_{n-k}}{\sum_{k=0}^n a_k}=\lim_{n\to\infty} b_n.$$

证明 由所设条件, 对任意给定$N\ge 1$, 有

$$\lim_{n\to\infty}\frac{\sum_{k=n-N+1}^n a_k}{\sum_{k=0}^n a_k}=0.$$

若b_n极限有限, 记作b, 那么对任意$\varepsilon>0$, 存在N, 使得当$n\ge N$时,

$$|b_n-b|<\varepsilon.$$

于是存在$B<\infty$使得$|b_n-b|<B$对一切n 成立. 进而

$$\Big|\sum_{k=0}^n a_k(b_{n-k}-b)\Big|\le\varepsilon\sum_{k=0}^{n-N}a_k+B\sum_{k=n-N+1}^n a_k.$$

两边同除以$\sum_{k=0}^n a_k$, 不难看出此时结论成立.

若b_n的极限为无穷, 不妨设是正无穷, 那么对任意$M>0$, 存在N, 使得当$n\ge N$时, $b_n\ge M$. 于是

$$\sum_{k=0}^n a_k b_{n-k}\ge M\sum_{k=0}^{n-N}a_k+(\min_{0\le n\le N}b_n)\sum_{k=n-N+1}^n a_k.$$

同样两边同除以$\sum_{k=0}^n a_k$, 不难看出此时结论也成立. □

定理3.4.2的证明 按n之前最后离开状态i的时刻可将随机事件$\{X_n=j\}$分解为不相容事件的并集,

$$\{X_n=j\}=\bigcup_{k=0}^{n-1}\{X_n=j,X_k=i,X_v\ne i,v=k+1,\cdots,n-1\}.$$

从而由马氏性及(3.2.4)可得

$$\begin{aligned}p_{i,j}^{(n)}=P_i(X_n=j)&=P_i(\bigcup_{k=0}^{n-1}\{X_n=j,X_k=i,X_v\ne i,v=k+1,\cdots,n-1\})\\&=\sum_{k=0}^{n-1}p_{i,i}^{(k)}P_i(X_n=j,X_v\ne i,v=k+1,\cdots,n-1|X_k=i)\\&=\sum_{k=0}^{n-1}p_{i,i}^{(k)}e_{i,j}^{(n-k)},\end{aligned}\tag{3.4.1}$$

这里$e_{i,j}^{(n-k)}$为(3.2.4)定义的禁忌概率. 进而

$$\begin{aligned}n&=\sum_{j\in S}\sum_{m=1}^n p_{i,j}^{(m)}=\sum_{j\in S}\sum_{m=1}^n\sum_{k=0}^{m-1}p_{i,i}^{(k)}e_{i,j}^{(m-k)}\\&=\sum_{j\in S}\sum_{k=0}^{n-1}p_{i,i}^{(k)}\sum_{m=1}^{n-k}e_{i,j}^{(m)}=\sum_{k=0}^{n-1}p_{i,i}^{(k)}\sum_{m=1}^{n-k}\sum_{j\in S}e_{i,j}^{(m)}.\end{aligned}$$

注意到

$$\sum_{j\in S} e_{i,j}^{(m)} = P_i(\bigcup_{j\in S}\{X_m=j, X_v\neq i, v=1,\cdots,m\})$$
$$= P_i(X_v\neq i, v=1,\cdots,m-1) = P_i(\tau_i\geq m).$$

因此

$$n = \sum_{k=0}^{n-1} p_{i,i}^{(k)} \sum_{m=1}^{n-k} P_i(\tau_i\geq m).$$

记$b_n=\sum\limits_{m=1}^{n} P_i(\tau_i\geq m)$, 那么

$$n = \sum_{k=0}^{n-1} p_{i,i}^{(k)} b_{n-k}.$$

若i正常返, $\lim\limits_{n\to\infty} b_n = m_{i,i} < \infty$; 若$i$零常返或非常返, $\lim\limits_{n\to\infty} b_n = \infty$. 由引理3.4.3可知定理成立. □

由定理3.4.2, 容易得到如下结果, 证明简单, 此略.

♦**推论3.4.4.** i正常返当且仅当$\limsup\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n} p_{i,i}^{(k)} > 0$.

♦**推论3.4.5.** 设状态i是正常返的, 如果$i\to j$, 那么j也是正常返的且$j\in C(i)$.

证明 由i常返可知i本质进而$j\in C(i)$, 因此存在$m,n>0$使得$p_{j,i}^{(m)}>0$, $p_{i,j}^{(n)}>0$. 由C-K方程,

$$\lim_{k\to\infty}\frac{1}{k}\sum_{v=1}^{k} p_{j,j}^{(v)} = \lim_{k\to\infty}\frac{1}{k}\sum_{v=m+n+1}^{k} p_{j,j}^{(v)} \geq p_{j,i}^{(m)}p_{i,j}^{(n)}\lim_{k\to\infty}\frac{1}{k}\sum_{v=1}^{k-n-m} p_{i,i}^{(v)} > 0,$$

因此j正常返. □

推论3.4.5表明若i正常返那么$C(i)$中所有状态都正常返, 即正常返性也是互通类中元素共有的性质, 此时称$C(i)$为正常返类. 结合此前关于常返与非常返的讨论, 我们可将互通类按其中状态的常返性分成正常返类, 零常返类和非常返类这三种情形. 类似, 我们也可以定义正常返、零常返和非常返马氏链, 若它的每个状态都是正常返、零常返和非常返的.

►**例3.4.6.** 直线上(q,p)-随机游动W当$p=q=1/2$时是零常返的.

解 已知W是不可约的, 而且$p=q=1/2$时常返. 此时对$n=2k$, $2k+1$, 当$k\geq 1$充分大时由Stirling公式可知

$$\sum_{m=0}^{n} p_{0,0}^{(m)} = 1+\sum_{m=1}^{k}\frac{C_{2m}^{m}}{4^m} = 1+\sum_{m=1}^{k}\frac{(2m)!}{m!m!4^m}$$
$$\asymp \sum_{m=1}^{k}\frac{\sqrt{4\pi m}(2m)^{2m}\mathrm{e}^{2m}}{4^m\mathrm{e}^{2m}(\sqrt{2\pi m}m^m)^2} = \sum_{m=1}^{k}\frac{1}{\sqrt{\pi m}}$$

$$\asymp \int_0^k \frac{\mathrm{d}x}{\sqrt{\pi x}} = 2\sqrt{\frac{k}{\pi}} = \sqrt{\frac{4k}{\pi}} \asymp \sqrt{\frac{2n}{\pi}}.$$

因此

$$\frac{1}{n}\sum_{m=0}^{n} p_{0,0}^{(m)} \asymp \sqrt{\frac{2}{\pi n}} \to 0.$$

这表明状态0为零常返的, 进而W零常返. □

★**定理3.4.7.** 对任意$i,j \in S$以及$u \in [0,1]$,

$$\lim_{n\to\infty} \frac{\sum_{k=1}^n p_{i,j}^{(k)} u^k}{\sum_{k=0}^n p_{j,j}^{(k)} u^k} = F_{i,j}(u),$$

其中$F_{i,j}(u)$是由(*3.3.1*)定义的幂级数.

证明 显然$u=0$时结果成立. 下设$u \in (0,1]$. 由(3.2.5)可得对任意$n \geq 1$,

$$\begin{aligned}\sum_{k=1}^n p_{i,j}^{(k)} u^k &= \sum_{k=1}^n \sum_{r=1}^k f_{i,j}^{(r)} p_{j,j}^{(k-r)} u^k \\ &= \sum_{r=1}^n f_{i,j}^{(r)} u^r \sum_{k=0}^{n-r} p_{j,j}^{(k)} u^k \\ &= \sum_{k=0}^{n-1} p_{j,j}^{(k)} u^k \sum_{r=1}^{n-k} f_{i,j}^{(r)} u^r.\end{aligned}$$

令

$$a_n = p_{j,j}^{(n)} u^n, \quad b_n = \sum_{r=1}^n f_{i,j}^{(r)} u^r.$$

那么$b_n \to F_{i,j}(u)$, 而且不论j是否常返引理3.4.3的条件都成立. 由此可知定理3.4.7成立. □

♦**推论3.4.8.** 对任意$i,j \in S$, 当$n \to \infty$时$\frac{1}{n}\sum_{k=1}^n p_{i,j}^{(k)}$的极限存在, 而且

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n p_{i,j}^{(k)} = \begin{cases} f_{i,j}/m_{j,j}, & j\text{正常返}, \\ 0, & j\text{为其他状态}. \end{cases}$$

证明 由定理3.4.2与定理3.4.7(令$u=1$)立刻可得. □

♠**注记3.4.9.** 对任意$i \in S$, 若j是非周期正常返的, 那么$\lim\limits_{n\to\infty} p_{i,j}^{(n)} = f_{i,j}/m_{j,j}$; 若$j$是零常返的, 那么$\lim\limits_{n\to\infty} p_{i,j}^{(n)} = 0$; 若$j$是非常返的, $\lim\limits_{n\to\infty} p_{i,j}^{(n)} = 0$. 前两结论我们将在第五章第二节给出解释, 最后一个结论是定理*3.2.20*和定理*3.4.7*的推论.

对任意$i,j \in S$, 令$P_{i,j}(s) = \sum\limits_{n=1}^{\infty} p_{i,j}^{(n)} s^n$. 由定理3.4.7可得

♦**推论3.4.10.** 对任意$i,j\in S$, $u\in[0,1)$, $P_{i,j}(u)=F_{i,j}(u)/(1-F_{j,j}(u))$. 此外当$u=1$时

$$\sum_{n=1}^{\infty}p_{i,j}^{(n)}=P_{i,j}(1-)=\begin{cases}f_{i,j}/(1-f_{j,j}), & f_{j,j}<1;\\ 0, & f_{i,j}=0\text{且}f_{j,j}=1;\\ \infty, & f_{i,j}>0\text{且}f_{j,j}=1.\end{cases}$$

3.4.2 平稳分布

♣**定义3.4.11.** 称不恒为零的非负有限数列$\mathbf{u}=(u_i,\ i\in S)$ 为马氏链的不变测度, 若对任意$i\in S$

$$u_i=\sum_{k\in S}u_kp_{k,i}. \tag{3.4.2}$$

当$\sum_{i\in S}u_i=1$, 即$(u_i,\ i\in S)$是概率分布时, 称$\mathbf{u}$是X的平稳分布.

若分布$\mathbf{u}=(u_i,i\in S)$为X的初始分布, 即对任意$j\in S$, $P(X_0=j)=u_j$, 那么对任意一个时刻$n\geq 0$, 由定理3.1.16可知, X_n的分布为

$$P(X_n=i)=\sum_{j\in S}u_jp_{j,i}^{(n)},\qquad i\in S.$$

由此我们看出, 若$\mathbf{u}$是平稳分布, 则由C-K方程可得

$$\begin{aligned}P(X_n=i)&=\sum_{j\in S}u_jp_{j,i}^{(n)}=\sum_{j\in S}u_j\sum_{k\in S}p_{j,k}p_{k,i}^{(n-1)}\\&=\sum_{k\in S}\Big(\sum_{j\in S}u_jp_{j,k}\Big)p_{k,i}^{(n-1)}=\sum_{k\in S}u_kp_{k,i}^{(n-1)}\\&=\cdots=\sum_{k\in S}u_kp_{k,i}=u_i.\end{aligned}$$

这表明以平稳分布为初始分布的马氏链在任意一个时刻的分布都是一样的.

▶**例3.4.12.** 设X的状态空间$S=\mathbb{N}$, 对任意$k\geq 0$, 转移概率$p_{0,k}=p_k$, 其中$\sum_{k=0}^{\infty}p_k=1$, 而当$n\geq 1$时$p_{n,n-1}=1$. 已知$\sum_{k=1}^{\infty}kp_k<\infty$, 求$X$的平稳分布.

解 记X的平稳分布为$(\pi_j,j\geq 0)$, 那么根据定义, π_j要满足以下方程.

$$\begin{cases}\sum_{j=0}^{\infty}\pi_j=1,\\ \pi_j=\sum_{k=0}^{\infty}\pi_kp_{k,j}=p_j\pi_0+\pi_{j+1},\ \ j\geq 0.\end{cases}$$

由后一个方程可得, 对任意$j\geq 0$

$$\begin{aligned}\pi_{j+1}&=\pi_j-p_j\pi_0=\pi_{j-1}-p_{j-1}\pi_0-p_j\pi_0=\cdots\\&=(1-p_0-\cdots-p_j)\pi_0=\pi_0\sum_{k=j+1}^{\infty}p_k.\end{aligned}$$

将该关系式代入前面方程组中第一个方程可得

$$1=\sum_{j=0}^{\infty}\pi_j=\pi_0+\sum_{j=1}^{\infty}\sum_{k=j}^{\infty}p_k\pi_0=\pi_0\left(1+\sum_{k=1}^{\infty}kp_k\right).$$

由条件$\sum_{k=1}^{\infty}kp_k<\infty$可知

$$\pi_0=\frac{1}{1+\sum_{k=1}^{\infty}kp_k},$$

进而对任意$j\geq 1$

$$\pi_j=\pi_0\sum_{k=j}^{\infty}p_k=\frac{\sum_{k=j}^{\infty}p_k}{1+\sum_{k=1}^{\infty}kp_k}. \qquad \square$$

对只有有限个状态的马氏链, 我们可以借助线性方程组, 方便地找到它可能的平稳分布. 设X是有n个状态的马氏链, 转移概率矩阵$\mathbf{P}=(p_{i,j})$, 记其可能的平稳分布为

$$\pi=(\pi_1,\pi_2,\cdots,\pi_n)^T,$$

那么π是下列方程组的解

$$\begin{cases}p_{1,1}\pi_1+p_{2,1}\pi_2+\cdots+p_{n,1}\pi_n=\pi_1;\\ \cdots\\ p_{1,n-1}\pi_1+p_{2,n-1}\pi_2+\cdots+p_{n,n-1}\pi_n=\pi_{n-1};\\ p_{1,n}\pi_1+p_{2,n}\pi_2+\cdots+p_{n,n}\pi_n=\pi_n;\\ \pi_1+\cdots+\pi_n=1.\end{cases}$$

由于转移概率矩阵满足$\sum_{j=1}^{n}p_{i,j}=1$, 因此前$n-1$个方程加到第n个方程会得到恒等式

$$\pi_1+\cdots+\pi_n=\pi_1+\cdots+\pi_n.$$

这意味着该方程组与把第n个方程去掉的方程组

$$\begin{cases}p_{1,1}\pi_1+p_{2,1}\pi_2+\cdots+p_{n,1}\pi_n=\pi_1;\\ \cdots\\ p_{1,n-1}\pi_1+p_{2,n-1}\pi_2+\cdots+p_{n,n-1}\pi_n=\pi_{n-1};\\ \pi_1+\cdots+\pi_n=1,\end{cases}$$

同解, 也即我们只须求解如下的矩阵方程

$$\mathbf{Q}^T\pi=\pi',$$

其中矩阵$\mathbf{Q}$是将转移概率矩阵$\mathbf{P}$最后一列替换为$(1,1,\cdots,1)^T$后所得到的矩阵, 而π' 是将π最后一个分量替换为1后所得向量.

▶**例3.4.13.** 设X的转移概率矩阵为$\begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 1/4 & 3/4 \\ 1 & 0 & 0 \end{pmatrix}$, 求$X$的平稳分布.

解 令$\mathbf{Q} = \begin{pmatrix} 0 & 1/2 & 1 \\ 0 & 1/4 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, 解方程组$\mathbf{Q}^T\pi = \pi'$, 即

$$\begin{cases} \pi_3 = \pi_1; \\ \pi_1/2 + 3\pi_2/4 = \pi_2; \\ \pi_1 + \pi_2 + \pi_3 = 1, \end{cases}$$

可得唯一非负解$(\pi_1, \pi_2, \pi_3) = (3/8, 1/4, 3/8)$. 此即为$X$的平稳分布. □

不是所有马氏链都有平稳分布. 不可约马氏链的平稳分布与回访周期相关.

★**定理3.4.14.** *若X为不可约马氏链, 那么X是正常返的当且仅当X存在平稳分布. 不可约正常返马氏链X的平稳分布$(\pi_i, i \in S)$是唯一的, 而且*

$$\pi_i = \frac{1}{m_{i,i}} > 0, \quad i \in S.$$

证明 先证前一结论的必要性. 为此我们只需验证$(1/m_{j,j}, j \in S)$是平稳分布. 由C-K方程可知

$$\frac{1}{n}\sum_{k=1}^{n} p_{j,j}^{(k)} = \frac{1}{n}\sum_{k=1}^{n}\sum_{i\in S} p_{j,i}^{(k-1)} p_{i,j} = \sum_{i\in S}\Big[\frac{1}{n}\sum_{k=1}^{n} p_{j,i}^{(k-1)}\Big] p_{i,j}. \tag{3.4.3}$$

因为X是不可约正常返链, 对任意$i, j \in S$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} p_{j,j}^{(k)} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} p_{i,j}^{(k)} = \frac{1}{m_{j,j}} > 0.$$

在(3.4.3)式两边令$n \to \infty$, 得

$$\frac{1}{m_{j,j}} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} p_{j,j}^{(k)} \ge \lim_{N\to\infty}\sum_{i\in A_N}\frac{p_{i,j}}{m_{i,i}} = \sum_{i\in S}\frac{p_{i,j}}{m_{i,i}}, \tag{3.4.4}$$

其中A_N表示S中元素个数为N的有限子集, 而且

$$A_N \subset A_{N+1} \subset \cdots, \qquad \bigcup_N A_N = S.$$

因此

$$\sum_{j\in S}\frac{1}{m_{j,j}} \ge \sum_{j\in S}\sum_{i\in S}\frac{p_{i,j}}{m_{i,i}} = \sum_{i\in S}\sum_{j\in S}\frac{p_{i,j}}{m_{i,i}} = \sum_{i\in S}\frac{1}{m_{i,i}}. \tag{3.4.5}$$

另一方面, 注意到

$$1 = \sum_{i\in S}\Big[\frac{1}{n}\sum_{k=1}^{n} p_{j,i}^{(k)}\Big].$$

用证明(3.4.4)的相同方法可得
$$1 \geq \sum_{i \in S} \frac{1}{m_{i,i}}.$$
这表明(3.4.5)中不等号应该为等号, 从而对任意$j \in S$,
$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{p_{i,j}}{m_{i,i}}.$$
由此进一步可得对任意$k \geq 1$,
$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{p_{i,j}^{(k)}}{m_{i,i}}.$$
因此
$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{1}{m_{i,i}} \Big[\frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)}\Big].$$
注意到$1 \geq \sum_{i \in S} \frac{1}{m_{i,i}}$, 由一致收敛级数求和与极限可交换性质可得, 当$n \to \infty$时,
$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{1}{m_{i,i} m_{j,j}}. \tag{3.4.6}$$
这表明
$$\sum_{i \in S} \frac{1}{m_{i,i}} = 1.$$
因此$(1/m_{j,j}, j \in S)$为X 的平稳分布.

再证前一结论的充分性. 设$(\pi_j, j \in S)$是X的任一平稳分布. 由
$$p_j = \sum_{i \in S} \pi_i p_{i,j} = \cdots = \sum_{i \in S} \pi_i p_{i,j}^{(k)}, \qquad j \in S, \ \ k \geq 1,$$
可知对任意$j \in S$,
$$\pi_j = \sum_{i \in S} \pi_i \Big[\frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)}\Big].$$
注意到上式右边关于n是一致收敛的, 由推论3.4.8可得
$$\pi_j = \sum_{i \in S} \pi_i \Big[\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)}\Big] = \begin{cases} [\sum\limits_{i \in S} \pi_i f_{i,j}]/m_{j,j}, & j\text{正常返}, \\ 0, & \text{其他}. \end{cases} \tag{3.4.7}$$
另一方面, 由$\sum\limits_{j \in S} \pi_j = 1$可知存在$j_0 \in S$使得$\pi_{j_0} > 0$. 从而$j_0$正常返, 进而由$X$不可约可知$X$是正常返的.

由于对任意$i, j \in S$, $f_{i,j} = 1$, 由(3.4.7)知$0 < \pi_j = 1/m_{j,j}$. 这表明平稳分布必唯一. 因此定理后一个结论也成立. □

对不可约的马氏链而言, 由定理3.4.14可知, 只有正常返链才有平稳分布, 这时我们可以选择这个平稳分布为初始分布使得对应的马氏链为严平稳过程. 更有意思的是:

★**定理3.4.15.** 若X是不可约正常返马氏链, 记其平稳分布为$\pi=(\pi_i, i\in S)$. 那么对任意$i,j\in S$,

$$\pi_j=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}p_{i,j}^{(k)}=\frac{1}{m_{j,j}}. \tag{3.4.8}$$

进一步, 对任意初始分布,

$$\lim_{n\to\infty}E\Big(\frac{1}{n}\sum_{k=1}^{n}1_{\{X_k=j\}}\Big)=\pi_j.$$

证明 由于X是不可约常返的, 由性质3.2.19可知, 对任意$i,j\in S$, $f_{i,j}=1$. 由推论3.4.8和定理3.4.14可知前半部分结论成立.

进一步, 设X的初始分布为$\mu=(\mu_i, i\in S)$. 那么

$$\begin{aligned}E(\frac{1}{n}\sum_{k=1}^{n}1_{\{X_k=j\}})&=\frac{1}{n}\sum_{k=1}^{n}E(1_{\{X_k=j\}})=\frac{1}{n}\sum_{k=1}^{n}\sum_{i\in S}\mu_i p_{i,j}^{(k)}\\&=\sum_{i\in S}\mu_i\Big[\frac{1}{n}\sum_{k=1}^{n}p_{i,j}^{(k)}\Big].\end{aligned}$$

由于$\sum_{i\in S}\mu_i=1$, 级数$\sum_{i\in S}\mu_i p_{i,j}^{(n)}$ 对n一致收敛. 因此

$$\lim_{n\to\infty}E(\frac{1}{n}\sum_{k=1}^{n}1_{\{X_k=j\}})=\sum_{i\in S}\mu_i\lim_{n\to\infty}\Big[\frac{1}{n}\sum_{k=1}^{n}p_{i,j}^{(k)}\Big]=\sum_{i\in S}\mu_i\pi_j=\pi_j. \qquad \square$$

注意到$N_n(j)=\sum_{k=1}^{n}1_{\{X_k=j\}}$表示过程$X$在时刻$n$之前(含$n$)到达状态$j$的次数, $N_n(j)/n$ 表示X在时刻n之前(含n)到达状态j比率(频率), 而$E(N_n(j)/n)$则表示到达状态j的平均比率. 因此定理3.4.15表明平稳分布还是马氏链在长时间内访问各状态的比率极限.

▶**例3.4.16.** 在研究人们消费习惯时常会调查人们对具有相同功能不同品牌的产品的消费情况. 假设有三个品牌的手机$1,2,3$, 令X_n表示消费者在第n次购买时选择的品牌. 顾客购买这三个品牌后下次选择相同品牌的概率分别为$0.8, 0.6, 0.8$, 若他们改变品牌则会在另外两个品牌中随机挑选一个. 问顾客长期消费后在各个品牌上的平均消费比例?

解 容易看出$X=\{X_n\}$是一个转移概率矩阵为

$$\mathbf{P}=\begin{pmatrix}0.8 & 0.1 & 0.1\\0.2 & 0.6 & 0.2\\0.1 & 0.1 & 0.8\end{pmatrix}$$

的时齐马氏链. X是不可约非周期的. 令

$$\mathbf{Q}=\begin{pmatrix}0.8 & 0.1 & 1\\0.2 & 0.6 & 1\\0.1 & 0.1 & 1\end{pmatrix}.$$

解方程$\mathbf{Q}^T\pi=\pi'$ 可得唯一非负解

$$(\pi_1,\pi_2,\pi_3)=(2/5,1/5,2/5).$$

此即为顾客在长期消费当中各个品牌的消费比例. □

与定理3.4.15的后半部分证明一样, 由注3.4.9可得

★定理3.4.17. 若X是非周期不可约正常返的, 那么对任意初始分布$\mu=(\mu_i,i\in S)$, 随着$n\to\infty$, X_n的分布收敛到π, 即对任意$j\in S$, $P(X_n=j)\to\pi_j$.

▶例3.4.18. (例3.1.13续) 问如果社会阶层的转换机制不改变, 长时间后社会阶层的比例是多少.

解 由转移概率矩阵可知, 表示社会阶层变化的马氏链X是不可约非周期的. 令

$$\mathbf{Q}=\begin{pmatrix}0.8 & 0.1 & 1\\ 0.3 & 0.4 & 1\\ 0.05 & 0.1 & 1\end{pmatrix}.$$

解方程$\mathbf{Q}^T\pi=\pi'$得唯一非负解

$$(\pi_1,\pi_2,\pi_3)=(12/35,5/35,18/35).$$

此即为长时间后社会阶层的比例. □

由定理3.4.14可知, 不可约链X的平稳分布中每一分量都为正数. 由此我们得到一个判别马氏链是否正常返的方法.

♦推论3.4.19. 不可约马氏链X是正常返的当且仅当方程组

$$u_i=\sum_{k\in S}u_k p_{k,i}$$

存在一个全不为零的非负解$(u_i,\ i\in S)$使得$\sum\limits_{i\in S}u_i<\infty$.

证明 此时$(u_i/\sum_{i\in S}u_i, i\in S)$为平稳分布, 从而推论成立. □

▶例3.4.20. 讨论例3.3.7中马氏链的正常返性.

解 由例3.3.7 可知X常返$\Leftrightarrow\sum\limits_{k=1}^{\infty}\frac{q_1\cdots q_k}{p_1\cdots p_k}=\infty$. 记$\rho_k=\frac{q_1\cdots q_k}{p_1\cdots p_k}$. 解方程组

$$z_i=\sum_{k=0}^{\infty}z_k p_{k,i},\quad i\geq 0.$$

由$r_0+p_0=1$, $r_n+p_n+q_n=1$, $n\geq 1$, 可得

$$\begin{cases}q_1z_1-p_0z_0=0\\ q_{n+1}z_{n+1}-p_nz_n=q_nz_n-p_{n-1}z_{n-1}, n\geq 1.\end{cases}$$

因此对任意$n \geq 0$, $q_{n+1}z_{n+1} = p_n z_n$, 继而

$$z_{n+1} = \frac{p_n \cdots p_0}{q_{n+1} \cdots q_1} z_0 = \frac{p_0}{p_{n+1}\rho_{n+1}} z_0.$$

由推论3.4.19, X正常返当且仅当

$$\sum_{n=0}^{\infty} z_i < \infty \Leftrightarrow z_0 \sum_{n=1}^{\infty} \frac{p_0}{p_n \rho_n} < \infty \Leftrightarrow \sum_{n=1}^{\infty} \frac{1}{p_n \rho_n} < \infty.$$

此即为X正常返的条件. □

练习题3.4

1 对有限状态马氏链X, 证明(1)X没有状态零常返, (2)若X不可约, 那么X是正常返的.

2 设马氏链X存在状态j使得方程组

$$z_i = \sum_{k \neq j} p_{i,k} z_k + f_{i,j}, \qquad i \neq j$$

存在一组满足条件$\sum\limits_{k \neq j} p_{j,k} u_k < \infty$的非负有限解$\{u_i; i \neq j\}$. 证明若$j$是常返的, 那么$j$ 是正常返的.

3 设X是一个不可约非常返的马氏链, 令

$$l_j = \sup\{n \geq 0, X_n = j, X_k \neq j, k > n\},$$

即l_j是X最后一次到达状态j的时间. 假定X 从j 出发, 求l_j的分布和均值.

4 设X的转移概率矩阵为$\begin{pmatrix} 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{pmatrix}$, 求$X$的平稳分布.

5 一个月后小明有两天假期, 他就打算只要假期里有一天的天气晴朗就出去旅行. 为此他查阅了气象资料, 发现每日天气晴朗(0)与非晴朗(1)之间的变化可以近似用时齐马氏链来刻画, 转移概率矩阵为$\begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}$. 试估计小明能出去旅行的概率.

6 某人有M把伞并在办公室和家之间往返: 出门上班时下雨并家里有伞就带把伞去办公室; 下班回家时下雨且办公室有伞就带把伞回家. 其它时候都不带伞. 假设每天上下班时是否下雨是独立同分布的, 下雨的概率为p, 不下雨的概率为$1-p$, 问此人出门上班时碰到下雨且没有雨伞的概率.

7 设X的状态空间$S = \mathbb{N}$, 对任意$n \geq 0$, 转移概率$p_{n,n+1} = p_n$, $p_{n,0} = 1 - p_n$, 其中$p_n \in (0,\ 1)$. 求(1) 能保证0是常返状态的条件; (2) 能保证0是正常返状态的条件; (3)平稳分布.

8 设X的转移概率满足: 对任意$i, j \in S = \mathbb{N}$, $p_{i,i} = \lambda_i + (1-\lambda_i)u_i$; 而$i \neq j$时,

$p_{i,j} = (1-\lambda_i)u_j$, 其中$u_j > 0$, $\sum\limits_{j=0}^{\infty} u_j = 1$, $0 < \lambda_i < 1$. 证明

(1) X是常返的; (2) X正常返当且仅当$\sum\limits_{i=0}^{\infty} u_i/(1-\lambda_i) < \infty$.

9 设X为不可约正常返马氏链, 证明对任意$i, j \in S$, $\pi_j = \pi_i e_{i,j}$, 其中$(\pi_i, i \in S)$ 为X的平稳分布, $e_{i,j} = \sum\limits_{n=1}^{\infty} e_{i,j}^{(n)}$.

10* 设不可约正常返马氏链X的转移概率矩阵为$(p_{i,j})_{i,j\in S}$, $\pi = (\pi_i, i \in S)$为平稳分布. 令$Y_n = (X_n, X_{n+1})$. 证明$Y = \{Y_n; n \geq 0\}$ 是状态空间为$\mathcal{M} = \{(i,j); i,j \in S, p_{i,j} > 0\}$的不可约正常返马氏链, 平稳分布为$(\pi_i p_{i,j}, (i,j) \in \mathcal{M})$.

11* 证明不可约链X正常返的充要条件是存在非负数列$\{z_i;\ i \in S\}$及一个状态(或任一状态)j使得对任意$i \neq j$, $\sum\limits_{k\in S} p_{i,k} z_k \leq z_i - 1$, 而且$\sum\limits_{k\in S} p_{j,k} z_k < \infty$.

12* 证明不可约正常返马氏链的不变测度μ必是有限测度(即$\sum\limits_{i\in S} \mu(i) < \infty$).

13* 证明任意不可约常返马氏链都有不变测度, 而且不变测度除一个常数因子外是唯一的.

3.5 遍历性定理

遍历性定理是马氏链理论中最重要的结果之一, 也是用马氏链模型解决应用问题时的理论基础. 这一节我们将介绍最基本的遍历性结论, 并结合例子说明遍历性定理的应用.

3.5.1 遍历性定理

在上一节定理3.4.15中我们证明了, 对不可约正常返马氏链X,

$$\lim_{n\to\infty} E\Big(\frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k=j\}}\Big) = \pi_j,$$

并通过一些例子展示了这个结论的应用. 注意到上式左边是数学期望, 从概率统计的角度看, 这刻画的是总体的规律性, 是具有相同变化规律的动态随机现象的平均行为. 更直观地说是对随机过程大量观察之后的平均. 然而对许多现实中的动态随机现象而言, 是不可能做大量观察的, 比如, 研究某地气候的变化时, 我们在每个时刻就只能有一个观察结果. 面对这种不可重复或不可大量重复观察的动态随机现象, 一个自然的问题是能否使上述总体规律对个体(样本)仍成立? 准确地说能否将上述结果加强为

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k=j\}} = \pi_j, \qquad a.s.$$

如果能做到这点, 那就意味着, 即使人们只对动态变化的随机现象做了一次长时间的观察, 我们也能以概率1保证从这个观察结果的时间平均中发现这个动态变化随机现象的平稳分布.

下面就不可约正常返马氏链X我们给出一个更一般的结果.

★定理3.5.1. 记不可约马氏链$\{X_n;\ n\geq 0\}$的平稳分布为$\pi=(\pi_i,\ i\in S)$. 若S上函数g满足$\sum_{j\in S}|g(j)|\pi_j<\infty$, 那么

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g(X_k) = \sum_{j\in S} g(j)\pi_j = E_\pi(g), \qquad a.s.$$

证明 由于$g=g^+-g^-=g\vee 0-(-g)\vee 0$. 不妨设$g$非负. 定义$\tau_i(0)=0$,

$$\tau_i(k)=\inf\{n>\tau_i(k-1), X_n=i\}, \quad W_i(k)=\tau_i(k)-\tau_i(k-1), \quad k\geq 1.$$

由定理3.2.26易知$\{W_i(k); k\geq 2\}$为i.i.d序列, 且

$$E(W_i(2))=m_{i,i}<\infty.$$

由强大数定律(定理1.4.12)可得

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}W_i(k)=\lim_{n\to\infty}\frac{1}{n}\tau_i(n)=m_{i,i},\qquad a.s.$$

再令

$$Y_k=\sum_{n=\tau_i(k)+1}^{\tau_i(k+1)}g(X_m).$$

仍由定理3.2.26可知Y_k, $k\geq 1$, 为独立同分布随机变量. 注意到

$$\begin{aligned}E(Y_1)&=E\Big(\sum_{m=\tau_i(k)+1}^{\tau_i(k+1)}g(X_m)\Big)=E_i\Big(\sum_{m=1}^{\tau_i(1)}g(X_m)\Big)=E_i\Big(\sum_{m=1}^{\tau_i}g(X_m)\Big)\\&=g(i)+\sum_{n=1}^{\infty}P_i(\tau_i=n)E_i\Big(\sum_{m=1}^{n-1}g(X_m)|\tau_i=n\Big)\\&=g(i)+\sum_{n=1}^{\infty}P_i(\tau_i=n)\sum_{m=1}^{n-1}\sum_{j\neq i}g(j)P_i(X_m=j|\tau_i=n)\\&=g(i)+\sum_{n=1}^{\infty}\sum_{m=1}^{n-1}\sum_{j\neq i}g(j)P_i(X_m=j,\tau_i=n).\end{aligned}$$

由于是非负项求和, 我们可以交换n,m的求和次序, 得

$$\begin{aligned}E(Y_1)&=g(i)+\sum_{m=1}^{\infty}\sum_{j\neq i}g(j)P_i(X_m=j,\tau_i\geq m+1)\\&=g(i)+\sum_{m=1}^{\infty}\sum_{j\neq i}g(j)e_{i,j}^{(m)}=g(i)+\sum_{j\neq i}g(j)e_{i,j}.\end{aligned}$$

由本章第四节习题9及定理3.4.15可知

$$E(Y_1)=\sum_{j\in S}g(j)\pi_j/\pi_i=m_{i,i}E_\pi(g)<\infty.$$

再次由强大数定律可知

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}Y_k=m_{i,i}E_\pi(g),\qquad a.s. \tag{3.5.1}$$

又由于i常返, 由命题3.2.25可知

$$N_n=\sum_{k=1}^{n}1_{\{X_k=i\}}\to\infty,\qquad \text{a.s.}$$

因此由第一章第四节习题1可得

$$\tau_i(N_n)/N_n\to m_{i,i}\quad a.s.\ \text{以及}\ \frac{1}{N_n}\sum_{k=1}^{N_n}Y_k=m_{i,i}E_\pi(g),\quad a.s.$$

注意到

$$\tau_i(N_n)\leq n<\tau_i(N_n+1).$$

我们有

$$\frac{\tau_i(N_n)}{N_n} \le \frac{n}{N_n} < \frac{\tau_i(N_n+1)}{N_n},$$

以及

$$\sum_{k=0}^{N_n-1} Y_k = \sum_{m=1}^{\tau_i(N_n)} g(X_m) \le \sum_{m=1}^{n} g(X_m) \le \sum_{m=1}^{\tau_i(N_n+1)} g(X_m) = \sum_{k=0}^{N_n} Y_k,$$

进而

$$\frac{N_n-1}{n}\Big[\frac{1}{N_n-1}\sum_{k=0}^{N_n-1} Y_k\Big] \le \frac{1}{n}\sum_{m=1}^{n} g(X_m) \le \frac{N_n}{n}\Big[\frac{1}{N_n}\sum_{k=0}^{N_n} Y_k\Big].$$

综合上面的分析, 令$n \to \infty$可知定理结论成立. □

由定理3.5.1直接可得

♦**推论3.5.2.** 设不可约马氏链$\{X_n;\ n \ge 0\}$的平稳分布为$\pi = (\pi_i,\ i \in S)$. 那么对任意$j \in S$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k=j\}} = \pi_j, \qquad a.s.$$

注意到不可约正常返马氏链平稳分布π中每个分量都是正数, 因此推论3.5.2还表明, 任意观察不可约正常返马氏链的一条轨道都能以概率1保证(或者更直观地说,对不可约马氏链做一次长时间的观察)状态空间S中的每个状态都能不可忽略地被观察到或者说能遍历S 中所有状态, 而且每个状态出现的频率稳定到对应的平稳分布. 通常我们称定理3.5.1或推论3.5.2为遍历性定理, 或者称它们为强遍历性定理而相应地称定理3.4.15为弱遍历性定理. 特别地, 对于非周期不可约正常返的马氏链, 由定理3.4.17我们还知道$P(X_n = j) \to \pi_j$. 这表明对非周期不可约正常返的马氏链而言, 无论是对其做一次长时间的观察, 还是大量观察它在足够长的时间之后某个时刻的随机表现, 直观上都能遍历状态空间中所有元素. 人们常称不可约非周期正常返马氏链为遍历马氏链, 相应地称非周期正常返状态为遍历状态.

由定理3.5.1还可得如下推论.

♦**推论3.5.3.** 设不可约马氏链$\{X_n;\ n \ge 0\}$的平稳分布为$\pi = (\pi_i,\ i \in S)$. 若g为非负函数, 那么

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g(X_k) = \sum_{i\in S} g(i)\pi_i, \quad a.s.$$

证明 若$\sum_{i\in S} g(i)\pi_i < \infty$, 推论3.5.3就是定理3.5.1 的直接应用. 若$\sum_{i\in S} g(i)\pi_i = \infty$, 那么对任意$N$, 定义函数$g_N$使得$g_N(i) = g(i) \wedge N$. 应用定理3.5.1 得

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g(X_k) \ge \limsup_{N\to\infty} \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g_N(X_k)$$

$$= \limsup_{N\to\infty} E_\pi g_N = \sum_{i\in S} g(i)\pi_i = \infty.$$

因此推论3.5.3成立. □

3.5.2 应用举例

下面我们结合几个具体的例子说明遍历性定理的应用.

▶**例3.5.4.** 观察一个动态变化的随机系统$X = \{X_n; n \geq 0\}$得到一组观察数据$(x_n, 0 \leq n \leq N)$. 若已知该随机动态系统构成状态空间$S = \{0, 1\}$的两状态时齐马氏链, 转移概率矩阵

$$\mathbf{P} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix},$$

其中$0 < p, q < 1$未知. 试给出p, q的极大似然估计量并验证该统计量的相合性. (记样本观察值$x_0, \cdots, x_N$中依次相邻地取值为$00, 01, 10, 11$ 的个数为$N_{0,0}, N_{0,1}, N_{1,0}, N_{1,1}$).

解 (1) 由马氏链的有限维分布计算可知随机变量$X_0, \cdots, X_N$取值为$x_0, \cdots, x_N$的概率为

$$\begin{aligned} p(x_0, \cdots, x_N) &= P(X_0 = x_0, \cdots, X_N = x_N) = \mu_{x_0} \prod_{k=1}^{N} p_{x_{k-1}, x_k} \\ &= \mu_{x_0} p^{N_{0,0}} (1-p)^{N_{0,1}} (1-q)^{N_{1,0}} q^{N_{1,1}}. \end{aligned}$$

由此可得对数似然函数

$$\begin{aligned} L(x_0, \cdots, x_N) &= \ln p(x_0, \cdots, x_N) \\ &= N_{0,0} \ln p + N_{0,1} \ln(1-p) + N_{1,1} \ln q + N_{1,0} \ln(1-q) + \ln \mu_{x_0}. \end{aligned}$$

由极大似然估计典型方法可得p, q的极大似然估计为

$$\hat{p}_{MLE} = \frac{N_{0,0}}{N_{0,0} + N_{0,1}}, \qquad \hat{q}_{MLE} = \frac{N_{1,1}}{N_{1,0} + N_{1,1}}.$$

(2) 对任意$n \geq 0$ 令$Y_n = (X_n, X_{n+1})$. 由本章第四节习题10可知$Y = \{Y_n; n \geq 0\}$是状态为$\{(i,j) : i, j = 0, 1\}$的不可约正常返马氏链, 平稳分布

$$(\pi_{0,0}, \pi_{0,1}, \pi_{1,0}, \pi_{1,1}) = (\pi_0 p, \pi_0(1-p), \pi_1(1-q), \pi_1 q),$$

其中(π_0, π_1)为X的平稳分布. 由推论3.5.2可知

$$\lim_{N\to\infty} \frac{N_{0,0}}{N} = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} 1_{\{Y_k = (0,0)\}} = \pi_{0,0} = \pi_0 p, \qquad a.s.$$

以及

$$\lim_{N\to\infty}\frac{N_{0,0}+N_{0,1}}{N}=\lim_{N\to\infty}\left[\frac{1}{N}\sum_{k=0}^{N-1}1_{\{Y_k=(0,0)\}}+\frac{1}{N}\sum_{k=0}^{N-1}1_{\{Y_k=(0,1)\}}\right]$$
$$=\pi_{0,0}+\pi_{0,1}=\pi_0 p+\pi_0(1-p)=\pi_0,\qquad a.s.$$

因此由几乎必然收敛的运算性质可得

$$\lim_{N\to\infty}\hat{p}_{MLE}=\lim_{N\to\infty}\frac{N_{0,0}}{N_{0,0}+N_{0,1}}=\lim_{N\to\infty}\frac{N_{0,0}/N}{(N_{0,0}+N_{0,1})/N}=\frac{\pi_0 p}{\pi_0}=p,\qquad a.s.$$

所以$\hat{p}_{MLE}$为p的相合估计. 类似证明可知$\hat{q}_{MLE}$也是q的相合估计. □

▶**例3.5.5.** 对零售商而言控制商品的仓储成本是非常重要的. 为此他们常常会根据商品的性质, 市场需求以及自身的实力选择最佳的采购策略. 一种比较简单但被广泛采用的策略如下: 先综合考虑确定一个商品库存量上限K, 然后确定一个补货标准k, 在每一个单位时间末期对商品库存量进行盘点, 当剩余商品数大于补货标准时, 则不用补充货源, 下一个单位时间内销售剩余商品; 当剩余商品数不超过补货标准时就及时采购货源使库存量达到上限K. 假定零售商使用这种采购策略, 采购补充的商品在下一个单位时间的期初到货. 以ξ_i 表示第i个单位时间内的市场需求, 并假定市场需求是独立同分布的. 在这样的设定下, 若以X_n 表示每个单位时间期末的零售商拥有的商品数, 可以检验$X=\{X_n;n\geq 0\}$ 就是例*3.1.7*中的模型, 通常我们称它为(k,K) 商品存储模型, 现在假定$K=3$, 表示需求的随机变量ξ 满足

$$P(\xi=0)=0.3,\ P(\xi=1)=0.4,\ P(\xi=2)=0.2,\ P(\xi=3)=0.1.$$

若每销售一件商品可以获利12元, 每件单位时间内未售出的商品需额外花费2元的存储费, 试确定最优的k使得从长期看平均收益最大.

解 为书写方便, 记$p_0=0.3$, $p_1=0.4$, $p_2=0.2$, $p_3=0.1$. 根据设定, 给定采购策略$k\in\{0,1,2\}$后X的转移概率矩阵$\mathbf{P}$与k有关. 记$\mathbf{P}=({}_kp_{i,j})_{4\times 4}$, 其中

$${}_kp_{i,j}=\begin{cases}P((3-\xi)\vee 0=j), & i\leq k\\ P((i-\xi)\vee 0=j), & k<i\leq K\end{cases}$$
$$=\begin{cases}p_{3-j}, & i\leq k,\\ p_{i-j}, & k<i\leq 3,\ 0<j\leq i,\\ p_i+\cdots+p_3, & k<i\leq 3,\ j=0,\\ 0, & k<i\leq 3,\ i<j\leq 3.\end{cases}$$

对任意$n \geq 1$, 在第n个单位时间获得的净利润$r_n^{(k)}$为

$$r_n^{(k)} = \begin{cases} 12(3-X_n) - 2X_n = 36 - 14X_n, & X_{n-1} \leq k, \\ 12(X_{n-1} - X_n) - 2X_n = 12X_{n-1} - 14X_n, & k < X_{n-1} \leq 3. \end{cases}$$

对任意$0 \leq i, j \leq 3$, 定义

$$g^{(k)}(i,j) = \begin{cases} 36 - 14j, & i \leq k \\ 12i - 14j, & k < i \leq 3. \end{cases}$$

那么$r_n^{(k)} = g^{(k)}(X_{n-1}, X_n)$. 显然无论$k$取何值, X都是有限状态不可约马氏链, 从而正常返, 记其平稳分布为

$$\pi = (\pi_0^{(k)}, \pi_1^{(k)}, \pi_2^{(k)}, \pi_3^{(k)}).$$

可知平均利润

$$\frac{1}{n}\sum_{m=1}^{n} r_m^{(k)} = \frac{1}{n}\sum_{m=1}^{n} g^{(k)}(X_{m-1}, X_m).$$

令$Y_n = (X_n, X_{n+1})$, 仍由本章第四节习题10 可知$Y = \{Y_n\}$在状态空间

$$\tilde{S} = \{(i,j); i, j \in S, p_{i,j} \neq 0\}$$

上仍是不可约正常返的, 平稳分布为

$$(\pi_i^{(k)}{}_k p_{i,j}, (i,j) \in \tilde{S}).$$

由定理3.5.1可知, 随$n \to \infty$,

$$\frac{1}{n}\sum_{m=1}^{n} r_m^{(k)} = \frac{1}{n}\sum_{m=1}^{n} g^{(k)}(Y_m) \to \sum_{i=0}^{3}\sum_{j=0}^{3} g^{(k)}(i,j)\pi_i^{(k)}{}_k p_{i,j}.$$

显然最后的极限就是采取(k, K)-存储策略的长期平均收益; 将其记作$r(k)$. 利用$g^{(k)}(i,j)$和${}_k p_{i,j}$的表达式, $r(k)$可简化为

$$\begin{aligned} r(k) &= \sum_{i=0}^{k}\sum_{j=0}^{3}(36 - 14j)\pi_i^{(k)} p_{3-j} + \sum_{i=k+1}^{3}\sum_{j=0}^{3}(12i - 14j)\pi_i^{(k)}{}_k p_{i,j}, \\ &= 9.4\sum_{i=0}^{k}\pi_i^{(k)} + 12\sum_{i=k+1}^{3} i\pi_i^{(k)} - 14\sum_{i=k+1}^{3}\left(\pi_i^{(k)}\sum_{j=1}^{i} jp_{i-j}\right). \end{aligned}$$

解决问题的关键就是找到使$r(\cdot)$取到最大值的k, 为此我们需要知道$\pi_i^{(k)}$.

当$k = 2$时直接计算可知,

$$r(2) = 9.4\sum_{i=0}^{2}\pi_i^{(2)} + 36\pi_3^{(2)} - 14\pi_3\sum_{j=1}^{3} jp_{3-j} = 9.4\sum_{i=0}^{3}\pi_i^{(2)} = 9.4.$$

当$k=1$时由转移概率矩阵

$$(p_{i,j}) = \begin{pmatrix} 0.1 & 0.2 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.2 & 0.4 & 0.3 \end{pmatrix}$$

可求得平稳分布

$$\pi^{(1)} = (19/110,\ \ 30/110,\ \ 40/110,\ \ 21/110).$$

此时平均收益极限

$$\begin{aligned} r(1) &= 9.4(\pi_0^{(1)} + \pi_1^{(1)}) + 12(2\pi_2^{(1)} + 3\pi_3^{(1)}) - 14(\pi_2^{(1)} + 1.9\pi_3^{(1)}) \\ &= 9.4 + 0.6\pi_2^{(1)} \approx 9.62. \end{aligned}$$

当$k=0$时由转移概率矩阵

$$(p_{i,j}) = \begin{pmatrix} 0.1 & 0.2 & 0.4 & 0.3 \\ 0.7 & 0.3 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.2 & 0.4 & 0.3 \end{pmatrix}$$

可求得平稳分布

$$\pi^{(0)} = (343/1070,\ 300/1070,\ 280/1070,\ 147/1070).$$

此时平均收益极限

$$\begin{aligned} r(0) &= 9.4\pi_0^{(0)} + 12(\pi_1^{(0)} + 2\pi_2^{(0)} + 3\pi_3^{(0)}) - 14(0.3\pi_1^{(0)} + \pi_2^{(0)} + 1.9\pi_3^{(0)}) \\ &= 9.4 - 1.6\pi_1^{(0)} + 0.6\pi_2^{(0)} \approx 9.11. \end{aligned}$$

因此$(1,3)$策略是长期平均利润准则下的最优策略. □

▶**例3.5.6.** *考虑例3.1.17中的马氏链$X=\{X_n; n\geq 0\}$, 并假设*

$$0 < p_0,\ p_0 + p_1 < 1,\ \rho = \sum_{k=1}^{\infty} k p_k = E(\xi_1) < 1.$$

求

$$(1)\ \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k, \qquad\qquad (2)\ \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k=0\}}.$$

解 由本节习题8可知给定条件下X是不可约正常返的马氏链. 记其平稳分布为$\pi=(\pi_k, k\geq 0)$. 令

$$A(s) = \sum_{k=0}^{\infty} p_k s^k, \qquad \pi(s) = \sum_{k=0}^{\infty} \pi_k s^k.$$

注意到对任意$i \in S$,

$$\pi_i = \sum_{k\in S} \pi_k p_{k,i} = \pi_0 p_i + \sum_{k=1}^{i+1} \pi_k p_{i+1-k}.$$

对任意$0 \le s \le 1$, 我们有

$$\begin{aligned}
\pi(s) &= \sum_{i=0}^{\infty} \pi_i s^i = \pi_0 \sum_{i=0}^{\infty} p_i s^i + \sum_{i=0}^{\infty}\sum_{k=1}^{i+1} \pi_k p_{i+1-k} s^i \\
&= \pi_0 \sum_{i=0}^{\infty} p_i s^i + \sum_{k=1}^{\infty}\sum_{i=k-1}^{\infty} \pi_k p_{i-k+1} s^i \\
&= \pi_0 A(s) + \sum_{k=1}^{\infty}\sum_{m=0}^{\infty} \pi_k p_m s^{m+k-1} \\
&= \pi_0 A(s) + \frac{1}{s}\sum_{k=1}^{\infty} \pi_k s^k \sum_{m=0}^{\infty} p_m s^m \\
&= \pi_0 A(s) + \frac{1}{s}(\pi(s) - \pi_0) A(s).
\end{aligned}$$

由此可得

$$\pi(s) = \frac{\pi_0 (s-1) A(s)}{s - A(s)}, \tag{3.5.2}$$

以及

$$\frac{A(s)-1}{s-1} = 1 - \pi_0 \frac{A(s)}{\pi(s)}.$$

令$s \to 1-$. 我们有

$$\rho = A'(1) = 1 - \pi_0.$$

因此由推论3.5.2可得

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{\{X_k=0\}} = \pi_0 = 1 - \rho.$$

另一方面, 将$\pi_0 = 1-\rho$代入(3.5.2)可得

$$\pi(s) = (1-\rho)\frac{(1-s)A(s)}{s - A(s)},$$

从而

$$\pi'(s) = (1-\rho)\Big\{\frac{(s-1)A'(s)}{s-A(s)} + \frac{A(s)[s - A(s) - (1-A'(s))(s-1)]}{(s-A(s))^2}\Big\}.$$

注意到

$$\lim_{s\to 1-} \frac{s - A(s)}{(1-\rho)(s-1)} = 1.$$

若还假设

$$A''(1-) = E(\xi_1^2) - E(\xi_1) < \infty,$$

那么

$$\pi'(1-) = \lim_{s\to 1-}(1-\rho)\Big\{\frac{(s-1)A'(s)}{(1-\rho)(s-1)} + \frac{A(s)[s-A(s)-(1-A'(s))(s-1)]}{(1-\rho)^2(s-1)^2}\Big\}$$
$$= A'(1) + \lim_{s\to 1-}\frac{A(s)[s-A(s)-(1-A'(s))(s-1)]}{(1-\rho)(s-1)^2}.$$

由微积分中洛必达法则可得

$$\lim_{s\to 1-}\frac{A(s)[s-A(s)-(1-A'(s))(s-1)]}{(1-\rho)(s-1)^2} = \lim_{s\to 1-}\frac{A(s)A''(s)}{2(1-\rho)} = \frac{A''(1-)}{2(1-\rho)}.$$

因此, 由推论3.5.3可得

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k = \sum_{k=0}^{\infty}k\pi_k = \pi'(1-) = \rho + \frac{A''(1-)}{2(1-\rho)}.$$ □

♠**注记3.5.7.** *例3.5.6与评估服务系统中窗口服务效率这样的实际问题有关. 评估时人们常关心像窗口排队长度、窗口空闲时间等问题. 在适当的假设条件下, 比如只考虑一个窗口, 设顾客到达服从强度为λ 的泊松过程, 顾客到达后按“先来先服务”的原则接受服务台的服务, 假设为每个顾客服务的时间服从相同的G分布并且相互独立, 而且与顾客到达的时间间隔也独立(称满足这些假定的服务系统为$M(\lambda)/G/1$ 系统). 为了简单, 假设服务时间服从概率密度函数为g 的连续分布. 以X_n 表示第n 个顾客服务完毕离开时, 在窗口等待服务的顾客数, 初始时刻取为第一个顾客到达时刻, 即$X_0=1$. 那么可验证$X=\{X_n;\ n\ge 0\}$ 为例3.1.17中马氏链. 事实上以U_n 表示第n个顾客接受服务的时长, ξ_n 表示在第n个顾客接受服务期间到达的顾客数, 由指数分布无记忆性可知$\{\xi_n;\ n\ge 1\}$ 是独立同分布的随机变量, 分布列*

$$p_k = P(\xi_n=k) = E(P(\xi_n=k|U_n)) = \int_0^\infty \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}g(t)\mathrm{d}t,\ \ k\ge 0,$$

而且

$$X_n = \begin{cases} X_{n-1}-1+\xi_n, & X_{n-1}>0,\\ \xi_n, & X_{n-1}=0.\end{cases}$$

一般地, 我们称X为排队系统的嵌入链. 显然例3.5.6就是所关心问题在嵌入链X上的体现.

练习题3.5

1 在例3.5.5中令$p_0=p_1=p_2=p_3=0.25$, 其他条件不变, 求最佳策略$(k,3)$.

2 对$M(\lambda)/G/1$嵌入链X, 若服务台的服务时间G服从参数为μ的指数分布且$\lambda<\mu$. 求长时间的平均队长和窗口无人的时间比例.

3 上班族小张每日若不下雨则乘地铁上班, 费用5元; 若下雨则乘出租车上班, 费用是80元. 已知上班时天气状态可以用一个两状态的时齐马氏链模

拟, 转移概率矩阵为$\begin{pmatrix} 0.6 & 0.4 \\ 0.8 & 0.2 \end{pmatrix}$, 其中第一行对应不下雨状态, 第二行对应下雨状态. 长期看, 小张上班的日平均交通费用是多少?

4 某种生产流水线分正常生产1和停机检修2两种状态, 在检修能及时进行条件下, 机器每日状态之间的变化可用转移概率矩阵为$\begin{pmatrix} 0.9 & 0.1 \\ 0.8 & 0.2 \end{pmatrix}$的时齐马氏链来刻画, 若检修时不能及时进行, 那么流水线保持停机检修状态直到检修后才正常运转. 每条流水线检修时需要1名专业维修工. 现在已知某工厂有2条这样的生产流水线, 流水线正常生产时独立运转. (1) 分别就工厂招募了k名维修工$(k=1,2)$ 给出每日停机检修流水线数量的转移概率矩阵. (2) 为了保证流水线90% 以上的时间能正常运转或得到及时检修, 最少需要招募几名维修工?

5 设X是$0,1$两状态的时齐马氏链, 转移概率矩阵$P=\begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix}$, 其中$0 \le p < 1$未知, $q=1-p$. 试利用观察值$x_0,\cdots,x_N$中状态0出现的次数N_0给出p的一个相合估计并对相合性给出证明.

6 若有限状态的马氏链X存在唯一平稳分布$\pi=(\pi_i,\ i\in S)$, 对状态空间S上的任意函数g. 证明

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} g(X_k)=\sum_{i\in S} g(i)\pi_i, \qquad a.s.$$

7* 设马氏链X的状态空间$S=\{1,2,3,4,5\}$, $X_0=5$, 转移概率矩阵为

$$\begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 3/4 & 0 \\ 0 & 0 & 3/4 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}.$$

令$\tau_1=\inf\{n\ge 1; X_n=1\}$, $\tau_3=\inf\{n\ge 1; X_n=3\}$ 并取定$\pi_1=4/7$, $\pi_2=3/7$, $\pi_3=\pi_4=1/2$, $\pi_5=0$.

(1) 对任意$j\in S$, 证明

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k=j\}}=\pi_j\left[1_{\{\tau_1<\infty\}}1_{\{1,2\}}(j)+1_{\{\tau_3<\infty\}}1_{\{3,4\}}(j)\right], \quad a.s.$$

(2) 对任意$j\in S$, 证明$\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n} P(X_k=j)=\pi_j/2$.

8* 利用本章第四节习题11证明例题3.5.6中的排队模型X是不可约正常返的.

第四章　马氏链应用模型

马氏链在实际问题中有着广泛应用. 上一章我们已经借助于一些例题对此做了部分展示. 这一章我们补充介绍可逆马氏链、隐马氏链和分枝过程这三个特殊马氏链模型, 并简要地介绍它们在理论与实践中的一些应用.

4.1　可逆马氏链与蒙特卡罗模拟

马氏链蒙特卡罗(MCMC)模拟是一种常见的随机模拟方法, 本节将在引入可逆马氏链基础上阐明MCMC方法的基本原理, 并介绍两个经典算法.

4.1.1　可逆马氏链

一个用马氏链描述的动态系统, 常需讨论它在长时间后的稳定表现. 由前面理论可知, 这个稳定表现体现为平稳分布. 一般地, 平稳分布可以通过求解一个线性方程组得到, 只是计算量会比较大. 但对一些特殊的马氏链而言, 我们可以用较简便的方法求出其平稳分布. 下面我们介绍其中的一种特殊马氏链——可逆马氏链, 它的定义如下.

♣**定义4.1.1.** *称以$\mathbf{P} = (p_{i,j})_{i,j\in S}$ 为转移矩阵的马氏链X为可逆的, 如果存在概率分布$\mu = (\mu_i, i \in S)$使得对任意$i, j \in S$,*

$$\mu_i p_{i,j} = \mu_j p_{j,i}. \tag{4.1.1}$$

μ称为X的可逆初分布, 简称可逆分布. 物理学中称(4.1.1)为细致平衡条件.

若X是可逆的, 并记可逆分布为μ, 那么由定义容易发现

$$\sum_i \mu_i p_{i,j} = \sum_i \mu_j p_{j,i} = \mu_j,$$

即μ也是X的平稳分布. 若进一步假定可逆分布μ是X的初始分布, 那么在任意时刻k, X_k的分布都是μ. 此时若以$Y = \{Y_n; 0 \le n \le N\}$表示$X$ 在时刻N及N之前

的“逆”时间过程, 即$Y_n = X_{N-n}$. 可以证明Y是初始分布为μ, 转移概率矩阵与X相同的马氏链(参见本节习题8).

由可逆分布一定是平稳分布可知, 如果可逆马氏链X还是不可约的, 那么X一定是正常返链, 此时唯一的平稳分布就是可逆分布, 进而可以通过方程(4.1.1) 求出平稳分布.

▶**例4.1.2.** *考察状态空间$S = \{0, 1, \cdots, M\}$, 转移概率为*

$$p_{i,i-1} = p_i, \quad p_{i,i} = r_i, \quad p_{i,i+1} = q_i, \quad i = 1, \cdots, M-1,$$

$$p_{0,0} = r_0; \ p_{0,1} = q_0; \ p_{M,M-1} = p_M; \quad p_{M,M} = r_M,$$

的马氏链X, 其中$1 \geq q_0, p_M > 0$, $r_0 = 1 - q_0, r_M = 1 - p_M$ 而且对任意$1 \leq i \leq M-1$, $0 < p_i, q_i < 1$, $p_i + r_i + q_i = 1$. 求X的平稳分布.

解 容易看出X是不可约的. 对任意$M > i \geq 0$, 令

$$\nu_i p_{i,i+1} = \nu_{i+1} p_{i+1,i},$$

即对任意$M > i \geq 0$, $\nu_i q_i = \nu_{i+1} p_{i+1}$. 由此可得, 对任意$M > i \geq 0$,

$$\nu_{i+1} = \nu_0 \prod_{k=1}^{i+1} \frac{q_{k-1}}{p_k}.$$

再令$\sum\limits_{i=0}^{M} \nu_i = 1$可得

$$\nu_0 = \frac{1}{1 + \sum_{i=1}^{M} \prod_{k=1}^{i} \frac{q_{k-1}}{p_k}}.$$

此时$\nu = (\nu_0, \nu_1, \cdots, \nu_M)$为$X$的可逆分布, 进而也是$X$的唯一平稳分布. □

当然, 并不是所有的马氏链都是可逆的. 比如, 对转移概率矩阵

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.2 & 0.7 \end{pmatrix}$$

的马氏链而言, 由于方程组

$$\mu_1 p_{1,2} = \mu_2 p_{2,1}; \quad \mu_2 p_{2,3} = \mu_3 p_{3,2}; \quad \mu_3 p_{3,1} = \mu_1 p_{1,3},$$

没有非零解. 即没有不全为零的μ_1, μ_2, μ_3使得

$$\mu_1 = \mu_2; \quad \mu_2 = 2\mu_3; \quad \mu_3 = \mu_1;$$

同时成立, 因此该马氏链不是可逆马氏链.

下面的命题给出了判断可逆马氏链的一个必要条件.

▲**命题4.1.3.** *若X为可逆不可约马氏链, 那么对任意$i, j \in S$, 转移概率$p_{i,j} = 0$当且仅当$p_{j,i} = 0$.*

证明 记X的可逆分布为$\mu = (\mu_i, i \in S)$. 那么μ也是X的平稳分布. 由于X不可约, 由定理3.4.14可知, 对任意$i \in S$, $\mu_i > 0$. 进而由(4.1.1)可知对任意$j \in S$, $p_{i,j} = 0$当且仅当$p_{j,i} = 0$. $\square$

一般地, 对于可逆不可约马氏链而言, 任取$i_0 \in S$, 对任意$j \in S$, 存在一条路径使得$i_0 \to j$. 不妨设该路径为

$$i_0 \hookrightarrow i_1 \hookrightarrow \cdots \hookrightarrow i_n \hookrightarrow j,$$

这里$\hookrightarrow$表示"一步可达", 那么

$$p_{i_0,i_1}p_{i_1,i_2}\cdots p_{i_n,j} > 0.$$

此时由命题4.1.3可知

$$p_{j,i_n}\cdots p_{i_2,i_1}p_{i_1,i_0} > 0.$$

再利用(4.1.1)可得, 对可逆分布$\mu = (\mu_i, i \in S)$,

$$\begin{aligned}\mu_{i_0}p_{i_0,i_1}p_{i_1,i_2}\cdots p_{i_n,j} &= \mu_{i_1}p_{i_1,i_2}\cdots p_{i_n,j}p_{i_1,i_0} \\ &= \mu_{i_2}p_{i_2,i_3}\cdots p_{i_n,j}p_{i_2,i_1}p_{i_1,i_0} \\ &= \cdots = \mu_j p_{j,i_n}\cdots p_{i_2,i_1}p_{i_1,i_0}.\end{aligned}$$

记

$$v(i_0,j) = \frac{p_{i_0,i_1}p_{i_1,i_2}\cdots p_{i_n,j}}{p_{j,i_n}\cdots p_{i_2,i_1}p_{i_1,i_0}} \in (0,\infty),$$

那么$v(i_0,i_0) = 1$且对任意$j \in S$, $\mu_j = \mu_{i_0}v(i_0,j)$. 由

$$1 = \sum_{j\in S}\mu_j = \sum_{j\in S}v(i_0,j)\mu_{i_0}$$

可知对任意$j \in S$

$$\mu_j = \frac{v(i_0,j)}{\sum_{j\in S}v(i_0,j)}. \tag{4.1.2}$$

下面定理为我们判断不可约正常返链是否可逆提供了一个更直观的方法.

★定理4.1.4. *若不可约正常返马氏链X满足条件: 对任意$i,j \in S$, $p_{i,j} = 0$当且仅当$p_{j,i} = 0$. 那么X是可逆马氏链的充要条件是从它任意一状态出发回到该状态的路径与它的反向路径有相同的概率. 即对任意$n \geq 1$以及任意状态$i, i_1, i_2, \cdots, i_n$, 若$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_n,i} > 0$, 那么*

$$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_n,i} = p_{i,i_n}p_{i_n,i_{n-1}}\cdots p_{i_1,i}. \tag{4.1.3}$$

证明 前面分析已给出了必要性, 下证充分性. 任取$i,j \in S$以及对任意$n \geq 1$, 由假设可知, 对任意$i_1, \cdots, i_n$,

$$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_n,j}p_{j,i} = p_{i,j}p_{j,i_n}\cdots p_{i_2,i_1}p_{i_1,i}.$$

关于所有$i_1, \cdots, i_n$求和得

$$p_{i,j}^{(n+1)}p_{j,i} = p_{i,j}p_{j,i}^{(n+1)}.$$

进一步关于n求和并求平均得

$$\lim_{k\to\infty}\frac{1}{k}\sum_{n=1}^{k}p_{i,j}^{(n+1)}p_{j,i}=p_{i,j}\lim_{k\to\infty}\frac{1}{k}\sum_{n=1}^{k}p_{j,i}^{(n+1)}.$$

由于X不可约且正常返, 由推论3.4.8以及定理3.4.14可知

$$\pi_j p_{j,i}=\pi_i p_{i,j},$$

其中$\pi=(\pi_i, i\in S)$是X的平稳分布. 因此X可逆. □

下面这个定理表明, 对任意离散分布, 都存在一个恰当的可逆不可约过程使得该过程的可逆分布,也即平稳分布, 就是给定的离散分布.

★定理4.1.5. *任取一个离散分布π,记其所有取值概率为正的点构成的集合为S, 即$\pi=(\pi_i, i\in S)$, 其中$\pi_i>0$且$\sum\limits_{i\in S}\pi_i=1$. 一定存在$S$上的可逆不可约马氏链$X$, 使得$\pi$是$X$的平稳分布.*

证明 任取S上不可约的马氏链Y使其转移概率矩阵$\mathbf{Q}=(q_{i,j})_{i,j\in S}$满足

$$q_{i,j}=0\Leftrightarrow q_{j,i}=0.$$

对任意$i\neq j\in S$, 令

$$\alpha_{i,j}=\begin{cases}\min\{\frac{\pi_j q_{j,i}}{\pi_i q_{i,j}},1\}, & q_{i,j}>0,\\ 1, & q_{i,j}=0,\end{cases}$$

以及

$$p_{i,j}=q_{i,j}\alpha_{i,j},\quad p_{i,i}=q_{i,i}+\sum_{j\neq i}q_{i,j}(1-\alpha_{i,j}).$$

容易检验$\mathbf{P}=(p_{i,j})_{i,j\in S}$是个随机矩阵而且由构造可知对任意$i\neq j$,

$$p_{i,j}>0\Leftrightarrow q_{i,j}>0.$$

因此由Y的不可约性可知, 以$\mathbf{P}$为转移概率矩阵的马氏链X也是不可约的. 进一步, 对任意$i\neq j$, 由$\pi_i p_{i,j}=\pi_i q_{i,j}\alpha_{i,j}$可知:

若$\alpha_{i,j}<1$, 则$\alpha_{i,j}=\frac{\pi_j q_{j,i}}{\pi_i q_{i,j}}$而且$\alpha_{j,i}=1$. 此时

$$\pi_i p_{i,j}=\pi_j q_{j,i}=\pi_j q_{j,i}\alpha_{j,i}=\pi_j p_{j,i}.$$

若$\alpha_{i,j}=1$且$\pi_i q_{i,j}=\pi_j q_{j,i}$, 则$\alpha_{j,i}=1$而且

$$\pi_i p_{i,j}=\pi_i q_{i,j}=\pi_j q_{j,i}=\pi_j q_{j,i}.$$

若$\alpha_{i,j}=1$且$\pi_i q_{i,j}>\pi_j q_{j,i}$, 则$\alpha_{j,i}<1$而且

$$\pi_i p_{i,j}=\pi_i q_{i,j}=\pi_j q_{j,i}\alpha_{j,i}=\pi_j q_{j,i}.$$

总之, 此时π为X的可逆分布, 从而为X的平稳分布. □

♠注记4.1.6. *若对不可约马氏链Y还要求存在某个i使得$p_{i,i}>0$, 那么X还是非周期的, 从而由定理3.4.17可知π还是X的极限分布, 即X_n的分布收敛到π.*

♠**注记4.1.7.** *通常称定理4.1.5证明中采用的构造转移矩阵*$\mathbf{P}$*进而得到可逆马氏链的方法为Hastings-Metropolis算法, 称转移矩阵*$\mathbf{Q}=(q_{i,j})_{i,j\in S}$*为预选矩阵.*

4.1.2 马氏链蒙特卡罗(MCMC)方法*

在复杂的统计问题中有许多问题很难通过直接地理论分析或计算给出准确回答, 需要借助随机模拟给出近似解答. 比如, 当我们直接计算诸如$E(h(X))$这样的算式有困难时, 一种自然的方法是借助大数定律, 通过生成一列与X同分布的随机数$x_1,x_2,\cdots,x_n$, 然后用$\sum_{i=1}^n h(x_i)/n$近似. 但当生成X的随机数有困难时, 这种方案实现起来就有一定难度. 马氏链蒙特卡罗(MCMC)方法是能解决这类问题的一种随机模拟方法. 该方法的主要思想是将特定分布设计成某个马氏链的平稳分布, 基于马氏链的极限定理(定理3.4.17)和遍历性定理(3.5.1和推论3.5.2), 通过模拟马氏链的轨道(实现), 近似生成特定分布的随机数, 进而估计所需要的各种统计量. 下面我们结合两种典型采样方法简单介绍MCMC方法.

(A) Metropolis采样

Metropolis采样就是基于Hastings-Metropolis算法生成一个可逆马氏链样本轨道的抽样方法. 下面是具体的算法解释和说明.

要对离散分布$\pi=(\pi_i,i\in S)$采样, 按照定理4.1.5及注记4.1.6, 我们可以构造一个非周期不可约可逆马氏链X使得π是X的平稳分布. 这样在足够长时间之后, 马氏链X的模拟值就可以近似看作是分布π的采样. 而要模拟出马氏链X关键是模拟它的转移机制$(p_{i,j})$. 由定理4.1.5 可知, 这要先模拟具有性质

$$q_{i,j}=0\Leftrightarrow q_{j,i}=0 \text{ 以及 } q_{i,i}>0$$

的转移机制$(q_{i,j})_{j\in S}$, 也即在给定$X_n=i$ 的条件下, 模拟分布列$(q_{i,j},j\in S)$. 然后注意到$j\neq i$ 时$p_{i,j}=q_{i,j}\alpha_{i,j}$. 为了体现因子$\alpha_{i,j}$, 我们可模拟一次均匀分布$U[0,1]$. 设$Y$是模拟分布列$(q_{i,j},j\in S)$所得的样本, Z是模拟$U[0,1]$所得的样本而且与Y独立. 令

$$W=\begin{cases}Y, & Z\le\alpha_{i,Y};\\ i, & \text{其它},\end{cases}$$

那么$j\neq i$时,

$$P(W=j)=P(Y=j)P(Z\le\alpha_{i,j})=q_{i,j}\alpha_{i,j}=p_{i,j},$$

而$P(W=i)=1-\sum_{j\neq i}p_{i,j}=p_{i,i}$. 即$W$是服从分布列$(p_{i,j},j\in S)$的随机变量. 因此可用如下算法实现模拟转移机制$(p_{i,j})$.

(s1) 设当前时刻$X_n=i$;

(s2) 产生一个分布列为$(q_{i,1}, q_{i,2}, \cdots)$的随机数, 假设为$j$;

(s3) 计算$\alpha_{i,j}$;

(s4) 独立地取一个服从$[0,1]$上均匀分布的随机数u, 如果$u \le \alpha_{i,j}$, 则将状态更新为$X_{n+1} = j$并回到(s1), 否则状态不更新, 即令$X_{n+1} = i$, 并回到(s1).

考虑到$\alpha_{i,j} = \frac{\pi_j q_{j,i}}{\pi_i q_{i,j}} \wedge 1$, 我们可以将上述算法中(s3),(s4)两步优化为

(s3) 若$\frac{\pi_j q_{j,i}}{\pi_i q_{i,j}} \ge 1$, 则将状态更新为$X_{n+1} = j$并回到(s1); 否则进行(s4);

(s4) 独立地取一个服从$[0,1]$上均匀分布的随机数u, 如果$u \le \frac{\pi_j q_{j,i}}{\pi_i q_{i,j}}$, 则将状态更新为$X_{n+1} = j$并回到(s1), 否则状态不更新, 即令$X_{n+1} = i$, 并回到(s1).

▶**例4.1.8.** *将所有n级排列所组成的集合$\mathfrak{F}$看作一个样本空间, 在$\mathfrak{F}$上可以定义一种离散分布π使得对任意$v \in \mathfrak{F}$, $\pi_v = T(v)/c$, 其中$T(v)$表示排列v中数n的位置, $c = (n+1)!/2$为所有$T(v)$的和. 试用Metropolis采样法生成$n = 50$时分布π的一个样本.*

算法设计: 为了利用Metropolis采样法, 我们需要设计一个预选矩阵, 为此对任意一个50级排列$v \in \mathfrak{F}$, 令

$$N(v) = \{u \in S, u\text{可由}v\text{最多经过一次相邻对换得到}\}.$$

对任意v, u, 令

$$q_{v,u} = \begin{cases} 0, & u \notin N(v), \\ 1/|N(v)| = 1/50, & u \in N(v), \end{cases}$$

其中$|N(v)|$表示集合$N(v)$中的元素个数. 由于任意两个不同的50级排列总可通过相邻对换相互转换, 因此可以验证以$(q_{v,u})$为转移概率矩阵的马氏链是不可约的, 而且

$$q(v,u) = 0 \Leftrightarrow q(u,v) = 0.$$

由于分布π 满足对任意$v \in \mathfrak{F}$, $\pi_v = cT(v)$, 其中c为常数. 因此当$q_{v,u} > 0$时,

$$\alpha_{v,u} = \min\left\{\frac{\pi_u q_{u,v}}{\pi_v q_{v,u}}, 1\right\} = \min\left\{\frac{T(u)}{T(v)}, 1\right\}.$$

注意到$q_{v,v} > 0$, 由Hastings-Metropolis算法得到的马氏链X是不可约非周期可逆马氏链, π是X的极限分布.

下面我们用Metropolis采样法模拟马氏链X在一个样本轨道下的序列$\{X_k\}$. 当k足够大时, X_k的值就是π的一个近似取样. 基本模拟流程如下

(s1) 给X_0任意赋予一个50级排列;

(s2) 设当前时刻$X_k = v$, v表示一个n级排列;

(s3) 产生一个在0到49整数点上均匀分布的随机数, 假设为j, 令u表示对换v中第j 和$j+1$位置上的数所得到的排列, 当$j = 0$时$u = v$;

(s4) 若$\frac{T(u)}{T(v)} \geq 1$, 则将状态更新为$X_{k+1} = u$并回到(s2); 否则进行(s5);

(s5) 独立地取一个$U[0,1]$的随机数U, 如果$U \leq \frac{T(u)}{T(v)}$, 则将状态更新为$X_{k+1} = u$并回到(s2), 否则状态不更新,即令$X_{k+1} = v$, 并回到(s2).

基于这些流程, 可以编写计算机程序实现对分布π的采样, 请读者自己完成.

(B) Gibbs 采样

在使用MCMC方法对多维随机变量采样时, 使用最为广泛的方法就是所谓Gibbs采样. 本质上它仍可以看成是一种采用特殊预选矩阵的Metropolis采样.

记离散多维随机变量$X = (X_1, \cdots, X_n)$. 其可能的取值集合为S. 记X的分布为π, 即对任意$\mathbf{x} = (x_1, x_2, \cdots, x_n) \in S$,

$$P(X = \mathbf{x}) = \pi_{\mathbf{x}} = g(x_1, \cdots, x_n) > 0.$$

对任意$\mathbf{x} = (x_1, \cdots, x_n) \in S$以及$i = 1, \cdots, n$, 记

$$b_i(\mathbf{x}) = P(X_j = x_j; j \neq i) = \sum_{u_i} g(x_1, \cdots, x_{i-1}, u_i, x_{i+1}, \cdots x_n).$$

那么

$$P(X_i = y_i | X_j = x_j; j \neq i) = \frac{\pi_{\mathbf{y}}}{b_i(\mathbf{x})}$$

表示给定X除i之外$n-1$个分量条件下, 第i个分量取值的条件分布, 其中$\mathbf{y} = (x_1, \cdots, x_{i-1}, y_i, x_{i+1}, \cdots, x_n)$. 注意到

$$\begin{aligned} P(X = \mathbf{x}) P(X_i = y_i | X_j = x_j; j \neq i) &= \pi_{\mathbf{x}} \frac{\pi_{\mathbf{y}}}{b_i(\mathbf{x})} \\ &= P(X = \mathbf{y}) P(X_i = x_i | X_j = x_j; j \neq i). \end{aligned} \tag{4.1.4}$$

这意味着如果对任意只有一个分量取值不同的$\mathbf{x}, \mathbf{y} \in S$, 定义

$$q_{\mathbf{x},\mathbf{y}} = a_i P(X_i = y_i | X_j = x_j; j \neq i), \tag{4.1.5}$$

其中a_i是正常数, i是$\mathbf{x}$与$\mathbf{y}$中值不同的分量序号. 那么我们总有

$$\pi_{\mathbf{x}} q_{\mathbf{x},\mathbf{y}} = \pi_{\mathbf{y}} q_{\mathbf{y},\mathbf{x}}.$$

受此启发, 构造状态空间S上的转移概率如下

$$p_{\mathbf{x},\mathbf{y}} = \begin{cases} 0, & \mathbf{x}\text{与}\mathbf{y}\text{至少有两个分量不同}, \\ q_{\mathbf{x},\mathbf{y}}, & \mathbf{x}\text{与}\mathbf{y}\text{只有一个分量不同}, \\ 1 - \sum_{\mathbf{x} \neq \mathbf{u}} p_{\mathbf{x},\mathbf{u}}, & \mathbf{x} = \mathbf{y}, \end{cases}$$

其中$\sum_{i=1}^n a_i = 1$(通常可将a_i取成$1/n$). 容易验证以$(p_{\mathbf{x},\mathbf{y}})$为转移概率的马氏链(记作$Y$)是不可约, 非周期的. 而且分布$\pi$就是该马氏链的可逆分布, 从而也是平稳分布. 因此当模拟马氏链Y足够多的步骤后, 即k足够大后, Y_k的值就可以看作分布π的一个取样点.

一般而言, 利用如上构造的转移概率矩阵$(p_{\mathbf{x},\mathbf{y}})$模拟生成非周期不可约可逆

马氏链Y, 并利用Y 的样本轨道对该马氏链的可逆平稳分布π进行采样的方法就是MCMC模拟中Gibbs采样法. 一种具体的生成马氏链Y的算法如下.

(s1) 假设目前的状态是$Y_k = \mathbf{x} = (x_1, x_2, \cdots, x_n)$;

(s2) 随机地在1到n中选取一个下标, 比如是i;

(s3) 固定所有$x_j, j \neq i$, 的值, 按条件分布$P(X_i = \cdot | X_j = x_j, j \neq i)$ 生成X_i的随机数, 比如$X_i = y$;

(s4) 将Y_{k+1}的值取成$\mathbf{y} = (x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_n)$并回到(s1).

♠**注记4.1.9.** *若在Hasting-Metropolis 算法中构建预选矩阵$\mathbf{Q} = (p_{\mathbf{x},\mathbf{y}})$, 那么算法中的函数*

$$\alpha_{\mathbf{x},\mathbf{y}} = \min\left\{\frac{\pi_{\mathbf{y}} p_{\mathbf{y},\mathbf{x}}}{\pi_{\mathbf{x}} p_{\mathbf{x},\mathbf{y}}}, 1\right\} = 1.$$

这表明Gibbs采样可以看作特殊预选矩阵下的Metropolis采样.

▶**例4.1.10.** *已知离散随机向量$X = (X_1, \cdots, X_{50})$的每一个分量的可能取值都是1和$-1$. 对$X$的每一个可能取值$\mathbf{x} = (x_1, x_2, \cdots, x_{50})$,*

$$P(X = \mathbf{x}) = c e^{-\frac{1}{2}\sum_{i=1}^{49} x_i x_{i+1}}.$$

试用Gibbs抽样法给出X分布的一个近似取样.

算法设计: 基本流程如下:

(s1) 给X_0任意赋予一个初值$(x_1^{(0)}, x_2^{(0)}, \cdots, x_{50}^{(0)})$;

(s2) 假设目前的状态是$Y_k = \mathbf{x} = (x_1, x_2, \cdots, x_{50})$;

(s3) 随机地在1到50中选取一个下标, 比如是i;

(s4) 固定所有$x_j, j \neq i$, 的值, 按条件分布$P(X_i = \cdot | X_j = x_j, j \neq i)$ 生成x_i的随机数, 比如$x_i = y$;

(s5) 将Y_{k+1}的值取成$\mathbf{y} = (x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_{50})$ 并回到(s2).

具体的程序编写, 请读者自己完成.

最后我们指出的是MCMC方法应用非常广泛, 不仅可以用来近似模拟离散分布的随机数, 还可以用来模拟连续分布的随机数; 不仅可以模拟随机数还可以结合Bayes方法估计参数以及求复杂样本空间上函数的极值, 等等. 对MCMC方法的系统介绍超出了本书范畴, 从略.

练习题4.1

1 以下转移概率矩阵中哪个矩阵对应的马氏链是可逆马氏链？请给出理由.

$$(A)\begin{pmatrix} 0.7 & 0.3 & 0 & 0 \\ 0.2 & 0.5 & 0.3 & 0 \\ 0 & 0.2 & 0.6 & 0.2 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}, \quad (B)\begin{pmatrix} 0.7 & 0.1 & 0.2 & 0 \\ 0.2 & 0.6 & 0.2 & 0 \\ 0.1 & 0.2 & 0.6 & 0.1 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}, \quad (C)\begin{pmatrix} 0.7 & 0.1 & 0.2 & 0 \\ 0.2 & 0.6 & 0.2 & 0 \\ 0 & 0.2 & 0.6 & 0.2 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}.$$

2 总共m个白球和m个黑球分别放在A,B两个坛子中, 每个坛子有m个球, 每次从这两个坛子中各随机地取一个球并把它们交换后放回. 以X_n表示经过n次后A坛中黑球个数. 求X_n的平稳分布.

3 证明当$q > p$时带一个反射壁的简单随机游动是可逆的, 并由此求$m_{0,0}$.

4 若X的转移概率矩阵为$(p_{i,j})$, $\pi = (\pi_i, i \in S)$是X的可逆分布. 证明对任意$n \geq 1$, $\pi_i p_{i,j}^{(n)} = \pi_j p_{j,i}^{(n)}$.

5 设不可约马氏链X的转移概率矩阵$\mathbf{P}$是双随机的, 即每行元素之和为1, 每列元素之和也为1. 证明(1) X是正常返的当且仅当X是有限状态马氏链; (2) X是可逆的当且仅当$\mathbf{P}$是有限阶的对称矩阵.

6 若对任意$i, j \in S$, $p_{i,j} > 0$, 那么以$(p_{i,j})_{i,j \in S}$为转移概率矩阵的正常返马氏链是可逆的当且仅当对任意的$i, j, k \in S$, $p_{i,j}p_{j,k}p_{k,i} = p_{i,k}p_{k,j}p_{j,i}$.

7 若用Hastings-Metropolis算法模拟离散分布$\pi = (\pi_i, 1 \leq i \leq N)$, 其中$\pi_i = 2i/[N(N+1)]$. 假设预选矩阵$\mathbf{Q} = (q_{i,j})_{1 \leq i,j \leq N}$, 其中

$$q_{i,i+1} = q_{i+1,i} = 1/2, \quad i = 1, \cdots, N-1,$$

而且$q_{1,1} = q_{N,N} = 1/2$. 试写出Hastings-Metropolis算法中的转移概率矩阵$\mathbf{P}$.

8 设X是以可逆分布μ为初始分布的可逆马氏链, 转移概率矩阵记作$\mathbf{P} = (p_{i,j})$. 对任意给定的$N > 1$以及任意$0 \leq n \leq N$, 定义$Y_n = X_{N-n}$. 证明$Y = \{Y_n; 0 \leq n \leq N\}$也是转移概率矩阵为$\mathbf{P}$, 初始分布为$\mu$ 的马氏链.

9 试分别写出例4.1.8和例4.1.10中模拟马氏链的转移概率矩阵.

4.2 隐马氏链

隐马氏链是一类具有广泛应用背景的随机模型. 本节我们简要介绍离散时间离散状态隐马尔可夫模型的有关概念和计算.

4.2.1 隐马氏链及其性质

在研究随机动态系统时人们常会发现所关心的随机状态是不可直接测量的, 能观测到的只是与状态有关的某种随机信号. 比如

▶**例4.2.1.** 评估正在生产的流水线是否正常工作时, 绝大部分情况下, 检测人员都不能直接观测到流水线是否正常, 能观察到的往往只是流水线生产的产品的质量. 显然产品的质量既受到机器状态的影响, 也会有其他干扰因素.

如何对这样的现象建模并帮助人们解决相关问题？一个简单的模型就是所谓的隐马尔可夫链.

♣**定义4.2.2.** 设$X_n, Y_n, n \geq 0$, 是分别取值于集合S和W上的离散型随机变量. 对所有$n \geq 0$, 令$Z_n = (X_n, Y_n)$. 若存在随机矩阵$\mathbf{P} = (p_{i,j})_{i,j\in S}$ 和$\mathbf{Q} = (q_{k,l})_{k\in S, l\in W}$ 使得对任意$n \geq 0$

$$P(Z_{n+1} = (i_{n+1}, j_{n+1})|Z_k = (i_k, j_k), 0 \leq k \leq n) = p_{i_n, i_{n+1}} q_{i_{n+1}, j_{n+1}}, \quad (4.2.1)$$

对任意$i_0, i_1, \cdots, i_{n+1} \in S$, $j_0, j_1, \cdots, j_{n+1} \in W$都成立, 并且对任意$i \in S$, $j \in W$

$$P(Y_0 = j|X_0 = i) = q_{i,j}, \quad (4.2.2)$$

那么我们称二维随机过程$Z = \{Z_n; n \geq 0\}$为隐马尔可夫链, 简称为隐马氏链.

在实际应用时, 模型Z 中$X = \{X_n; n \geq 0\}$表示隐含的、不可观测的随机状态. 而$Y = \{Y_n; n \geq 1\}$表示可观测、可记录的随机信号. 通常我们称X为状态序列, 称Y为观测序列.

由定义易知, 隐马氏链是一个向量值的时齐马氏链, 其转移概率为

$$p_{(i_n, j_n),(i_{n+1}, j_{n+1})} = p_{i_n, i_{n+1}} q_{i_{n+1}, j_{n+1}}.$$

事实上, 由全概率公式和(4.2.1)可知

$$\begin{aligned}
&P(Z_{n+1} = (i_{n+1}, j_{n+1})|Z_n = (i_n, j_n)) \\
&\quad = \frac{P(Z_{n+1} = (i_{n+1}, j_{n+1}), Z_n = (i_n, j_n))}{P(Z_n = (i_n, j_n))} \\
&\quad = \sum_{\substack{(i_k, j_k) \\ 0 \leq k < n}} \frac{P(Z_{n+1} = (i_{n+1}, j_{n+1}), Z_n = (i_n, j_n), Z_k = (i_k, j_k), 0 \leq k < n)}{P(Z_n = (i_n, j_n))}
\end{aligned}$$

$$= \sum_{\substack{(i_k,j_k)\\0\le k<n}} \frac{P(Z_k=(i_k,j_k),0\le k\le n+1)}{P(Z_k=(i_k,j_k),0\le k\le n)} \frac{P(Z_k=(i_k,j_k),0\le k\le n)}{P(Z_n=(i_n,j_n))}$$
$$= p_{i_n,i_{n+1}} q_{i_{n+1},j_{n+1}} \sum_{\substack{(i_k,j_k)\\0\le k<n}} \frac{P(Z_k=(i_k,j_k),0\le k\le n)}{P(Z_n=(i_n,j_n))}$$
$$= p_{i_n,i_{n+1}} q_{i_{n+1},j_{n+1}}.$$

因此

$$\begin{aligned} &P(Z_{n+1}=(i_{n+1},j_{n+1})|Z_k=(i_k,j_k),0\le k\le n)\\ &\quad = P(Z_{n+1}=(i_{n+1},j_{n+1})|Z_n=(i_n,j_n)). \end{aligned}$$

类似地, 由全概率公式以及(4.2.1)还可知

$$P(Z_{n+1}=(i_{n+1},j_{n+1})|X_n=i_n)=p_{i_n,i_{n+1}}q_{i_{n+1},j_{n+1}}.$$

因此

$$\begin{aligned} &P(Z_{n+1}=(i_{n+1},j_{n+1})|Z_k=(i_k,j_k),0\le k\le n)\\ &\quad = P(Z_{n+1}=(i_{n+1},j_{n+1})|X_n=i_n). \end{aligned} \tag{4.2.3}$$

这表明对于隐马氏链Z而言, 只要已知状态序列X在n时刻(现在)的值, Z在时刻n之前“历史”与n之后的“未来”独立.

对于状态序列$X=\{X_n;n\ge 0\}$, 利用全概率公式和(4.2.1), 我们同样可得

$$P(X_{n+1}=i_{n+1}|X_k=i_k,0\le k\le n)=P(X_{n+1}=i_{n+1}|X_n=i_n)=p_{i_n,i_{n+1}}.$$

这表明X是一个以$\mathbf{P}=(p_{i,j})$为转移概率矩阵的时齐马氏链, 因此我们也称X为状态马氏链.

此外, 对观测序列$Y=\{Y_n;n\ge 0\}$以及任意的$n\ge 1$,

$$\begin{aligned} &P(Y_n=j_n|X_n=i_n,X_k=i_k,Y_k=j_k,0\le k\le n-1)\\ &\quad = \frac{P(Y_n=j_nX_n=i_n,X_k=i_k,Y_k=j_k,0\le k\le n-1)}{P(X_n=i_n,X_k=i_k,Y_k=j_k,0\le k\le n-1)}\\ &\quad = \frac{P(X_k=i_k,Y_k=j_k,0\le k\le n)}{\sum_{j_n}P(X_k=i_k,Y_k=j_k,0\le k\le n)}\\ &\quad = \frac{p_{i_{n-1},i_n}q_{i_n,j_n}P(X_k=i_k,Y_k=j_k,0\le k<n)}{\sum_{j_n}p_{i_{n-1},i_n}q_{i_n,j_n}P(X_k=i_k,Y_k=j_k,0\le k<n)}=q_{i_n,j_n}. \end{aligned}$$

类似地, 我们还可得

$$\begin{aligned} &P(Y_n=j_n|X_n=i_n,X_k=i_k,0\le k\le n-1)\\ &\quad = P(Y_n=j_n|X_n=i_n,Y_k=j_k,0\le k\le n-1)\\ &\quad = P(Y_n=j_n|X_n=i_n)=q_{i_n,j_n}. \end{aligned}$$

由这些等式以及条件(4.2.2)可知, 在明确已知状态序列的条件下, $\{Y_n;n\ge 0\}$是

独立的; 它在任何时刻是一个依赖于同期X 取值的一个随机变量, 并且分布列由矩阵$\mathbf{Q}=(q_{k,l})$ 的行向量给出.

需要强调的是, 状态序列虽然构成时齐马氏链, 但在实际问题中是不可观测的, 是被隐含的, 因而真正能观测到的观测序列本身并不一定能表现出相互独立性, 甚至都不具有马氏性. 这也是我们称这种模型为隐马尔可夫链的主要原因.

下面我们再看两个隐马尔可夫链的具体例子.

▶**例4.2.3.** 以$Y=\{Y_n;n\geq 0\}$记录某只股票每日的涨跌幅度. Y是可观测的. 容易设想的是, 股票的涨跌与股票的"牛", "熊"密切相关, 即与股票是否在上升(1)、下跌(−1)或者横盘(0)等状态有关. 以$X=\{X_n;n\geq 0\}$ 表示股票每日所处的状态, 显然X是不可直接观测的. 一个简单的股票价格波动模型将(1)每日股票的涨跌看作是当日股票状态的随机表现, 即给定当日股票状态i之后, 股票涨跌服从某种具有参数i的随机分布, 记该分布列为$(q_{i,l})$; (2) 每日股票的状态过程X 看作是一个(时齐)马氏链, 转移概率矩阵为$(p_{i,j})$. 由此构建的模型$\{(X_n,Y_n);n\geq 0\}$就是一个刻画股票价格波动率的隐马氏链模型.

▶**例4.2.4.** 设X_n是取值于状态空间$S=\{1,2,\cdots,L\}$的马氏链, 其样本不能被实际测量得到, 而能测量到的是如下的Y_n 的样本,

$$Y_n=g(X_n)+w_n,$$

其中g是一个未知函数, w_n, $n\geq 0$, 是独立同分布的随机干扰, 只取有限个值, 且与随机过程$\{X_n;n\geq 0\}$独立. 那么$\{(X_n,Y_n);n\geq 0\}$就是一个隐马氏链模型.

记X_0的初始分布$\pi=(\pi_i,i\in S)$, 对任意$n\geq 0$ 以及$i_0,i_1,\cdots,i_n\in S$, $j_0,j_1,\cdots,j_n\in W$, 由马氏链的有限维分布表示可知

$$P(Z_k=(i_k,j_k),0\leq k\leq n)=\pi_{i_0}q_{i_0,j_0}\prod_{k=1}^{n}p_{i_{k-1},i_k}q_{i_k,j_k}. \tag{4.2.4}$$

这表明初始分布π以及随机矩阵$\mathbf{P}$, $\mathbf{Q}$ 完全确定了Z 的统计特征. 通常称$(\pi,\mathbf{P},\mathbf{Q})$为隐马氏链$Z$的三要素, 称$\pi$为状态马氏链的初始分布或简称为初始分布, 称$\mathbf{P}$为状态马氏链的转移概率矩阵或简称为转移概率矩阵, 而称$\mathbf{Q}$为观测概率矩阵.

4.2.2 观察结果出现的概率

对于隐马氏链模型$\{(X_n,Y_n);n\geq 0\}$, 当它的三要素$(\pi,\mathbf{P},\mathbf{Q})$ 给定后, 人们关心的第一个问题就是各种观测结果出现的概率. 在应用中也常把这类概率的计算问题称为是评估问题.

对于任意给定的$j_0,j_1,\cdots,j_n\in W$. 为了计算概率

$$P(Y_k=j_k,0\leq k\leq n),$$

我们对任意$0 \le m \le n$, 令

$$F_m(i) = P(X_m = i, Y_k = j_k, 0 \le k \le m), \qquad i \in S.$$

那么由(4.2.2), (4.2.3)和全概率公式可知

$$\begin{aligned} F_0(i) &= \pi_i q_{i,j_0}, \\ F_m(i) &= \sum_{i_{m-1} \in S} P(X_{m-1} = i_{m-1}, X_m = i, Y_k = j_k, 0 \le k \le m) \\ &= \sum_{i_{m-1} \in S} \Big[P(X_m = i, Y_m = j_m | X_{m-1} = i_{m-1}, Y_k = j_k, 0 \le k \le m-1) \\ &\qquad \times P(X_{m-1} = i_{m-1}, Y_k = j_k, 0 \le k \le m-1) \Big] \\ &= q_{i,j_m} \sum_{i_{m-1} \in S} p_{i_{m-1},i} F_{m-1}(i_{m-1}), \qquad m \ge 1. \end{aligned} \tag{4.2.5}$$

对任意$0 \le m < n$, 令

$$B_m(i) = P(Y_n = j_n, \cdots, Y_{m+1} = j_{m+1} | X_m = i),$$

类似地, 再由(4.2.3), 条件概率公式和全概率公式可得

$$B_m(i) = \sum_{i_{m+1} \in S} p_{i,i_{m+1}} q_{i_{m+1},j_{m+1}} B_{m+1}(i_{m+1}), \qquad m \le n-1. \tag{4.2.6}$$

其中约定$B_n(i) \equiv 1$.

基于公式(4.2.5)和(4.2.6), 我们可以分别通过向前和向后迭代得到所有的$F_m(i)$和$B_m(i)$, 其中$0 \le m \le n$. 事实上, 由$F_m(i)$和$B_m(i)$的定义可得

$$P(Y_k = j_k, 0 \le k \le n) = \sum_{i \in S} P(X_n = i, Y_k = j_k, 0 \le k \le n) = \sum_{i \in S} F_n(i), \tag{4.2.7}$$

以及

$$\begin{aligned} P(Y_k = j_k, 0 \le k \le n) &= \sum_{i \in S} P(X_0 = i, Y_k = j_k, 0 \le k \le n) \\ &= \sum_{i \in S} P(Y_k = j_k, 1 \le k \le n | X_0 = i, Y_0 = j_0) P(X_0 = i, Y_0 = j_0) \\ &= \sum_{i \in S} P(Y_k = j_k, 1 \le k \le n | X_0 = i) \pi_i q_{i,j_0} = \sum_{i \in S} B_0(i). \end{aligned} \tag{4.2.8}$$

通常我们分别称公式(4.2.7)和(4.2.8)为计算观察出现的向前和向后算法, 或者称为评估问题的向前和向后算法.

当然我们也可以通过同时计算$F_m(i)$和$B_m(i)$得到评估问题的混合计算方法. 对任意$0 \le m \le n$, 由马氏性(参见本节习题1),

$$\begin{aligned} P(Y_k = j_k, 0 \le k \le n) &= \sum_{i \in S} P(X_m = i, Y_k = j_k, 0 \le k \le n) \\ &= \sum_{i \in S} P(X_m = i, Y_k = j_k, 0 \le k \le m) \end{aligned}$$

$$\times P(Y_k=j_k, m<k\le n|X_m=i, Y_k=j_k, 0\le k\le m)$$
$$=\sum_{i\in S}F_m(i)P(Y_k=j_k, m<k\le n|X_m=i)=\sum_{i\in S}F_m(i)B_m(i). \quad (4.2.9)$$

在评估问题的基础上, 我们还可以计算未来观测值出现的概率, 仅以考虑未来一步的观测值为例, 更多步的情形请读者自己写出.

$$\begin{aligned}&P(Y_{n+1}=j|Y_k=j_k, 0\le k\le n)\\&\quad=\sum_{i,i_{n+1}\in S}P(Y_{n+1}=j, X_{n+1}=i_{n+1}, X_n=i|Y_k=j_k, 0\le k\le n)\\&\quad=\sum_{i,i_{n+1}\in S}\frac{p_{i,i_{n+1}}q_{i_{n+1},j}F_n(i)}{\sum_{i\in S}F_n(i)}.\end{aligned} \quad (4.2.10)$$

►**例4.2.5.** 假设例*4.2.3*中股票价格波动率分$-1,0,1$三个层次, 在股票状态只有$-1,1$两种情形. 在股票状态为1 时, 股票价格波动率为$-1,0,1$ 的概率分别是$0.2,0.3,0.5$, 而在股票状态为-1时, 股票价格波动率分$-1,0,1$ 的概率分别为$0.6,0.2,0.2$. 假定股票状态转移概率为

$$p_{-1,1}=1/4,\ p_{-1,-1}=3/4,\ p_{1,1}=4/5,\ p_{1,-1}=1/5.$$

若$P(X_0=1)=4/5$, (1) 问前*3*天的价格波动率恰为$Y_0=-1,Y_1=0,Y_2=1$的概率, (2) 若前三天价格波动恰为$Y_0=1,Y_1=0,Y_2=-1$, 求第四天最有可能出现的价格波动率.

解 (1) 已知前3天的波动恰为$Y_0=-1,Y_1=0,Y_2=1$, 因此给定$j_0=-1,j_1=0,j_2=1$, 套用公式(4.2.5), 计算可知

$$F_0(1)=\pi_1q_{1,-1}=0.8\times0.2=0.16,\quad F_0(-1)=\pi_{-1}q_{-1,-1}=0.2\times0.6=0.12,$$
$$F_1(1)=q_{1,0}[p_{-1,1}F_0(-1)+p_{1,1}F_0(1)]=0.0474,$$
$$F_1(-1)=q_{-1,0}[p_{-1,-1}F_0(-1)+p_{1,-1}F_0(1)]=0.0244,$$
$$F_2(1)=q_{1,1}[p_{-1,1}F_1(-1)+p_{1,1}F_1(1)]=0.02201,$$
$$F_2(-1)=q_{-1,1}[p_{-1,-1}F_1(-1)+p_{1,-1}F_1(1)]=0.005556.$$

因此由向前算法(4.2.7), 所求概率

$$P(Y_0=1,Y_1=0,Y_2=-1)=0.02201+0.005556=0.027566.$$

(2) 此时前3天$Y_0=1,Y_1=0,Y_2=-1$, 因此给定$j_0=1,j_1=0,j_2=-1$, 仍然使用向前算法可得

$$F_0(1)=\pi_1q_{1,1}=0.8\times0.5=0.4,\quad F_0(-1)=\pi_{-1}q_{-1,1}=0.2\times0.2=0.04,$$
$$F_1(1)=q_{1,0}[p_{-1,1}F_0(-1)+p_{1,1}F_0(1)]=0.099,$$
$$F_1(-1)=q_{-1,0}[p_{-1,-1}F_0(-1)+p_{1,-1}F_0(1)]=0.022,$$

$$F_2(1) = q_{1,-1}[p_{-1,1}F_1(-1) + p_{1,1}F_1(1)] = 0.01694,$$
$$F_2(-1) = q_{-1,-1}[p_{-1,-1}F_1(-1) + p_{1,-1}F_1(1)] = 0.02178.$$

为了计算方便, 我们令

$$a(1) = p_{1,1}q_{1,1} + p_{1,-1}q_{-1,1} = 0.44,\ a(-1) = p_{-1,1}q_{1,1} + p_{-1,-1}q_{-1,1} = 0.275;$$
$$b(1) = p_{1,1}q_{1,0} + p_{1,-1}q_{-1,0} = 0.28,\ b(-1) = p_{-1,1}q_{1,0} + p_{-1,-1}q_{-1,0} = 0.225;$$
$$c(1) = p_{1,1}q_{1,-1} + p_{1,-1}q_{-1,-1} = 0.28,\ c(-1) = p_{-1,1}q_{1,-1} + p_{-1,-1}q_{-1,-1} = 0.5.$$

由(4.2.10)可得(取$n = 2$)

$$P(Y_3 = 1|Y_0 = 1, Y_1 = 0, Y_2 = -1) = \frac{a(1)F_2(1) + a(-1)F_2(-1)}{F_2(1) + F_2(-1)} \approx 0.347;$$
$$P(Y_3 = 0|Y_0 = 1, Y_1 = 0, Y_2 = -1) = \frac{b(1)F_2(1) + b(-1)F_2(-1)}{F_2(1) + F_2(-1)} \approx 0.249;$$
$$P(Y_3 = -1|Y_0 = 1, Y_1 = 0, Y_2 = -1) = \frac{c(1)F_2(1) + c(-1)F_2(-1)}{F_2(1) + F_2(-1)} \approx 0.404.$$

因此第四个观察日股票价格波动最可能的层次为-1. □

4.2.3 状态估计与预测

对隐Markov 模型, 在应用中我们常碰到这样的问题: 已知模型参数$(\pi, \mathbf{P}, \mathbf{Q})$和一段观测序列$\{Y_k; k \le n\}$条件下, 估计状态马氏链$X_m$, $m \le n$或其中一部分的最合适取值. 应用中常称这类问题为解码问题.下面我们基于极大似然的思想讨论这类问题, 也即我们认为X_m的最合适的值是在给定观察序列条件下X_m以最大概率取到的值(为简便, 本小节我们总假定最大概率可以取到).

给定一段序列$\{Y_k; k \le n\}$的观察值$y_0, \cdots, y_n$. 当我们只估计n之前某一个时刻m的最佳值时, 由(4.2.9)可知

$$P(X_m = i|Y_k = j_k, 0 \le k \le n) = \frac{F_m(i)B_m(i)}{\sum_{i\in S} F_m(i)B_m(i)}.$$

因此X_m的最佳值就是使得$F_m(i)B_m(i)$取值最大的i.

若我们需要估计n及n之前所有时刻的最佳值, 那么由(4.2.1)可知

$$\begin{aligned}
&\max_{i_k\in S, 0\le k\le n} P(X_k = i_k, Y_k = j_k, 0 \le k \le n)\\
&= \max_{i_k\in S, 0\le k\le n} \Big[P(X_n = i_n, Y_n = j_n|X_k = i_k, Y_k = j_k, 0 \le k < n)\\
&\qquad\qquad \times P(X_k = i_k, Y_k = j_k, 0 \le k < n)\Big]\\
&= \max_{i_k\in S, 0\le k\le n} [p_{i_{n-1},i_n} q_{i_n,j_n} P(X_k = i_k, Y_k = j_k, 0 \le k < n)]\\
&= \max_{i_n\in S} \Big\{q_{i_n,j_n} \max_{i_k\in S, 0\le k<n} [p_{i_{n-1},i_n} P(X_k = i_k, Y_k = j_k, 0 \le k < n)]\Big\}.
\end{aligned} \tag{4.2.11}$$

令$M_0(i)=\pi_i$, 其中π是X的初始分布. 对任意$1\le m\le n$, 令

$$M_m(i)=\max_{i_k\in S,0\le k<m}\big[p_{i_{m-1},i}P(X_k=i_k,Y_k=j_k,0\le k\le m-1)\big],$$

那么对任意$m\ge 2$, 再次由(4.2.1)可得

$$\begin{aligned}M_m(i)&=\max_{\substack{i_k\in S\\0\le k<m}}\big[p_{i_{m-1},i}q_{i_{m-1},j_{m-1}}p_{i_{m-2},i_{m-1}}P(X_k=i_k,Y_k=j_k,0\le k\le m-2)\big]\\&=\max_{i_{m-1}\in S}[p_{i_{m-1},i}q_{i_{m-1},j_{m-1}}M_{m-1}(i_{m-1})],\end{aligned}\tag{4.2.12}$$

并且直接验证可知上式对$m=1$也成立.

利用(4.2.12)以及$M_0(i)$, 由向前递归计算可得任意的$M_m(i)$.

为了获得n及n之前所有时刻的X的最佳值, 由(4.2.11)我们可得

$$\max_{i_k\in S,0\le k\le n}P(X_k=i_k,Y_k=j_k,0\le k\le n)=\max_{i_n\in S}[q_{i_n,j_n}M_n(i_n)].$$

这表明X_n的最合适取值

(1) i_n是使函数$q_{i,j_n}M_n(i),i\in S$, 取到最大值的一个状态.

再由(4.2.12)可知, n之前X的最合适取值$i_0,i_1,\cdots,i_{n-1}$可按如下方式给定.

(2) 对任意$1\le m<n$, 给定i_{m+1}条件下, i_m就是使得函数

$$p_{i,i_{m+1}}q_{i,j_m}M_m(i),\qquad i\in S,$$

取最大值的一个状态.

对任意$0\le m<n$以及$u\in S$, 记$I_m(u)\in S$ 使得

$$p_{I_m(u),u}q_{I_m(u),j_m}M_m(I_m(u))=\max_{i\in S}p_{i,u}q_{i,j_m}M_m(i).$$

那么

$$i_{n-1}=I_{n-1}(i_n),i_{n-2}=I_{n-2}(i_{n-1}),\cdots,i_0=I_0(i_1).$$

给定了观测序列的观测值, 上面寻找最可能的状态序列的方法称为Viterbi算法.

▶**例4.2.6.** *(接例4.2.5) 若前三天价格波动恰为$Y_0=1,Y_1=0,Y_2=-1$,问前三个观察日该股票最可能的状态序列是什么?*

解 给定$j_0=1,j_1=0,j_2=-1$. 由假设可知$M_0(1)=0.8,M_0(-1)=0.2$,

$$M_1(1)=\max_{i\in\{-1,1\}}p_{i,1}q_{i,j_0}M_0(i)=\max_{i\in\{-1,1\}}p_{i,1}q_{i,1}M_0(i)=0.32,\quad I_0(1)=1$$

类似可知

$$\begin{aligned}M_1(-1)&=\max_{i\in\{-1,1\}}p_{i,-1}q_{i,j_0}M_0(i)=0.08,\quad I_0(-1)=1.\\M_2(1)&=\max_{i\in\{-1,1\}}p_{i,1}q_{i,j_1}M_1(i)=0.0768,\quad I_1(1)=1.\\M_2(-1)&=\max_{i\in\{-1,1\}}p_{i,-1}q_{i,j_1}M_1(i)=0.0192,\quad I_1(-1)=1.\end{aligned}$$

由于

$$\max_{i\in\{-1,1\}}q_{i,j_2}M_2(i)=q_{1,-1}M_2(1)=0.01536.$$

因此$I_2=1$进而前三个观察日该股票最可能的状态序列是

$$i_2=1, i_1=I_1(1)=1, i_0=I_0(1)=1. \qquad \square$$

4.2.4 隐马氏链的参数估计*

在前面的讨论中我们都假定隐马氏链的三要素是已知的. 但在应用时我们需要从一段观测序列$\{Y_k;k\le n\}$出发估计这些参数. 这类问题又称为学习问题, 属于参数估计范畴.

若假设$\{X_k;k\ge 0\}$的样本也已知, 此时作为马氏链$Z=\{Z_k;k\ge 0\}$是完全可观测的. 那么只要有充分长的样本, 对任意$i,j\in S$, 把X中从状态i到下一个时刻为状态j的频数记为$A_{i,j}$, 把状态i出现的频数记作A_i, 由定理3.2.26以及大数定律可以证明, 只要i常返就可以用$\hat{p}_{i,j}=A_{i,j}/A_i$ 作为$p_{i,j}$的相合估计, 同样若以$B_{i,l}$表示X取值i的同时Y取值l的频数, 则类似可证明$\hat{q}_{i,l}=B_{i,l}/A_i$ 是$q_{i,l}$相合估计. 然而这种方法只是"理想"的, 因为实际模型中X常不可测量.

下面我们仅给定Y的一组观测值$l_0,l_1,\cdots,l_N$. 极大似然原理告诉我们θ的极大似然估计$\hat{\theta}$满足

$$P(Y_k=l_k,0\le k\le N|\hat{\theta})=\max_{\theta}P(Y_k=l_k,0\le k\le N|\theta),$$

其中$P(\cdot|\theta)$表示给定参数θ下的概率, 这里θ指代的就是三要素$(\pi,\mathbf{P},\mathbf{Q})$.

一般而言, 直接求解该问题非常困难. 一种折中方案是通过构造一个关于参数θ的递推算法, 使之能逐步提高条件概率$P(Y_k=l_k,0\le k\le N|\theta)$ 的取值, 并将递推算法的最终结果作为参数θ的估计. 算法的具体步骤如下:

(s1) 给定参数$\theta_n=(\pi(n),\mathbf{P}(n)=(p_{i,j}(n))_{i,j\in S},\mathbf{Q}(n)=(q_{i,l}(n))_{i\in S,l\in W})$;

(s2) 对任意$\theta=(\pi,\mathbf{P}=(p_{i,j})_{i,j\in S},\mathbf{Q}=(q_{i,l})_{i\in S,l\in W})$, 令

$$\begin{aligned}T(\theta|\theta_n)=\sum_{i_k\in S,0\le k\le N}\Big[&P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)\\&\times\ln(P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta))\Big].\end{aligned}$$

(s3) 求θ^*使得$T(\theta^*|\theta_n)=\max_\theta T(\theta|\theta_n)$.

(s4) 令$\theta_{n+1}=\theta^*$ 并回到(s1).

按上述算法,

$$\begin{aligned}0&\le T(\theta_{n+1}|\theta_n)-T(\theta_n|\theta_n)\\&=\sum_{i_k\in S,0\le k\le N}\Big[P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)\\&\qquad\times\ln\Big(\frac{P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_{n+1})}{P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)}\Big)\Big].\end{aligned}$$

因为对任意$x \in (0,\infty)$, $\ln x \le x-1$, 由上式可得

$$P(Y_k=l_k,0\le k\le N|\theta_{n+1})=\sum_{i_k\in S,0\le k\le N}P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_{n+1})$$
$$\ge\sum_{i_k\in S,0\le k\le N}P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)=P(Y_k=l_k,0\le k\le N|\theta_n).$$

这表明上述算法得到的参数θ_n确实能使条件概率$P(Y_k=l_k,0\le k\le N|\theta_n)$随着算法次数的增加而优化(增加).

根据上述算法(s2)步中的定义, $T(\theta|\theta_n)$等于

$$\sum_{\substack{i_k\in S\\0\le k\le N}}\Big[P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)\ln(\pi_{i_0}q_{i_0,l_0}\prod_{n=1}^{N}p_{i_{n-1},i_n}q_{i_n,l_n})\Big]$$
$$=\sum_{i_k\in S,0\le k\le N}\Big[P(X_k=i_k,Y_k=l_k,0\le k\le N|\theta_n)$$
$$\times(\ln\pi_{i_0}+\sum_{n=0}^{N}\ln q_{i_n,l_n}+\sum_{n=1}^{N}\ln p_{i_{n-1},i_n})\Big]$$
$$=\sum_{i_0\in S}P(X_0=i_0,Y_k=l_k,0\le k\le N|\theta_n)\ln\pi_{i_0}$$
$$+\sum_{i_n\in S}\sum_{n=0}^{N}P(X_n=i_n,Y_k=l_k,0\le k\le N|\theta_n)\ln q_{i_n,l_n}$$
$$+\sum_{i_{n-1},i_n\in S}\sum_{n=1}^{N}P(X_{n-1}=i_{n-1},X_n=i_n,Y_k=l_k,0\le k\le N|\theta_n)\ln p_{i_{n-1},i_n}.$$

或重写为

$$T(\theta|\theta_n)=\sum_{i\in S}P(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)\ln\pi_i$$
$$+\sum_{\substack{i\in S\\l\in W}}\sum_{n=0}^{N}P(X_n=i,Y_k=l_k,0\le k\le N|\theta_n)1_{\{Y_n=l\}}\ln q_{i,l}$$
$$+\sum_{i,j\in S}\sum_{n=1}^{N}P(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)\ln p_{i,j}.$$

由本节习题5可知, 当$\theta=(\pi^*,\mathbf{P}^*,\mathbf{Q}^*)$时$T(\theta|\theta_n)$取到最大值, 其中$\pi^*=(\pi_i^*)$, $\mathbf{P}^*=(p_{i,j}^*)$, $\mathbf{Q}^*=(q_{i,l}^*)$满足: 对任意$i,j\in S,l\in W$,

$$\pi_i^*=\frac{P(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)}{\sum_{i\in S}P(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)},\tag{4.2.13}$$

$$p_{i,j}^*=\frac{\sum_{n=1}^{N}P(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)}{\sum_{i,j\in S}\sum_{n=1}^{N}P(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)},\tag{4.2.14}$$

$$q_{i,l}^* = \frac{P(X_n = i, Y_k = l_k, 0 \le k \le N|\theta_n)1_{\{Y_n=l\}}}{\sum_{\substack{i\in S\\ l\in W}} \sum_{n=0}^{N} P(X_n = i, Y_k = l_k, 0 \le k \le N|\theta_n)1_{\{Y_n=l\}}}. \qquad (4.2.15)$$

令$\theta_{n+1} = (\pi^*, \mathbf{P}^*, \mathbf{Q}^*)$, 我们得到了参数$\theta_{n+1}$与$\theta_n$ 之间的递推公式.

一般称(4.2.13)-(4.2.15)为Welch–Baum 公式. 上面的算法被称为EM 算法. 其中E和M分别指代步骤(s2)和(s3), 求数学期望与求最大值. EM算法是针对在测量数据不完全时求参数的一种近似于极大似然估计的统计方法. 隐马氏链参数的估计, 是EM 算法的一种典型运用.

评估问题、解码问题以及学习问题是隐马氏链应用中三个最基本的问题, 随着应用的不同, 人们还会关心许多其他问题, 比如对于一个特定的观测链$\{Y_k, k \le n\}$, 已知它可能是由已经学习好的若干模型之一所得的观测, 要决定此观测究竟是得自其中哪一个模型. 这类问题也称为识别问题. 对于这些问题, 超出了本书的范畴, 我们就不再介绍了.

练习题4.2

1 设$\{(X_n, Y_n);\ n \ge 0\}$是隐马氏链, 其中$\{X_n; n \ge 0\}$. $\{Y_n; n \ge 0\}$分别是状态序列和观察序列, 证明对任意$n \ge 0$, $m \ge 1$,

$$\begin{aligned} &P(Y_{n+k} = j_{n+k}, 1 \le k \le m|X_n = i, Y_l = j_l, 0 \le l \le n) \\ &= P(Y_{n+k} = j_{n+k}, 1 \le k \le m|X_n = i). \end{aligned}$$

2 假设例4.2.1中产品的质量分为1(合格)和0(不合格). 在机器状态为1 时, 产品质量为$0, 1$ 的概率分别为$0.05, 0.95$, 而在机器状态为2 时, 产品质量为$0, 1$ 的概率分别为$0.5, 0.5$. 假设在产品检测的间隔时间内机器状态转移概率为

$$p_{1,1} = 0.9,\ p_{1,2} = 0.1,\ p_{2,1} = 0,\ p_{2,2} = 1.$$

若初始机器状态为1, 求第4次检测才首次检测到不合格产品的概率并在此条件下计算第5次抽到合格品的概率.

3 仍考虑上面习题2的模型, 并仍假初始机器状态为1. 若规定从$n = 1$开始连续检测到两次不合格品时就对机器检修, 试证明机器检修决策正确(即检修时发现机器状态为2)的概率不小于1090/1171.

4 在例4.2.5设定下, 若前3天价格波动恰为$Y_0 = 1, Y_1 = 0, Y_2 = -1$, 求$X_2$的分布, 并预测$X_3$最可能的取值.

5* 对任意$i \in T$, 设$z_i > 0$而且$\sum_{i\in T} z_i < \infty$. 证明

$$\sum_{i\in T} z_i \ln\left(\frac{z_i}{\sum_{i\in T} z_i}\right) = \sup_{\substack{x_i>0, i\in T\\ \sum_i x_i=1}} \sum_{i\in T} z_i \ln x_i.$$

6* 试寻找一些具体问题, 用隐马氏链建模并做出分析和预测.

4.3 分枝过程

这一节我们介绍一类特殊的过程——分枝过程. 这类过程虽然仍是一种马氏链, 但作为一类非常基础的随机模型, 人们已经为它建立了非常丰富的理论并将它广泛应用于自然科学和社会科学的各个领域.

4.3.1 Galton-Watson分枝过程

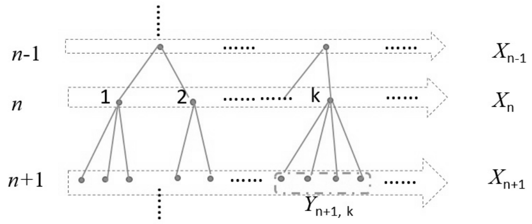
最经典的一类分枝过程是所谓Galton-Watson分枝过程(Branching Process). 这类过程源于19世纪Francis Galton爵士对英国家族姓氏消亡的统计调查. 模型的定义如下:

♣**定义4.3.1.** 设$\{Y_{n,k}; n \geq 1, k \geq 1\}$为一族独立同分布的非负整数值的随机变量, X_0是给定的非负整数值随机变量且与$\{Y_{n,k}; n \geq 1, k \geq 1\}$ 独立. 对任意$n \geq 0$, 递归地定义随机变量X_{n+1} 如下

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n+1,k},$$

其中约定$\sum_{k=1}^{0} \cdot = 0$. 称随机过程$X = \{X_n; n \geq 0\}$为*Galton-Watson*分枝过程, 简记为*G-W*过程.

直观地看, 该模型描述了如下图所示的分枝现象.



其中对任意的$n \geq 0$, X_n表示第n层粒子总数, $Y_{n+1,k}$则表示第n层中第k个粒子所分枝出来的粒子数量. 通常我们也会把第n层称为第n代, 把由某个粒子分枝出来的所有粒子统称为是该粒子的后代. 要强调的是, G-W 分枝过程理想地假设了每个粒子产生后代的行为是独立同分布的. 这为我们简化讨论带来了方便.

容易证明(参见本节习题1)G-W过程X是一个时齐马氏链, 转移概率

$$p_{i,j} = P(Y_{1,1} + \cdots + Y_{1,i} = j).$$

而且由模型定义可知0为吸收态. 即若某代粒子数为0, 那么此后每一代的粒子数都为零.

记$A=\cup_{n=1}^{\infty}\{X_n=0\}$, 通常我们称随机事件$A$为G-W过程的灭绝事件. 若G-W过程刻画的是某种生物种群数量的代系变化的话, 那么当A发生时, 意味着在某代之后, 再也没有该种群的新生代了, 从而该种群必将消亡. 对G-W过程而言, 人们关心的一些基本问题包括: *灭绝事件是否发生? 什么时候发生? 发生的概率多大?* 为了探讨这些问题, 我们记独立同分布随机变量$\{Y_{n,k};n\geq 1,k\geq 1\}$的共同分布为$\eta$, 并称其为后代分布. 对任意$i\geq 0$,

$$\eta_i=P(Y_{n,k}=i).$$

为避免平凡, 以下总设$\eta_1<1$. 令

$$\tau_0=\inf\{n\geq 1,X_n=0\},$$

称其为X的灭绝时间. 显然$A=\{\tau_0<+\infty\}$. 记$P_1(A)$为q, 那么q表示初始时只有一个粒子时灭绝事件发生的概率,

$$q=P_1(\tau_0<+\infty)=f_{10}=\lim_{n\to\infty}p_{1,0}^{(n)}.$$

注意到若初始时刻有k个粒子, 要使灭绝事件发生, 意味着每个粒子及其后代都得消亡, 又由于每个粒子及其后代的表现是相互独立的, 因此

$$f_{k,0}=P_k(\tau_0<+\infty)=f_{10}^k=q^k,\qquad p_{k,0}^{(n)}=P_k(X_n=0)=(p_{1,0}^{(n)})^k.$$

以m表示每个粒子分枝出后代的平均数量, 即

$$m=\sum_{i=1}^{\infty}i\eta_i=E(Y_{n,k}).$$

以$\eta(s)$表示分布η的概率母函数, 即对任意$s\in[0,1]$,

$$\eta(s)=\sum_{i=0}^{\infty}s^i\eta_i=E(s^{Y_{n,k}}).$$

那么$\eta'(1)=m$.

对于分枝过程的灭绝事件, 我们有如下结果.

▲**命题4.3.2.** *若$m\leq 1$, 那么X 必然消亡, 即$q=1$.*

证明 由定理3.3.1可知

$$q=f_{1,0}=\sum_{k=1}^{\infty}p_{1,k}f_{k,0}+p_{1,0}=\sum_{k=1}^{\infty}\eta_kq^k+\eta_0.$$

那么q满足方程$q=\eta(q)$. 由于$0\leq q\leq 1$, 而由第一章第三节习题6可知该方程在$[0,1)$内无解. 直接验算可知$q=1$. □

▲**命题4.3.3.** *若$m>1$, 那么X以正概率不消亡, 即$q<1$.*

证明 在命题4.3.2证明中我们已指出q是$\eta(s)=s$在区间$[0,1]$上一个解. 注意

到0是吸收状态, 对任意$n \geq 2$

$$p_{1,0}^{(n)} = \sum_{k=1}^{\infty} p_{1,k} p_{k,0}^{(n-1)} + p_{1,0} = \sum_{k=1}^{\infty} p_k (p_{1,0}^{(n-1)})^k + p_0 = \eta(p_{1,0}^{(n-1)}),$$

其中$p_{1,0}^{(1)} = p_0$. 对任意$n \geq 1$, 记

$$\eta^{(n)}(s) = \underbrace{\eta(\eta(\cdots(\eta}_{n}(s))))$$

为η的n次复合. 对方程$s = \eta(s)$的任意非负解u以及任意n, 都满足

$$\begin{aligned} u &= \eta^{(n)}(u) = \eta^{(n-1)}(\eta(u)) \\ &\geq \eta^{(n-1)}(\eta(p_0)) = \eta^{(n-1)}(p_{1,0}^{(2)}) \\ &\geq \cdots \geq \eta(p_{1,0}^{(n)}) = p_{1,0}^{(n+1)}. \end{aligned}$$

令$n \to \infty$, 得$q = f_{1,0} \leq u$. 这表明q是$s = \eta(s)$的最小非负解. 因此, 由第一章第三节习题6可知当$m > 1$时, $\eta(s) = s$在$[0,1)$内有唯一非负解, 从而$q < 1$. □

下面我们通过一个例子来说明如何利用命题4.3.2和命题4.3.3计算灭绝概率.

▶**例4.3.4.** 设G-W过程X的后代服从概率母函数为

$$\eta(s) = a + bs + cs^2$$

的分布, 其中a, b, c非负, 而且$a + b + c = 1$. 求$X_0 = 1$条件下X的灭绝概率q.

解 直接计算可知, 当$c \leq a$时, 平均后代数

$$m = \eta'(1) = b + 2c \leq 1.$$

由命题4.3.2可知$q = 1$. 当$a < c$时, $m > 1$. 由命题4.3.3的证明可知q是方程

$$s = a + bs + cs^2$$

的最小非负解. 将$1 = a + b + c$代入得

$$cq(1-q) = a(1-q).$$

这表明$q = a/c$. □

利用后代分布的概率母函数$\eta(s)$以及它的复合函数$\eta^{(n)}(s)$, 我们还能非常方便地写出分枝过程X在任意一个时刻k的概率母函数. 我们有如下的结果.

▲**命题4.3.5.** 记分枝过程X的后代分布概率母函数为$\eta(s)$, 那么对任意$k \geq 1$, 在$X_0 = 1$条件下, X_k的概率母函数为$\eta^{(k)}(s)$.

证明 记$X_0 = 1$条件下, X_k的概率母函数为$g_k(s)$. 显然$k = 1$时,

$$g_1(s) = E_1(s^{X_1}) = E(s^{Y_{1,1}} | X_0 = 1) = E(s^{Y_{1,1}}) = \eta(s).$$

那么对$k > 1$,

$$g_k(s) = E_1(s^{X_k}) = E_1(E_1(s^{\sum_{i=1}^{X_{k-1}} Y_{k,i}} | X_{k-1}))$$

$$
\begin{aligned}
&= \sum_{l=0}^{\infty} E_1(s^{\sum_{i=1}^{X_{k-1}} Y_{k,i}} | X_{k-1} = l) P_1(X_{k-1} = l) \\
&= \sum_{l=0}^{\infty} E_1(s^{\sum_{i=1}^{l} Y_{k,i}} | X_{k-1} = l) P_1(X_{k-1} = l) \\
&= \sum_{l=0}^{\infty} E(s^{\sum_{i=1}^{l} Y_{k,i}} | X_{k-1} = l, X_0 = 1) P_1(X_{k-1} = l) \\
&= \sum_{l=0}^{\infty} E(s^{\sum_{i=1}^{l} Y_{k,i}}) P_1(X_{k-1} = l) \\
&= \sum_{l=0}^{\infty} [\eta(s)]^l P_1(X_{k-1} = l) = g_{k-1}(\eta(s)).
\end{aligned}
$$

由此关系式迭代可得

$$g_k(s) = g_{k-1}(\eta(s)) = g_{k-2}(\eta(\eta(s))) = \cdots = g_1(\underbrace{\eta(\eta \cdots (\eta}_{k-1}(s)))) = \eta^{(k)}(s).$$

因此对任意$k \geq 1$命题成立. □

注意到0是吸收状态, 对任意$k \geq 1$, $P_1(\tau_0 \leq k) = p_{1,0}^{(k)}$, 由命题4.3.5立即可得

$$P_1(\tau_0 \leq k) = \eta^{(k)}(0),$$

进而, 对任意$k \geq 1$,

$$P_1(\tau_0 = k) = \eta^{(k)}(0) - \eta^{(k-1)}(0). \tag{4.3.1}$$

上式给出了灭绝时间分布的表达式.

▶**例4.3.6.** 设分枝过程X的后代分布概率母函数为$\eta(s) = \frac{p}{1-(1-p)s}$, 即对任意$k \geq 0$, $\eta_k = p(1-p)^k$, 其中$0 < p < 1$. 对任意$n \geq 1$, 求$P_1(\tau_0 = n)$.

解 记分布η的数学期望$(1-p)/p$为α. 注意到$\eta(s) = s$有两个解$s_1 = 1$, $s_2 = 1/\alpha$, 当$p \neq 1/2$ 即$\alpha \neq 1$时,

$$\frac{\eta(s) - \eta(s_1)}{\eta(s) - \eta(s_2)} = \frac{\eta(s) - 1}{\eta(s) - \frac{1}{\alpha}} = \alpha \frac{s-1}{s - \frac{1}{\alpha}}.$$

对任意$s \in [0, 1/(1-p))$以及$s \neq 1/\alpha$成立. 由此可得

$$\frac{\eta^{(n)}(s) - 1}{\eta^{(n)}(s) - \frac{1}{\alpha}} = \alpha^n \frac{s-1}{s - \frac{1}{\alpha}}.$$

解之得: 对任意$s \in [0, 1]$,

$$\eta^{(n)}(s) = \frac{\alpha s - 1 - \alpha^n(s-1)}{\alpha s - 1 - \alpha^{n+1}(s-1)}. \tag{4.3.2}$$

因此

$$\eta^{(n)}(0) = \frac{\alpha^n - 1}{\alpha^{n+1} - 1}.$$

进而对任意$n \geq 1$,

$$P_1(\tau_0 = n) = \frac{\alpha^n - 1}{\alpha^{n+1} - 1} - \frac{\alpha^{n-1} - 1}{\alpha^n - 1}.$$

当$p = 1/2$即$\alpha = 1$时, $\eta(s) = 1/(2-s)$. 用数学归纳法直接验算可知

$$\eta^{(n)}(s) = \frac{n-(n-1)s}{n+1-ns}.$$

此时对任意$n \geq 1$,

$$P_1(\tau_0 = n) = \frac{n}{n+1} - \frac{n-1}{n} = \frac{1}{n(n+1)}. \qquad \square$$

更一般地我们有下面的结论.

▲**命题4.3.7.** *设分枝过程初值$X_0 = 1$. 若后代分布均值$m < 1$, 那么平均灭绝时间是有限的.*

证明 $m < 1$时$P_1(\tau_0 < \infty) = 1$. 平均灭绝时间为

$$\begin{aligned} E_1(\tau_0) &= \sum_{n=1}^{\infty} P_1(\tau_0 \geq n) = 1 + \sum_{n=2}^{\infty} (1 - P_1(\tau_0 \leq n-1)) \\ &= 1 + \sum_{n=1}^{\infty} (1 - \eta^{(n)}(0)) \leq 1 + \sum_{n=1}^{\infty} [\eta^{(n)}(s)]'_{s=1}, \end{aligned}$$

其中不等号用到了微分中值定理以及导函数$[\eta^{(n)}(s)]'$单调增的性质. 注意到

$$\begin{aligned} [\eta^{(n)}(s)]'_{s=1} &= [\eta(\eta^{(n-1)}(s))]'_{s=1} = \eta'(\eta^{n-1}(1))[\eta^{(n-1)}(s)]'_{s=1} \\ &= \eta'(1)[\eta^{(n-1)}(s)]'_{s=1} = m[\eta^{(n-1)}(s)]'_{s=1} \\ &= \cdots = m^{n-1}\eta'(1) = m^n. \end{aligned} \tag{4.3.3}$$

因此

$$E_1(\tau_0) \leq 1 + \sum_{n=1}^{\infty} [\eta^{(n)}(s)]'_{s=1} \leq 1 + \sum_{n=1}^{\infty} m^n = \frac{1}{1-m} < \infty. \qquad \square$$

♠**注记4.3.8.** *在例4.3.6中, 大家可以看到$\alpha = 1$时$P_1(\tau_0 \geq n)$与$1/n$是同阶无穷小. 事实上, 在分布η的均值为1, 方差存在的条件下这个结论也成立(具体的结论感兴趣的同学可以参阅Athreya和Ney的专著 [7]), 此时平均灭绝时间是无穷的.*

♠**注记4.3.9.** *由(4.3.3)可知对任意$n \geq 1$, $E_1(X_n) = [\eta^{(n)}(s)]'_{s=1} = m^n$.*

对于一个灭绝的分枝过程而言, 人们有时还会问, 曾经存活过的粒子平均有多少? 注意到$n \geq \tau_0$时$X_n = 0$. 曾经存活过的粒子总数为

$$\begin{aligned} Z &= \sum_{n=0}^{\infty} X_n = \sum_{n=0}^{\infty} X_n 1_{\{n<\tau_0\}} \\ &= \sum_{n=0}^{\infty} X_n 1_{\{n+1 \leq \tau_0 < +\infty\}} + \sum_{n=0}^{\infty} X_n 1_{\{\tau_0 = +\infty\}}. \end{aligned}$$

因此在灭绝条件下存活过的粒子平均数为(假定$X_0=1$),

$$
\begin{aligned}
E_1(Z|A) &= E_1(Z|\tau_0<+\infty) = E_1(\sum_{n=0}^{\infty} X_n 1_{\{\tau_0\geq n+1\}}|\tau_0<+\infty) \\
&= \frac{\sum_{n=0}^{\infty} E_1(X_n 1_{\{n+1\leq\tau_0<+\infty\}})}{P_1(\tau_0<+\infty)}. \qquad (4.3.4)
\end{aligned}
$$

利用分枝过程的马氏性和时齐性可得

$$
\begin{aligned}
& E_1(X_n 1_{\{+\infty>\tau_0\geq n+1\}}) = E_1(E_1(X_n 1_{\{+\infty>\tau_0\geq n+1\}}|X_n)) \\
&= E_1(X_n E_1(1_{\{+\infty>\tau_0\geq n+1\}}|X_n)) \\
&= \sum_{k=1}^{\infty} k E_1(1_{\{+\infty>\tau_0\geq n+1\}}|X_n=k) P_1(X_n=k) \\
&= \sum_{k=1}^{\infty} k E(1_{\{+\infty>\tau_0\geq 1\}}|X_0=k) P_1(X_n=k). \qquad (4.3.5)
\end{aligned}
$$

在$X_0=k$条件下要让分枝过程灭绝, 意味其中任何一个粒子及其后代最终都要灭绝, 由于分枝过程中每个粒子及其后代都是独立的, 因此

$$
E(1_{\{+\infty>\tau_0\geq 1\}}|X_0=k) = [P_1(\tau_0<+\infty)]^k.
$$

将其代入(4.3.5)得

$$
E_1(X_n 1_{\{+\infty>\tau_0\geq n+1\}}) = \sum_{k=1}^{\infty} k[P_1(\tau_0<+\infty)]^k P_1(X_n=k).
$$

再将该式代入(4.3.4)得

$$
E_1(Z|A) = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} k[P_1(\tau_0<+\infty)]^{k-1} P_1(X_n=k).
$$

由此, 我们容易得到如下结论.

▲**命题4.3.10.** *设分枝过程初值$X_0=1$. 若后代分布均值$m<1$, 那么存活过的粒子平均数为$1/(1-m)$; 若$m=1$, 那么存活过的粒子总数为无穷; 如果$m>1$, 那么在灭绝条件下存活过的粒子平均数为$1/(1-\eta'(q))<+\infty$, 其中q为灭绝概率.*

证明 当$m\leq 1$时$P_1(A)=P_1(\tau_0<+\infty)=1$, 因此存活过的粒子平均数就是

$$
E_1(Z|A) = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} kP_1(X_n=k) = \sum_{n=0}^{\infty} E(X_n) = \sum_{n=0}^{\infty} m^n.
$$

显然$m<1$时该求和式为$1/(1-m)$, 而$m=1$时求和式为无穷.

当$m>1$时, 由命题4.3.3可知$q=P_1(\tau_0<+\infty)<1$. 因此灭绝条件下存活过的粒子平均数

$$
E_1(Z|A) = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} kq^{k-1}P_1(X_n=k) = 1+\sum_{n=1}^{\infty}[\eta^{(n)}(s)]'_{s=q}.
$$

注意到$\eta(q)=q$, 以及类似于(4.3.3)的推导可得

$$[\eta^{(n)}(s)]'_{s=q}=\eta'(\eta^{n-1}(q))[\eta^{(n-1)}(s)]'_{s=q}=\eta'(q)[\eta^{(n-1)}(s)]'_{s=q}=[\eta'(q)]^n.$$

由简单的数学分析可知(参见本节习题11)$\eta'(q)<1$. 因此

$$E_1(Z|A)=1+\sum_{n=1}^{\infty}[\eta'(q)]^n=\frac{1}{1-\eta'(q)}<+\infty. \qquad \square$$

♠**注记4.3.11.** 通常我们把后代分布均值$m<1$, $m=1$和$m>1$的分枝过程分别称为下临界, 临界和上临界分枝过程.

4.3.2 带移民的Galton-Watson分枝过程*

对G-W模型而言, 0是吸收态而所有非零状态都是非常返的(参见本节后习题15.2), 因此若X不灭绝则意味着X_n要趋向正无穷. 这种两极化现象显然与人们对许多随机现象的观察不符. 一种对G-W 模型简单的修正是引进所谓的“移民”, 得到如下的带移民G-W 过程.

♣**定义4.3.12.** 设$\{Y_{n,k};n\geq 1,k\geq 1\}$为一族独立同分布的非负整数值随机变量, 而$\{I_n;n\geq 1\}$ 是一列独立同分布的非负整数值随机变量且与$\{Y_{n,k};n\geq 1,k\geq 1\}$ 独立, 称随机过程$\{X_n;n\geq 0\}$ 为带移民的G-W 过程, 如果

$$X_{n+1}=\sum_{k=1}^{X_n}Y_{n+1,k}+I_{n+1}, \quad n\geq 0,$$

其中X_0是非负整数值随机变量且与$\{Y_{n,k}\}$以及$\{I_n\}$都独立.

与G-W过程一样我们仍称$Y_{n,k}$的共同分布为后代分布, 记作$\eta=(\eta_i,i\geq 0)$. 不同的是, 在带移民的G-W过程中每一代(比如第n代)粒子中都有(I_n 个)粒子, 它们不是第$n-1$代粒子的后代, 而是作为“移民”新加入到这个随机系统的. 这些粒子加入后与其他粒子一样产生下一代粒子. 模型中假定移入的粒子数量与n 之前的粒子数无关, 与第n 代新生的粒子数也无关. 我们称I_n的共同分布为移民分布, 记为$\mu=(\mu_i,i\geq 0)$.

与G-W过程一样, 由独立性假设, 容易验证带移民G-W过程X仍为时齐马氏链. 为叙述方便, 这一小节我们约定$1>\eta_0>0$并且对所有$k\geq 0$, $\mu_k>0$. 在此约定下, 对任意$i\geq 0$,

$$p_{i,0}=P(Y_{11}+\cdots+Y_{1i}+I_1=0)=\eta_0^i\mu_0>0, \quad p_{0,i}=P(I_1=i)=\mu_i>0.$$

由此可知X是不可约非周期的(见本节习题8).

以下总记$m=E(Y_{1,1})$, $\sigma^2=Var(Y_{1,1})$, $\lambda=E(I_1)$, $\beta^2=Var(I_1)$. 它们分别表示后代分布的均值和方差以及移民分布的均值和方差. 由本节习题9可知

▲**命题4.3.13.** 对带移民的G-W分枝过程X, 如下的条件矩公式成立.

(1) $E(X_n|X_{n-1}) = mX_{n-1} + \lambda$;

(2) $E(X_n^2|X_{n-1}) = (mX_{n-1} + \lambda)^2 + \sigma^2 X_{n-1} + \beta^2$;

其中$X_{n-1} = 0$时, 我们约定$mX_{n-1} = 0$, $\sigma^2 X_{n-1} = 0$.

带移民的分枝过程在一定条件下是正常返的. 我们有如下简单结论.

▲**命题4.3.14.** 若带移民的G-W分枝过程X的后代分布的均值$m \in (0,1)$以及移民分布的均值$\lambda \in (0,\infty)$, 那么X正常返.

证明 由前面的约定可知X是不可约非周期马氏链. 若X不是正常返的, 那么由注3.4.9可知, 对任意$i \geq 0$, $p_{1,i}^{(n)} \to 0$. 这意味着对正常数$M = \frac{2\lambda}{1-m} + 2$, 存在充分大的$n$, 使得对任意$0 \leq k < M$, $p_{1,k}^{(n)} < \frac{1}{2M}$. 因此

$$P_1(X_n \geq M) = 1 - \sum_{k=0}^{M-1} p_{1,k}^{(n)} > 1/2,$$

从而$E_1(X_n) > 1 + \frac{\lambda}{1-m}$. 但另一方面, 由命题4.3.13中(1)可知

$$\begin{aligned} E_1(X_n) &= E_1(E(X_n|X_{n-1})) = mE_1(X_{n-1}) + \lambda \\ &= \cdots = \lambda(1 + \cdots + m^{n-1}) + m^n \leq \frac{\lambda}{1-m} + 1. \end{aligned}$$

矛盾表明X一定正常返. □

▲**命题4.3.15.** 设X是带移民的G-W过程. 分别记后代分布、移民分布的概率母函数为$\eta(s)$, $\mu(s)$. 若X存在平稳分布, 那么平稳分布的概率母函数$\pi(s)$满足:

$$\pi(s) = \mu(s) \prod_{k=1}^{\infty} \mu(\eta^{(k)}(s)).$$

证明 以π为X的初始分布, 由π是平稳分布可得

$$\begin{aligned} \pi(s) &= E(s^{X_1}) = E(E(s^{X_1}|X_0)) = E(E(s^{\sum_{i=1}^{X_0} Y_{1,i} + I_1}|X_0)) \\ &= E(\mu(s)[\eta(s)]^{X_0}) = \mu(s)\pi(\eta(s)). \end{aligned}$$

由此可知, 对任意$n \geq 1$

$$\pi(s) = \mu(s)\pi(\eta^{(n+1)}(s)) \prod_{k=1}^{n} \mu(\eta^{(k)}(s)). \tag{4.3.6}$$

若$m > 1$(包括$m = +\infty$), 那么由$\eta(s) = s$在$[0,1)$内存在唯一解q, 以及当$s \in [q,1)$时, $\eta^{(n)}(s)$随n的增加单调不增, 而对$s \in [0,q)$时, $\eta^{(n)}(s)$随n的增加单调不降等事实可知对任意$s \in [0,1)$, $\lim_{n\to\infty} \eta^{(n)}(s) = q$. 由于$\mu(q) < 1$, 在(4.3.6)中令$n \to \infty$, 我们可知对任意$s \in [0,1)$, $\pi(s) = 0$. 这与$\pi(s)$为概率母函数矛盾. 因

此$m<1$. 此时由简单的数学分析讨论可知$n\to\infty$时$\eta^{(n+1)}(s)\to 1$. 因此

$$\pi(s)=\lim_{n\to\infty}\left[\mu(s)\pi(\eta^{n+1}(s))\prod_{k=1}^{n}\mu(\eta^{(k)}(s))\right]=\mu(s)\prod_{k=1}^{\infty}\mu(\eta^{(k)}(s)).\qquad\square$$

▶**例4.3.16.** 设带移民分枝过程X的后代分布与移民分布的概率母函数均为$\psi(s)=2/(3-s)$, 求X的平稳分布π.

解: 由命题4.3.15可知, π的概率母函数

$$\pi(s)=\prod_{k=1}^{\infty}\psi^{(k)}(s).$$

注意到$\psi(s)$对应分布的均值$\alpha=1/2$. 由(4.3.2)可知

$$\pi(s)=\lim_{n\to\infty}\prod_{k=1}^{n}\psi^{(k)}(s)=\lim_{n\to\infty}\frac{\alpha s-1-\alpha(s-1)}{\alpha s-1-\alpha^n(s-1)}=\frac{1}{2-s}.$$

因此, 对任意$k\geq 0$, $\pi_k=1/2^{k+1}$. $\square$

当一个随机动态系统分布稳定时, 人们可以考虑这个随机系统的统计推断问题. 对带移民的G-W分枝过程, 统计推断也是人们关注的问题之一. 以参数估计为例. 我们在后代分布方差σ^2和移民分布方差β^2有限条件下考虑后代分布均值m 和移民均值λ 的参数估计.

总设$0<m<1$, $0<\lambda<\infty$. 记此时的平稳分布为$\pi=(\pi_i,i\geq 0)$. 注意到已知X_{n-1} 条件下, X_n 的最佳估计是条件期望

$$E(X_n|X_{n-1})=mX_{n-1}+\lambda.$$

在给定样本$X_0,\cdots,X_n$条件下, 通过使得表示误差的函数

$$Q(m,\lambda)=\sum_{k=1}^{n}(X_k-mX_{k-1}-\lambda)^2$$

达到最小值, 我们可以得到参数m,λ的某种表示, 称其为(m,λ)的最小二乘估计. 直接计算可知(m,λ)的最小二乘估计为

$$\hat{m}_n=\frac{n\sum_{k=1}^{n}X_{k-1}X_k-\sum_{k=1}^{n}X_{k-1}\sum_{k=1}^{n}X_k}{n\sum_{k=1}^{n}X_{k-1}^2-(\sum_{k=1}^{n}X_{k-1})^2},$$

以及

$$\hat{\lambda}_n=\frac{1}{n}(\sum_{k=1}^{n}X_k-\hat{m}_n\sum_{k=1}^{n}X_{k-1}).$$

注意到

$$\hat{m}_n=\frac{\frac{\sum_{k=1}^{n}X_{k-1}X_k}{n}-\frac{\sum_{k=1}^{n}X_{k-1}}{n}\frac{\sum_{k=1}^{n}X_k}{n}}{\frac{\sum_{k=1}^{n}X_{k-1}^2}{n}-(\frac{\sum_{k=1}^{n}X_{k-1}}{n})^2}.$$

由推论3.5.3以及第三章第四节习题10可知

$$\hat{m}_n\to\frac{\sum_{i=0}^{\infty}i\pi_i\sum_{j=0}^{\infty}jp_{i,j}-(\sum_{i=0}^{\infty}i\pi_i)^2}{\sum_{i=0}^{\infty}i^2\pi_i-(\sum_{i=0}^{\infty}i\pi_i)^2}$$

$$= \frac{\sum_{i=0}^{\infty} i(mi+\lambda)\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2}{\sum_{i=0}^{\infty} i^2\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2}, \quad a.s. \tag{4.3.7}$$

下面我们计算$\sum\limits_{i=0}^{\infty} i\pi_i$以及$\sum\limits_{i=0}^{\infty} i^2\pi_i$. 为此我们假设$X$的初始分布为$\pi$, 那么$X_0$, X_1的分布都是π, 从而由命题4.3.13中结论(1)可知,

$$\sum_{i=0}^{\infty} i\pi_i = E_\pi(X_0) = E_\pi(X_1) = E_\pi(E_\pi(X_1|X_0)) = E_\pi(mX_0+\lambda)$$

因此

$$\sum_{i=0}^{\infty} i\pi_i = \frac{\lambda}{1-m}.$$

同样由命题4.3.13中结论(2)可知,

$$\begin{aligned}\sum_{i=0}^{\infty} i^2\pi_i &= E_\pi(X_0^2) = E_\pi(X_1^2) = E_\pi(E_\pi(X_1^2|X_0)) \\ &= m^2E_\pi(X_0^2) + (2\lambda m+\sigma^2)E_\pi(X_0) + \lambda^2 + \beta^2,\end{aligned}$$

可得

$$\sum_{i=0}^{\infty} i^2\pi_i = \left[\frac{\sigma^2\lambda}{1-m} + \beta^2\right]\frac{1}{1-m^2} + \frac{\lambda^2}{(1-m)^2}.$$

因此

$$\begin{aligned}\sum_{i=0}^{\infty} i^2\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2 &= \left[\frac{\sigma^2\lambda}{1-m} + \beta^2\right]\frac{1}{1-m^2}, \\ \sum_{i=0}^{\infty} i(mi+\lambda)\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2 &= \left[\frac{\sigma^2\lambda}{1-m} + \beta^2\right]\frac{m}{1-m^2}.\end{aligned}$$

将以上结果代入(4.3.7)得

$$\hat{m}_n \to m, \quad a.s.$$

由于

$$\hat{\lambda}_n = \frac{1}{n}\sum_{k=1}^{n} X_k - \hat{m}_n - \frac{1}{n}\sum_{k=1}^{n} X_{k-1}$$

由定理3.5.1以及几乎必然收敛的运算性质可得

$$\hat{\lambda}_n \to \sum_{i=0}^{\infty} i\pi_i - m\sum_{i=0}^{\infty} i\pi_i = \frac{\lambda}{1-m} - m\frac{\lambda}{1-m} = \lambda, \qquad a.s.$$

即$\hat{m}_n$, $\hat{\lambda}_n$分别是m和λ在几乎必然收敛意义下的相合估计.

分枝过程在理论和应用方面还有许多重要内容, 限于本书的篇幅, 我们对此不再深入. 感兴趣的同学可以通过参阅 [7]学习一些经典的结果, 也可以通过查阅研究论文了解最新发展.

练习题4.3

1 证明G-W过程$X = \{X_n; n \geq 0\}$ 是一个时齐马氏链, 转移概率为

$$p_{i,j} = P(Y_{1,1} + \cdots + Y_{1,i} = j).$$

2 证明, 只要$\eta_1 < 1$, Galton-Watson过程所有非零状态都是非常返的.

3 假设在红细胞培养试验中红细胞的生存时间是1分钟, 一个红细胞死亡时以1/4的概率生成两个红细胞, 以2/3的概率产生一个红细胞一个白细胞, 以1/12 的概率产生两个白细胞, 白细胞死亡后不会再生. 再生的红细胞又按前面的概率再生,且彼此互不干扰. 初始时只有一个红细胞. (1) 求在$n + 0.5$分钟的培养过程中都没有白细胞的概率. (2) 在整个细胞培养过程中,细胞最终灭绝, 求该现象发生的概率.

4 假设分枝过程X的初值为$m \geq 1$, 后代分布的概率母函数为$\eta(s) = 1-p+ps$, 其中$0 < p < 1$, 求该分枝过程灭绝时间的分布.

5 已知分枝过程X的后代分布概率母函数$\eta(s) = 1/(3-2s)$, 而且$X_0 = k$, 求在过程灭绝条件下存活过的粒子平均数.

6 假设X是初值为1的Galton-Watson分枝过程. 后代分布的概率母函数为

$$g(s) = \frac{1}{2} + \frac{1}{3}s + \frac{1}{6}s^2.$$

以M表示该过程存活过但没有后代的粒子数量, 求$E(M)$.

7 设X为G-W分枝过程, τ_0是灭绝时间. 证明$E_k(s^{\tau_0}) \geq [E_1(s^{\tau_0})]^k$.

8 对带移民分枝过程, 若$1 > p_i, q_i > 0$, $i = 0, 1$, 证明它是非周期不可约的,

9 证明命题4.3.13并求其中X_n的数学期望和方差.

10 设带移民分枝过程X的后代分布为取值为$0, 1$的两点分布, 移民分布是强度为λ的泊松分布. 求X的平稳分布.

11 记非退化的非负整数值随机变量ξ的概率母函数为$\eta(s)$. 若q为方程$\eta(s) = s$在$[0, 1)$内的解, 那么$\eta'(q) < 1$.

第五章　更新过程

马氏过程需要条件独立性, 要求在已知现在状态的条件下, 未来与过去无关. 本章介绍的更新过程是泊松过程的推广, 它一般不具有马氏性, 但在某些特定的时刻却有“ 更新”的特征.

5.1　更新过程与更新方程

给定概率空间$(\Omega, \mathfrak{F}, P)$及其上计数过程$N = \{N(t); t \geq 0\}$. 由第二章第三节可知, 对任意整数$k \geq 1$,

$$T_k = \inf\{t \geq 0; N(t) \geq k\} \qquad \text{以及} \qquad W_k = T_k - T_{k-1},$$

其中$T_0 = 0$, 分别为第k次(随机事件)发生时刻和第k个间隔时间. 计数过程N与到达时间序列$\{T_k;\ k \geq 1\}$满足如下关系

$$N(t) = \sup\{k;\ T_k \leq t\} = \sum_{k=1}^{\infty} 1_{\{T_k \leq t\}},$$

而且对任意$k \geq 0$以及$t \geq 0$,

$$\{N(t) \geq k\} = \{T_k \leq t\}, \tag{5.1.1}$$

$$\{N(t) = k\} = \{T_k \leq t < T_{k+1}\}, \tag{5.1.2}$$

$$N(t) + 1 = \inf\{k; T_k > t\}. \tag{5.1.3}$$

5.1.1　更新过程与更新性

♣**定义5.1.1.** *若计数过程$N = \{N(t);\ t \geq 0\}$具有独立同分布的时间间隔序列, 即存在独立同分布的非负随机变量序列$\{W_k;\ k \geq 1\}$, 使得*

$$N(t) = \sup\{n \geq 0;\ \sum_{k=1}^{n} W_k \leq t\}. \tag{5.1.4}$$

则称N为更新过程. 称W_k的分布(函数)F为间隔时间分布.

显然泊松过程是一个更新过程, 因此更新过程是泊松过程的一种推广. 为了避免$N(t) \equiv \infty$这种平凡情形, 此后我们总设$F(0) < 1$.

对任意$t \geq 0$以及到达时刻T_n,

$$N(t+T_n) - n = \sum_{k=n+1}^{\infty} 1_{\{T_k - T_n \leq t\}} = \sum_{k=n+1}^{\infty} 1_{\{\sum_{i=n+1}^{k} W_i \leq t\}}.$$

注意到$W_i, i \geq 1$, 是独立同分布的随机变量, $T_n = \sum_{k=1}^{n} W_k$只依赖于n之前的W_i, 因此$N(t+T_n) - n$与T_n独立, 而且

$$\{N(t+T_n) - n;\ t \geq 0\} \overset{f.d.}{=} \left\{ \sum_{k=1}^{\infty} 1_{\{\sum_{i=1}^{k} W_i \leq t\}};\ t \geq 0 \right\} = \{N(t);\ t \geq 0\},$$

其中符号$\overset{f.d.}{=}$表示它两边的过程具有相同的有限维分布族. 这个事实表明, 从恰好发生第n事件的时刻T_n往后看, 更新过程增量与T_n独立且增量的随机变化就像是从零时刻开始的更新过程. 直观地说, 在T_n之后, 过程$N(t)$重"新"开始. 为了引用方便, 我们把这种性质称为更新性, 我们也常称T_n 为(第n次)更新时刻.

▶**例5.1.2.** *假设$X = \{X_n;\ n \geq 1\}$为伯努利过程, N_n表示n次试验之前(含第n次)成功的次数, 并规定$N_0 = 0$. 对任意$t \geq 0$, 令$N(t) = N_{\lfloor t \rfloor}$. 那么$N = \{N(t);\ t \geq 0\}$ 就是一个更新过程. 此时到达时刻间隔W_i 服从几何分布:*

$$P(W_1 = k) = P(X_1 = 0, \cdots, X_{k-1} = 0, X_k = 1) = (1-p)^{k-1} p.$$

▶**例5.1.3.** *某型商品专卖店采取例3.5.5中的(k, K)-式进货仓储策略, 假设开始时商店有K件商品而且补充的商品能立即进货. 若顾客按强度为λ 的泊松过程$N(t)$到达, 而且每个顾客购买的商品数是独立同分布的随机变量ξ_i. 令$C(t)$表示$(0, t]$时段内的补货次数. 假定$\{\xi_i, i \geq 1\}$与$\{N(t); t \geq 0\}$独立, 那么$C(t)$就是一个更新过程, 间隔时间W分布为*

$$P(W > t) = P(\sum_{i=1}^{N(t)} \xi_i < K - k).$$

比如, 若$k = 2$, $K = 4$, 并设

$$P(\xi_i = 1) = p, \qquad P(\xi_i = 2) = q = 1 - p.$$

那么我们可得

$$\begin{aligned} P(W > t) &= P(\{(0, t]\text{内没有顾客}\} \cup \{(0, t]\text{内只有1个顾客且只买了1 件商品}\}) \\ &= \mathrm{e}^{-\lambda t} + p\lambda t \mathrm{e}^{-\lambda t}. \end{aligned}$$

注意到$T_k = \sum_{i=1}^{k} W_i$为独立同分布的随机变量之和, 由强大数定律可知

$$T_n / n \to \mu = E(W_1), \quad \text{a.s.}$$

因此$T_n \to \infty$, a.s. 此外利用分布的卷积表示, 我们知道

$$P(T_n \leq t) = F^{*n}(t),$$

其中$F^{*0}(t)=\begin{cases}1, & t\ge 0,\\ 0, & t<0.\end{cases}$ 进而$N(t)$的分布列为

$$P(N(t)=k)=P(T_k\le t)-P(T_{k+1}\le t)=F^{*k}(t)-F^{*(k+1)}(t),$$

并且对一切$t\ge 0$, $n\in\mathbb{N}$,

$$P(N(t)<\infty)=\lim_{n\to\infty}P(N(t)<n)=\lim_{n\to\infty}P(T_n>t)=1,$$

$$\lim_{t\to\infty}P(N(t)<n)=\lim_{t\to\infty}P(T_n>t)=0.$$

因此, $N(t)$取有限值且由其单调非降性可得$t\to\infty$时$N(t)\to\infty$, a.s.

★**定理5.1.4.** *更新过程$N=\{N(t)\}$的任意阶矩存在, 即对任意$t\ge 0$, $k\ge 1$,*

$$E[(N(t))^k]<+\infty,$$

特别地,

$$m(t)=E(N(t))=\sum_{n=1}^{\infty}F^{*n}(t).$$

证明 对任意固定的$t\ge 0$,

$$E[(N(t))^k]=\sum_{n=1}^{\infty}n^kP(N(t)=n)\le\sum_{n=1}^{\infty}n^kP(N(t)\ge n)$$
$$=\sum_{n=1}^{\infty}n^kP(T_n\le t)=\sum_{n=1}^{\infty}n^kF^{*n}(t).$$

由$\lim\limits_{n\to\infty}P(T_n\le t)=0$可知, 存在正整数$m$使得

$$a=F^{*m}(t)<1.$$

又因为对一切$n\ge 0$,

$$F^{*(m+n)}(t)=\int_0^t F^{*m}(t-s)\mathrm{d}F^{*n}(s)\le F^{*m}(t)F^{*n}(t).$$

所以对任意$l=0,1,2\cdots$, $r=0,1,2,\cdots,m-1$,

$$F^{*(lm+r)}(t)\le(F^{*m}(t))^l=a^l.$$

因此

$$E[(N(t))^k]\le\sum_{n=1}^{\infty}n^kF^{*n}(t)=\sum_{l=0}^{\infty}\sum_{r=0}^{m-1}(lm+r)^kF^{*(lm+r)}(t)$$
$$\le m^{k+1}\sum_{l=0}^{\infty}(l+1)^ka^l<+\infty. \tag{5.1.5}$$

此时

$$m(t)=E(N(t))=\sum_{n=1}^{\infty}nP(N(t)=n)=\sum_{n=1}^{\infty}n(F^{*n}(t)-F^{*(n+1)}(t))$$
$$=F(t)+\sum_{n=2}^{\infty}[n-(n-1)]F^{*n}(t)=\sum_{n=1}^{\infty}F^{*n}(t). \qquad \square$$

特别地, 若N是泊松过程, 容易检验
$$m(t)=\sum_{n=1}^{\infty}F^{*n}(t)=\sum_{n=1}^{\infty}\int_0^t\frac{\lambda^n s^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda s}\mathrm{d}s=\int_0^t\lambda\mathrm{e}^{-\lambda s}\sum_{n=0}^{\infty}\frac{\lambda^n s^n}{n!}\mathrm{d}s=\lambda t,$$
这与命题2.3.5相关结果一致.

5.1.2 更新函数与更新方程

♣**定义5.1.5.** 称$m(t)=E(N(t))$为更新过程$N=\{N(t);\ t\geq 0\}$的更新函数. 当N的间隔时间分布为F时, 也称m是分布F的更新函数.

♠**注记5.1.6.** 为了叙述方便, 若无特别说明, 本章此后更新过程的间隔时间分布F都是离散格子点分布(参见本节习题7) 或者$(0,+\infty)$上的连续分布且具有局部有界(即在任一有界区间上都有界) 的概率密度函数.

当F是步长为d的离散格子点分布时, 此时由本节习题8可知$m(t)$为阶梯函数, 每个阶梯的跳跃点都可表示成nd的形式.

当F连续时, 假设概率密度函数为f, 那么对任意$k\geq 1$, 由
$$F^{*k}(t)=\int_0^t F^{*(k-1)}(t-s)f(s)\mathrm{d}s,$$
可知$F^{*k}(t)$的概率密度函数存在, 记其为$f_k(t)$, 其中$f_1(t)\equiv f(t)$且
$$f_k(t)=\int_0^t f_{k-1}(t-s)f(s)\mathrm{d}s.$$
定理5.1.4的证明过程表明级数$\sum_{n=1}^{\infty}F^{*n}(t)$关于$t$在任意有界集上是一致收敛的. 因此, 当概率密度函数f局部有界时, 由
$$\begin{aligned}\sum_{n=1}^{\infty}f_n(t)&=\Big[f(t)+\sum_{n=2}^{\infty}\int_0^t f_{n-1}(t-s)f(s)\mathrm{d}s\Big]\\&\leq\sup_{0\leq s\leq t}f(s)\Big[1+\sum_{n=1}^{\infty}\int_0^t f_n(s)\mathrm{d}s\Big]\\&=\sup_{0\leq s\leq t}f(s)\Big[1+\sum_{n=1}^{\infty}F^{*n}(t)\Big]<\infty,\end{aligned}$$
可知
$$\frac{\mathrm{d}m(t)}{\mathrm{d}t}=\sum_{n=1}^{\infty}\frac{\mathrm{d}F^{*n}(t)}{\mathrm{d}t}=\sum_{n=1}^{\infty}f_n(t).$$
即$m(t)$是连续可微的.

♣**定义5.1.7.** 设$b(t)$是$[0,\infty)$上的局部有界函数, 称如下形式的方程
$$B(t)=b(t)+\int_0^t B(t-s)\mathrm{d}F(s),\tag{5.1.6}$$
为更新方程, 其中$B(t)$为未知函数, $F(s)$是非负随机变量的分布函数.

♠**注记5.1.8.** *方程5.1.6右边的积分*$\int_0^t B(t-s)\mathrm{d}F(s)$*是所谓的Riemann-Stieltjes积分. 对此积分的具体定义和性质, 本书不展开介绍. 但我们指出: 在注记5.1.6条件下该积分分别对应为离散求和与牛顿-莱布尼茨积分: 当*F *是离散分布时, 该积分是和式*

$$\sum_{0\le x_k\le t} B(t-x_k)p_k,$$

其中x_k*为*F*支撑集(参见本节习题7)中的元素,* $p_k = F(x_k) - F(x_k-)$*; 当*F*是概率密度函数为*f*的连续分布时, 该积分为牛顿- 莱布尼茨积分*

$$\int_0^t B(t-s)f(s)\mathrm{d}s.$$

除特别说明外, 此后不论F*是离散还是连续分布, 我们统一用积分*$\int_0^\cdot \phi(s)\mathrm{d}F(s)$*表示对函数*$\phi$*的求和或牛顿-莱布尼茨积分.*

★**定理5.1.9.** *更新函数*$m(t)$*满足如下积分方程*

$$m(t) = F(t) + \int_0^t m(t-s)\mathrm{d}F(s). \tag{5.1.7}$$

证明 由重期望公式可得

$$m(t) = E(N(t)) = E(E(N(t)|T_1)) = \int_0^\infty E(N(t)|T_1=s)\mathrm{d}F(s).$$

当$s=T_1>t$时, $N(t)=0$; 当$s=T_1<t$时, 由更新性可知

$$N(t) = N(t-T_1+T_1) - 1 + 1 \overset{d}{=} 1 + N(t-T_1) = 1 + N(t-s),$$

其中符号$\overset{d}{=}$表示等号两边的随机变量具有相同概率分布. 因此

$$E(N(t)|T_1=s) = \begin{cases} 0, & s>t, \\ 1+E(N(t-s)) = 1+m(t-s), & s<t. \end{cases}$$

由此可得

$$m(t) = \int_0^t [1+m(t-s)]\mathrm{d}F(s) = F(t) + \int_0^t m(t-s)\mathrm{d}F(s). \qquad \square$$

♠**注记5.1.10.** *人们常把利用重期望公式和更新性建立更新方程的方法称为更新技巧(方法).*

►**例5.1.11.** *设更新过程*$N(t)$*的间隔时间分布*F*为*$[0,1]$*上的均匀分布, 求对应的更新函数*$m(t)$*, 其中*$0<t\le 1$*.*

解 对任意$0<t\le 1$, 所求更新函数满足的更新方程为

$$m(t) = F(t) + \int_0^t m(t-s)\mathrm{d}F(s) = t + \int_0^t m(s)\mathrm{d}s.$$

两边关于t求导得$m'(t) = 1 + m(t)$. 因此$\frac{\mathrm{d}m(t)}{1+m(t)} = \mathrm{d}t$. 由此可解得

$$\ln(1+m(t)) = c + t \Rightarrow m(t) = C\mathrm{e}^t - 1.$$

由$m(0)=0$可知$C=1$, 即$m(t)=\mathrm{e}^t-1$. □

利用更新函数可以给出更新方程唯一局部有界解的显式表示, 证明省略.

★定理5.1.12. *设$b(t)$为局部有界函数, 更新方程(5.1.6)有唯一局部有界解*

$$B(t)=b(t)+\int_0^t b(t-s)\mathrm{d}m(s), \tag{5.1.8}$$

其中$m(s)$是F的更新函数.

♠注记5.1.13. (1) *方程(5.1.8)中积分$\int_0^t b(t-s)\mathrm{d}m(s)$仍是Riemann-Stieltjes积分. 在注5.1.6条件下, 当F是具有局部有界概率密度函数的连续分布时, $m(s)$连续可微, 该积分与如下牛顿-莱布尼茨积分一致*

$$\int_0^t b(t-s)m'(s)\mathrm{d}s;$$

当F是步长为d的格子点分布时, $m(s)$是跳跃点在kd的阶梯函数, 该积分是离散求和,

$$\sum_{kd\le t} b(t-kd)\big(m(kd)-m((k-1)d)\big),$$

其中约定$m(-d)=0$.

(2) *由$N(t)$ 的单调非降和右连续性可知$m(t)$ 是单调非降右连续函数,*

$$\mathrm{d}m(t)=\mathrm{d}\Big(\sum_{n=1}^{\infty}F^{*n}(t)\Big)\approx\sum_{n=1}^{\infty}P(t<T_n\le t+\mathrm{d}t).$$

因此$\mathrm{d}m(t)$可直观地理解为N在时间区间$(t,t+\mathrm{d}t]$发生跳跃的概率.

由定理5.1.9与定理5.1.12, 我们可以得到更新函数$m(t)$的另一种表示:

$$m(t)=F(t)+\int_0^t F(t-s)\mathrm{d}m(s). \tag{5.1.9}$$

►例5.1.14. *对任意$t,s\ge 0$, 求$E(N^2(t))$.*

解 记$K(t)=E(N^2(t))$. 由重期望公式

$$K(t)=E(E(N^2(t)|T_1))=\int_0^{\infty}E(N^2(t)|T_1=u)\mathrm{d}F(u).$$

由更新性,

$$E(N^2(t)|T_1=u)=\begin{cases}0, & u>t,\\ E((1+N(t-u))^2), & u<t.\end{cases}$$

即
$$E(N^2(t)|T_1=u)=\begin{cases}0, & u>t,\\ 1+2m(t-u)+K(t-u), & u<t.\end{cases}$$

所以$K(t)$满足方程

$$K(t)=\int_0^t(1+2m(t-u)+K(t-u))\mathrm{d}F(u)$$

$$= F(t) + 2\int_0^t m(t-u)\mathrm{d}F(u) + \int_0^t K(t-u)\mathrm{d}F(u)$$
$$= 2m(t) - F(t) + \int_0^t K(t-u)\mathrm{d}F(u).$$

由定理5.1.4易知$K(t)$是局部有界函数. 由定理5.1.12以及(5.1.9)可得

$$K(t) = 2m(t) - F(t) + \int_0^t (2m(t-s) - F(t-s))\mathrm{d}m(s)$$
$$= m(t) + 2\int_0^t m(t-s)\mathrm{d}m(s). \qquad \square$$

►**例5.1.15.** 求如下方程的局部有界解

$$g(t) = t + \int_0^t g(t-s)2\mathrm{e}^{-2s}\mathrm{d}s.$$

解 令$\mathrm{d}F(s) = 2\mathrm{e}^{-2s}\mathrm{d}s$, 即$F$是参数为2的指数分布. 再令$b(t) = t$. 由定理5.1.12, 所求方程的唯一局部有界解为

$$g(t) = b(t) + \int_0^t b(t-s)\mathrm{d}m(s),$$

其中$m(s)$是分布F的更新函数. 由于以F为间隔时间的更新过程是强度为2的泊松过程, 由命题2.3.5可知$m(s) = 2s$. 因此

$$g(t) = t + \int_0^t (t-s)2\mathrm{d}s = t^2 + t. \qquad \square$$

5.1.3 更新过程的几个统计量

♣**定义5.1.16.** 设$N = \{N(t);\ t \geq 0\}$为更新过程. 令

$$A(t) = t - T_{N(t)}, \quad R(t) = T_{N(t)+1} - t, \quad L(t) = T_{N(t)+1} - T_{N(t)}.$$

分别称$A(t), R(t), \beta(t)$为更新过程N在时刻t的年龄,剩余寿命和总寿命.

显然, $0 \leq A(t) \leq t$, $t \geq 0$ 且对任意$x \geq 0$, $0 \leq y \leq t$,

$$\{R(t) > x, A(t) \geq y\} = \{3(t-y, t+x]\text{中}N\text{没有跳跃}\}$$
$$= \{R(t-y) > x+y\}$$
$$= \{A(t+x) \geq x+y\}.$$

取$y = 0$得

$$\{R(t) > x\} = \{A(t+x) \geq x\}.$$

★**定理5.1.17.** 若N的间隔时间分布F均值有限, 则对任意$x, t \geq 0$,

$$P(R(t) > x) = \bar{F}(t+x) + \int_0^t \bar{F}(t-s+x)\mathrm{d}m(s), \tag{5.1.10}$$

$$E(R(t))=\int_t^{\infty}\bar{F}(t)\mathrm{d}t+\int_0^t\int_{t-s}^{\infty}\bar{F}(u)\mathrm{d}u\mathrm{d}m(s),\tag{5.1.11}$$

其中$\bar{F}(t)=1-F(t)$.

证明 注意到对任意$x\geq 0$,

$$P(R(t)>x|T_1=s)=\begin{cases}1, & s>t+x,\\ 0, & t<s\leq t+x,\\ P(R(t-s)>x), & s\leq t.\end{cases}$$

利用更新技巧得

$$\begin{aligned}P(R(t)>x) &= E(P(R(t)>x|T_1))\\ &=\int_0^{\infty}P(R(t)>x|T_1=s)\mathrm{d}F(s)\\ &=\int_{t+x}^{\infty}\mathrm{d}F(s)+\int_0^t P(R(t-s)>x)\mathrm{d}F(s)\\ &=\bar{F}(t+x)+\int_0^t P(R(t-s)>x)\mathrm{d}F(s).\end{aligned}$$

由定理5.1.12可知(5.1.10)成立. 再注意到对任意$t\geq 0$,

$$E(R(t)|T_1=s)=\begin{cases}s-t, & s>t,\\ E(R(t-s)), & s\leq t.\end{cases}$$

由更新技巧可得

$$\begin{aligned}E(R(t)) &= E(E(R(t)|T_1))=\int_0^{\infty}E(R(t)|T_1=s)\mathrm{d}F(s)\\ &=\int_t^{\infty}(u-t)\mathrm{d}F(u)+\int_0^t E(R(t-s))\mathrm{d}F(s)\\ &=\int_t^{\infty}\bar{F}(u)\mathrm{d}u+\int_0^t E(R(t-s))\mathrm{d}F(s).\end{aligned}\tag{5.1.12}$$

此外由(5.1.10)可得

$$P(R(t)>x)\leq\bar{F}(x)+\bar{F}(x)m(t).$$

因此

$$\begin{aligned}E(R(t)) &= \int_0^{\infty}P(R(t)>x)\mathrm{d}x\\ &\leq\int_0^{\infty}\bar{F}(x)\mathrm{d}x(1+m(t))\\ &=\mu(1+m(t)).\end{aligned}$$

这表明$E(R(\cdot))$是局部有界函数. 由(5.1.12)以及定理5.1.12可得(5.1.11). □

对于$A(t),L(t)$也可计算类似问题, 做为习题请读者自己完成.

值得指出的是, 对任意$x>0$, 若$A(t)>x$, 那么

$$P(L(t)>x)=1;$$

若$0\le s=A(t)\le x$, 那么$t-s$恰好是一个更新时刻, 此时由更新性可得

$$\begin{aligned}P(L(t)>x|A(t)=s) &= P(L(s)>x|A(s)=s)\\ &= P(W>x|W>s)\\ &\ge P(W>x),\end{aligned}$$

其中W表示更新过程的一个间隔时间. 因此不论t时刻的年龄如何, 总有

$$P(L(t)>x|A(t))\ge P(W>x),$$

从而

$$P(L(t)>x)\ge P(W>x). \tag{5.1.13}$$

这表明我们在某个固定t时刻观察到的间隔时间会比通常“理论上”的间隔时间长些——对于泊松过程我们可以给出更清楚的计算(参见第二章第三节习题3). 这种现象被称为长度偏离取样(length-biased sampling) 或检验悖论. 它在统计意义上的直观解释是: 由于长的时间间隔比短的时间间隔更可能包含给定的时刻t, 因此t 时刻观察到的间隔时间会偏长.

练习题5.1

1 设更新过程N的间隔时间分布F的概率密度函数为$f(t)=te^{-t}$, $t>0$,

(1) 对任意$0<t_1<t_2$以及正整数n,m, 求$P(N(t_1)=n,N(t_2)=m)$.

(2) 对任意$t>0$, 求更新函数$m(t)$.

2 求例5.1.11中$m(t)$在$t\in(1,2]$时的值.

3 求积分方程$g(t)=t+\int_0^t g(t-s)3e^{-2s}ds$ 的局部有界解.

4 在等待时间均值有限条件下, 求$P(A(t)\ge x)$, $E(A(t))$以及$P(L(t)\ge x)$, $E(L(t))$.

5 设N是强度为λ的泊松过程, 试计算其在时刻t时年龄$A(t)$、剩余寿命 $R(t)$以及总寿命$L(t)$的数学期望.

6 设N是更新过程, 间隔时间分布期望为μ, 证明$E(T_{N(t)+1})=\mu(1+m(t))$.

7* 记离散型随机变量X的分布为F, 称$S=\{x;P(X=x)>0\}$为分布F 的支撑集. 若存在非负实数$d>0$使得$S\subset\{nd;n\in\mathbb{Z}\}$, 则称$F$ 为格子点分布或X服从格子点分布, 此时称

$$\sup\{d>0;\ S\subset\{nd;n\in\mathbb{Z}\}\}$$

为格子点分布F的步长. 证明若F的步长为d, 那么$S\subset\{nd;n\in\mathbb{Z}\}$而且存

在正整数m以及互素的整数$n_k, k=1,2,\cdots,m$, 使得

$$\{n_k d;\ k=1,2,\cdots,m\} \subset S.$$

8* 设等待时间分布F是步长为d的格子点分布, $m(t)$是F的更新函数, 证明$m(t)$是阶梯函数, 跳跃点t必能被d整除, 而且存在充分大的N, 当$n>N$时nd是跳跃点.

5.2 更新极限定理

上一节我们对更新过程与更新函数的趋势作了简单介绍, 下面我们研究更新过程极限的更精细性质并简要说明这些性质的应用.

5.2.1 大数定律与中心极限定理

★**定理5.2.1.** *对更新过程$N=\{N(t);\ t\geq 0\}$, 我们有*

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{\mu},\quad a.s.$$

其中$\mu=\infty$时$1/\mu$ 为0.

证明 由T_n的定义可知$T_{N(t)}\leq t<T_{N(t)+1}$, 从而

$$\frac{T_{N(t)}}{N(t)}\leq\frac{t}{N(t)}\leq\frac{T_{N(t)+1}}{N(t)}=\frac{T_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}.\tag{5.2.1}$$

注意到$N(t)\to\infty$, a.s.以及$T_n/n\to\mu$, a.s. 由第一章第四节习题1可知

$$\lim_{t\to\infty}\frac{T_{N(t)}}{N(t)}=\mu,\quad a.s.$$

将此结果应用于(5.2.1)两边得$t/N(t)\to\mu$, a.s. 从而$N(t)/t\to 1/\mu$, a.s. □

通常称$1/\mu$为更新过程的速率.

▶**例5.2.2.** *假设伯努利实验中成功的概率为p, 失败的概率为$q=1-p$, 连续做若干次实验直至连续成功k次或连续失败k次为止, 问平均实验次数.*

[**分析**] 以连续成功或连续失败k次为一次计数事件A, 假定伯努利实验在计数事件A发生后重新计算连续成功或连续失败, 以$N(n)$表示到第n次实验时事件A 出现的次数, 那么N 是一个更新过程. 类似, 若以连续成功k次为一次计数事件B, 以$N_1(n)$ 表示到第n次实验时事件B出现的次数, 那么N_1也是一个更新过程; 以连续失败k次为一次计数事件C, 以$N_2(n)$表示到第n次实验时事件C 出现的次数, 那么N_2 也是一个更新过程. 显然$N=N_1+N_2$. 由定理5.2.1, N 的更新速率为N_1, N_2的更新速率之和.

解: 如上述分析建立更新过程N, N_1, N_2. 由第一章第二节习题4可知N_1, N_2的平均等待时长分别为

$$\mu_1=\frac{1-p^k}{p^k(1-p)},\qquad \mu_2=\frac{1-q^k}{q^k(1-q)}.$$

因此N的更新速率为

$$\frac{1}{\mu}=\frac{p^k(1-p)}{1-p^k}+\frac{q^k(1-q)}{1-q^k}.$$

所以平均实验次数为

$$\mu = \frac{(1-p^k)(1-q^k)}{pq(p^{k-1}+q^{k-1}-p^{k-1}q^{k-1})}. \qquad \square$$

★**定理5.2.3.** *若更新过程$N=\{N(t);\ t\geq 0\}$满足条件*

$$E(W_1)=\mu<\infty, \mathbb{V}ar(W_1)=\sigma^2<\infty,$$

那么对一切实数y,

$$\lim_{t\to\infty} P\Big(\frac{N(t)-t/\mu}{\sigma\sqrt{t/\mu^3}}\leq y\Big)=\Phi(y),$$

其中$\Phi(y)$是标准正态分布函数.

证明 令$r_t=t/\mu+y\sigma\sqrt{t/\mu^3}$, $\hat{r}_t=\lfloor r_t\rfloor+1$, 则

$$P\Big(\frac{N(t)-t/\mu}{\sigma\sqrt{t/\mu^3}}\leq y\Big)=P(N(t)\leq\lfloor r_t\rfloor)=P(N(t)<\hat{r}_t)$$
$$=P(T_{\hat{r}_t}>t)=P\Big(\frac{T_{\hat{r}_t}-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}>\frac{t-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}\Big).$$

注意到$\hat{r}_t\to\infty$,

$$\lim_{t\to\infty}\frac{t-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}=\lim_{t\to\infty}\frac{-y\sigma\sqrt{t/\mu}}{\sigma\sqrt{t/\mu+y\sigma\sqrt{t/\mu^3}}}=-y,$$

以及

$$\frac{T_{\hat{r}_t}-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}=\frac{1}{\sigma\sqrt{\hat{r}_t}}\sum_{k=1}^{\hat{r}_t}(W_k-\mu).$$

由中心极限定理可知结论成立. $\square$

▶**例5.2.4.** *一台机器相继不间断地加工某种零件. 假定加工一件零件的时间服从1到3 分钟的均匀分布, 估计100个小时内以0.95的概率至少可加工的零件个数.*

解: 以分为时间单位, 加工一件零件时间的均值为2, 方差为1/3. 以$N(t)$ 表示到t时刻加工完成的零件数. 由定理5.2.3, 当t很大时$N(t)$ 近似服从均值为$t/2$, 方差为$t/24$的正态分布. 设所求零件个数为x, 那么

$$0.95=P(N(6000)>x)=P\Big(\frac{N(6000)-3000}{\sqrt{6000/24}}>\frac{x-3000}{5\sqrt{10}}\Big)$$
$$\approx 1-\Phi\Big(\frac{x-3000}{5\sqrt{10}}\Big).$$

由此可知$x\approx 3000-1.64\times 5\sqrt{10}\approx 2974$. $\square$

5.2.2 更新定理

下面的结论本质上是所谓的关键更新定理, 它是我们研究更新方程局部有界

解极限的重要工具.

★**定理5.2.5.** 设F为非负随机变量的分布函数, 期望为μ (可以是∞), $m(t)$是F的更新函数. $|h(t)|$ 在 $[0,\infty)$上可积且单调不增. B 是更新方程的局部有界解,

$$B(t)=h(t)+\int_0^t B(t-s)\mathrm{d}F(s)=h(t)+\int_0^t h(t-s)\mathrm{d}m(s). \qquad (5.2.2)$$

(1) 如果F是步长为d的格子点分布, 那么对任意$0\le c<d$,

$$\lim_{n\to\infty}B(c+nd)=\begin{cases} d\sum_{n=0}^{\infty}h(c+nd)/\mu, & \mu<\infty,\\ 0, & \mu=\infty.\end{cases}$$

(2) 若F的概率密度函数局部有界, 那么

$$\lim_{t\to\infty}B(t)=\lim_{t\to\infty}\int_0^t h(t-s)\mathrm{d}m(s)=\frac{1}{\mu}\int_0^{\infty}h(s)\mathrm{d}s,$$

其中若$\mu=\infty$, 上式右端理解为0.

证明* (1) 当F是步长为d的格子点分布时, 由更新方程, 对任意$n\ge 1$,

$$B(c+nd)=h(c+nd)+\sum_{k=0}^{n}B(c+(n-k)d)\nu(kd),$$

其中$\nu(kd)=F(kd)-F((k-1)d)=P(X=kd)$. 为方便, 对任意$n\ge 0$, 记

$$a_n=B(c+nd),\quad b_n=h(c+nd),\quad u_n=\nu(nd).$$

则

$$a_n=b_n+\sum_{k=0}^{n}a_{n-k}u_k. \qquad (5.2.3)$$

进一步, 由本章第一节习题7可知, 存在$m_1,m_2,\cdots,m_r$使得

$$m_1,\cdots,m_r\text{互素而且对任意 } i=1,2,\cdots r,\ u_{m_i}=P(X=m_id)>0.$$

此时由第一章第一节习题10可知, 存在正整数J, 当整数$j\ge J$时存在非负整数$l_1,\cdots,l_r$, 使得

$$j=l_1m_1+l_2m_2+\cdots+l_rm_r. \qquad (5.2.4)$$

下面我们将证明分三步.

第一步, 证明$\{a_n\}$为有界数列. 注意到$\sum_{k=0}^{\infty}u_k=1$, $u(0)<1$, 令

$$c_n=\frac{\sum_{k=0}^{n}|b_k|}{1-u_0},\qquad n\ge 0,\quad c_{-1}=0.$$

由$|h|$在$[0,+\infty)$可积且单调不增的假设容易知道

$$\sum_{k=0}^{\infty}|b_k|=\sum_{k=0}^{\infty}|h(c+kd)|<\infty.$$

注意到c_k为单调增加数列, 由(5.2.3)及数学归纳法容易证明, 对任意$n\ge 0$,

$$|a_n|\le\frac{|b_n|}{1-u_0}+c_{n-1}=c_n.$$

因此$\{a_n\}$为有界数列. 记$M = \sup_{n\geq 0} |a_n|$.

第二步, 证明$\limsup_{n\to\infty} a_n \leq \begin{cases} d\sum_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$

为此记$l = \limsup_{n\to\infty} a_n$. 此时存在$n_i \to \infty$使得$a_{n_i} \to l$. 任取$m_k$, $1 \leq k \leq r$, 若$n_i \to \infty$时$a_{n_i - m_k} \not\to l$, 那么存在$l' < l$以及无穷多个i使得$a_{n_i - m_k} \leq l'$. 令$\varepsilon = u_{m_k}(l - l')/4$. 由$\sum_{k=0}^{\infty} u_k = 1$ 可知存在$N > m_k$, 当$n \geq N$ 时

$$\sum_{k=n}^{\infty} u_k < \varepsilon/M.$$

进而存在$n_j \geq N$使得

$$a_{n_j} > l - \varepsilon, \quad a_{n_j - m_k} < l', \quad |b_{n_j}| \leq \varepsilon,$$

而且对一切$n \geq n_j - N$, $a_n \leq l + \varepsilon$成立. 此时

$$\begin{aligned} a_{n_j} &\leq \sum_{k=0}^{n_j} a_{n_j - k} u_k + \varepsilon < \sum_{k=0}^{N} a_{n_j - k} u_k + M \sum_{k=N+1}^{n_j} u_k + \varepsilon \\ &< (1 - u_{m_k})(l + \varepsilon) + u_{m_k} l' + 2\varepsilon \\ &< l + 3\varepsilon - u_{m_k}(l - l') = l - \varepsilon. \end{aligned}$$

矛盾表明

$$\lim_{n_i\to\infty} a_{n_i - m_k} = l.$$

反复使用以上方法可知, 对任意非负整数$l_1, \cdots, l_k$,

$$\lim_{n_i\to\infty} a_{n_i - \sum_{k=1}^{r} l_k m_k} = l.$$

由(5.2.4)可知, 对任意$j \geq J$, $\lim_{n_i\to\infty} a_{n_i - j} = l$. 记$n_i' = (n_i - J) \vee 1$, 对任意$j \geq 0$,

$$\lim_{n_i'\to\infty} a_{n_i' - j} = l.$$

令$U_n = \sum_{k=n+1}^{\infty} u_k$, 由关系式$u_k = U_{k-1} - U_k$以及$U_{-1} = 1$, (5.2.3)可改写为

$$U_0 a_n + U_1 a_{n-1} + \cdots + U_n a_0 = U_0 a_{n-1} + U_1 a_{n-2} + \cdots + U_{n-1} a_0 + b_n, \quad n \geq 1.$$

记$A_n = \sum_{k=0}^{n} U_k a_{n-k}$, 那么$A_0 = b_0$且$A_n = A_{n-1} + b_n$, 进而$A_n = \sum_{k=0}^{n} b_k$. 因此对任意$N \leq n_i'$,

$$U_0 a_{n_i'} + \cdots + U_N a_{n_i' - N} \leq A_{n_i'} \leq \sum_{k=0}^{\infty} |b_k|.$$

先令$n_i' \to \infty$再令$N \to \infty$并注意到$\sum\limits_{k=0}^{\infty} U_k = \sum\limits_{k=1}^{\infty} k u_k = \mu/d$, 我们可得

$$l \le \frac{\sum\limits_{k=0}^{\infty} b_k}{\sum\limits_{k=1}^{\infty} k u_k} = \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

第三步, 证明$\liminf\limits_{n\to\infty} a_n \ge \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$

令$\underline{l} = \liminf\limits_{n\to\infty} a_n$. 由第二步的类似讨论同样可知存在$n_i$ 使得对任意$k \ge 0$,

$$a_{n_i - k} \to \underline{l}.$$

记$g(N) = \sum\limits_{k=N+1}^{\infty} U_k$, 由$\sum\limits_{k=0}^{\infty} U_k = \mu/d < \infty$ 可知$g(N) \to 0$. 注意到

$$\sum_{k=0}^{n_i} b_k = \sum_{k=0}^{n_i} U_k a_{n_i-k} \le \sum_{k=0}^{N} U_k a_{n_i-k} + g(N)M.$$

先令$n_i \to \infty$再令$N \to \infty$得

$$\underline{l} \ge \frac{\sum\limits_{k=0}^{\infty} b_n}{\sum\limits_{k=1}^{\infty} k u_k} = \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

综合第二、三步的结果可知(1)成立.

(2) 此时证明更复杂, 此略. □

♠**注记5.2.6.** 事实上(2)对一切非格子点分布都成立.

♠**注记5.2.7.** 在(1)的证明中我们得到了如下一般的数列极限结果:

若数列$\{a_n\}$, $\{b_n\}$和非负数列$\{c_n\}$满足

(1) $\sum\limits_{n=0}^{\infty} |b_n| < \infty$,

(2) $\sum\limits_{n=0}^{\infty} c_n = 1$ 且数集$\{n \ge 1; c_n > 0\}$的最大公因子为1,

(3) 对任意$n \ge 0$, $a_n = b_n + \sum\limits_{k=0}^{n} a_{n-k} c_k$.

那么

$$\lim_{n\to\infty} a_n = \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=1}^{\infty} n c_n},$$

其中商式分母为$+\infty$时, 极限取值为0.

▶**例5.2.8.** 设X是以$\mathbf{P} = (p_{i,j})_{i,j\in S}$为转移概率的马氏链, 若$j$非周期常返, 证

明

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \begin{cases} 0, & j\text{是零常返}, \\ f_{i,j}/m_{j,j}, & j\text{是正常返}. \end{cases}$$

证明 注意到对任意$n \geq 1$,

$$\begin{aligned} p_{i,j}^{(n)} &= P_i(X_n = j) = P_i(X_n = j, X_{n-1} = j) + P_i(X_n = j, X_{n-1} \neq j) \\ &= \sum_{k=1}^{n} P_i(X_n = j, X_k = j, X_v \neq j, v = k+1, \cdots, n-1) \\ &\qquad + P_i(X_n = j, X_v \neq j, v = 1, \cdots, n-1) \\ &= f_{i,j}^{(n)} + \sum_{k=1}^{n-1} p_{i,j}^{(k)} f_{j,j}^{(n-k)} = f_{i,j}^{(n)} + \sum_{k=1}^{n-1} p_{i,j}^{(n-k)} f_{j,j}^{(k)}. \end{aligned}$$

对$n \geq 1$, 令$a_n = p_{i,j}^{(n)}$, $b_n = f_{i,j}^{(n)}$, $c_n = f_{j,j}^{(n)}$, 并约定$a_0 = b_0 = c_0 = 0$, 那么由j非周期常返可知$\sum\limits_{n=0}^{\infty} c_n = 1$以及集合$\{n; c_n > 0\}$中所有元素的最大公因子为1(参见第三章第二节习题16). 容易验证注记5.2.7 中其它两条件对$\{a_n\}$, $\{b_n\}$也成立, 因此

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \lim_{n\to\infty} a_n = \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=1}^{\infty} nc_n} = \frac{\sum\limits_{n=0}^{\infty} f_{i,j}^{(n)}}{\sum\limits_{n=1}^{\infty} nf_{j,j}^{(n)}} = \begin{cases} 0, & j\text{是0常返}, \\ f_{i,j}/m_{j,j}, & j\text{是正常返}. \end{cases} \qquad \square$$

♦**推论5.2.9.** (初等更新定理) $\lim\limits_{t\to\infty} m(t)/t = 1/\mu$.

证明 当F是连续分布时, 对任意$a > 0$, 令$h(s) = 1_{[0,a]}(s)$, 那么$h(s)$为非负单调不增函数, 由关键更新定理可知

$$\lim_{t\to\infty}(m(t) - m(t-a)) = \lim_{t\to\infty}\int_0^t h(t-s)dm(s) = \frac{a}{\mu}.$$

从而对任意非负整数N,

$$\lim_{N\to\infty}\frac{m(N)}{N} = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}(m(k) - m(k-1)) = \frac{1}{\mu}.$$

又因为$m(t)$单调不降, 由

$$\frac{m([t])}{t} \leq \frac{m(t)}{t} \leq \frac{m([t]+1)}{t},$$

可得推论成立.

当F是格子点分布时, 利用本节习题8的结论, 同样可证明推论成立. 细节请读者自己补充. $\square$

▸**例5.2.10.** 设更新过程N的间隔时间分布F连续, 且均值μ与方差σ^2都有限. $L(t)$ 是N 在时刻t 的总寿命. 求$\lim\limits_{t\to\infty} E(L(t))$.

解 由本章第一节习题4的结果可知

$$E(L(t))=\int_t^\infty s\mathrm{d}F(s)+\int_0^t\int_{t-s}^\infty u\mathrm{d}F(u)\mathrm{d}m(s).$$

令$h(t)=\int_t^\infty s\mathrm{d}F(s)$. 显然$h(t)$是单调不增的, 而且

$$\int_0^\infty h(t)\mathrm{d}t=\int_0^\infty\int_t^\infty u\mathrm{d}F(u)\mathrm{d}t=\int_0^\infty u^2\mathrm{d}F(u)=\sigma^2+\mu^2<\infty.$$

因此由关键更新定理可得

$$\lim_{t\to\infty}E(L(t))=\frac{\int_0^\infty h(t)\mathrm{d}t}{\mu}=\mu+\frac{\sigma^2}{\mu}.\qquad\square$$

上例计算结果表明, 随着$t\to\infty$, t时刻总寿命的均值收敛到一个严格大于平均间隔时间的极限. 这从另外一个角度验证了我们在上节中提到的长度偏离取样(检验悖论)现象.

▶**例5.2.11.** *设顾客按强度为λ的泊松过程来到某个只有一个服务窗口的服务台, 设服务台服务每个顾客所需时间是独立同分布的连续型随机变量, 分布为G, 期望为$1/\mu$, 其中$\mu>\lambda$. 若顾客到达发现需要等待则立刻离开, 否则接受服务; 问长时间后系统空闲的比例.*

[分析] 从0时间开始观察, 服务台空闲、忙碌再空闲, 这样周而复始. 记第i个周期中的忙碌时长为Y_i, 空闲时长为U_i. 由假设可知Y_i 就是第i 个顾客的服务时间, 相互独立且$E(Y_i)=1/\mu$; U_i就是从第$i-1$个顾客服务完到第i个顾客到达的时间间隔, 由指数分布的无记忆性, U_i服从参数为λ 的指数分布, $E(U_i)=1/\lambda$. 若以一个空闲与忙碌为周期, 那么观察到的服务系统就是一个更新过程. 该更新系统的等待时长$W_i=Y_i+U_i$, 平均等待间长为$1/\lambda+1/\mu$. 令$A_t=\{t\text{时刻系统空闲}\}$, 所求比例为$\lim\limits_{t\to\infty}P(A_t)$.

解 如上定义Y_i,U_i,W_i. W_i服从均值为$1/\mu+1/\lambda$的连续分布(记作F). 那么

$$\begin{aligned}f(t)=P(A_t)&=E(P(A_t|T_1))=\int_0^\infty P(A_t|T_1=s)\mathrm{d}F(s)\\&=\int_t^\infty P(U_1>t|W_1=s)\mathrm{d}F(s)+\int_0^t P(A_{t-s})\mathrm{d}F(s).\end{aligned}$$

注意到$\int_t^\infty P(U_1>t|W_1=s)\mathrm{d}F(s)=P(U_1>t,W_1>t)=P(U_1>t)$, 因此

$$f(t)=P(U_1>t)+\int_0^t f(t-s)\mathrm{d}F(s).$$

由关键更新定理,

$$\lim_{t\to\infty}f(t)=\frac{\int_0^\infty P(U_1>t)\mathrm{d}t}{1/\lambda+1/\mu}=\frac{\mu}{\lambda+\mu}.\qquad\square$$

练习题5.2

1 两台机器独立不间断地加工零件, 若其中一台加工零件的时间服从参数

为$1/2$ 的指数分布, 另外一台服从$(0,4)$ 上的均匀分布, 求到时间$t=100$ 时两台机器一起至少加工90 个零件的概率估计.

2 设N是间隔时间分布为连续分布的更新过程, 平均间隔时间为μ, $A(t)$表示在时刻t的年龄, 对任意$c>0$, 问长时间后事件$\{A(t)<c\}$ 发生的概率.

3 令$L(t)$表示时刻t更新过程的总寿命, 若平均间隔时间$\mu<\infty$, 证明

$$\lim_{t\to\infty} L(t)/t=0, \qquad a.s.$$

4 设更新过程N的间隔时间分布F连续且均值μ与方差σ^2都有限, 证明

$$\lim_{t\to\infty}(m(t)-t/\mu)=(\sigma^2-\mu^2)/2\mu^2.$$

5 设更新过程N的间隔时间分布F连续且均值μ与方差σ^2都有限. $A(t)$是N在时刻t的年龄, 求$\lim\limits_{t\to\infty} E(A(t))$.

6 设X是状态空间为S, 周期为d的不可约正常返时齐马氏链, 记平稳分布为$\pi=(\pi_i, i\in S)$. 任取$i\in S$. 对任意$0\le k<d$, 求$\lim\limits_{n\to\infty} p_{i,i}^{(nd+k)}$.

7* 设$g(t)$是积分方程$G(t)=\mathrm{e}^{-t}+\int_0^t G(t-s)\mathrm{e}^{-2s}\mathrm{d}s$的最小非负解, 求$\lim\limits_{t\to\infty} g(t)$.

8* 设F是步长为d的格子点分布, 数学期望为μ, $m(t)$是以F为间隔时间分布的更新函数, 证明

$$\lim_{n\to\infty}[m(nd)-m((n-1)d)]=d/\mu.$$

5.3 两类广义更新过程

更新过程可做很多推广, 本节简要介绍更新报酬过程与可终止更新过程.

5.3.1 更新报酬过程

♣**定义5.3.1.** 设$\{(W_n, Y_n); n \geq 1\}$为独立同分布的二维随机变量序列, $W_n \geq 0$, $N = \{N(t); t \geq 0\}$是以$\{W_n\}$为时间间隔序列的更新过程. 对任意$t \geq 0$, 令

$$X(t) = \sum_{n=1}^{N(t)} Y_n,$$

称$X = \{X(t); t \geq 0\}$为更新报酬过程或有偿更新过程.

容易看出, 若N为泊松过程且$\{Y_n; n \geq 1\}$与$\{W_n; n \geq 1\}$独立, 则X就是复合泊松过程.

★**定理5.3.2.** 设X为更新报酬过程, $E(W_1) < \infty$, $E|Y_1| < \infty$. 那么

$$\lim_{t\to\infty} \frac{X(t)}{t} = \frac{E(Y_1)}{E(W_1)}, \quad a.s. \text{ 而且 } \lim_{t\to\infty} \frac{E(X(t))}{t} = \frac{E(Y_1)}{E(W_1)}.$$

证明 由强大数定理及$N(t) \to \infty$, a.s. 可知

$$\frac{X(t)}{t} = \frac{\sum_{k=1}^{N(t)} Y_k}{N(t)} \frac{N(t)}{t} \to \frac{E(Y_1)}{E(W_1)}, \quad a.s.$$

定理前半部分极限成立. 对后半部分极限, 注意到

$$\begin{aligned}
E\Big(\sum_{k=1}^{N(t)+1} Y_k\Big) &= \sum_{n=1}^{\infty} E\Big(\sum_{k=1}^{n} Y_k 1_{\{N(t)+1=n\}}\Big) \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{n} E\Big(Y_k 1_{\{\sum\limits_{i=1}^{n-1} W_i \leq t, \sum\limits_{i=1}^{n} W_i > t\}}\Big) \\
&= \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} E\Big(Y_k 1_{\{\sum\limits_{i=1}^{n-1} W_i \leq t, \sum\limits_{i=1}^{n} W_i > t\}}\Big) \\
&= \sum_{k=1}^{\infty} E\Big(Y_k - Y_k \sum_{n=1}^{k-1} 1_{\{\sum\limits_{i=1}^{n-1} W_i \leq t, \sum\limits_{i=1}^{n} W_i > t\}}\Big).
\end{aligned}$$

由$\{(W_n, Y_n); n \geq 1\}$的独立同分布性可得

$$\begin{aligned}
E\Big(\sum_{k=1}^{N(t)+1} Y_k\Big) &= \sum_{k=1}^{\infty} E(Y_k)\Big(1 - \sum_{n=1}^{k-1} P(\sum_{i=1}^{n-1} W_i \leq t, \sum_{i=1}^{n} W_i > t)\Big) \\
&= \sum_{k=1}^{\infty} E(Y_k)(1 - P(N(t)+1 \leq k-1))
\end{aligned}$$

$$= \sum_{k=1}^{\infty} E(Y_k) P(N(t)+1 \geq k)$$
$$= E(Y_1) \sum_{k=1}^{\infty} P(N(t)+1 \geq k)$$
$$= E(Y_1)(1+m(t)). \tag{5.3.1}$$

进而

$$\lim_{t\to\infty} \frac{1}{t} E\Big(\sum_{i=1}^{N(t)+1} Y_i\Big) = \lim_{t\to\infty} \frac{g(t)}{t} = E(Y_1) \lim_{t\to\infty} \frac{1+m(t)}{t} = \frac{E(Y_1)}{E(W_1)}.$$

另一方面,

$$\begin{aligned} E(Y_{N(t)+1}) &= \int_0^{\infty} E(Y_{N(t)+1} | T_1 = s) \mathrm{d}F(s) \\ &= \int_t^{\infty} E(Y_1 | W_1 = s) \mathrm{d}F(s) + \int_0^t E(Y_{N(t-s)+1}) \mathrm{d}F(s). \end{aligned}$$

记$B(t) = E(Y_{N(t)+1})$, $b(t) = \int_t^{\infty} E(Y_1 | W_1 = s)\mathrm{d}F(s)$, 则

$$B(t) = b(t) + \int_0^t B(t-s) \mathrm{d}F(s).$$

由

$$|b(t)| \leq \int_0^{\infty} E(|Y_1| | W_1 = s) \mathrm{d}F(s) = E(|Y_1|) < \infty,$$

可知$b(t)$为局部有界函数, 且由(5.3.1)易知$B(t)$也是局部有界函数, 因此

$$B(t) = b(t) + \int_0^t b(t-s) \mathrm{d}m(s).$$

又由$t \to \infty$时$b(t) \to 0$可知, 对任意$\varepsilon > 0$, 存在T, 当$t > T$时, $|b(t)| \leq \varepsilon$. 所以

$$\begin{aligned} 0 \leq \frac{1}{t} \int_0^t |b(t-s)| \mathrm{d}m(s) &= \frac{1}{t} \Big[\int_0^{t-T} + \int_{t-T}^t \Big] |b(t-s)| \mathrm{d}m(s) \\ &\leq \varepsilon \frac{m(t-T)}{t} + E(|Y_1|) \frac{m(t) - m(t-T)}{t}. \end{aligned}$$

先令$t \to \infty$再令$\varepsilon \to 0$, 由推论5.2.9可得

$$\lim_{t\to\infty} \frac{B(t)}{t} = 0.$$

最后注意到$E(X(t)) = E\Big(\sum_{i=1}^{N(t)+1} Y_i\Big) - E(Y_{N(t)+1})$, 综合以上结果可得

$$\lim_{t\to\infty} \frac{E(X(t))}{t} = \frac{E(Y_1)}{E(W_1)}.$$

□

▶**例5.3.3.** 设手机寿命为Z年, Z服从连续分布H. 某人使用手机, 若原来手机用了T年还没坏则花C元换一部新手机,若在T年前损坏, 则用D元换新手机. 问长时间后, 此人平均每年多少钱用于手机更换?

解 可用更新报酬过程来描述这一问题. 令

$$W_n = Z_n \wedge T = \min\{Z_n, T\},$$

其中Z_n表示第n部手机的寿命, N为以$\{W_n\}$为更新时间的更新过程. 那么

$$Y_n = C1_{\{W_n=T\}} + D1_{\{W_n<T\}}$$

为第n次手机更新的费用. 在时刻t之前总的手机更新费用为

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

平均费用为

$$\lim_{t\to\infty} \frac{X(t)}{t} = \frac{E(Y_1)}{E(W_1)} = \frac{C + (D-C)H(T-)}{\int_0^T \bar{H}(x)dx},$$

其中$H(T-) = P(W_n < T)$. □

▶**例5.3.4.** 设顾客按强度为λ的泊松过程来到某个只有一个服务窗口的服务台, 设服务台服务每个顾客所需时间是独立同分布的随机变量, 分布为G, 期望为$1/\mu$, $\mu > \lambda$, 不论顾客到达后是否需要等待, 顾客都在接受服务后才离开. 从长期看, 顾客到达即可接受服务的平均概率是多少?

解 可用更新报酬过程来描述所讨论问题. 假定系统从第一个顾客到达开始, 显然服务台按"工作"和"空闲"两种状态更新循环运行. 以X_i表示第i 次系统忙的时长, Y_i表示第i次系统闲的时长, 那么$W_i = X_i + Y_i$就是系统一个服务周期的时长. 按假定, (X_i, Y_i)为独立同分布的二维随机变量. 特别地, Y_i 为第i个忙期结束到下一个顾客到达的时间, 由题设, 服从参数为λ 的指数分布. 在$[0,t]$时段内, 顾客到达即能接受服务的时长为

$$T(t) = \sum_{k=1}^{N(t)} Y_i + (t - T_{N(t)} - X_{N(t)+1})1_{\{t>T_{N(t)}+X_{N(t)+1}\}},$$

其中N为以$\{W_i\}$为更新时间的更新过程, 上式右边后一项表示第$N(t)+1$个服务周期中在t之前的空闲时间. 显然

$$\sum_{i=1}^{N(t)} Y_i \le T(t) \le \sum_{i=1}^{N(t)+1} Y_i.$$

因此所求平均概率总有

$$\lim_{t\to\infty} \frac{T(t)}{t} = \frac{E(Y_1)}{E(W_1)} = \frac{1/\lambda}{E(X_1) + 1/\lambda}.$$

以U表示服务第一个顾客的时间, V表示服务第一个顾客时到达的顾客数. 显然U服从分布G, 而V为强度为λ的泊松过程在$[0,U]$时间内粒子到达数. 所以

$$E(U) = \frac{1}{\mu}, \quad E(V) = \int_0^\infty \lambda u \mathrm{d}G(u) = \frac{\lambda}{\mu}.$$

由于服务顾客是独立同分布的, $E(X_1|U,V)=U+VE(X_1)$, 进而

$$E(X_1)=E(U)+E(X_1)E(V)=\frac{1}{\mu}+E(X_1)\frac{\lambda}{\mu}.$$

由此可得$E(X_1)=1/(\mu-\lambda)$. 所以

$$\lim_{t\to\infty}\frac{T(t)}{t}=1-\frac{\lambda}{\mu}.$$ □

▶**例5.3.5.** *设保险公司理赔事件的发生构成强度为λ的泊松过程N, 每次理赔金额Y_i服从参数为μ的指数分布而且相互独立. 进一步, 假定理赔金额序列$\{Y_i\}$与泊松过程N独立. 设公司初始资本为x, 单位时间内的保费收入为c. 问该公司永不破产的概率? (公司破产就是指公司净值产为负) 这里假定$c>\lambda/\mu$.*

解 在时间区间$[0,t]$内, 公司需支付的理赔费用为$\sum_{i=1}^{N(t)}Y_i$, 从而t时刻公司资产为

$$W(t)=x+ct-\sum_{i=1}^{N(t)}Y_i.$$

公司破产当且仅当事件$A=\{\omega;\text{存在}t>0\text{使得}W(t,\omega)<0\}$ 发生. 注意到W的轨道的右连续性质, 由有理数在实数上的稠密性可得

$$A=\bigcup_{t\text{为非负有理数}}\Big\{\omega;W(t,\omega)<0\Big\}.$$

显然对任意$t\geq 0$, $\{\omega;W(t,\omega)<0\}$都是一个随机事件, 进而A也是一个随机事件. 记公司初始资本为x时永不破产的概率为$B(x)$, 那么当$x\geq 0$时,

$$B(x)=1-P(A)=P(x+ct-\sum_{i=1}^{N(t)}Y_i\geq 0,t\geq 0),$$

并约定$B(x)=0,x<0$. 对任意$x\geq 0$, 由重期望公式

$$B(x)=\int P(W(t)\geq 0,t\geq 0|W_1=s,Y_1=y)\mathrm{d}H(s,y),$$

其中$H(s,y)$为(W_1,Y_1)的联合分布. 若已知$W_1=s,Y_1=y$, 那么在s之前公司资本是逐渐累加的, 在s时刻由于理赔支出, 资本变成$x+cs-y$. 显然

$$P(W(t)\geq 0,t\geq 0|W_1=s,Y_1=y)=\begin{cases}0, & y>x+cs,\\ B(x+cs-y), & y\leq x+cs.\end{cases}$$

由假设$\mathrm{d}H(s,y)=\lambda\mathrm{e}^{-\lambda s}\mu\mathrm{e}^{-\mu y}\mathrm{d}s\mathrm{d}y$, 因此,

$$\begin{aligned}B(x)&=\int_0^\infty\lambda\mathrm{e}^{-\lambda s}\mathrm{d}s\int_0^{x+cs}B(x+cs-y)\mu\mathrm{e}^{-\mu y}\mathrm{d}y\\&=\int_x^\infty\frac{\lambda}{c}\mathrm{e}^{-\frac{\lambda}{c}(u-x)}\mathrm{d}u\int_0^u B(u-y)\mu\mathrm{e}^{-\mu y}\mathrm{d}y.\end{aligned}$$

两边关于x求导得

$$B'(x)=\frac{\lambda}{c}B(x)-\frac{\lambda}{c}\int_0^x B(y)\mu\mathrm{e}^{-\mu(x-y)}\mathrm{d}y, \tag{5.3.2}$$

进一步关于x求导, 整理得$(\frac{\lambda}{c}-\mu)B'(x)=B''(x)$. 由此可得

$$B(x)=b+\frac{a}{\frac{\lambda}{c}-\mu}\mathrm{e}^{(\frac{\lambda}{c}-\mu)x},$$

其中a,b为待定常数. 由定理14.1可知

$$\lim_{t\to\infty}\frac{W(t)}{t}=c-\frac{\lambda}{\mu}>0,$$

从而$\lim\limits_{x\to\infty}B(x)=1$(证明省略). 因此

$$B(x)=1+\frac{a}{\frac{\lambda}{c}-\mu}\mathrm{e}^{(\frac{\lambda}{c}-\mu)x}.$$

将其代入方程(5.3.2)得$a=-\frac{\lambda}{c\mu}(\frac{\lambda}{c}-\mu)$. 所以

$$B(x)=1-\frac{\lambda}{c\mu}\mathrm{e}^{(\frac{\lambda}{c}-\mu)x}.$$

□

5.3.2 可终止更新过程

♣**定义5.3.6.** *若更新过程间隔时间分布F以正概率取$+\infty$, 那么我们称此更新过程为可终止更新过程.*

为了叙述简单, 与注记5.1.6一样我们要求间隔时间分布F限制在$[0,\infty)$上时, 或者是离散格子点分布, 或者具有局部有界的概率密度函数.

▶**例5.3.7.** *设j是马氏链X的非常返状态, $N(t)$表示从j出发的马氏链X到时刻t为止回到状态j的次数, 那么$\{N(t);t\geq 0\}$就是可终止更新过程.*

由定义可知可终止更新过程仍有更新性, 其更新函数$m(t)=E(N(t))$仍满足更新方程

$$m(t)=F(t)+\int_0^t m(t-s)\mathrm{d}F(s),$$

并且更新方程

$$B(t)=b(t)+\int_0^t B(t-s)\mathrm{d}F(s)$$

的局部有界解仍可表示成

$$B(t)=b(t)+\int_0^t b(t-s)\mathrm{d}m(s).$$

但对可终止更新过程而言, 事件发生若干次之后会以正概率不再发生. 记这最后一次事件发生的时刻为ξ, 即

$$\xi=\sup\{T_n;T_n<\infty,\ n\geq 0\}.$$

那么$N(\infty)=N(\xi)$就是更新事件发生的总次数.

★**定理5.3.8.** 设$p = F(\infty)$, 即$1-p$是F在$+\infty$的概率. 对以F为间隔时间分布的可终止更新过程N而言,

$$P(\xi \le t) = (1-p)(1+m(t)), \quad P(N(\infty)=k) = p^k(1-p), \ \ k \ge 0,$$

以及

$$E(\xi) = \frac{1}{1-p}\int_0^\infty (p - F(t))\mathrm{d}t, \quad m(\infty) = E(N(\infty)) = \frac{p}{1-p}.$$

证明 记N的第n个间隔时间为W_n. 显然

$$\{N(\infty)=k\} = \{W_1 < \infty, \cdots, W_k < \infty, W_{k+1} = \infty\}.$$

因此$P(N(\infty)=k) = p^k(1-p)$, 进而$E(N(\infty)) = p/(1-p)$.

注意到$W_1 = \infty$时$\xi = 0$, 因此对任意$t \ge 0$, 由更新技巧可得

$$P(\xi \le t) = E(P(\xi \le t|W_1)) = 1 - p + \int_0^\infty P(\xi \le t|W_1 = s)\mathrm{d}F(s)$$
$$= 1 - p + \int_0^t P(\xi \le t - s)\mathrm{d}F(s).$$

因此$P(\xi \le t) = (1-p)(1+m(t))$. 同理

$$E(\xi) = E(E(\xi|W_1)) = \int_0^\infty (s + E(\xi))\mathrm{d}F(s) = \int_0^\infty s\mathrm{d}F(s) + pE(\xi).$$

因此

$$E(\xi) = \frac{\int_0^\infty s\mathrm{d}F(s)}{1-p} = \frac{1}{1-p}\int_0^\infty (p - F(t))\mathrm{d}t. \qquad \square$$

由于$N(t)$单调不降地几乎必然收敛到$N(\infty)$, 由单调收敛定理(定理1.4.23)可知, $m(t)$单调不降地收敛到$m(\infty) = p/(1-p)$, 这与通常更新过程$m(t) \to \infty$完全不同.

■**性质5.3.9.** 设F为非格子点分布, $p = F(\infty) < 1$且存在$\lambda > 0$使得

$$\int_0^\infty \mathrm{e}^{\lambda t}\mathrm{d}F(t) = 1.$$

若存在$M > 0$使得$\mathrm{e}^{\lambda t}(p - F(t))$在$t \in [M, \infty)$上单调不增, $m(t)$是以F为间隔时间分布的更新函数, 那么

$$\lim_{t\to\infty} e^{\lambda t}[m(t) - m(\infty)] = -\frac{1}{\lambda}\Big[\int_0^\infty t\mathrm{e}^{\lambda t}\mathrm{d}F(t)\Big]^{-1}.$$

证明 由$m(t) = F(t) + \int_0^t m(t-s)\mathrm{d}F(s)$以及

$$m(\infty) = \frac{p}{1-p} = \frac{p}{1-p}(1 - F(t)) + \int_0^t \frac{p}{1-p}\mathrm{d}F(s),$$

可知

$$\mathrm{e}^{\lambda t}[m(\infty) - m(t)] = \mathrm{e}^{\lambda t}\frac{p - F(t)}{1-p} + \int_0^t \mathrm{e}^{\lambda(t-s)}(m(\infty) - m(t-s))\mathrm{e}^{\lambda s}\mathrm{d}F(s).$$

注意到

$$\int_0^\infty \mathrm{e}^{\lambda t}\frac{p-F(t)}{1-p}\mathrm{d}t=-\frac{1}{\lambda}\frac{p}{1-p}+\int_0^\infty\frac{1}{\lambda(1-p)}\mathrm{e}^{\lambda t}\mathrm{d}F(s)=\frac{1}{\lambda},$$

由关键更新定理可知

$$\lim_{t\to\infty}\mathrm{e}^{\lambda t}[m(\infty)-m(t)]=\frac{1}{\int_0^\infty\lambda s\mathrm{e}^{\lambda s}\mathrm{d}F(s)}.$$ □

▶**例5.3.10.** 设动物穿过某一条保护区的公路时构成一个更新过程, 其更新间隔$W_i, i\geq 1$的公共分布为F, 在该保护区行驶的汽车碰到动物穿过马路时一定要停车等待, 直到前后动物间隔时间大于τ 时才能通过. 假定某汽车恰好碰到动物穿过马路, 求它需要等待的平均时间.

解 令$\tilde{W}_i=\begin{cases}W_i, & W_i\leq\tau,\\ +\infty, & W_i>\tau.\end{cases}$ $\tilde{N}$是以$\tilde{W}_i$为间隔时间的可终止更新过程, ξ为该可终止更新过程的最后更新时刻, 那么所求平均时间

$$t=E(\xi)+\tau=\frac{1}{1-F(\tau)}\int_0^\tau(F(\tau)-F(t))\mathrm{d}t+\tau=\frac{1}{1-F(\tau)}\int_0^\tau t\mathrm{d}F(t)+\tau.$$ □

练习题5.3

1 假设当计数器接收到一个粒子后将封闭长为τ的时段, 若在这个期间又有粒子到来, 从粒子到达时刻起, 计数器还将封闭τ时段, 直至在长为τ的时段内没有粒子到达, 此时计数器才重新开放计算. 设粒子到达服从等待分布为F的更新过程, 求计数器在一个周期内被封闭的平均时长.

2 三个射手轮流射击一个目标, 射手1射击直到他未中, 然后射手2射击直到他未中, 然后射手3射击直到他未中, 然后又回到射手1, 如此循环. 假设三个射手射中的概率分别为p_1,p_2,p_3, 确定每个射手射击次数的长程比例.

3 设N是间隔时间分布为F的更新过程, 设F的均值和方差都有限, 分别是μ, σ^2. $A(t)$表示t时刻的年龄, 求$\lim\limits_{t\to\infty}\int_0^t A(s)\mathrm{d}s/t$.

4* 更新报酬过程X满足$E(W_i)<\infty$, $E|Y_i|<\infty$, 其中W_i 服从连续分布. 令$m(t)=E(X(t))$, 证明对任意$a>0$,

$$\lim_{t\to\infty}(m(t+a)-m(t))=\frac{aE(Y_1)}{E(W_1)}.$$

第六章　布朗运动

作为一种物理现象, 布朗运动由19世纪20年代英国植物学家Robert Brown在观察悬浮在液体表面的花粉粒子的运动轨迹时发现. 1905年Einstein给出了布朗运动的物理学解释和方程刻画, 其深层次的数学理论基础与构造则由Wiener在1923年得到. 到目前, 布朗运动已成为运用最广泛的数学模型之一, 在自然科学、金融学、社会科学等领域有大量应用.

6.1　布朗运动及其分布

6.1.1　布朗运动

我们首先定义布朗运动如下:

♣**定义6.1.1.** 记$I=[0,+\infty)$或$[0,M]$, 其中M为正实数. 称实值随机过程$B=\{B(t);\ t\in I\}$ 为布朗运动或*Wiener*过程, 若它满足

(a) 独立增量性: $B(0)=0$ 且B是独立增量过程, 即对任意$n\geq 2$以及$0\leq t_1<t_2<\cdots<t_n\in I$, 增量
$$B(t_1),B(t_2)-B(t_1),\cdots,B(t_n)-B(t_{n-1})$$
相互独立.

(b) 正态性: 对任意$0\leq s<t\in I$, $B(t)-B(s)\sim N(0,\sigma^2(t-s))$, 其中$\sigma^2$ 为正常数.

(c) 连续性: B具有连续轨道, 即对任意ω, $t\mapsto B(t,\omega)$是连续函数.

通常我们称正常数σ^2为布朗运动B的方差参数. 称方差参数为1的布朗运动为标准布朗运动. 容易看出, 若B是方差参数为σ^2的布朗运动, 那么B/σ为标准布朗运动. 若$B_x(t)=x+B(t)$, 称$B_x=\{B_x(t);\ t\in I\}$为从x出发的布朗运动.

在一定意义上, 布朗运动B可以通过简单对称随机游动在时间和空间上采取一定方式连续化后得到. 正如我们在第二章第二节所指出的, 简单对称随机游

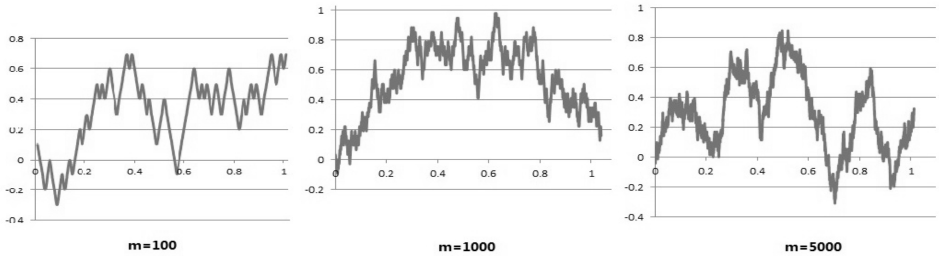
动$S(n)=\sum\limits_{k=1}^{n}X_k$可以用来表示每单位时间向左或向右等可能地跳一个单位的粒子的轨迹变化. 令

$$S_m(t)=\sum_{k=1}^{[mt]}\frac{1}{\sqrt{m}}X_k.$$

那么$S_m(t)$记录了每$1/m$单位时间向左或向右(各1/2可能)跳$1/\sqrt{m}$单位距离的粒子的轨迹变化. $m\to\infty$则意味着, 粒子跳跃的越来越频繁, 跳跃幅度也越来越小. 对任意固定的时刻t, 中心极限定理保证了

$$S_m(t)=\frac{1}{\sqrt{m}}\sum_{k=1}^{[mt]}X_k=\frac{\sqrt{[mt]}}{\sqrt{m}}\frac{1}{\sqrt{[mt]}}\sum_{k=1}^{[mt]}X_k\to W(t),$$

其中$W(t)$是一个服从正态分布$N(0,t)$的随机变量. 更严格的数学分析表明, 随机过程$\{S_m(t);t\geq 0\}$随着$m\to\infty$会收敛到标准布朗运动$\{B(t);t\geq 0\}$. 下面分别是$m=100,1000,5000$时, S_m的轨道局部($t\in[0,1]$). 当m比较大时, 可以把S_m看作是标准布朗运动的近似.



由定义中的(a),(b)两条件可知, 布朗运动B是平稳独立增量过程, 而且增量服从均值为0的正态分布. 直接计算可知

■**性质6.1.2.** *方差参数为σ^2的布朗运动B的均值函数$m(t)\equiv 0$; 方差函数$D(t)=\sigma^2 t$; 协方差函数$Cov(s,t)=(s\wedge t)\sigma^2$.*

证明: 只需验证协方差函数. 不妨设$s<t$,

$Cov(s,t)=E(B(s)B(t))=E(B(s)(B(t)-B(s))+E(B^2(s))=\sigma^2 s.$ □

▶**例6.1.3.** *$t_1,t_2,t_3\in I$且互不相同, 求$E(B(t_1)B(t_2)B(t_3))$, 其中$\{B(t);t\in I\}$是布朗运动.*

解 不妨设$0\leq t_1<t_2<t_3$. 由独立增量与正态性可知

$$\begin{aligned}&E(B(t_1)B(t_2)B(t_3))\\&\quad=E\big(B(t_1)\big(B(t_1)+B(t_2)-B(t_1)\big)\big(B(t_2)+B(t_3)-B(t_2)\big)\big)\\&\quad=E\big(B(t_1)B(t_1)B(t_2)\big)+E\big(B(t_1)\big(B(t_2)-B(t_1)\big)B(t_2)\big)\\&\quad=E\big(B^3(t_1)\big)+E\big(B(t_1)\big(B(t_2)-B(t_1)\big)^2\big)=0.\end{aligned}$$ □

★**定理6.1.4.** 若$\{B(t);t\geq 0\}$为标准布朗运动, 那么以下过程也是标准布朗运动.

(1) $\{-B(t);\ t\geq 0\}$.

(2) $\{B(t+s)-B(s);\ t\geq 0\}$, 其中$s$为任意给定非负数.

(3) $\{cB(t/c^2);\ t\geq 0\}$ 其中c为任意给定正数.

(4) $\{B(u)-B(u-t);\ 0\leq t\leq u\}$ 其中u为给定正数.

证明 按布朗运动定义, 不难验证上述结论成立. 下面以(4)为例解释如下: 令

$$X(t)=B(u)-B(u-t),\quad 0\leq t\leq u.$$

(a) 显然$X(0)=0$且X轨道连续.

(b) 对任意$0\leq t_1<t_2<\cdots<t_n\leq u$,

$$\begin{aligned}&(X(t_1),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1}))\\&=(B(u)-B(u-t_1),B(u-t_1)-B(u-t_2),\cdots,B(u-t_{n-1})-B(u-t_n)).\end{aligned}$$

由布朗运动独立增量性质, 显然相互独立.

(c) 对任意$0\leq s<t\leq u$,

$$X(t)-X(s)=B(u-s)-B(u-t)\sim N(0,t-s).$$

因此$\{X(t);0\leq t\leq u\}=\{B(u)-B(u-t);0\leq t\leq u\}$为标准布朗运动. $\square$

♠**注记6.1.5.** (2), (3), (4) 分别表示对标准布朗运动做起点变换, 尺度变换以及时间反向变换后仍是标准布朗运动. 此外我们还可以证明$\{tB(1/t);\ t\geq 0\}$ (其中$t=0$时取值为0) 也是标准布朗运动.

6.1.2 布朗运动的分布

下面我们研究布朗运动的有限维分布. 不失一般性, 本小节我们总设$B=\{B(t);\ t\geq 0\}$为标准布朗运动, 并记

$$\phi(x,t)=\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{x^2}{2t}}.$$

那么对任意$a\in\mathbb{R}$, $t>0$, $\phi(x-a,t)$是正态分布$N(a,t)$的概率密度函数.

★**定理6.1.6.** 对任意$n\geq 1$, $0\leq t_1<t_2<\cdots<t_n$, $(B(t_1),B(t_2),\cdots,B(t_n))$是

均值为$\mathbf{0}$, 协方差矩阵为

$$\mathbf{D}=(t_i\wedge t_j)_{i,j=1,\cdots n}=\begin{pmatrix} t_1 & t_1 & t_1 & \cdots & t_1\\ t_1 & t_2 & t_2 & \cdots & t_2\\ t_1 & t_2 & t_3 & \cdots & t_3\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ t_1 & t_2 & t_3 & \cdots & t_n \end{pmatrix}$$

的n元正态随机变量. 因此$(B(t_1),\cdots,B(t_n))$的联合概率密度函数为

$$\rho(x_1,\cdots,x_n;t_1,\cdots,t_n)=\phi(x_1,t_1)\prod_{i=2}^{n}\phi(x_i-x_{i-1},t_i-t_{i-1}).$$

证明 对任意$0\le t_1<t_2<\cdots<t_n$, $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$, 令$t_0=0$. 由布朗运动平稳独立的正态增量性质可知

$$\begin{aligned}
&E\Big(\exp\Big\{\mathrm{i}\sum_{k=1}^{n}\lambda_k B(t_k)\Big\}\Big)\\
&=E\Big(\exp\Big\{\mathrm{i}\sum_{k=1}^{n}\lambda_k\sum_{j=1}^{k}[B(t_j)-B(t_{j-1})]\Big\}\Big)\\
&=E\Big(\exp\Big\{\mathrm{i}\sum_{k=1}^{n}(B(t_k)-B(t_{k-1}))\sum_{j=k}^{n}\lambda_j\Big\}\Big)\\
&=\prod_{k=1}^{n}E\Big(\exp\Big\{\mathrm{i}(B(t_k)-B(t_{k-1}))\sum_{j=k}^{n}\lambda_j\Big\}\Big)\\
&=\exp\Big\{-\sum_{k=1}^{n}\frac{t_k-t_{k-1}}{2}\Big(\sum_{j=k}^{n}\lambda_j\Big)^2\Big\}.
\end{aligned}$$

直接验证可知

$$\begin{aligned}
\sum_{k=1}^{n}\frac{t_k-t_{k-1}}{2}\Big(\sum_{j=k}^{n}\lambda_j\Big)^2&=\frac{1}{2}\Big[\sum_{k=1}^{n}t_k\Big(\sum_{j=k}^{n}\lambda_j\Big)^2-\sum_{k=1}^{n-1}t_k\Big(\sum_{j=k+1}^{n}\lambda_j\Big)^2\Big]\\
&=\frac{1}{2}\Big[\sum_{k=1}^{n}2t_k\lambda_k\sum_{j=k+1}^{n}\lambda_j+\sum_{k=1}^{n}t_k\lambda_k^2\Big]\\
&=\frac{(\lambda_1,\cdots,\lambda_n)\mathbf{D}(\lambda_1,\cdots,\lambda_n)^T}{2}.
\end{aligned}$$

因此$(B(t_1),\cdots,B(t_n))$服从均值为$\mathbf{0}$, 协方差矩阵为$\mathbf{D}$的正态分布, 联合概率密度函数

$$\rho(x_1,\cdots,x_n;t_1,\cdots,t_n)=\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{D}|}}\mathrm{e}^{-\frac{(x_1,\cdots,x_n)\mathbf{D}^{-1}(x_1,\cdots,x_n)^T}{2}}.$$

由矩阵求逆计算可知$\mathbf{D}^{-1}$可表示为

$$\begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{t_1} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{t_2-t_1} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{t_{n-1}-t_{n-2}} & 0 \\ 0 & 0 & \cdots & 0 & \frac{1}{t_n-t_{n-1}} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

令$x_0 = 0$, $t_0 = 0$, 将$\mathbf{D}^{-1}$的上述表达式代入联合概率密度函数得

$$\begin{aligned} \rho(x_1, \cdots, x_n; t_1, \cdots, t_n) &= \frac{1}{\prod_{i=1}^n \sqrt{2\pi(t_i - t_{i-1})}} \mathrm{e}^{-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}} \\ &= \phi(x_1, t_1) \prod_{i=2}^n \phi(x_i - x_{i-1}, t_i - t_{i-1}). \end{aligned}$$ □

有时候我们需要在一定条件下讨论布朗运动的分布, 为此我们先给出一个关于多元正态分布的一般性结果.

★定理6.1.7. 设$X = (X_1, \cdots, X_m, X_{m+1}, \cdots, X_n)$服从$n$元正态分布$N(\mu, \mathbf{D})$其中$\mu \in \mathbb{R}^n$, $\mathbf{D}$为n阶正定矩阵. 那么在给定

$$(X_{m+1}, \cdots, X_n) = \bar{y}^T = (\bar{y}_{m+1}, \cdots, \bar{y}_n)$$

条件下, $(X_1, \cdots, X_m)$ 服从m 元正态分布. 事实上, 记

$$\mu = \begin{pmatrix} \tilde{\mu} \\ \nu \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix},$$

其中$\tilde{\mu} \in \mathbb{R}^m$, $\nu \in \mathbb{R}^{n-m}$, $\mathbf{A}$, $\mathbf{C}$分别为m和$n-m$阶正定矩阵, 那么

$$(X_1, \cdots, X_m) \sim N(\tilde{\mu} + \mathbf{B}\mathbf{C}^{-1}(\bar{y} - \nu), \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T).$$

证明 为表示方便, 记

$$\Xi = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T, \quad \hat{\mu} = \tilde{\mu} + \mathbf{B}\mathbf{C}^{-1}(\bar{y} - \nu).$$

我们首先指出Ξ是正定矩阵. 事实上Ξ显然是对称矩阵. 若Ξ不是正定的, 那么存在$0 \neq z \in \mathbb{R}^m$ 使得$z^T \Xi z \leq 0$, 也即

$$z^T \mathbf{A} z \leq z^T \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T z.$$

此时令$x = \binom{z}{-\mathbf{C}^{-1}\mathbf{B}^T z}$, 那么

$$\begin{aligned} x^T \mathbf{D} x &= \begin{pmatrix} z \\ -\mathbf{C}^{-1}\mathbf{B}^T z \end{pmatrix}^T \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} \begin{pmatrix} z \\ -\mathbf{C}^{-1}\mathbf{B}^T z \end{pmatrix} \\ &= z^T \mathbf{A} z - 2 z^T \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T z + z^T \mathbf{B}\mathbf{C}^{-1}\mathbf{C}\mathbf{C}^{-1}\mathbf{B}^T z \\ &= z^T \mathbf{A} z - z^T \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T z \leq 0. \end{aligned}$$

这与$\mathbf{D}$正定矛盾. 因此Ξ是正定的.

由分块矩阵乘法可知

$$\begin{pmatrix} I & -\mathbf{B}\mathbf{C}^{-1} \\ 0 & \mathbf{C}^{-1} \end{pmatrix} \mathbf{D} = \begin{pmatrix} I & -\mathbf{B}\mathbf{C}^{-1} \\ 0 & \mathbf{C}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} = \begin{pmatrix} \Xi & 0 \\ \mathbf{C}^{-1}\mathbf{B}^T & I \end{pmatrix},$$

以及

$$\begin{pmatrix} \Xi^{-1} & -\Xi^{-1}\mathbf{B}\mathbf{C}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}^T\Xi^{-1} & \mathbf{C}^{-1}\mathbf{B}^T\Xi^{-1}\mathbf{B}\mathbf{C}^{-1} + \mathbf{C}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

由此可得$|\Xi| = |\mathbf{D}|/|\mathbf{C}|$而且

$$\mathbf{D}^{-1} = \begin{pmatrix} \Xi^{-1} & -\Xi^{-1}\mathbf{B}\mathbf{C}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}^T\Xi^{-1} & \mathbf{C}^{-1}\mathbf{B}^T\Xi^{-1}\mathbf{B}\mathbf{C}^{-1} + \mathbf{C}^{-1} \end{pmatrix}.$$

注意到n元随机向量X的概率密度函数为

$$f(x) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{D}|}} \exp\Big\{ -\frac{(x-\mu)^T\mathbf{D}^{-1}(x-\mu)}{2} \Big\}.$$

记$z = (x_1, \cdots, x_m)^T, y = (x_{m+1}, \cdots, x_n)^T$, 那么

$$\begin{aligned}
(x-\mu)^T\mathbf{D}^{-1}(x-\mu) &= \tbinom{z-\tilde{\mu}}{y-\nu}^T \mathbf{D}^{-1} \tbinom{z-\tilde{\mu}}{y-\nu} \\
&= (z-\tilde{\mu})^T\Xi^{-1}(z-\tilde{\mu}) - 2(z-\tilde{\mu})^T\Xi^{-1}\mathbf{B}\mathbf{C}^{-1}(y-\nu) \\
&\quad + (y-\nu)^T\mathbf{C}^{-1}\mathbf{B}^T\Xi^{-1}\mathbf{B}\mathbf{C}^{-1}(y-\nu) + (y-\nu)^T\mathbf{C}^{-1}(y-\nu) \\
&= (z-c(y))^T\Xi^{-1}(z-c(y)) + (y-\nu)^T\mathbf{C}^{-1}(y-\nu),
\end{aligned}$$

其中$c(y) = \tilde{\mu} + \mathbf{B}\mathbf{C}^{-1}(y-\nu)$. 因此$Y = (X_{m+1}, \cdots, X_n)$的边际概率密度函数

$$\begin{aligned}
g(y) &= \int_{\mathbb{R}^m} f(z, y)\mathrm{d}z \\
&= \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{C}|}} \exp\Big\{ -\frac{(y-\nu)^T\mathbf{C}^{-1}(y-\nu)}{2} \Big\} \\
&\quad \times \int_{\mathbb{R}^m} \frac{1}{\sqrt{|\Xi|}} \exp\Big\{ -\frac{(z-c(y))^T\Xi^{-1}(z-c(y))}{2} \Big\}\mathrm{d}z \\
&= \frac{1}{(2\pi)^{(n-m)/2}\sqrt{|\mathbf{C}|}} \exp\Big\{ -\frac{(y-\nu)^T\mathbf{C}^{-1}(y-\nu)}{2} \Big\}.
\end{aligned}$$

从而$Z = (X_1, \cdots, X_m)$在$Y = \bar{y}$条件下的条件概率密度为

$$h(z) = \frac{f(z, \bar{y})}{g(\bar{y})} = \frac{1}{(2\pi)^{m/2}\sqrt{|\Xi|}} \exp\Big\{ -\frac{(z-c(\bar{y}))^T\Xi^{-1}(z-c(\bar{y}))}{2} \Big\}.$$

由于$c(\bar{y}) = \hat{\mu}$, $h(z)$为m元正态分布$N(\hat{\mu}, \Xi)$的概率密度函数, 因此$(X_1, \cdots, X_m)$在$Y = \bar{y}$条件下服从m元正态分布. □

♠**注记6.1.8.** *在定理6.1.7中为了叙述方便我们假设了已知取值条件的随机变量为最后几个. 事实上, 条件分布仍为正态分布的结论并不依赖这一假定, 因为对任意位置的随机变量, 我们总可以通过调整次序使得它们排在最后.*

▶**例6.1.9.** 记$\mu=(1,2,3)^T$, $\mathbf{D}=\begin{pmatrix}1&1&1\\1&4&2\\1&2&9\end{pmatrix}$. 若$(X_1,X_2,X_3)\sim N(\mu,\mathbf{D})$, 求$X_3=12$时$(X_1,X_2)$的联合分布.

解 当$X_3=12$时, 由定理6.1.7可知, (X_1,X_2)服从正态分布

$$N(\tilde{\mu}+\mathbf{BC}^{-1}(12-\nu),\mathbf{A}-\mathbf{BC}^{-1}\mathbf{B}^T),$$

其中$\nu=3,\tilde{\mu}=(1,2)^T,\mathbf{B}=(1,2)^T,\mathbf{C}=(9),\mathbf{A}=\begin{pmatrix}1&1\\1&4\end{pmatrix}$. 直接计算可得

$$\tilde{\mu}+\mathbf{BC}^{-1}(12-\nu)=(2,4)^T,$$

以及

$$\mathbf{A}-\mathbf{BC}^{-1}\mathbf{B}^T=\begin{pmatrix}1&1\\1&4\end{pmatrix}-\frac{1}{9}\begin{pmatrix}1\\2\end{pmatrix}(1,2)=\begin{pmatrix}\frac{8}{9}&\frac{7}{9}\\\frac{7}{9}&\frac{32}{9}\end{pmatrix}.$$

因此$X_3=12$时(X_1,X_2)服从正态分布$N\left(\begin{pmatrix}2\\4\end{pmatrix},\begin{pmatrix}\frac{8}{9}&\frac{7}{9}\\\frac{7}{9}&\frac{32}{9}\end{pmatrix}\right)$. □

由于正态分布由均值向量和协方差矩阵唯一确定. 将定理6.1.7用到标准布朗运动可得如下固定一个时刻取值时布朗运动的有限维分布. 证明省略.

♦**推论6.1.10.** 对任意$n\ge 1$, $0\le t_1<t_2<\cdots<t_n$, 当$B(t_k)=x$时,

$$(B(t_1),\cdots,B(t_{k-1}),B(t_{k+1}),\cdots,B(t_n))$$

服从$n-1$维的正态分布$N(\mu,\Xi)$, 其中均值参数

$$\mu=(\frac{t_1x}{t_k},\cdots,\frac{t_{k-1}x}{t_k},x,\cdots,x)^T,$$

协方差矩阵$\Xi=(d_{i,j})_{i,j\ne k}$, 其中

$$d_{i,j}=\begin{cases}t_{i\wedge j}-\frac{t_it_j}{t_k}, & i\vee j<k,\\ 0, & i\wedge j<k<i\vee j,\\ t_{i\wedge j}-t_k, & k<i\wedge j.\end{cases}$$

▶**例6.1.11.** 在甲乙两人的自行车比赛中, 以$Y(t)$表示当完成$100t\%$的比赛路程时甲领先的时间(以秒记). 假设$Y(t)$是方差参数为σ^2的布朗运动, 若甲以σ秒领先赢得比赛, 问甲在路程中点领先的概率是多大?

解 注意到$B(t)=Y(t)/\sigma$ 为标准布朗运动, 所求概率为

$$p=\mathbb{P}(B(1/2)>0|B(1)=1).$$

由推论6.1.10可知在$B(1)=1$条件下, $B(1/2)$服从正态分布$N(1/2,1/4)$, 因此

$$p=\int_0^\infty\phi(x-1/2,1/4)\mathrm{d}x=\int_{-1}^\infty\phi(x,1)\mathrm{d}x=\Phi(1)\approx 0.8413,$$

其中Φ为标准正态分布函数. □

进一步, 利用定理6.1.7我们还可以得到给定布朗运动在区间的两个端点取值条件下它的有限维分布.

♦**推论6.1.12.** $B(s)=x, B(t)=y$时$(B(t_1),\cdots,B(t_n))$ 是均值为

$$\left(x+\frac{y-x}{t-s}(t_i-s)\right)_{1\le i\le n},$$

协方差矩阵

$$\mathbf{D}=\left(\frac{(t-t_{i\vee j})(t_{i\wedge j}-s)}{t-s}\right)_{i,j=1,\cdots,n}$$

的n元正态随机变量, 其中$s\le t_1<\cdots<t_n\le t$.

证明 由定理6.1.6可知$(B(t_1),\cdots,B(t_n),B(s),B(t))$服从$n+2$维的正态分布. 由性质6.1.2可知, 该正态分布的均值向量为$\mathbf{0}$, 协方差矩阵

$$\mathbf{D}=\begin{pmatrix} t_1 & t_1 & \cdots & t_1 & s & t_1\\ t_1 & t_2 & \cdots & t_2 & s & t_2\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\ t_1 & t_2 & \cdots & t_n & s & t_n\\ s & s & \cdots & s & s & s\\ t_1 & t_2 & \cdots & t_n & s & t \end{pmatrix}.$$

令$\mathbf{A}=(t_{i\wedge j})_{1\le i,j\le n}$, $\mathbf{B}=\begin{pmatrix} s & s & \cdots & s\\ t_1 & t_2 & \cdots & t_n\end{pmatrix}^T$, $\mathbf{C}=\begin{pmatrix} s & s\\ s & t\end{pmatrix}$. 那么

$$\begin{aligned}\mathbf{B}\mathbf{C}^{-1}\begin{pmatrix}x\\ y\end{pmatrix} &= \frac{1}{s(t-s)}\mathbf{B}\begin{pmatrix} t & -s\\ -s & s\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\\ &=\left(\frac{(t-t_i)x+(t_i-s)y}{t-s}\right)_{1\le i\le n}\\ &=\left(x+\frac{y-x}{t-s}(t_i-s)\right)_{1\le i\le n},\end{aligned}$$

并且

$$\begin{aligned}\mathbf{A}-\mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T &= \mathbf{A}-\frac{1}{s(t-s)}\mathbf{B}\begin{pmatrix} t & -s\\ -s & s\end{pmatrix}\begin{pmatrix} s & s & \cdots & s\\ t_1 & t_2 & \cdots & t_n\end{pmatrix}\\ &=\mathbf{A}-\frac{1}{(t-s)}\mathbf{B}\begin{pmatrix} t-t_1 & t-t_2 & \cdots & t-t_n\\ t_1-s & t_2-s & \cdots & t_n-s\end{pmatrix}\\ &=\left(t_{i\wedge j}-\frac{st-st_j+t_it_j-st_i}{t-s}\right)_{1\le i,j\le n}\\ &=\left(\frac{(t-t_{i\vee j})(t_{i\wedge j}-s)}{t-s}\right)_{1\le i,j\le n}.\end{aligned}$$

由定理6.1.7可知推论成立. □

▶**例6.1.13.** 求$B(1)=0,B(3)=u$时, 事件$\{B(2)>u,B(4)>u\}$的概率.

解 用推论6.1.12类似的计算可知$B(3)=u,B(1)=0$时$(B(2),B(4))$服从正态分布$N(\mu,\Xi)$, 其中均值$\mu=(u/2,u)^T$, 协方差矩阵$\Xi=\mathrm{diag}(1/2,1)$. 这表明在给定条件下, $B(2)\sim N(u/2,1/2)$, $B(4)\sim N(u,1)$ 而且两者独立. 因此所求概率

$$
\begin{aligned}
&P(B(2)>u,B(4)>u|B(3)=u,B(1)=0)\\
&=\int_u^\infty \phi(x-u/2,1/2)\mathrm{d}x\int_u^\infty \phi(x-u,1)\mathrm{d}x\\
&=\frac{1}{2}\int_u^\infty \phi(x-u/2,1/2)\mathrm{d}x=\frac{1}{2}\int_{u/2}^\infty \phi(x,1/2)\mathrm{d}x\\
&=\frac{1}{2}\int_{u/\sqrt{2}}^\infty \phi(x,1)\mathrm{d}x=\frac{1-\Phi(u/\sqrt{2})}{2}.
\end{aligned}
$$
□

利用布朗运动的独立增量性, 我们还能得到如下的条件独立性质.

★**定理6.1.14.** 任取$u,t>0$, $0\le t_1<\cdots<t_n<t$以及$x_1,\cdots,x_n,x,y\in\mathbb{R}$,

$$
\begin{aligned}
&\mathbb{P}(B(t+u)\le y|B(t)=x,B(t_n)=x_n,\cdots,B(t_1)=x_1)\\
&\qquad=\mathbb{P}(B(t+u)\le y|B(t)=x). \qquad (6.1.1)
\end{aligned}
$$

进而对任意随机变量$f(B(t+u))$,

$$
E\big(f(B(t+u))|B(t),B(t_n),\cdots,B(t_1)\big)=E\big(f(B(t+u))|B(t)\big). \qquad (6.1.2)
$$

证明 由标准布朗运动增量的独立正态分布性质可知, 在给定

$$
B(t)-B(t_n)=x-x_n,\cdots,B(t_2)-B(t_1)=x_2-x_1,B(t_1)=x_1
$$

条件下, $B(t+u)-B(t)\sim N(0,u)$. 因此,

$$
\begin{aligned}
&\mathbb{P}(B(t+u)\le y|B(t)=x,B(t_n)=x_n,\cdots,B(t_1)=x_1)\\
&=\mathbb{P}(B(t+u)-B(t)\le y-x|B(t)-B(t_n)=x-x_n,\cdots,\\
&\qquad B(t_2)-B(t_1)=x_2-x_1,B(t_1)=x_1)\\
&=\int_{-\infty}^{y-x}\phi(z,u)dz=\int_{-\infty}^{y}\phi(z-x,u)dz.
\end{aligned}
$$

这表明在$B(t)=x,B(t_n)=x_n,\cdots,B(t_1)=x_1$条件下$B(t+u)$服从正态分布$N(x,u)$. 同理可知在只给定$B(t)=x$条件下$B(t+u)$也服从正态分布$N(x,u)$. 由此可知(6.1.1)成立.

对任意ω, 记$x=B(t,\omega)$, $x_i=B(t_i,\omega)$, $i=1,\cdots,n$. 由条件数学期望定义以及上面证明的结果可知, 对几乎所有的ω,

$$
E(f(B(t+u))|B(t),B(t_n),\cdots,B(t_1))(\omega)
$$

$$
\begin{aligned}
&= E(f(B(t+u))|B(t)=x, B(t_n)=x_n, \cdots, B(t_1)=x_1) \\
&= \int_{\mathbb{R}} f(y)\phi(y-x,u)\mathrm{d}y \\
&= E(f(B(t+u))|B(t)=x) = E(f(B(t+u))|B(t))(\omega).
\end{aligned}
$$

因此(6.1.2)成立. □

♠**注记6.1.15.** *定理6.1.14本质上是布朗运动马氏性的一种表示, 因此布朗运动是连续时间马氏过程的一个具体例子. 记$B(t)=x$ 条件下$B(t+u)$的概率密度函数为$p(t,u+t,x,y)$, 显然$p(t,s+t,x,y)=\phi(y-x,u)$. 在马氏过程领域, 常称$p(t,t+u,x,y)$ 为从t 时刻状态x转移到$t+u$时刻状态y的转移概率密度(函数). 关于连续时间连续状态马氏过程的更多介绍超出了本书范围, 从略.*

♠**注记6.1.16.** *对连续时间随机过程还可以考虑这样的问题: 将某个时间t及t之前的所有过程变量做为条件变量讨论条件概率或条件期望等问题. 数学上可把这样的问题记成*

$$E(\cdot|X(s), 0\le s\le t), \tag{6.1.3}$$

其中$X=\{X(t);t\ge 0\}$为连续时间随机过程. 表达式(6.1.3)的准确含义需要更多测度论知识, 我们这里不展开. 但为后文需要, 我们指出, 对标准布朗运动B而言, $E(\cdot|B(s),0\le s\le t)$可以直观理解为$n\to\infty$后$E(\cdot|B(t),B(t_1),\cdots,B(t_n))$的极限, 其中$\{t_1,\cdots,t_n,\cdots\}$是$[0,t)$中任意一个可数稠密子集. 按照这一解释, 由定理6.1.14可得: 对任意$u>0$以及随机变量$f(B(t+u))$,

$$
\begin{aligned}
E(f(B(t+u))|B(s),0\le s\le t) &= \lim_{n\to\infty} E(f(B(t+u))|B(t),B(t_1),\cdots,B(t_n)) \\
&= E(f(B(t+u))|B(t)).
\end{aligned} \tag{6.1.4}
$$

6.1.3 轨道性质*

对刻画随机现象动态变化规律的随机过程模型, 掌握它的有限维分布(族)是非常重要的, 因为它能帮助我们认识动态变化现象的随机规律. 但在布朗运动的定义中, 我们还要求对所有的基本事件ω布朗运动的轨道都连续. 这种“额外”要求既来自于人们对相关随机现象的实际观察, 也有充分的理论保证.

★**定理6.1.17.** *布朗运动$B=\{B(t);t\ge 0\}$存在.*

定理的证明依赖于如下的柯尔莫哥洛夫连续修正定理, 该定理的证明从略.

Kolmogorov 连续修正定理 *设$\{X(t);t\ge 0\}$为实值随机过程, 若在某个参数区间$[a,b]$ ($[a,b]$可以是$[0,\infty]$)上存在正常数α, ε 和C, 使得对任意$t,t+h\in$*

$[a,b]$,

$$E(|X(t)-X(t+h)|^\alpha)\le C\cdot h^{1+\varepsilon},$$

那么$\{X(t),t\in[a,b]\}$存在一个轨道连续的修正.

定理6.1.17的证明 设$B=\{B(t);t\in I\}$满足布朗运动定义中的条件(a)与条件(b). 那么对任意$t,t+h\in I$, 由$B(t+h)-B(t)\sim N(0,h)$可知

$$E(|B(t+h)-B(t)|^4)=3h^2.$$

由前面的连续修正定理可知存在B的一个修正, 使得轨道连续条件(c)成立. □

有趣的是布朗运动轨道函数还提供了连续但处处不可微的函数实例. 我们以标准布朗运动为例解释如下.

★**定理6.1.18.** 标准布朗运动$B=\{B(t);\ t\ge 0\}$的几乎所有轨道处处不可微.

证明 只考虑$t\in[0,1]$的情形. 对任意ω, 若$B(\cdot,\omega)$在s 可微, 则存在$\delta\in(0,1/2)$ 与整数$l\ge 1$使得当$|t-s|<\delta$时

$$|B(t,\omega)-B(s,\omega)|<l|t-s|.$$

对任意$n>4/\delta$, 令$i=\lfloor ns\rfloor+1\le n+1$, 对$j=i\pm1,i\pm2,i\pm3$都应有

$$\begin{aligned}&|B(j/n,\omega)-B((j-1)/n,\omega)|\\ &\le|B(j/n,\omega)-B(s,\omega)|+|B((j-1)/n,\omega)-B(s,\omega)|\\ &\le l|j/n-s|+l|(j-1)/n-s|\le 7l/n.\end{aligned}\tag{6.1.5}$$

令随机事件$A_{n,l}^i=A_{n,l}^{i,+}\cup A_{n,l}^{i,-}$, 其中随机事件

$$A_{n,l}^{i,+}=\{\omega;\text{对给定的 }l,i,\ (6.1.5)\text{ 对 }j=i+1,i+2,i+3\text{同时成立}\},$$

$$A_{n,l}^{i,-}=\{\omega;\text{对给定的 }l,i,\ (6.1.5)\text{ 对 }j=i-1,i-2,i-3\text{同时成立}\}.$$

显然随着l 的增加, $A_{n,l}^i$ 单调不降. 特别地, 取$\delta=1/m$, $m=2,3,\cdots$, 则

$$\varOmega_1=\{\omega;\text{存在}s\in[0,1]\text{使得}B(\cdot,\omega)\text{在}s\text{点可微}\}\subset\bigcup_{l\ge1,m\ge2}\bigcap_{n\ge4m}\bigcup_{i\le n+1}A_{n,l}^i.$$

此时, 对任意i, 由布朗运动的平稳独立增量性可得

$$\begin{aligned}P(A_{n,l}^i)&\le P(A_{n,l}^{i,+})+P(A_{n,l}^{i,-})=2\Big[\mathbb{P}(|B((i+1)/n)-B(i/n)|\le 7l/n)\Big]^3\\ &=2\Big[\int_{-7l/n}^{7l/n}\frac{1}{\sqrt{2\pi/n}}\mathrm{e}^{-ny^2/2}dy\Big]^3\le 2(14l/\sqrt{2\pi n})^3.\end{aligned}$$

当$n\to\infty$时,

$$\mathbb{P}\Big(\bigcup_{i\le n+1}A_{n,l}^i\Big)\le(n+1)\max_{1\le i\le n+1}\mathbb{P}(A_{n,l}^i)\le 2(n+1)(14l/\sqrt{2\pi n})^3\to 0.$$

因此

$$\mathbb{P}(\varOmega_1)\le\lim_{l\to\infty}\lim_{n\to\infty}\mathbb{P}\Big(\bigcup_{i\le n+1}A_{n,l}^i\Big)=0,$$

即对几乎必然的ω, $B(\cdot,\omega)$处处不可微. □

布朗运动的轨道还有其它很多有趣的性质, 涉及更深刻的数学基础, 许多内容已超出本书的范畴. 对此感兴趣的同学可以参考文献 [8].

练习题6.1 (以下总设$B=\{B(t);t\geq 0\}$为标准布朗运动.)

1 设$(X_1,\cdots,X_n)$服从n元正态分布, 证明(1) 对任意$m>1$, $(X_1,\cdots,X_m)$服从m元正态分布. (2)对任意不全为0的常数$c_1,\cdots,c_n$, 线性组合$c_1X_1+\cdots+c_nX_n$服从一元正态分布.

2 对任意$0\leq s<t$, 求$B(s)+B(t)$的概率密度函数.

3 在例6.1.9假设下, (1)求$X_2=2$时(X_1,X_3)的条件联合分布; (2)求$X_2=2$, $X_3=3$时X_1的条件分布.

4 已知$B(1)=x$, 对任意$0\leq s_1<s_2\leq s_3<s_4\leq 1$, 证明

(1) $E((B(s_4)-B(s_3))(B(s_2)-B(s_1)))=(s_4-s_3)(s_2-s_1)(x^2-1)$,

(2) $E[(B(s_4)-B(s_1))^2]=(s_4-s_1)+(s_4-s_1)^2(x^2-1)$.

5 设B是方差参数$\sigma^2=3$的布朗运动, 已知$B(1)=1$, $B(4)=4$, 求$E(B(2)B(3))$和$E(B(3)B(5))$.

6 在例6.1.11中若甲在一半赛程时领先乙σ秒, 问比赛结束时甲领先的概率.

7 令$Y(t)=B^2(t)-t$, $R(t)=\mathrm{e}^{cB(t)-c^2t/2}$, 其中$c>0$为常数. 证明对任意$0\leq t<s$, $E(Y(s)|B(t))=Y(t)$以及$E(R(s)|B(t))=R(t)$.

6.2 反射原理与极值分布

上一节我们考察了布朗运动的有限维分布以及在给定若干时刻取值条件下的有限维分布, 其中涉及的随机事件本质上都只与布朗运动在有限个时刻点的随机变化有关. 这一节我们讨论布朗运动最大值的分布, 这能帮助我们讨论布朗运动在一段时间内的表现.

6.2.1 反射原理

若无特别说明, 本节我们讨论的布朗运动B都是标准布朗运动. 为方便, 我们仍以$\phi(x,t)$表示正态分布$N(0,t)$的概率密度函数. 注意对任意$a,b\in\mathbb{R}, s,t>0$,

$$\phi(x-a,s)\phi(x-b,t)=\phi(a-b,s+t)\phi(x-\frac{ta+sb}{s+t},\frac{st}{s+t}). \qquad (6.2.1)$$

对任意$a\in\mathbb{R}$, 令

$$\tau_a=\inf\{t\geq 0,\ B(t)=a\}=\begin{cases}\inf\{t\geq 0,\ B(t)\geq a\}, & a\geq 0;\\ \inf\{t\geq 0,\ B(t)\leq a\}, & a<0.\end{cases}$$

称τ_a为B首次到达状态a的时刻, 简称为a的首达时. 利用布朗运动的轨道连续性我们知道(以$a>0$为例),

$$\begin{aligned}\{\tau_a>t\}&=\{B(s)<a,0\leq s\leq t\}=\bigcup_{n=1}^{\infty}\{B(s)\leq a-\frac{1}{n},s\in[0,t]\}\\ &=\bigcup_{n=1}^{\infty}\{B(s)\leq a-\frac{1}{n},s\in[0,t]\text{为有理数}\}\\ &=\bigcup_{n=1}^{\infty}\bigcap_{s\in[0,t]\text{为有理数}}\{B(s)\leq a-\frac{1}{n}\}.\end{aligned}$$

由于可数多个随机事件的交和并仍然是随机事件, 所以$\{\tau_a>t\}$为随机事件, 进而τ_a是一个随机变量(若以正概率取到$+\infty$,则是广义随机变量).

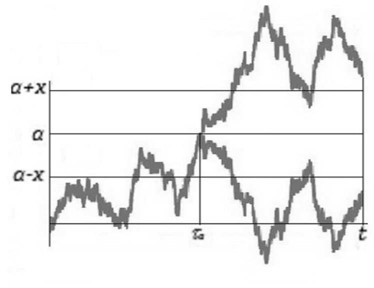
♠**注记6.2.1.** 由上面的分析我们还可以知道, *事件$\{\tau_a>t\}$(从而事件$\{\tau_a\leq t\}$)是否发生只跟布朗运动B在t以及t之前的取值有关*. 这其实是以连续时间随机过程为参照的所谓"停时"具有的特性, 因此我们也称τ_a 是布朗运动B 的停时.

对于首达时τ_a, 利用布朗运动的对称性和独立增量性, 可得如下反射原理.

▼**引理6.2.2.** (反射原理) 对任意$a\in\mathbb{R}$及$x\geq 0$,

$$\mathbb{P}(B(t)\geq a+x,\tau_a\leq t)=\mathbb{P}(B(t)<a-x,\tau_a\leq t). \qquad (6.2.2)$$

引理6.2.2的严格数学证明从略. 它的直观含义为: 将$B(t)$在时刻τ_a之后的轨道以直线$t=\tau_a$为镜面做反射, 那么由增量分布的对称性可知, $B(t)-B(\tau_a)$以相同的概率落在$(-\infty,-x)$与$(x,+\infty)$ 这两个区间. 图形解释如下:



♦**推论6.2.3.** 对任意$a\in\mathbb{R}$, $\mathbb{P}(B(t)\geq a|\tau_a\leq t)=1/2$.

证明 由引理6.2.2可知

$$\mathbb{P}(B(t)\geq a,\tau_a\leq t)=\mathbb{P}(B(t)<a,\tau_a\leq t).$$

因此

$$\begin{aligned}\mathbb{P}(\tau_a\leq t)&=\mathbb{P}(B(t)\geq a,\tau_a\leq t)+\mathbb{P}(B(t)<a,\tau_a\leq t)\\&=2\mathbb{P}(B(t)\geq a,\tau_a\leq t).\end{aligned}$$

所以

$$\mathbb{P}(B(t)\geq a|\tau_a\leq t)=\frac{\mathbb{P}(B(t)\geq a,\tau_a\leq t)}{\mathbb{P}(\tau_a\leq t)}=1/2. \qquad \square$$

6.2.2 极值与首达时分布

布朗运动的轨道函数是连续的, 因此在任意给定的一个时间区间内, 布朗运动存在最大值、最小值. 显然掌握最大值、最小值的规律能帮助我们从整体上把握布朗运动在这段时间的变化. 下面我们以最大值为例, 讨论它的分布规律.

记$B^*(t)=\max\limits_{0\leq u\leq t}B(u)$. 对任意$z\geq 0$, 由于$B$的轨道是连续的, 而且$B(0)=0$, 我们可在最大值与首达时之间建立如下的等式

$$\{B^*(t)\geq z\}=\{\tau_z\leq t\}. \tag{6.2.3}$$

利用这个等式以及反射原理, 容易得到如下结果.

★**定理6.2.4.** 对任意$z\geq 0$, $x\leq z$,

$$\mathbb{P}(B^*(t)\geq z,B(t)<x)=\mathbb{P}(B(t)>2z-x)=\int_{2z-x}^{\infty}\phi(y,t)\mathrm{d}y.$$

因此$(B^*(t),B(t))$的联合概率密度函数为

$$h(z,x)=2\frac{\partial\phi(2z-x,t)}{\partial x}=\frac{2(2z-x)}{\sqrt{2\pi t^3}}\mathrm{e}^{-\frac{(2z-x)^2}{2t}},\quad x\leq z,z\geq 0.$$

证明 由(6.2.3)以及反射原理(6.2.2)可得

$$\begin{aligned}&\mathbb{P}(B^*(t)\geq z, B(t)<x)=\mathbb{P}(\tau_z\leq t, B(t)<x)\\&=\mathbb{P}(B(t)<z-(z-x),\tau_z\leq t)=\mathbb{P}(B(t)>z+(z-x),\tau_z\leq t)\\&=\mathbb{P}(B(t)>2z-x)=\int_{2z-x}^{\infty}\phi(y,t)\mathrm{d}y.\end{aligned}$$

上式两边关于z,x求二阶混合偏导并取相反数后得联合概率密度函数$h(z,x)$. □

▶**例6.2.5.** 求$P(B^*(t)>x|B(t)=B^*(t))$.

解 由概率论知识可知$\{B(t)=B^*(t)\}$条件下$B^*(t)$的边际概率密度函数为

$$g(z)=\frac{h(z,z)}{\int_0^\infty h(z,z)\mathrm{d}z}=\frac{z\mathrm{e}^{-\frac{z^2}{2t}}}{\int_0^\infty z\mathrm{e}^{-\frac{z^2}{2t}}\mathrm{d}z}=\frac{z}{t}\mathrm{e}^{-\frac{z^2}{2t}}.$$

因此

$$P(B^*(t)>x|B(t)=B^*(t))=\int_x^\infty g(z)\mathrm{d}z=\mathrm{e}^{-\frac{x^2}{2t}}.$$ □

♦**推论6.2.6.** 对任意$z\geq 0$,

$$\mathbb{P}(B^*(t)\geq z)=2\mathbb{P}(B(t)\geq z)=\mathbb{P}(|B(t)|\geq z).$$

因此$B^*(t)$与$|B(t)|$ 同分布, $B^*(t)$ 的概率密度函数

$$g(z)=\frac{2}{\sqrt{2\pi t}}\mathrm{e}^{-z^2/2t}.$$

证明 对任意$z\geq 0$,

$$\begin{aligned}\mathbb{P}(B^*(t)\geq z)&=\mathbb{P}(B^*(t)\geq z, B(t)\geq z)+\mathbb{P}(B^*(t)\geq z, B(t)<z)\\&=\mathbb{P}(B(t)\geq z)+\mathbb{P}(B(t)>z)=2\mathbb{P}(B(t)\geq z)\\&=2\int_z^\infty\frac{1}{2\pi t}\mathrm{e}^{-u^2/2t}\mathrm{d}u=\mathbb{P}(|B(t)|\geq z).\end{aligned}$$ □

▶**例6.2.7.** 假定一只股票价格按初值为发行价的标准布朗运动变化. 某人以价格$b+c$(单位: 元)买入, 而现在的价格恰好为b, 若此人计划在股票价格回到$b+c$或最多再观望T 时间后将股票卖出, 求股票卖出价格不是$b+c$的概率.$(c>0)$.

解 从现在开始看股票的价格$S(t)$是一个从b出发的标准布朗运动, 即$S(t)=b+B(t)$. 股票卖出价格不是$b+c$对应于在时刻T之前$S(t)$最大值小于$b+c$, 即

$$B_T^*=\max_{0\leq s\leq T}B(s)<c.$$

由推论6.2.6可知所求概率为

$$p=1-P(B^*(T)\geq c)=1-P(|B(T)|\geq c)=2\Phi(c/\sqrt{T})-1.$$ □

♦**推论6.2.8.** 对任意$t,s>0$, $(B^*(t),B(t+s))$的联合概率密度函数

$$g(z,x)=\int_{-\infty}^{z}\frac{2z-u}{\pi t\sqrt{ts}}\mathrm{e}^{-\frac{(2z-u)^2}{2t}-\frac{(x-u)^2}{2s}}\mathrm{d}u,\qquad z>0,x\in\mathbb{R}.$$

证明 对任意$z>0,x\in\mathbb{R}$, 由标准布朗运动的轨道连续性可知对$[0,t]$上任意可数稠密子集$\{t_1,t_2,\cdots,t_n,\cdots\}$,

$$\{B^*(t)\le z\}=\lim_{n\to\infty}\{B(t_1)\le z,\cdots,B(t_n)\le z\}.$$

由布朗运动的马氏性(定理6.1.14)和独立增量性可得

$$\begin{aligned}
&P(B^*(t)\le z,B(t+s)\le x)\\
&=\lim_{n\to\infty}P\big(B(t_1)\le z,\cdots,B(t_n)\le z,B(t+s)\le x\big)\\
&=\lim_{n\to\infty}E\Big(E\big(1_{\{B(t_1)\le z,\cdots,B(t_n)\le z\}}1_{\{B(t+s)\le x\}}|B(t_1),\cdots,B(t_n),B(t)\big)\Big)\\
&=\lim_{n\to\infty}E\Big(1_{\{B(t_1)\le z,\cdots,B(t_n)\le z\}}E\big(1_{\{B(t+s)\le x\}}|B(t)\big)\Big)\\
&=E\Big(1_{\{B^*(t)\le z\}}E\big(1_{\{B(t+s)\le x\}}|B(t)\big)\Big)\\
&=E\Big(1_{\{B^*(t)\le z\}}P\big(B(t+s)-B(t)\le x-B(t)\big)\Big)\\
&=E\Big(1_{\{B^*(t)\le z\}}\int_{-\infty}^{x-B(t)}\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{w^2}{2s}}\mathrm{d}w\Big).
\end{aligned}$$

将$(B^*(t),B(t))\sim h(z,x)=\frac{2(2z-x)}{\sqrt{2\pi t^3}}\mathrm{e}^{-\frac{(2z-x)^2}{2t}}$代入上式得,

$$\begin{aligned}
&P(B^*(t)\le z,B(t+s)\le x)\\
&=\int_0^z\mathrm{d}v\int_{-\infty}^{v}\Big[\frac{2(2v-u)}{\sqrt{2\pi t^3}}\mathrm{e}^{-\frac{(2v-u)^2}{2t}}\int_{-\infty}^{x-u}\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{w^2}{2s}}\mathrm{d}w\Big]\mathrm{d}u.
\end{aligned}$$

上式两边关于z,x求二阶混合偏导即得推论. □

▸**例6.2.9.** *求例6.1.11中在比赛后半程甲都不落后于乙的概率.*

解 依题设$Y(t)$是方差参数为σ^2的布朗运动, $B(t)=Y(t)/\sigma$是标准布朗运动. 令

$$B_*=\min\{B(u),\ 1/2\le u\le 1\}.$$

已知$B(1)=1$, 所求概率为

$$p=P(B_*\ge 0|B(1)=1).$$

注意到$W(t)=B(1)-B(1-t)$, $0\le t\le 1$, 仍为标准布朗运动, 且$W(1)=B(1)$.

$$W^*(1/2)=\max\{W(t),\ 0\le u\le 1/2\}=1-B_*.$$

因此$p=P(W^*(1/2)\le 1|W(1)=1)$. 由推论6.2.8(其中$t=s=1/2$)可得

$$p=\frac{\int_0^1 g(z,1)\mathrm{d}z}{\int_0^\infty g(z,1)\mathrm{d}z}.$$

直接计算可知

$$
\begin{aligned}
\int_0^1 g(z,1)\mathrm{d}z &= \int_0^1 \mathrm{d}z\int_{-\infty}^{z}\frac{4(2z-u)}{\pi}\mathrm{e}^{-(2z-u)^2-(1-u)^2}\mathrm{d}u\\
&= \int_0^1 \mathrm{d}z\int_{z}^{\infty}\frac{4y}{\pi}\mathrm{e}^{-2(y+\frac{1-2z}{2})^2-\frac{(1-2z)^2}{2}}\mathrm{d}y \quad (\text{积分变换}y=2z-u)\\
&= \int_{-1}^1 \mathrm{d}u\int_{\frac{u+1}{2}}^{\infty}\frac{2y}{\pi}\mathrm{e}^{-2(y-\frac{u}{2})^2-\frac{u^2}{2}}\mathrm{d}y \quad (\text{积分变换}u=2z-1)\\
&= \int_{-1}^1 \mathrm{d}u\int_{\frac{1}{2}}^{\infty}\frac{2(v+u/2)}{\pi}\mathrm{e}^{-2v^2-\frac{u^2}{2}}\mathrm{d}v \quad (\text{积分变换}v=y-u/2)\\
&= \int_{-1}^1 \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u\int_{\frac{1}{2}}^{\infty}\frac{4v}{\sqrt{2\pi}}\mathrm{e}^{-2v^2}\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-1/2}\int_{-1}^1\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u.
\end{aligned}
$$

类似(其实更容易)可得

$$\int_0^\infty g(z,1)\mathrm{d}z = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-1/2}.$$

因此

$$p=\int_{-1}^1\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u = 2\Phi(1)-1. \qquad \square$$

♦**推论6.2.10.** 对任意$z>0$, τ_z的概率密度函数

$$p(t)=\frac{z}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t}.$$

因此$\mathbb{P}(\tau_z<\infty)=1$, $E(\tau_z)=\infty$.

证明 对任意$z\geq 0$, 由(6.2.3)可知

$$\mathbb{P}(\tau_z\leq t)=\mathbb{P}(B^*(t)\geq z)=2\int_z^\infty\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-u^2/2t}\mathrm{d}u=\int_{z/\sqrt{t}}^\infty\frac{2}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u.$$

两边关于t求导得

$$p(t)=\frac{z}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t}.$$

即

$$\mathbb{P}(\tau_z\leq t)=\int_0^t\frac{z}{\sqrt{2\pi s^3}}\mathrm{e}^{-z^2/2s}\mathrm{d}s.$$

因此

$$\mathbb{P}(\tau_z<\infty)=\int_0^\infty\frac{z}{\sqrt{2\pi s^3}}\mathrm{e}^{-z^2/2s}\mathrm{d}s=\int_0^\infty\frac{2}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u=1,$$

而且

$$E(\tau_z)=\int_0^\infty\frac{z}{\sqrt{2\pi s}}\mathrm{e}^{-z^2/2s}\mathrm{d}s=\infty. \qquad \square$$

注意对任意$z\geq 0$, τ_z与τ_{-z}分布相同(参见本节习题4). 因此由推论6.2.10可

知, 对任意$z \in \mathbb{R}$, τ_z的概率密度函数
$$p(t) = \frac{|z|}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t},$$
而且布朗运动必然在有限时间内到达这一点,只是到达的平均时长是无穷的.

♠**注记6.2.11.** 利用关系式$\min\limits_{0\le s\le t} B(s) = -\max\limits_{0\le s\le t}[-B(s)]$以及$-B$仍是布朗运动的事实, 我们很容易把上面关于极大值$\max\limits_{0\le s\le t} B(s)$的结果类推到极小值$\min\limits_{0\le s\le t} B(s)$. 具体的结论请读者自己完成. 我们也可以给出$(\max\limits_{0\le s\le t} B(s), \min\limits_{0\le s\le t} B(s))$的联合分布. 感兴趣的读者可参阅[4]中第二章第八节的相关内容.

6.2.3 零点与极大值点*

这一小节, 我们讨论布朗运动轨道在一段时间内零点个数以及极大值点的分布问题. 我们有如下的一些结论.

★**定理6.2.12.** 从$x \neq 0$ 出发的标准布朗运动$B_x(s) = x + B(s)$在$[0,t]$至少有一个零点的概率为
$$p_t(x) = \frac{|x|}{\sqrt{2\pi}}\int_0^t s^{-3/2}\mathrm{e}^{-\frac{x^2}{2s}}\mathrm{d}s.$$

证明 先设$x < 0$, 那么
$$\begin{aligned}p_t(x) &= \mathbb{P}(\max_{0\le u\le t} B(u) + x \ge 0) = \mathbb{P}(\max_{0\le u\le t} B(u) \ge -x) = \mathbb{P}(\tau_{-x} \le t)\\ &= \mathbb{P}(\tau_{|x|} \le t) = \frac{|x|}{\sqrt{2\pi}}\int_0^t s^{-3/2}\mathrm{e}^{-\frac{x^2}{2s}}\mathrm{d}s.\end{aligned}$$
再设$x > 0$, 注意到$-B$也是标准布朗运动,
$$\begin{aligned}p_t(x) &= \mathbb{P}(\min_{0\le u\le t} B(u) + x \le 0) = \mathbb{P}(\min_{0\le u\le t} B(u) \le -x)\\ &= \mathbb{P}(\max_{0\le u\le t}(-B(u)) \ge x) = \mathbb{P}(\tau_x \le t)\\ &= \frac{x}{\sqrt{2\pi}}\int_0^t s^{-3/2}\mathrm{e}^{-\frac{x^2}{2s}}\mathrm{d}s.\end{aligned}$$
综合两种情况可知定理成立. □

♦**推论6.2.13.** 设$0 < s < t$, B在时间区间(s,t)内至少存在一个零点的概率
$$q = \frac{2}{\pi}\arccos\sqrt{\frac{s}{t}}.$$

证明 由定理6.2.12可知$p_t(x) = \int_0^t \frac{|x|}{s}\phi(x,s)\mathrm{d}s$. 因此
$$q = E(p_{t-s}(B(s))) = \int_{-\infty}^{\infty}\phi(x,s)\int_0^{t-s}\frac{|x|}{u}\phi(x,u)\mathrm{d}u$$

$$= 2\int_0^{t-s}\mathrm{d}u\int_0^{\infty}\phi(x,s)\phi(x,u)\frac{x}{u}\mathrm{d}x.$$

由(6.2.1)以及积分变量代换$v=\sqrt{u/s}$得

$$\begin{aligned} q &= 2\int_0^{t-s}\frac{\phi(0,s+u)}{u}\mathrm{d}u\int_0^{\infty}x\phi(x,\frac{su}{s+u})\mathrm{d}x \\ &= \frac{1}{\pi}\int_0^{t-s}\frac{1}{s+u}\sqrt{\frac{s}{u}}\mathrm{d}u = \frac{2}{\pi}\int_0^{\sqrt{(t-s)/s}}\frac{1}{1+v^2}\mathrm{d}v \\ &= \frac{2}{\pi}\arctan\sqrt{\frac{t-s}{s}} = \frac{2}{\pi}\arccos\sqrt{\frac{s}{t}}. \end{aligned}$$

推论得证. □

►**例6.2.14.** 设$c>0$为常数，求$p=P($存在$t_0\in(1,2)$使得$B(t_0)=ct_0|B(3)=3c)$.

解 令$W(t)=tB(1/t)$, 那么$W(t)$仍为标准布朗运动, 此时

$$\{\text{存在}t_0\in(1,2)\text{使得}B(t_0)=ct_0\}=\{\text{存在}t_0\in(1/2,1)\text{使得}W(t_0)=c\}.$$

进而由推论6.2.13可知

$$\begin{aligned} p &= P(\text{存在}t_0\in(1/2,1)\text{使得}W(t_0)=c|W(1/3)=c) \\ &= P(\text{存在}t_0\in(1/6,2/3)\text{使得}W(t_0)=0) = \frac{2}{\pi}\arccos\sqrt{\frac{1}{4}} = \frac{2}{3}. \end{aligned}$$ □

★**定理6.2.15.** $v_t=\sup\{s;0\le s\le t, B(s)=B^*(t)\}$的概率密度函数为

$$\nu(s)=\frac{1}{\pi\sqrt{s(t-s)}},\qquad 0<s<t.$$

证明 注意到对任意$0<s<t$,

$$\{v_t\le s\}=\big\{\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)\big\}.$$

因此

$$\begin{aligned} \mathbb{P}(v_t\le s) &= \mathbb{P}\big(\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)\big) \\ &= E\big(\mathbb{P}(\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)|B^*(s),B(s))\big) \\ &= E\big(\mathbb{P}(B^*(s)-B(s)>\max_{s<u\le t}(B(u)-B(s))|B^*(s),B(s))\big). \end{aligned}$$

由平稳独立增量性可得

$$\begin{aligned} \mathbb{P}(v_t\le s) &= E\Big(\mathbb{P}\big(B^*(s)-B(s)>\max_{0\le u\le t-s}W(u)\big)\Big) \\ &= E\Big(\int_0^{B^*(s)-B(s)}2\phi(u,t-s)\mathrm{d}u\Big), \end{aligned}$$

其中W表示与B独立的标准布朗运动. 由定理6.2.4以及等式(6.2.1),

$$\mathbb{P}(v_t\le s)=\int_0^{\infty}\mathrm{d}z\int_0^{\infty}4\phi(z+y,s)\frac{z+y}{s}\mathrm{d}y\int_0^{y}\phi(u,t-s)\mathrm{d}u$$

$$
\begin{aligned}
&= \int_0^\infty \mathrm{d}z \int_0^\infty 4\phi(z+u,s)\phi(u,t-s)\mathrm{d}u \\
&= \int_0^\infty 4\phi(z,t)\mathrm{d}z \int_0^\infty \phi\Big(u+\frac{t-s}{t}z, \frac{(t-s)s}{t}\Big)\mathrm{d}u \\
&= \int_0^\infty 4\phi(z,t)\mathrm{d}z \int_{z\sqrt{\frac{(t-s)}{ts}}}^\infty \phi(u,1)\mathrm{d}u.
\end{aligned}
$$

两边求导得

$$
\nu(s) = \int_0^\infty 4\phi(z,t)\phi\Big(z\sqrt{\frac{t-s}{ts}},1\Big)\frac{\sqrt{t}z}{2\sqrt{s^3(t-s)}}\mathrm{d}z.
$$

直接计算可知

$$
\phi(z,t)\phi\Big(z\sqrt{\frac{t-s}{ts}},1\Big) = \frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-\frac{z^2}{2t}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{z^2(t-s)}{2ts}} = \frac{1}{2\pi\sqrt{t}}\mathrm{e}^{-\frac{z^2}{2s}}.
$$

从而

$$
\nu(s) = \int_0^\infty \frac{z}{\pi s\sqrt{s(t-s)}}\mathrm{e}^{-\frac{z^2}{2s}}\mathrm{d}z = \frac{1}{\pi\sqrt{s(t-s)}}. \qquad \square
$$

练习题6.2 (以下各题中过程B都是标准布朗运动.)

1 求$E(B^*(t) \mid B(t)=x)$.

2 求$B^*(t)-B(t)$的概率密度函数.

3 求$P(B(1)\le x|B(u)\ge 0, 0\le u\le 1)$.

4 证明对任意$z\ge 0$, τ_z与τ_{-z}同分布.

5 对任意$\lambda\ge 0$, 求$E(\mathrm{e}^{-\lambda\tau_z})$.

6 当$0<z\to\infty$时, 证明$\tau_z\to+\infty$, a.s.

7* 令$\gamma_t=\sup\{s; 0\le s\le t, B(s)=0\}$, 证明

$$
\mathbb{P}(\gamma_t\le x) = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{t}}, \qquad x\le t.
$$

8* 对任意$0<s<t$ 求$(B^*(t), B(s))$的联合概率密度函数.

6.3 几何布朗运动与Black-Scholes公式

几何布朗运动在风险资产价格建模中被广泛运用, 其中一个最著名的应用就是期权定价理论中的Black-Scholes公式.

6.3.1 几何布朗运动

♣**定义6.3.1.** *若$B=\{B(t);t\geq 0\}$为标准布朗运动, μ和σ是两任意给定的常数, 称由$Y(t)=\sigma B(t)+\mu t$定义的随机过程$\{Y(t);t\geq 0\}$ 为具有漂移系数为μ和方差参数为σ^2的布朗运动, 简称为带漂移的布朗运动.*

显然, 对具有漂移系数μ和方差参数σ^2的布朗运动Y而言:

(1) $Y(0)=0$.　　(2) $\{Y(t);t\geq 0\}$有平稳独立增量.

(3) $Y(t)$服从均值为μt, 方差为$\sigma^2 t$的正态分布.

(4) 定理6.1.14和注记6.1.16中的标准布朗运动B可替换为带漂移布朗运动Y.

♣**定义6.3.2.** *若$Y=\{Y(t);t\geq 0\}$为带漂移的布朗运动, 那么称由$X(t)=\mathrm{e}^{Y(t)}$定义的随机过程$X=\{X(t);t\geq 0\}$为几何布朗运动(也称为指数布朗运动).*

对几何布朗运动X, 利用上面漂移布朗运动的性质(4)可得, 对任意$t>u$以及随机变量$g(X(t))$,

$$E(g(X(t))|X(s),0\leq s\leq u)=E(g(X(t)))|X(u)). \tag{6.3.1}$$

这里条件$X(s),0\leq s\leq u$, 表示已知u之前所有的$X(s)$. 因此,

$$\begin{aligned}E\big(g(X(t))|X(s),0\leq s\leq u\big)&=E\big(g(\mathrm{e}^{Y(t)})|\mathrm{e}^{Y(u)}\big)=E\big(g(\mathrm{e}^{Y(u)}\mathrm{e}^{Y(t)-Y(u)})|\mathrm{e}^{Y(u)}\big)\\&=\int_{-\infty}^{+\infty}g(\mathrm{e}^y\mathrm{e}^x)\phi(x-\mu(t-u),\sigma^2(t-u))\mathrm{d}x\Big|_{y=Y(u)}\\&=\int_{-\infty}^{+\infty}g(X(u)\mathrm{e}^x)\phi(x-\mu(t-u),\sigma^2(t-u))\mathrm{d}x,\end{aligned} \tag{6.3.2}$$

进而,

$$\begin{aligned}E(X(t)|X(s),0\leq s\leq u)&=X(u)\int_{-\infty}^{+\infty}\mathrm{e}^x\phi(x-\mu(t-u),\sigma^2(t-u))\mathrm{d}x\\&=X(u)\mathrm{e}^{(t-u)(\mu+\sigma^2/2)}.\end{aligned} \tag{6.3.3}$$

6.3.2 期权定价原理

对现代金融市场而言, 期权无处不在. 所谓期权是指一种合约, 该合约赋予持有人在某一特定日期或该日之前的任何时间以固定价格购进或售出一种资产

的权利. 期权定义的要点如下: 1、期权是一种权利. 期权合约至少涉及购买人和出售人两方. 持有人享有权力但不承担相应的义务. 2、期权的标的物. 它包括股票、政府债券等基础证券, 也包括货币、股票指数、商品期货等对象. 期权是这些标的物"衍生"的, 因此也被称为衍生金融工具或衍生证券. 值得注意的是, 期权出售人不一定拥有标的物. 期权是可以"卖空"的. 期权购买人也不定真的想购买资产标的物. 因此, 期权到期时双方不一定进行标的物的实物交割, 而只需按价差补足价款即可. 3、到期日. 双方约定的期权到期的那一天称为"到期日", 如果该期权只能在到期日执行, 则称为欧式期权; 如果该期权可以在到期日或到期日之前的任何时间执行, 则称为美式期权. 4、期权的执行. 依据期权合约购进或售出标的资产的行为称为"执行". 在期权合约中约定的, 期权持有人据以购进或售出标的资产的固定价格, 称为"执行价格"或"敲定价格". 按合约购进/卖出的期权也称为看涨/看跌期权.

若期权敲定价格为K, 期权标的物在到期日的价格为P, 那么依据期权合约, 一份看涨/看跌期权在到期日的价值分别为$(P-K)^+$和$(K-P)^+$. 显然期权对期权合约的买入方是一项有利的经济权利, 因此为了获得这份权利, 期权合约买入方需要给予卖出方一定的经济补偿. 确定补偿的合理值就是对期权定价. 通常该定价过程遵循两个原则: 无套利市场假设和风险中性原理.

套利是指通过市场交易的买卖差价获取利益. 所谓无套利市场, 直观而言, 就是市场上任何交易都不能获得无风险的套利. 在市场有效条件下, 金融市场的套利机会总是暂时的. 因此, 无套利市场假设被用于对金融产品进行定价, 要求金融产品的价格不能给市场带来无风险套利机会, 这也称为无风险套利定价原理或者简称为无套利定价原理.

在无套利市场条件下, 人们进一步提出风险中性定价原理: 假设经济环境是风险中性的, 即处在该经济环境中的投资者总体而言对待风险的态度是中性的, 他们在投资风险资产时自己承担风险, 并不要求任何的风险补偿或风险报酬. 基于风险中性假设, 在市场不存在任何套利可能性的条件下, 风险资产的期望收益率都恰好等于无风险利率. 因此, 若期权的价格依赖于可交易的基础证券, 那么该期权的价格与具体投资者的风险态度无关, 而是期权到期日的价值经过无风险利率贴现的结果.

就数学模型而言, 无套利市场和风险中性假设使得用来刻画市场统计规律的概率空间$(\Omega,\mathcal{F},P)$对任何可投资的风险资产$X(t)$都满足:

$$E(X(t+s)|\mathcal{F}_t)=X(t)\mathrm{e}^{rs},$$

其中r为常数, 表示该市场按复利计算的无风险利率(本节剩下内容中的无风险利率都按此理解), $\mathcal{F}_t$表示t及t之前的所有市场信息; 而$t+s$时刻价格为$Y(t+s)$的

衍生证券在t时刻的价格

$$Y(t)=E(\mathrm{e}^{-rs}Y(t+s)|\mathcal{F}_t).$$

如果市场只有一种风险资产$X(t)$和一种无风险资产$r(t)$, 那么t及t之前的所有市场信息就由该风险资产的信息构成. 此时,

$$E(X(t+s)|X(u),0\le u\le t)=X(t)\mathrm{e}^{rs}. \tag{6.3.4}$$

若以该风险资产为标的物的期权在$t+s$时刻价格为$Y(t+s)$, 那么在t时刻的价格就是

$$Y(t)=E(\mathrm{e}^{-rs}Y(t+s)|X(u),0\le u\le t). \tag{6.3.5}$$

6.3.3 Black-Scholes公式

为了简单, 现在我们假定有一个无套利风险中性市场, 市场中只有一种无风险资产, 该资产的收益率, 即无风险利率为r, 一种股票, 其价格变化可用过程$X=\{X_t;t\ge 0\}$表示, 和一种以该股票为基础证券的欧式看涨期权, 期权的到期日为T, 敲定价格为K.

按期权合约, 期权在到期日的价值为$(X_T-K)^+$. 因此在无套利及风险中性假设下, (6.3.5)表明该期权在$0\le t\le T$时的价格C_t为

$$C_t=E(\mathrm{e}^{-r(T-t)}(X_T-K)^+|X_s,0\le s\le t).$$

假定股价价格变化过程X是一个几何布朗运动$X(t)=\mathrm{e}^{Y(t)}$, 其中$Y(t)$是一个漂移系数为μ、方差为σ^2的布朗运动. 由(6.3.2)可得

$$\begin{aligned}C_t&=\mathrm{e}^{-r(T-t)}E((X_T-K)^+|X_s,0\le s\le t)\\&=\mathrm{e}^{-r(T-t)}\int_{-\infty}^{+\infty}(y\mathrm{e}^x-K)^+\phi(x-\mu(T-t),\sigma^2(T-t))\mathrm{d}x\Big|_{y=X(t)}\\&=\mathrm{e}^{-r(T-t)}\int_{\ln(K/y)}^{+\infty}(y\mathrm{e}^x-K)\phi(x-\mu(T-t),\sigma^2(T-t))\mathrm{d}x\Big|_{y=X(t)}.\end{aligned}$$

计算可得

$$\begin{aligned}&\int_{\ln(K/y)}^{+\infty}y\mathrm{e}^x\phi(x-\mu(T-t),\sigma^2(T-t))\mathrm{d}x\\&=y\mathrm{e}^{(T-t)(\mu+\sigma^2/2)}\int_{\ln(K/y)}^{+\infty}\phi(x-(\mu+\sigma^2)(T-t),\sigma^2(T-t))\mathrm{d}x\\&=y\mathrm{e}^{(T-t)(\mu+\sigma^2/2)}\int_{[\ln(K/y)-(\mu+\sigma^2)(T-t)]/(|\sigma|\sqrt{T-t})}^{+\infty}\phi(u,1)\mathrm{d}u\\&=y\mathrm{e}^{(T-t)(\mu+\sigma^2/2)}\Big(1-\Phi\Big(-|\sigma|\sqrt{T-t}-b(K,y,T,t)\Big)\Big)\\&=y\mathrm{e}^{(T-t)(\mu+\sigma^2/2)}\Phi\Big(|\sigma|\sqrt{T-t}+b(K,y,T,t)\Big),\end{aligned}$$

其中
$$b(K,y,T,t)=\frac{\mu(T-t)-\ln(K/y)}{|\sigma|\sqrt{T-t}}.$$
另外, 直接计算可知
$$\begin{aligned}&\int_{\ln(K/y)}^{+\infty}K\phi(x-\mu(T-t),\sigma^2(T-t))\mathrm{d}x\\&=K\int_{[\ln(K/y)-\mu(T-t)]/(|\sigma|\sqrt{T-t})}^{+\infty}\phi(u,1)\mathrm{d}u=K\Phi\big(b(K,y,T,t)\big).\end{aligned}$$
那么C_t为
$$\mathrm{e}^{-r(T-t)}\Big[y\mathrm{e}^{(T-t)(\mu+\frac{\sigma^2}{2})}\Phi(|\sigma|\sqrt{T-t}+b(K,y,T,t))-K\Phi(b(K,y,T,t))\Big]_{y=X(t)}.$$
另一方面, 由(6.3.3)和(6.3.4)可知
$$E[X(t+s)|X(u),0\le u\le t]=X(t)\mathrm{e}^{(\mu+\sigma^2/2)s}=X(t)\mathrm{e}^{rs}.$$
因此$r=\mu+\sigma^2/2$. 将其代入到上面运算结果, 可得
$$C_t=y\Phi\big(|\sigma|\sqrt{T-t}+b(K,y,T,t)\big)-K\mathrm{e}^{-r(T-t)}\Phi\big(b(K,y,T,t)\big)\Big|_{y=X(t)}.\quad(6.3.6)$$
公式(6.3.6)就是著名的欧式看涨期权的Black-Scholes定价公式. 对看跌的欧式期权有类似的定价公式, 我们留作练习请同学自己推算.

由(6.3.6)可知, 欧式看涨期权在t时刻的价格依赖于股票在t时刻的价格y, 期权执行的时间T, 敲定价格K, 市场无风险利率r和股票价格模型中的方差参数σ^2, 并且依据定价原理, 该价格是无套利的. 在实际操作中, 容易保证y, T 和K 都是已知的, 但参数σ^2往往是隐含的. 为了得到σ^2, 注意到以几何布朗运动为价格的股票, 对充分小的Δ,
$$\frac{X(t+\Delta)-X(t)}{X(t)}\approx Y(t+\Delta)-Y(t)=\mu\Delta+\sigma[B(t+\Delta)-B(t)].$$
而$\frac{X(t+\Delta)-X(t)}{X(t)}$表示股票在$[t,t+\Delta]$时间内的收益率. 近似地看, 在不相交的时间区间上股票收益率是独立的, 若不相交的时间区间还是等长的, 收益率还可以看作是同分布的; σ^2为收益率的波动率. 因此人们可用统计方法估计该波动率并将其视为σ^2的值. 一种较自然地获得σ^2估计的方法如下: 假设获得n个时长为h的收益率数据$x_1,\cdots,x_n$. 此时数据总体看作均值为μh和方差为$\sigma^2 h$ 的正态分布, 从而样本方差
$$\frac{1}{n-1}\sum_{k=1}^{n}(x_i-\bar{x})^2$$
是$\sigma^2 h$ 的无偏估计, 其中$\bar{x}=(x_1+\cdots+x_n)/n$表示样本均值.

▶**例6.3.3.** 已知某股票现价为100元, 市场无风险利率约为0.15(单位: 年). 某公司创建一种1年后到期的欧式看涨期权, 执行价格设为120, 试给出该期权现在的每份价格. (假定股票价格服从几何布朗运动, 而且过去*12*个月的月收益率分

别为0.03, 0.09, −0.06, −0.03, 0.10, 0.02, −0.07, 0.04, 0.03, −0.05, 0.02, 0.0)

解 我们先估计σ^2. 由该股票12个月的收益率数据可得月收益率的方差的无偏估计为

$$\frac{1}{11}\sum_{i=1}^{12}(x_i-\bar{x})^2=0.003,$$

其中x_i表示第i个月的收益率, $\bar{x}$表示月平均收益率. 从而$\hat{\sigma}^2=12\times 0.003=0.036$. 将$y=100, K=120, T=1, t=0, r=0.15, \sigma^2=0.036$代入(6.3.6)可得,

$$C_0=100\Phi\big(\sqrt{0.036}+b(120,100,1,0)\big)-120\mathrm{e}^{-0.15}\Phi\big(b(120,100,1,0)\big),$$

其中

$$b(120,100,1,0)=\frac{0.15-\ln 1.2}{\sqrt{0.036}}-\frac{\sqrt{0.036}}{2}=-0.26522.$$

因此

$$C_0\approx 100\Phi(-0.075)-103.285\Phi(-0.265)\approx 47-40.85=6.15.$$

所以该期权现在的每份价格为6.15元. □

我们还可以利用无套利市场下风险中性定价原理给出其他类型期权的价格. 下面这个例子本质上解释了所谓回望看涨期权的定价过程.

▶**例6.3.4.** 假设某股票的价格$X(t)=\mathrm{e}^{B(t)}$, 其中$B=\{B(t);t\geq 0\}$为标准布朗运动. 甲乙两人围绕该股票订立合约, 规定: 乙在时刻s向甲支付d元, 而甲在$s+T$时刻向乙支付$[s,s+T]$时段内该股票最大价格与敲定价格e^K之间的溢价$(\max\limits_{s\leq u\leq s+T}X(u)-\mathrm{e}^K)^+$. 假定无风险利率为1/2, 股票在$s$时刻价格为$\mathrm{e}^x$. 试利用风险中性定价原理确定$d$的值.

解 按风险中性定价原理, d值应该为乙在$s+T$时刻预期收益在s时按无风险利率的折现, 即

$$d=\mathrm{e}^{-T/2}E((\max_{s\leq u\leq T+s}X(u)-\mathrm{e}^K)^+|X(u),0\leq u\leq s)|_{X(s)=\mathrm{e}^x}.$$

由于$X(t)$由标准布朗运动$B(t)$唯一确定, 因此

$$\begin{aligned}&E((\max_{s\leq u\leq T+s}X(u)-\mathrm{e}^K)^+|X(u),0\leq u\leq s)|_{X(s)=\mathrm{e}^x}\\&\qquad=E((\max_{s\leq u\leq T+s}X(u)-\mathrm{e}^K)^+|B(u),0\leq u\leq s)|_{B(s)=x}.\end{aligned}$$

由注记6.1.16以及布朗运动的独立增量性可知

$$\begin{aligned}&E((\max_{s\leq u\leq T+s}X(u)-\mathrm{e}^K)^+|B(u),0\leq u\leq s)|_{B(s)=x}\\&=\lim_{n\to\infty}E((\max_{s\leq u\leq T+s}\mathrm{e}^{B(s)+B(u)-B(s)}-\mathrm{e}^K)^+|B(s),B(t_n),\cdots,B(t_1))|_{B(s)=x}\\&=\lim_{n\to\infty}E((\max_{s\leq u\leq T+s}\mathrm{e}^{B(s)+B(u)-B(s)}-\mathrm{e}^K)^+)|_{B(s)=x}\end{aligned}$$

$$= E((\max_{s \le u \le T+s} \mathrm{e}^{B(s)+B(u)-B(s)} - \mathrm{e}^K)^+)|_{B(s)=x}$$
$$= E((\mathrm{e}^{B(s)} \max_{s \le u \le T+s} \mathrm{e}^{B(u)-B(s)} - \mathrm{e}^K)^+)|_{B(s)=x}$$
$$= E(\mathrm{e}^{B(s)}(\mathrm{e}^{B^*(T)} - \mathrm{e}^K)^+)|_{B(s)=x} = \mathrm{e}^x E((\mathrm{e}^{B^*(T)} - \mathrm{e}^K)^+),$$

其中$B^*(T)$表示B在$[0,T]$上的最大值. 由推论6.2.6, $B^*(T)$的概率密度函数为$2\phi(y,T)$, 从而

$$E((\mathrm{e}^{x+B^*(T)} - \mathrm{e}^K)^+) = \int_{(K-x)^+}^{\infty} (\mathrm{e}^{x+y} - \mathrm{e}^K) 2\phi(y,T)\mathrm{d}y$$
$$= 2\mathrm{e}^x \int_{(K-x)^+}^{\infty} \mathrm{e}^y \phi(y,T)\mathrm{d}y - 2\mathrm{e}^K \int_{(K-x)^+}^{\infty} \phi(y,T)\mathrm{d}y$$
$$= 2\mathrm{e}^{x+T/2} \int_{((K-x)^+-T)/\sqrt{T}}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} \mathrm{d}y - 2\mathrm{e}^K \int_{(K-x)^+/\sqrt{T}}^{\infty} \phi(y,1)\mathrm{d}y$$
$$= 2\big[\mathrm{e}^{x+T/2}\Phi((T-(K-x)^+)/\sqrt{T}) - \mathrm{e}^K\Phi(-(K-x)^+/\sqrt{T})\big].$$

因此$d = 2\big[\mathrm{e}^x\Phi((T-(K-x)^+)/\sqrt{T}) - \mathrm{e}^{K-T/2}\Phi(-(K-x)^+/\sqrt{T})\big]$. □

练习题6.3

1 股票A,B的价格比服从随机过程$X(t) = \mathrm{e}^{B(t)}$其中$B(t)$为标准布朗运动, 已知在时刻$t = 1, t = 2$时价格比值分别是$\mathrm{e}^{0.5}$和e, 试估算在$(1,2)$时间段内B股票价格曾达到或超过A价格的概率.

2 在无套利假设和风险中性原理下推导欧式看跌期权的定价公式.

6.4 高斯过程与积分布朗运动*

高斯过程是一类由正态随机变量构成的随机过程. 布朗运动是一类特殊的高斯过程. 固定两个端点值的布朗运动以及积分布朗运动也是高斯过程.

6.4.1 高斯过程与布朗桥

♣**定义6.4.1.** 设$X = \{X(t);\ t \in T\}$是概率空间$(\Omega, \mathfrak{F}, P)$ 上的随机过程. 称X为高斯过程, 若对任意的正整数n以及参数$t_1, \cdots, t_n \in T$,

$$(X(t_1), X(t_2), \cdots, X(t_n))$$

服从n元正态分布(包括退化情形), 亦即存在n维向量$\mu = \{\mu_1, \cdots, \mu_n\}^T$以及$n$阶半正定或正定矩阵$\mathbf{D}$使得对任意$\lambda = (\lambda_1, \cdots, \lambda_n)^T \in \mathbb{R}^n$,

$$E(\mathrm{e}^{\mathrm{i}\sum_{k=1}^n \lambda_k X(t_k)}) = \exp\left\{\mathrm{i}\lambda^T\mu - \frac{1}{2}\lambda^T\mathbf{D}\lambda\right\}.$$

要注意的是, 一般而言上述定义中的向量μ和矩阵$\mathbf{D}$与参数$(t_1, t_2, \cdots, t_n)$有关, 而且对给定的$(t_1, t_2, \cdots, t_n)$, 对应的μ和$\mathbf{D}$就是$(X(t_1), X(t_2), \cdots, X(t_n))$ 的均值向量和协方差矩阵. 由定义还可知, 高斯过程的有限维分布族由它的均值函数和协方差函数唯一确定. 进一步地, 由定理6.1.4可知, 布朗运动B是一个高斯过程, 而且由推论6.1.12可知在$B(s), B(t)$给定的条件下, 布朗运动$\{B(u);\ s < u < t\}$ 也是高斯过程.

★**定理6.4.2.** $X = \{X(t);\ t \in T\}$是高斯过程当且仅当其中任意有限个随机变量的线性组合都服从一元正态分布.

证明 “必要性”: 任取有限个随机变量$X(t_1), \cdots, X(t_n)$以及任意$\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, 线性组合$\sum_{k=1}^n \lambda_k X(t_k)$的特征函数

$$E(\mathrm{e}^{\mathrm{i}u(\sum_{k=1}^n \lambda_k X(t_k))}) = E(\mathrm{e}^{\mathrm{i}(\sum_{k=1}^n u\lambda_k X(t_k))}) = \exp\left\{\mathrm{i}u\lambda^T\mu - \frac{u^2}{2}\lambda^T\mathbf{D}\lambda\right\}.$$

由此可知$\sum_{k=1}^n \lambda_k X(t_k)$服从正态分布$N(\lambda^T\mu, \lambda^T\mathbf{D}\lambda)$.

“充分性”: 任取$X(t_1), \cdots, X(t_n)$以及$\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, 由于$\sum_{k=1}^n \lambda_k X(t_k)$服从正态分布,

$$\begin{aligned}
E(\mathrm{e}^{\mathrm{i}\sum_{k=1}^n \lambda_k X(t_k)}) &= \exp\left\{\mathrm{i}E(\sum_{k=1}^n \lambda_k X(t_k)) - \frac{1}{2}Var(\sum_{k=1}^n \lambda_k X(t_k))\right\} \\
&= \exp\left\{\mathrm{i}\sum_{k=1}^n \lambda_k E(X(t_k)) - \frac{1}{2}\sum_{k,j=1}^n \lambda_k\lambda_j Cov(X(t_k), X(t_j))\right\} \\
&= \exp\left\{\mathrm{i}\lambda^T\mu - \frac{1}{2}\lambda^T\mathbf{D}\lambda\right\},
\end{aligned}$$

其中$\mu=(E(X(t_1)),\cdots,E(X(t_n)))^T$, $\mathbf{D}=\big(Cov(X(t_i),X(t_j))\big)_{i,j}$正定或半正定. 因此$(X(t_1),\cdots,X(t_n))$ 服从均值为μ, 协方差矩阵为$\mathbf{D}$的n元正态分布. □

▶**例6.4.3.** 设W为标准布朗运动, 对任意$x,y\in\mathbb{R}$, $0\le s<t$, 以及$u\in[s,t]$, 令

$$W_x^y(u)=W(u)+\Big(\frac{t-u}{t-s}(x-W(s))+\frac{u-s}{t-s}(y-W(t))\Big).$$

一般称$\{W_x^y(u)\}_{s\le u\le t}$为布朗桥. 任取$s\le u_1<\cdots<u_n\le t$, $\lambda_1,\cdots,\lambda_n\in\mathbb{R}$,

$$\sum_{k=1}^n\lambda_kW_x^y(u_k)=\sum_{k=1}^n\lambda_k\Big[W(u_k)+\Big(\frac{t-u_k}{t-s}(x-W(s))+\frac{u_k-s}{t-s}(y-W(t))\Big)\Big]$$
$$=\sum_{k=1}^n\lambda_kW(u_k)-\Big(\sum_{k=1}^n\frac{\lambda_k(t-u_k)}{t-s}\Big)W(s)-\Big(\sum_{k=1}^n\frac{\lambda_k(u_k-s)}{t-s}\Big)W(t)+C,$$

其中常数$C=\sum_{k=1}^n\lambda_k[x+\frac{u_k-s}{t-s}(y-x)]$. 由于标准布朗运动是高斯过程, 定理*6.4.2*表明布朗桥也是高斯过程.

注意到$W_x^y(s)\equiv x, W_x^y(t)\equiv y$,

$$E(W_x^y(u))=x+\frac{u-s}{t-s}(y-x),$$

而且直接计算可知对任意$u\le v\in[s,t]$,

$$\begin{aligned}&Cov(W_x^y(u),W_x^y(v))\\&=Cov\Big(W(u)-\frac{t-u}{t-s}W(s)-\frac{u-s}{t-s}W(t),W(v)-\frac{t-v}{t-s}W(s)-\frac{v-s}{t-s}W(t)\Big)\\&=Cov\Big(W(u),W(v)-\frac{t-v}{t-s}W(s)-\frac{v-s}{t-s}W(t)\Big)=\frac{(t-v)(u-s)}{t-s}.\end{aligned}$$

由于高斯过程有限维分布族由它的均值和协方差函数确定, 对照推论6.1.10,可知

$$\{W_x^y(u);\ u\in[s,t]\}$$

与给定$B(s)=x,B(t)=y$条件下的布朗运动$\{B(u);s\le u\le t\}$有相同的有限维分布族, 从而可看作是它的一个版本. 不同的是, 在$\{W_x^y(u)\}$ 的构造中, 我们并没有对标准布朗运动W提出额外的要求.

▲**命题6.4.4.** 若轨道连续的高斯过程$\{X(t);t\ge0\}$对任意$0\le s\le t$都有

$$E(X(s))=0 \text{ 且 } E(X(s)X(t))=s,$$

那么$X=\{X(t);\ t\ge0\}$是标准布朗运动.

证明 由条件可知, 只需验证布朗运动的前两个条件(a),(b).

(a) 由条件$E\big([X(0)]^2\big)=0$可知$X(0)=0$ a.s. 对任意$0\le s_1<t_1\le s_2<t_2$,

$$\begin{aligned}&E[(X(t_2)-X(s_2))(X(t_1)-X(s_1))]\\&=E[X(t_2)X(t_1)]-E[X(t_2)X(s_1)]-E[X(s_2)X(t_1)]+E[X(s_2)X(s_1)]=0.\end{aligned}$$

这表明$X(t_2)-X(s_2)$与$X(t_1)-X(s_1)$不相关. 另一方面, 由X是高斯过程可知, $(X(t_2)-X(s_2),X(t_1)-X(s_1))$服从二元正态分布. 因此$X(t_2)-X(s_2)$与$X(t_1)-X(s_1)$相互独立. 由此进一步可得, 对任意$0\le t_1<t_2<\cdots<t_n$,

$$X(t_1),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1})$$

的协方差矩阵为对角矩阵以及$(X(t_1),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1}))$服从$n$元正态分布. 因此$(X(t_1),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1})$ 相互独立.

(b) 对任意$0\le s<t$, 由$E(X(t)-X(s))=0$,

$$E[(X(t)-X(s))^2]=E[X^2(t)+X^2(s)-2X(s)X(t)]=t-s,$$

以及$X(t)-X(s)$服从正态分布可知$X(t)-X(s)\sim N(0,t-s)$. □

▲**命题6.4.5.** 设ξ_n, $n\ge1$, 是一列正态随机变量. 若$\xi_n\xrightarrow{P}X$, 那么X也是正态随机变量, 而且

$$E(X)=\lim_{n\to\infty}E(\xi_n),\quad Var(X)=\lim_{n\to\infty}Var(\xi_n).$$

证明 由于ξ_n, $n\ge1$, 是正态随机变量, 存在μ_n, σ_n^2 使得对任意$\lambda\in\mathbb{R}$,

$$E(\mathrm{e}^{\mathrm{i}\lambda\xi_n})=\exp\{\mathrm{i}\lambda\mu_n-\frac{\lambda^2}{2}\sigma_n^2\}.$$

由$\xi_n\xrightarrow{P}X$可知$P(|X|<\infty)=1$而且

$$E(\mathrm{e}^{\mathrm{i}\lambda\xi_n})\to\Psi(\lambda)=E(\mathrm{e}^{\mathrm{i}\lambda X}).$$

因此

$$|E(\mathrm{e}^{\mathrm{i}\lambda\xi_n})|=\mathrm{e}^{-\frac{\lambda^2}{2}\sigma_n^2}\to|\Psi(\lambda)|,\qquad \mathrm{i}\lambda\mu_n-\frac{\lambda^2}{2}\sigma_n^2\to\ln\Psi(\lambda).$$

注意到$\Psi(0)=1$, 由$\Psi(\lambda)$连续可知, 存在$\delta>0$使得对任意$\lambda\in(-\delta,\delta)$,

$$|\Psi(\lambda)|>0.$$

这表明$n\to\infty$ 时, σ_n^2 极限存在且有限, 记作σ^2. 注意到

$$E(\mathrm{e}^{\mathrm{i}\lambda(\xi_n-\mu_n)})=\exp\{-\frac{\lambda^2}{2}\sigma_n^2\}\to\exp\{-\frac{\lambda^2}{2}\sigma^2\}.$$

因此$\xi_n-\mu_n$依分布收敛到一个服从正态分布$N(0,\sigma^2)$的随机变量. 记

$$\liminf_{n\to\infty}\mu_n=a,\ \ \limsup_{n\to\infty}\mu_n=b,$$

由$\xi_n\xrightarrow{P}X$可知$X-a$与$X-b$都服从正态分布$N(0,\sigma^2)$. 这表明$a=b<\infty$, 进而μ_n极限存在, 记作μ. 所以$X\sim N(\mu,\sigma^2)$为正态随机变量. □

由命题6.4.5容易得到以下推论, 证明请读者自己完成.

♦**推论6.4.6.** 设$\{X(t);\ t\in T\}$是高斯过程, $t_l^k\in T$, $(k=1,2,\cdots,m,l=1,2,\cdots)$. 若对任何固定的$k$, $X(t_l^k)$依概率收敛于$Y(k)$ 那么

$$\{X(t);\ t\in T\}\bigcup\{Y(k);\ k=1,2,\cdots,m\}$$

仍是高斯过程, 而且$Cov(Y(k),Y(n))=\lim\limits_{l\to\infty}Cov(X(t_l^k),X(t_l^n))$.

6.4.2 与布朗运动有关的简单积分

在本章的最后, 我们介绍两类特殊情形下与布朗运动有关的积分.

设B是布朗运动,由于$B(t)$关于t连续, 黎曼积分$\int_0^t B(u)\mathrm{d}u$ 总有意义. 一般地称随机过程$\{\int_0^t B(u)\mathrm{d}u, t\geq 0\}$ 为积分布朗运动.

更一般地, 设$f(t)$是具有连续一阶导数的非随机函数, 对任意$0\leq a\leq b$, 黎曼积分

$$\int_a^b B(u)\mathrm{d}f(u)=\int_a^b B(u)f'(u)\mathrm{d}u=\lim_{\delta\to 0}\sum_{\substack{i=1\\ \delta=\max\{t_i-t_{i-1}\}}}^n B(t_i)f'(t_i)(t_i-t_{i-1})$$

总有意义, 其中$a=t_0<t_1<\cdots<t_n=b$为$[a,b]$的任一分割. 注意上式右边极限中每个部分和都是正态随机变量, 因此由命题6.4.5可知$\int_a^b B(u)\mathrm{d}f(u)$ 服从正态分布. 进一步由推论6.4.6 还可知随机过程

$$\Big\{\int_a^t B(u)\mathrm{d}f(u),\ t\geq a\Big\}$$

是高斯过程. 直接计算可知对任意$s,t\geq a$,

$$E\Big(\int_a^t B(u)\mathrm{d}f(u)\Big)=\int_0^t E(B(u))\mathrm{d}f(u)=0,$$

$$\begin{aligned}Cov\Big(\int_a^t B(u)\mathrm{d}f(u),\int_a^s B(v)\mathrm{d}f(v)\Big)&=\int_a^t\int_a^s E(B(u)B(v))\mathrm{d}f(u)\mathrm{d}f(v)\\&=\int_a^s\int_a^t (s\wedge t)\mathrm{d}f(s)\mathrm{d}f(u).\end{aligned}$$

由上述分析可知, 当B是标准布朗运动时, 积分布朗运动$\int_0^t B(u)\mathrm{d}u$是均值为0、协方差(函数)为

$$\begin{aligned}R(s,t)&=\int_0^s\int_0^t(u\wedge v)\mathrm{d}u\mathrm{d}v=\int_0^s\int_0^s(u\wedge v)\mathrm{d}u\mathrm{d}v+\int_0^s\int_s^t(u\wedge v)\mathrm{d}u\mathrm{d}v\\&=\frac{s^3}{3}+\frac{(t-s)s^2}{2}=s^2(\frac{t}{2}-\frac{s}{6}),\quad s\leq t,\end{aligned}$$

的高斯过程, 而且$\int_0^t B(u)\mathrm{d}u\sim N(0,t^3/3)$.

对布朗运动还可以讨论另外一种形式的积分. 这里我们简单介绍$f(t)$是具有连续一阶导数的非随机函数的情形. 对任意$0\leq a\leq b$, 定义区间$[a,b]$上随机积分

$$\int_a^b f(s)\mathrm{d}B(s)=\lim_{\delta\to 0}\sum_{i=1}^n f(t_{i-1})[B(t_i)-B(t_{i-1})],\tag{6.4.1}$$

其中$a=t_0<t_1<\cdots<t_n=b$为$[a,b]$的任一分割, $\delta=\max\{t_i-t_{i-1};1\leq i\leq n\}$. 注意到

$$\sum_{i=1}^n f(t_{i-1})[B(t_i)-B(t_{i-1})]=f(b)B(b)-f(a)B(a)-\sum_{i=1}^n B(t_i)(f(t_i)-f(t_{i-1})).$$

因此

$$\int_a^b f(s)\mathrm{d}B(s) = f(b)B(b) - f(a)B(a) - \int_a^b B(s)\mathrm{d}f(s).$$

通常我们称该式为随机积分的分部积分公式. 由(6.4.1)可知, 当$\phi(t)$为非随机函数时, 极限中的部分和仍为正态随机变量, 进而$\int_a^b f(s)\mathrm{d}B(s)$也是正态随机变量. 进一步, 由推论6.4.6可知随机过程

$$\Big\{\int_a^t f(u)\mathrm{d}B(u),\ t \geq a\Big\}$$

也是高斯过程. 当B是标准布朗运动时, 对任意$a \leq s \leq t$ 以及$[0,t]$中的分割$0 = t_0 < t_1 < \cdots < t_m = s < \cdots < t_n = t$,

$$\begin{aligned}
E\Big(\int_a^t f(u)\mathrm{d}B(u)\Big) &= \lim_{\delta\to 0}\sum_{i=1}^n f(t_{i-1})(E[B(t_i)] - E[B(t_{i-1})]) \\
&= \int_a^t f(u)\mathrm{d}E(B(u)) = 0,
\end{aligned}$$

$$\begin{aligned}
&Cov\Big(\int_a^t f(u)\mathrm{d}B(u), \int_a^s f(u)\mathrm{d}B(u)\Big) \\
&\quad = \lim_{\delta\to 0} E\Big(\sum_{i=1}^n f(t_{i-1})[B(t_i) - B(t_{i-1})]\sum_{i=1}^m f(t_{i-1})[B(t_i) - B(t_{i-1})]\Big) \\
&\quad = \lim_{\delta\to 0} E\Big(\sum_{i=1}^m f^2(t_{i-1})[B(t_i) - B(t_{i-1})]^2\Big) \\
&\quad = \lim_{\delta\to 0}\sum_{i=1}^m f^2(t_{i-1})(t_i - t_{i-1}) = \int_a^s f^2(u)\mathrm{d}u,
\end{aligned}$$

其中倒数第二个等号用到了布朗运动的独立增量性.

▶**例6.4.7.** *设B是标准布朗运动, 试求$\int_0^1 u\mathrm{d}B(u)$在$B(1) = 1$条件下的分布.*

解 由于$B(1) = 1$条件下的布朗运动仍是高斯过程, 由(6.4.1)可知, 此时$\int_0^1 u\mathrm{d}B(u)$仍是正态随机变量. 由推论6.1.10以及本章第一节习题2, 仿上面的计算过程可知

$$E(\int_0^1 u\mathrm{d}B(u)|B(1) = 1) = \int_0^1 u\mathrm{d}E(B(u)|B(1) = 1) = \int_0^1 u\mathrm{d}u = \frac{1}{2},$$

$$\begin{aligned}
E\Big[\Big(\int_0^1 u\mathrm{d}B(u)\Big)^2\Big|B(1) = 1\Big] &= \lim_{\delta\to 0}\sum_{i=1}^m t_{i-1}^2 E\Big([B(t_i) - B(t_{i-1})]^2\Big|B(1) = 1\Big) \\
&= \lim_{\delta\to 0}\sum_{i=1}^m t_{i-1}^2(t_i - t_{i-1}) = \int_0^1 t^2\mathrm{d}t = 1/3.
\end{aligned}$$

由此可知$B(1) = 1$时$\int_0^1 u\mathrm{d}B(u) \sim N(1/2, 1/12)$. □

要指出的是在积分表达式$\int f(t)\mathrm{d}B(t)$中, 虽然用了微分符号$\mathrm{d}B(t)$, 但它并

不是$B(t)$的真实微分, 因为由定理6.1.18可知布朗运动B 是处处不可微的. 通常我们称$\mathrm{d}B(t)$ 为白噪声(或将$\mathrm{d}B(t)$记作$\dot{B}(t)\mathrm{d}t$, 而称$\dot{B}(t)$为白噪声).

此外, 当函数$f(t)$并不连续可微, 甚至$f(t)$也是一个随机过程时, 我们仍然可以在一定条件和一定意义下定义积分$\int f(t)\mathrm{d}B(t)$, 其中最重要的积分就是所谓的伊藤清(Itô)积分. Itô积分是现代随机分析中的重要内容, 也有着广泛应用. 但这些内容已超出本书范畴, 对此感兴趣的读者可以参阅有关书籍(比如 [4, 9]).

练习题6.4

1 设$B=\{B(t);t\geq 0\}$为标准布朗运动, 求$Y=\int_0^1 B(s)ds$的分布.若$B(1)=x$, 试在此条件下再求Y的分布.

2 设B是标准布朗运动, 试求$\int_0^1 u\mathrm{d}B(u)$在$B(1)=x$条件下的分布.

3* 证明推论6.4.6.

第七章　离散时间鞅

鞅(Martingale)这个概念来自赌博策略研究, 其构词法来自法文中的首字母缩写. 鞅自Lévy和Doob 等人将其引入概率论后, 目前已成为研究随机问题的一个强有力工具. 本章将简要地介绍离散时间鞅的基本理论.

7.1　鞅与停时

7.1.1　鞅的定义及举例

♣**定义7.1.1.** *设$X=\{X_n;n\geq 0\}$, $Y=\{Y_n;n\geq 0\}$是两个随机过程, 称随机过程X关于Y是鞅(下鞅, 上鞅), 若对每个$n\geq 0$,*

(1) $E(|X_n|)<\infty$;

(2) X_n是$Y_0,\cdots,Y_n$的函数;

(3) $E(X_{n+1}|Y_0,\cdots,Y_n)=X_n$(相应地 $\geq X_n, \leq X_n$).

♠**注记7.1.2.** *当X关于Y是鞅时, 由条件数学期望性质, 定义中条件(2)自然成立.*

从定义上看, 所谓(下, 上)鞅就是在已知$Y_0,\cdots,Y_n$条件下, X_{n+1}的条件期望与X_n相同(不小于X_n, 不大于X_n). 更形象些, 若将时刻n理解为现在, $Y_0,\cdots,Y_n$理解为从开始到现在累积的已知信息, 那么(下, 上)鞅就是在掌握现在所有已知信息的条件下预测未来与现在相同(优于现在, 劣于现在). 若$Y=X$或由上下文Y不言自明时, 我们也简称$\{X_n;n\geq 0\}$为鞅(下鞅, 上鞅).

作为(下, 上)鞅最早的例子, 出现在赌徒问题研究中.

▶**例7.1.3.** *假定某赌徒参与某种轮盘赌, 设赌徒初始赌本为a, 以X_i表示赌徒每局赢或输的筹码, 以Y_0表示赌局的初始信息, Y_i表示赌徒从第i个赌局中所能获得的所有信息, 令$S_0=a$, $S_n=S_{n-1}+X_n=a+\sum_{k=1}^n X_k$, 那么$S_n$表示第$n$局*

后该赌徒拥有的资本, 可以看作是$Y_0,\cdots,Y_n$的函数. 假定$E(|S_n|)<+\infty$对一切$n\geq 1$都成立.

(1) 若对任意$n\geq 0$, $E(X_{n+1}|Y_0,\cdots,Y_n)=0$, 那么

$$\begin{aligned}E(S_{n+1}|Y_0,\cdots,Y_n)&=E(S_n+X_{n+1}|Y_0,\cdots,Y_n)\\&=E(S_n|Y_0,\cdots,Y_n)+E(X_{n+1}|Y_0,\cdots,Y_n)=S_n,\end{aligned}$$

即$S=\{S_n;n\geq 0\}$关于$Y=\{Y_n;n\geq 0\}$是鞅.

(2) 若对任意$n\geq 0$, $E(X_{n+1}|Y_0,\cdots,Y_n)>0$, 那么

$$\begin{aligned}E(S_{n+1}|Y_0,\cdots,Y_n)&=E(S_n+X_{n+1}|Y_0,\cdots,Y_n)\\&=E(S_n|Y_0,\cdots,Y_n)+E(X_{n+1}|Y_0,\cdots,Y_n)>S_n,\end{aligned}$$

即S关于Y是下鞅, 但不是殃.

(3) 若对任意$n\geq 0$, $E(X_{n+1}|Y_0,\cdots,Y_n)<0$, 那么

$$\begin{aligned}E(S_{n+1}|Y_0,\cdots,Y_n)&=E(S_n+X_{n+1}|Y_0,\cdots,Y_n)\\&=E(S_n|Y_0,\cdots,Y_n)+E(X_{n+1}|Y_0,\cdots,Y_n)<S_n,\end{aligned}$$

即S关于Y是上鞅, 但不是鞅.

直观而言, 在两人博弈的赌博模型中, 若S为鞅, 则赌局(设置)是公平的; 若S不是鞅, 那么赌局(设置)是不公平的, 上鞅则有利于对手, 下鞅则有利于自己. 由赌徒模型还可以看出, 在鞅的定义中引进"信息"过程Y在有些时候是很有必要的, 因为赌本S_n的变化并不总能完全反映赌场信息.

下面的例子表明, 鞅(现象)在随机过程中可以自然地出现.

▶**例7.1.4.** 设$Y=\{Y_n;n\geq 0\}$是状态空间S上时齐马氏链, 转移概率矩阵$\mathbf{P}=(p_{i,j})$. 对S上任意有界函数g, 令$\phi(i)\triangleq\sum_{j\in S}p_{i,j}g(j)$, $i\in S$, 定义

$$X_0=g(Y_0),\quad X_n=g(Y_n)-\sum_{k=0}^{n-1}[\phi(Y_k)-g(Y_k)],\quad n\geq 1.$$

显然X_n是$Y_0,\cdots,Y_n$的函数而且$E(|X_n|)<+\infty$. 由条件期望性质与马氏性, 对任意$n\geq 1$,

$$\begin{aligned}E(X_{n+1}|Y_0,\cdots,Y_n)&=E(g(Y_{n+1})|Y_n)-\sum_{k=0}^{n}[\phi(Y_k)-g(Y_k)]\\&=\sum_{j\in S}p_{i,j}g(j)|_{i=Y_n}-\sum_{k=0}^{n}[\phi(Y_k)-g(Y_k)]\\&=\phi(Y_n)-\sum_{k=0}^{n}[\phi(Y_k)-g(Y_k)]\end{aligned}$$

$$= g(Y_n) - \sum_{k=0}^{n-1}[\phi(Y_k) - g(Y_k)] = X_n.$$

这表明$\{X_n; n \geq 0\}$关于Y为鞅. 特别地, 若对任意$i \in S$, $\phi(i) = (\geq, \leq)g(i)$, 那么$\{g(Y_n); n \geq 0\}$为鞅(下鞅, 上鞅).

▶**例7.1.5.** 设$X = \{X_n; n \geq 1\}$为独立同分布随机变量.

(1) 若$E(X_n) = \mu$, 令$S_0 = 0$, $S_n = \sum\limits_{k=1}^{n} X_i - n\mu$, 则$\{S_n; n \geq 0\}$为鞅.

(2) 若存在$\lambda \in \mathbb{R}$使得$\phi(\lambda) = E(\mathrm{e}^{\lambda X_n}) < \infty$, 令

$$Y_n = [\phi(\lambda)]^{-n}\mathrm{e}^{\lambda \sum\limits_{i=1}^{n} X_i}, \quad n \geq 0, \ (\text{规定}Y_0 = 1).$$

则$\{Y_n; n \geq 0\}$为鞅. 特别地, 若$\phi(\lambda_0) = 1$, 则$\left\{\mathrm{e}^{\lambda_0 \sum\limits_{i=1}^{n} X_i}; n \geq 0\right\}$为鞅.

(3) 设f_0, f_1为两概率密度函数, 令

$$L_n = \prod_{i=1}^{n} \frac{f_1(X_i)}{f_0(X_i)}, \qquad n \geq 0, \ (\text{规定}L_0 = 1).$$

若以f_0为X_n的概率密度函数为假设检验中的原假设, 那么$\{L_n; n \geq 0\}$为鞅.

证明 (1) 对任意$n \geq 0$,

$$\begin{aligned} E(S_{n+1}|S_0, X_1, \cdots, X_n) &= E(S_n + X_{n+1} - \mu|S_0, X_1, \cdots, X_n) \\ &= S_n + E(X_{n+1} - \mu|S_0, X_1, \cdots, X_n) = S_n. \end{aligned}$$

所以$\{S_n; n \geq 0\}$为鞅.

(2)对任意$n \geq 0$,

$$\begin{aligned} E(Y_{n+1}|Y_0, X_1, \cdots, X_n) &= E(Y_n\phi^{-1}(\lambda)\mathrm{e}^{\lambda X_{n+1}}|Y_0, X_1, \cdots, X_n) \\ &= Y_n\phi^{-1}(\lambda)E(\mathrm{e}^{\lambda X_{n+1}}) = Y_n. \end{aligned}$$

所以$\{Y_n; n \geq 0\}$为鞅.

(3)在X_n的概率密度函数为f_0的条件假设下, 对任意$n \geq 0$,

$$\begin{aligned} E(L_{n+1}|L_0, X_1, \cdots, X_n) &= E\Big(L_n\frac{f_1(X_{n+1})}{f_0(X_{n+1})}\Big|L_0, X_1, \cdots, X_n\Big) \\ &= L_nE\Big(\frac{f_1(X_{n+1})}{f_0(X_n)}\Big) = L_n\int_{-\infty}^{\infty} f_1(x)dx = L_n. \end{aligned}$$

所以$\{L_n; n \geq 0\}$为鞅. □

▶**例7.1.6.** (1) 设$N(t)$是强度为λ的泊松过程, 对任意非负递增数列$\{t_n\}$, 令

$$M_n = N(t_n) - \lambda t_n.$$

那么由条件数学期望性质与独立增量性,

$$\begin{aligned} &E(M_{n+1}|M_0, \cdots, M_n) = M_n + E(M_{n+1} - M_n|M_0, \cdots, M_n) \\ &\quad = M_n + E(N(t_{n+1}) - N(t_n) - \lambda(t_{n+1} - t_n)|N(t_0), \cdots, N(t_n)) \end{aligned}$$

$$= M_n + E(N(t_{n+1}) - N(t_n) - \lambda(t_{n+1} - t_n)) = M_n.$$

因此$\{N(t_n) - \lambda t_n; n \geq 0\}$ 是鞅.

(2)设$B(t)$是方差为σ^2的布朗运动, 对任意$0 \leq t_0 < t_1 < \cdots < t_n < \cdots$, 令

$$M_n = \exp\big\{\lambda B(t_n) - \frac{\lambda\sigma^2 t_n}{2}\big\}.$$

那么同样由条件数学期望性质,独立增量性与正态性,

$$\begin{aligned}
&E(M_{n+1}|M_0, \cdots, M_n)\\
&= M_n E\Big(\exp\Big\{\lambda(B(t_{n+1}) - B(t_n)) - \frac{\lambda\sigma^2(t_{n+1}-t_n)}{2}\Big\}\Big|B(t_0), \cdots, B(t_n)\Big)\\
&= M_n E\Big(\exp\Big\{\lambda(B(t_{n+1}) - B(t_n)) - \frac{\lambda\sigma^2(t_{n+1}-t_n)}{2}\Big\}\Big) = M_n.
\end{aligned}$$

因此$\big\{\exp\big\{\lambda B(t_n) - \frac{\lambda\sigma^2 t_n}{2}\big\}; n \geq 0\big\}$ 也是鞅.

♠**注记7.1.7.** 上面两例只验证了鞅定义中的性质(*3*). 严格而言, 这是不充分的, 我们还需验证鞅定义中的条件(*1*)和(*2*). 具体验证过程简单, 请读者自己完成.

7.1.2 鞅的简单性质

由定义7.1.1可知, 在给定“信息”过程条件下, $X = \{X_n; n \geq 0\}$为上鞅当且仅当$-X$为下鞅; X为鞅当且仅当它既是上鞅又是下鞅. 由条件数学期望的平滑性还可知, 若随机过程X关于$\{Z_n; n \geq 0\}$是鞅(下鞅), 那么对任意$n \geq 0$以及$k \geq 1$,

$$\begin{aligned}
E(X_{n+k}|Z_0, \cdots, Z_n) &= E(E(X_{n+k}|Z_0, \cdots, Z_{n+k-1})|Z_0, \cdots, Z_n)\\
&= (\geq) E(X_{n+k-1}|Z_0, \cdots, Z_n) = (\geq) X_n. \qquad (7.1.1)
\end{aligned}$$

■**性质7.1.8.** 若随机过程$\{X_n; n \geq 0\}$, $\{Y_n; n \geq 0\}$关于$\{Z_n; n \geq 0\}$都是上鞅(下鞅), 那么

(1) $E(X_n), E(Y_n)$ 是非增(降)数列.

(2) 对任意非负实数α, β, $\{\alpha X_n + \beta Y_n; n \geq 0\}$是上鞅(下鞅).

(3) $\{X_n \wedge Y_n; n \geq 0\}$ 是上鞅($\{X_n \vee Y_n; n \geq 0\}$ 是下鞅).

证明 以上鞅为例, 只证结论(3). 对任意$n \geq 0$, 由假设易知

$$E(|X_n \wedge Y_n|) \leq E(|X_n|) + E(|Y_n|) < +\infty.$$

又由

$$E(X_{n+1} \wedge Y_{n+1}|Z_0, \cdots, Z_n) \leq \begin{cases} E(Y_{n+1}|Z_0, \cdots, Z_n) \leq Y_n, \\ E(X_{n+1}|Z_0, \cdots, Z_n) \leq X_n, \end{cases}$$

可知

$$E(X_{n+1}\wedge Y_{n+1}|Z_0,\cdots,Z_n)\le \min\{X_n,Y_n\}=X_n\wedge Y_n.$$

所以$\{X_n\wedge Y_n;n\ge 0\}$ 仍是上鞅. □

▼**引理7.1.9.** (*Jensen*不等式) 设$\psi(x)$ 是$\mathbb{R}$上的凸函数, X为随机变量. 若

$$E(|X|)<\infty \text{ 且 } E(|\psi(X)|)<\infty,$$

那么$\psi(E(X))\le E(\psi(X))$. 进一步, 任取随机变量序列$\{Y_n;n\ge 0\}$, 对任意$n\ge 0$

$$\psi(E(X|Y_0,\cdots,Y_n))\le E(\psi(X)|Y_0,\cdots,Y_n).$$

证明 只证后一个不等式. 由于ψ是凸函数, 由微积分知识可知ψ的左右导数都存在且对任意$x,y\in\mathbb{R}$,

$$\psi(y)-\psi(x)\ge \psi'_+(x)(y-x),$$

其中$\psi'_+(y)$表示y的右导数. 因此

$$\psi(x)-\psi(E(X|Y_0,\cdots,Y_n))\ge \psi'_+(E(X|Y_0,\cdots,Y_n))(X-E(X|Y_0,\cdots,Y_n)).$$

两边关于$Y_0,\cdots,Y_n$求条件期望得

$$\begin{aligned}&E(\psi(x)-\psi(E(X|Y_0,\cdots,Y_n))|Y_0,\cdots,Y_n)\\&\quad\ge \psi'_+(E(X|Y_0,\cdots,Y_n))E(X-E(X|Y_0,\cdots,Y_n)|Y_0,\cdots,Y_n)\\&\quad\ge \psi'_+(E(X|Y_0,\cdots,Y_n))[E(X|Y_0,\cdots,Y_n)-E(X|Y_0,\cdots,Y_n)]=0.\end{aligned}$$

因此$E(\psi(x)|Y_0,\cdots,Y_n)\ge \psi(E(X|Y_0,\cdots,Y_n))$. □

由于$r>1$时$|x|^r$在$\mathbb{R}$上为凸函数, 由Jensen不等式可得Hölder不等式

$$E(|X|^r)\ge [E(|X|)]^r \quad 即 \quad E(|X|)\le [E(|X|^r)]^{1/r}.$$

■**性质7.1.10.** 设$\{X_n;n\ge 0\}$关于$\{Y_n;n\ge 0\}$为(下)鞅, $\psi(x)$为$\mathbb{R}$上的(非降)凸函数. 若对所有$n\ge 0$, $E(|\psi(X_n)|)$都有限, 那么$\{\psi(X_n);n\ge 0\}$关于$\{Y_n\}$是下鞅.

证明 由Jensen不等式, 对任意$n\ge 0$,

$$E(\psi(X_{n+1})|Y_0,\cdots,Y_n)\ge \psi(E(X_{n+1}|Y_0,\cdots,Y_n)=\psi(X_n)(\ge\psi(X_n)).$$

因此$\{\psi(X_n);n\ge 0\}$是下鞅. □

利用性质7.1.10我们容易构造出下鞅. 若$\{X_n;n\ge 0\}$是鞅, 只要下面的随机变量序列有意义且可积, 那么

$$\{|X_n|;n\ge 0\},\quad \{\mathrm{e}^{X_n};n\ge 0\},\quad \{-\ln(X_n);n\ge 0\}$$

都是下鞅, 而当$r\ge 1$时$\{|X_n|^r;n\ge 0\}$ 也是下鞅.

■**性质7.1.11.** 若$\{X_n;n\ge 0\}$关于$\{Y_n;n\ge 0\}$是下鞅, 那么存在一个关于$\{Y_n;n\ge 0\}$的鞅$\{M_n;n\ge 0\}$使得对任意$n\ge 0$, $X_n=M_n+N_n$, 其中$\{N_n;n\ge 0\}$是一个单调不降的随机序列, $N_0=0$且N_n由$Y_0,\cdots,Y_{n-1}$完全给定.

证明 对任意$n \geq 1$, 令

$$N_n = \sum_{k=0}^{n-1}[E(X_{k+1}|Y_0, \cdots, Y_k) - X_k].$$

由条件数学期望与下鞅定义可知, N_n是$Y_0, \cdots, Y_{n-1}$的函数,而且$\{N_n; n \geq 1\}$是一个单调不降的非负随机序列. 对任意$n \geq 0$, 再令$M_n = X_n - N_n$, 其中$N_0 = 0$. 那么对任意$n \geq 0$,

$$E(|M_n|) \leq E(N_n) + E(|X_n|) = E(X_n) - E(X_0) + E(|X_n|) < +\infty,$$

而且

$$\begin{aligned}
&E(M_{n+1}|Y_0, \cdots, Y_n) \\
&= E(X_{n+1}|Y_0, \cdots, Y_n) - \sum_{k=0}^{n}[E(X_{k+1}|Y_0, \cdots, Y_k) - X_k] \\
&= X_n - \sum_{k=0}^{n-1}[E(X_{k+1}|Y_0, \cdots, Y_k) - X_k] = M_n.
\end{aligned}$$

由此可知M_n为鞅. □

一般地, 我们称N_n这种任意n步时的值能由n步以前信息$\{Y_0, \cdots, Y_{n-1}\}$唯一确定的随机过程为可料的(Predictable)或可容许的(Admissible).

▶**例7.1.12.** *以X_i表示某赌徒第i局的输赢情况: 若其赢则$X_i = 1$, 否则$X_i = -1$. 假设每局胜负结果相互独立, 而且$P(X_i = 1) \leq 1/2$. 问该赌徒能否通过倍增策略(即首局投注1, 后面每局投注都是前一局的两倍直到该赌徒赢得一局, 然后重新执行这种投注策略)实现稳赢? (约定$X_0 = 1$).*

解 以H_n表示第n局的投注量, 显然H_n的可能取值为$\{1, 2, \cdots, 2^{n-1}\}$而且对任意$k \leq n-1$,

$$\{H_n = 2^k\} = \{X_{n-k-1} = 1, X_{n-k} = -1, \cdots, X_{n-1} = -1\}.$$

因此H_n为$X_0, X_1, \cdots, X_{n-1}$的函数, 即存在函数$\phi_n \leq 2^n$使得

$$H_n = \phi_n(X_0, X_1, \cdots, X_{n-1}).$$

到第n局, 该赌徒赢得的赌注为

$$M_n = \sum_{k=1}^{n} H_k X_k.$$

令$M_0 = 0$, 由条件期望性质可得

$$\begin{aligned}
E(M_n|M_0, X_0, \cdots, X_{n-1}) &= \sum_{k=1}^{n} E(H_k X_k|M_0, X_0, \cdots, X_{n-1}) \\
= \sum_{k=1}^{n-1} H_k X_k + H_n E(X_n|M_0, X_0, \cdots, X_{n-1}) &\leq \sum_{k=1}^{n-1} H_k X_k = M_{n-1}.
\end{aligned}$$

即M_n关于X_n为上鞅, 从而$E(M_n)$为非增序列. 这表明倍增策略不能稳赢. □

7.1.3 停时及其简单性质

离散随机过程的停时定义已在第二章第一节中给出. 在前面的章节中, 我们也引入了一些具体的停时, 比如马氏链的首回时就是一个例子. 下面我们再看两个例子:

▶**例7.1.13.** 若T为常数,那么T一定是停时, 此时事件$\{T=n\}$或者为不可能事件$\emptyset$或者为必然事件Ω.

▶**例7.1.14.** 设N是以$\{W_i; i\geq 1\}$为时间间隔序列的更新过程, T_n为第n次更新发生的时间, 那么对任意$t\geq 0$, $N(t)$不是关于$T=\{T_n; n\geq 0\}$的停时, 但$N(t)+1$是关于T的停时. 事实上, 对任意$n\geq 0$, 由

$$\{N(t)=n\}=\{T_n\leq t, T_{n+1}>t\}, \quad \{N(t)+1=n\}=\{T_n>t, T_{n-1}\leq t\}$$

可知, 判断$\{N(t)=n\}$是否发生不仅需要知道T在n及之前的信息还要知道T在n之后的信息, 而判断事件$\{N(t)+1=n\}$是否发生, 则只需要知道T在n及n之前的信息.

对于停时我们有如下的运算性质.

■**性质7.1.15.** 若S, T是关于$\{X_n; n\geq 0\}$的停时, 那么$S+T, S\wedge T, S\vee T$都是停时.

证明 由于S, T均为停时, 对任意$k\geq 0$, 存在ϕ_k, ψ_k使得

$$1_{\{S=k\}}=\phi_k(X_0,\cdots,X_k), \quad 1_{\{T=k\}}=\psi_k(X_0,\cdots X_k).$$

对任意$n\geq 0$, 由$\{S+T=n\}=\cup_{k=0}^n\{S=k, T=n-k\}$可知

$$1_{\{S+T=n\}}=\sum_{k=0}^n 1_{\{S=k\}}1_{\{T=n-k\}}=\sum_{k=0}^n\phi_k(X_0,\cdots,X_k)\psi_{n-k}(X_0,\cdots,X_{n-k}).$$

因此$S+T$为停时. 类似地由

$$\{S\wedge T=n\}=(\{S=n\}\cup\{T=n\})\setminus(\{S=n, T<n\}\cup\{S<n, T=n\}),$$

可知

$$\begin{aligned}1_{\{S\wedge T=n\}}&=\phi_n(X_0,\cdots,X_n)+\psi_n(X_0,\cdots,X_n)\\&\quad-\sum_{k=0}^{n-1}\phi_n(X_0,\cdots,X_n)\psi_k(X_0,\cdots,X_k)\\&\quad-\sum_{k=0}^{n}\phi_k(X_0,\cdots,X_k)\psi_n(X_0,\cdots,X_n)\end{aligned}$$

仍是$X_0,\cdots,X_n$的函数, 因此$S\wedge T$是停时. 同理, $S\vee T$也是停时, 细节省略. □

★定理7.1.16. ($Wald$等式) 设$\{X_n;n\geq 1\}$为独立随机变量序列, 均值相同且有限. 若T是关于$\{X_n;n\geq 1\}$的停时且$E(T)<\infty$. 那么

$$E(\sum_{i=1}^{T}X_i)=E(X_1)E(T). \tag{7.1.2}$$

证明 只证X为非负随机变量序列的情形. 由于T为停时, 对任意$k\geq 1$, 存在函数ϕ_k使得$1_{\{T=k\}}=\phi_k(X_1,\cdots,X_k)$. 由引理1.4.22以及第一章第一节习题8可得

$$\begin{aligned}E\Big(\sum_{k=1}^{T}X_k\Big)&=\sum_{k=1}^{\infty}E\Big(\sum_{i=1}^{k}X_i1_{\{T=k\}}\Big)=\sum_{i=1}^{\infty}\sum_{k=i}^{\infty}E\Big(X_i1_{\{T=k\}}\Big)\\&=\sum_{i=1}^{\infty}E\Big(X_i-X_i\sum_{k=1}^{i-1}\phi_k(X_1,\cdots,X_k)\Big).\end{aligned}$$

利用条件数学期望性质并注意到$\{X_n,n\geq 1\}$独立且均值相同, 我们有

$$\begin{aligned}E\Big(\sum_{k=1}^{T}X_k\Big)&=\sum_{i=1}^{\infty}E\Big(E(X_i-X_i\sum_{k=1}^{i-1}\phi_k(X_1,\cdots,X_k)|X_1,\cdots,X_{i-1})\Big)\\&=\sum_{i=1}^{\infty}E(X_i|X_1,\cdots,X_{i-1})\Big(1-E(\sum_{k=1}^{i-1}\phi_k(X_1,\cdots,X_k))\Big)\\&=E(X_1)\sum_{i=1}^{\infty}(1-E(\sum_{k=1}^{i-1}1_{\{T=k\}}))=E(X_1)\sum_{i=1}^{\infty}P(T\geq i).\end{aligned}$$

注意到$\sum\limits_{i=1}^{\infty}P(T\geq i)=E(T)$, (7.1.2)成立. □

练习题7.1

1 设$Y=\{Y_n;n\geq 0\}$是状态空间S上时齐马氏链, 令$X_n=g(Y_n)=f_{Y_n,j_0}$, 其中f_{i,j_0}表示Y从i出发有限时间内到达(回到)某个给定状态j_0 的概率. 验证X关于Y是上鞅.

2 设$X_1,X_2,\cdots$为独立同分布随机变量满足

$$P(X_1=1)=p,\ \ P(X=-1)=q=1-p,\qquad 0<p<1.$$

对任意$n\geq 0$, 令$Y_0=1$以及

$$Y_n=\Big(\frac{q}{p}\Big)^{S_n},\ 其中\ \ S_n=\sum_{k=1}^{n}X_k,\ \ n\geq 1.$$

证明$\{Y_n;n\geq 0\}$关于X为鞅.

3 设$X_1,X_2,\cdots$为独立同分布随机变量, $E(X_n)=0$, $E(X_n^2)=a^2$. 对任

意$n \geq 0$, 令$Y_0 = 0$,

$$Y_n = (\sum_{k=1}^{n} X_k)^2 - na^2, \qquad n \geq 1.$$

证明$\{Y_n; n \geq 0\}$关于X为鞅.

4 设$\tau = \min\{n \geq 0, X_n = a或b\}$, $\eta = \min\{n \geq 0, X_n \leq a\}$. 证明$\tau, \eta$都是关于过程$X = \{X_n; n \geq 0\}$ 的停时.

5 设$\{M_n; n \geq 0\}$为鞅, 证明对任意$n > m \geq k \geq 0$, $E[(M_n - M_m)M_k] = 0$.

6 证明Lyapunov不等式: 对任意随机变量X, 若$0 < s < t$, 则

$$(E(|X|^s))^{1/s} \leq (E(|X|^t))^{1/t}.$$

7 设T是关于过程$Y = \{Y_k; k \geq 1\}$的停时, 其中Y_k, $k \geq 1$, 是均值为0, 方差为σ^2的独立随机变量. 证明: 若$E(T) < \infty$, 那么

$$E\Big[(\sum_{k=1}^{T} Y_k)^2\Big] = \sigma^2 E(T).$$

8* 证明若$\psi(x)$是区间(a,b)内的凸函数,则函数ψ在(a,b)内左右导数都存在且对任意$x, y \in (a,b)$, $\psi(y) - \psi(x) \geq \psi'_+(x)(y - x)$, 其中$\psi'_+(x)$表示在$x$的右导数.

7.2 鞅的停时定理

本节总设$X = \{X_n; n \geq 0\}$关于某个信息过程$Y = \{Y_n; n \geq 0\}$为(下)鞅, 所有停时也以该信息过程为参考(若X是上鞅, 我们可以通过下鞅$-X$讨论, 故不赘述). 由(7.1.1)可知, 对任意给定的确定时间$n \leq n + k$, 我们总有

$$E(X_{n+k}) = (\geq) E(X_n). \tag{7.2.1}$$

若将时间用停时替换, 上式是否仍然成立? 本节对此问题做些初步介绍.

7.2.1 有界停时定理

一般而言, 将确定时间替换为停时后(7.2.1)并不能总是成立(具体的例子, 留着练习, 请读者自己构造), 但当停时是有界的, 我们会得到肯定的回答. 为了解释这一点, 我们先给出如下的引理.

▼**引理7.2.1.** 设$\{X_n; n \geq 0\}$是鞅(下鞅), T是一个停时, 那么对任意$0 \leq k \leq n$,

$$E(X_n 1_{\{T=k\}}) = (\geq) E(X_k 1_{\{T=k\}}).$$

证明 由条件数学期望性质以及(7.1.1)可得

$$\begin{aligned} E(X_n 1_{\{T=k\}}) &= E(E(X_n 1_{\{T=k\}} | Y_0, \cdots, Y_k)) \\ &= E(1_{\{T=k\}} E(X_n | Y_0, \cdots, Y_k)) = (\geq) E(X_k 1_{\{T=k\}}). \end{aligned}$$ □

▼**引理7.2.2.** 设$\{X_n; n \geq 0\}$是鞅(下鞅). 若T为停时且存在正整数$n < m$使得$n \leq T \leq m$, 那么

$$E(X_T | Y_0, \cdots, Y_n) = (\geq) X_n.$$

证明 注意到X_n是$Y_0, Y_1, \cdots, Y_n$的函数, 由条件数学期望的平滑性,

$$\begin{aligned} E(X_T | Y_0, \cdots, Y_n) &= E(E(X_T | Y_0, \cdots, Y_{m-1}) | Y_0, \cdots, Y_n) \\ &= E(\sum_{k=n}^{m} E(X_k 1_{\{T=k\}} | Y_0, \cdots, Y_{m-1}) | Y_0, \cdots, Y_n) \\ &= E(\sum_{k=n}^{m-1} X_k 1_{\{T=k\}} | Y_0, \cdots, Y_n) \\ &\quad + E(E(X_m 1_{\{T=m\}} | Y_0, \cdots, Y_{m-1}) | Y_0, \cdots, Y_n). \end{aligned}$$

注意到$1_{\{T=m\}} = 1 - 1_{\{T \leq m-1\}}$ 是$Y_0, \cdots, Y_{m-1}$的函数. 由条件分布以及鞅(下鞅)性质可知

$$E(E(X_m 1_{\{T=m\}} | Y_0, \cdots, Y_{m-1}) | Y_0, \cdots, Y_n)$$

$$= E(1_{\{T=m\}}E(X_m|Y_0,\cdots,Y_{m-1})|Y_0,\cdots,Y_n)$$
$$= (\geq)E(X_{m-1}1_{\{T=m\}}|Y_0,\cdots,Y_n).$$

因此

$$\begin{aligned}E(X_T|Y_0,\cdots,Y_n) &= (\geq)E(\sum_{k=n}^{m-1}X_k1_{\{T=k\}}|Y_0,\cdots,Y_n)\\ &\qquad +E(X_{m-1}1_{\{T=m\}}|Y_0,\cdots,Y_n)\\ &= E(X_{T\wedge(m-1)}|Y_0,\cdots,Y_n).\end{aligned}$$

重复上面的分析(将$T\wedge(m-1)$看作新的T并以此类推)我们可得

$$E(X_T|Y_0,\cdots,Y_n) = (\geq)E(X_{T\wedge n}|Y_0,\cdots,Y_n) = X_n.$$ □

★**定理7.2.3.** (有界停时定理) 设$\{X_n;n\geq 0\}$是鞅(下鞅), 若S,T为两有界停时且$S\leq T$ *a.s.* 那么$E(X_T)=(\geq)E(X_S)$.

证明 不妨设$S,T\leq m$. 注意到$S\leq T$, 由引理7.2.2,

$$\begin{aligned}E(X_T) &= \sum_{n=0}^{m}E(X_T1_{\{S=n\}}) = \sum_{n=0}^{m}E(X_{T\vee n}1_{\{S=n\}})\\ &= \sum_{n=0}^{m}E(E(X_{T\vee n}1_{\{S=n\}}|Y_0,\cdots,Y_n))\\ &= \sum_{n=0}^{m}E(1_{\{S=n\}}E(X_{T\vee n}|Y_0,\cdots,Y_n))\\ &= (\geq)\sum_{n=0}^{m}E(1_{\{S=n\}}X_n) = E(X_S).\end{aligned}$$

定理得证. □

7.2.2 一般停时定理

对一般停时, 使(7.2.1)成立的条件要复杂些, 下面我们介绍一些具体条件.

▼**引理7.2.4.** 设随机变量W满足$E(|W|)<\infty$, 若$P(T<\infty)=1$, 那么

$$\lim_{n\to\infty}E(W1_{\{T>n\}})=0 \quad 且 \quad \lim_{n\to\infty}E(W1_{\{T\leq n\}})=E(W).$$

证明 由正项级数的收敛性, 当$n\to\infty$时,

$$E(|W|)\geq E(|W|1_{\{T\leq n\}}) = \sum_{k=0}^{n}E(|W|1_{\{T=k\}}) \to \sum_{k=0}^{\infty}E(|W|1_{\{T=k\}}).$$

由于$P(T<\infty)=1$, 由引理1.4.22可知

$$E(|W|1_{\{T\geq n+1\}}) = \sum_{k=n+1}^{\infty}E(|W|1_{\{T=k\}}) \to 0.$$

从而

$$0 \le |E(W) - E(W1_{\{T\le n\}})| \le |E(W1_{\{T\ge n+1\}})| \le E(|W|1_{\{T\ge n+1\}}) \to 0.$$

由此可得引理结论. □

★定理7.2.5. 设$\{X_n; n\ge 0\}$为鞅, 停时$T<+\infty$ *a.s.* 若

$$E(\sup_{n\ge 0}|X_{T\wedge n}|)<\infty,$$

那么$E(X_T)=E(X_0)$.

证明 令$W=\sup\limits_{n\ge 0}|X_{T\wedge n}|$. 由于$P(T<\infty)=1$,

$$X_T=\sum_{k=0}^{\infty}X_k1_{\{T=k\}}=\sum_{k=0}^{\infty}X_{T\wedge k}1_{\{T=k\}}.$$

从而$|X_T|\le W$. 因此$E(|X_T|)\le E(W)<\infty$. 由引理7.2.4可知

$$|E(X_{T\wedge n})-E(X_T)|\le E(|X_{T\wedge n}-X_T|1_{\{T\ge n\}})\le 2E(W1_{\{T\ge n\}})\to 0.$$

因为对任意$n\ge 1$, $T\wedge n$总是有界停时. 由有界停时定理可得

$$E(X_T)=\lim_{n\to\infty}E(X_{T\wedge n})=E(X_0).$$ □

♦推论7.2.6. 设$\{X_n; n\ge 0\}$为鞅. 停时$T<+\infty$ *a.s.* 且期望有限. 若存在非负有限常数K使得

$$E(|X_{n+1}-X_n||Y_0,\cdots,Y_n)\le K,$$

那么$E(X_T)=E(X_0)$.

证明 令$Z_0=|X_0|$, $Z_n=|X_n-X_{n-1}|$, $W=Z_0+\cdots+Z_T$. 则$W\ge|X_T|$, 而且

$$E(W)=\sum_{n=0}^{\infty}\sum_{k=0}^{n}E(Z_k1_{\{T=n\}})=\sum_{k=0}^{\infty}E(Z_k1_{\{T\ge k\}}).$$

注意到$1_{\{T\ge k\}}=1-1_{\{T\le k-1\}}$为$Y_0,\cdots,Y_{k-1}$的函数, 由条件假设可知

$$\begin{aligned}E(Z_k1_{\{T\ge k\}})&=E(1_{\{T\ge k\}}E(Z_k|Y_0,\cdots,Y_{k-1}))\\&=E\big(1_{\{T\ge k\}}E\big(|X_k-X_{k-1}|\big|Y_0,\cdots,Y_{k-1}\big)\big)\le KP(T\ge k),\end{aligned}$$

从而

$$E(W)\le\sum_{k=0}^{\infty}KP(T\ge k)\le K(1+E(T))<\infty.$$

因为$|X_{T\wedge n}|\le W$对一切n成立, 由定理7.2.5 可知结论成立. □

★定理7.2.7. 设$\{X_n; n\ge 0\}$为鞅. 停时$T<+\infty$ *a.s.* 若$E(|X_T|)<\infty$且

$$\lim_{n\to\infty}E(X_n1_{\{T\ge n\}})=0,$$

那么$E(X_T)=E(X_0)$.

证明 注意到

$$\begin{aligned}E(X_T) &= E(X_T 1_{\{T<n\}}) + E(X_T 1_{\{T\geq n\}}) = E(X_{T\wedge n} 1_{\{T<n\}}) + E(X_T 1_{\{T\geq n\}})\\ &= E(X_{T\wedge n}) - E(X_{T\wedge n} 1_{\{T\geq n\}}) + E(X_T 1_{\{T\geq n\}}).\end{aligned}$$

由条件 $\lim\limits_{n\to\infty} E(X_{T\wedge n} 1_{\{T\geq n\}}) = 0$及引理7.2.4可知 $\lim\limits_{n\to\infty} E(X_T 1_{\{T\geq n\}}) = 0$. 因此

$$E(X_T) = \lim_{n\to\infty} E(X_{T\wedge n}) = E(X_0),$$

其中最后等号由有界停时定理可得. □

★**定理7.2.8.** 设$\{X_n; n\geq 0\}$为下鞅, 停时$T<\infty$, *a.s.* 若存在随机变量W使得$E(|W|)<\infty$且对任意$n\geq 0$, $X_{T\wedge n}\leq W$成立, 那么$E(X_0)\leq E(X_T)$.

证明 任意取定一个正整数N, 令$X_n^N = X_n \vee (-N)$. 那么$\{X_n^N; n\geq 0\}$ 仍是下鞅. 由下鞅有界停时定理可得$E(X_0^N)\leq E(X_{T\wedge n}^N)$. 由条件假设可知

$$|X_{T\wedge n}^N| \leq N\vee |W| < |W| + N \quad 以及\quad X_T^N = \lim_{n\to\infty} X_{T\wedge n}^N, \quad a.s.$$

因此$|X_T^N|\leq N+|W|$. 当$n\to\infty$时, 由引理7.2.4可得,

$$E|X_{T\wedge n}^N - X_T^N| \leq E(|X_{T\wedge n}^N - X_T^N| 1_{\{T>n\}}) \leq 2E((N+|W|)1_{\{T>n\}}) \to 0.$$

所以 $\lim\limits_{n\to\infty} E(X_{T\wedge n}^N) = E(X_T^N)$. 由此可知

$$E(X_0) \leq E(X_0^N) \leq E(X_T^N).$$

因为$E(X_T^N)$随着N的增加而单调减少, $\lim\limits_{N\to\infty} E(X_T^N)$存在. 由极限的保序性,

$$E(X_0) \leq \lim_{N\to\infty} E(X_T^N). \tag{7.2.2}$$

由$X_T = \lim\limits_{n\to\infty} X_{T\wedge n}$及条件假设$X_{T\wedge n}\leq W$ 可知: $X_T\leq W$以及

$$E(X_T 1_{\{X_T\geq 0\}}) \leq E(|W|) < \infty.$$

进而由全期望公式,

$$\begin{aligned}E(X_T) &= E(X_T 1_{\{X_T\geq 0\}}) + E(X_T 1_{\{X_T<0\}})\\ &= E(X_T 1_{\{X_T\geq 0\}}) + \sum_{k=1}^{+\infty} E(X_T 1_{\{-k<X_T\leq -k+1\}})\\ &= E(X_T 1_{\{X_T\geq 0\}}) + \lim_{N\to\infty}\sum_{k=1}^{N} E(X_T 1_{\{-k<X_T\leq -k+1\}})\\ &\geq E(X_T 1_{\{X_T\geq 0\}}) + \lim_{N\to\infty} E(X_T^N \wedge 0) = \lim_{N\to\infty} E(X_T^N).\end{aligned} \tag{7.2.3}$$

综合(7.2.2)与(7.2.3)两式可得$E(X_0)\leq E(X_T)$. □

♦**推论7.2.9.** 若$\{X_n\}$为非正下鞅, 停时$T<\infty$ *a.s.* 那么$E(X_0)\leq E(X_T)$.

7.2.3 停时定理应用举例

下面我们看几个应用停时定理的例子.

▶**例7.2.10.** 设$a < b$ 为正整数, S_n是从位置a出发的简单(p, q)随机游动, 定义

$$\tau_0 = \inf\{n > 0; S_n = 0\},\ \tau_b = \inf\{n > 0; S_n = b\} \text{ 以及 } T = \tau_0 \wedge \tau_b.$$

求$E(T)$, $E(S_T)$以及$P(\tau_b < \tau_0)$.

解 以$X = \{X_i; i \geq 0\}$表示随机游动中每步的步长, 那么X_i独立同分布且

$$P(X_i = 1) = p, \quad P(X_i = -1) = q.$$

显然T是关于$\{S_n; n \geq 0\}$的停时. 对任意$k > b$, 存在$m \geq 1$以及$r = 1, \cdots, b$, 使得$k = mb + r$. 将k步按每连续b 步组合成$m + 1$个区段

$$\{1, 2, \cdots, b\}, \cdots, \{(m-1)b + 1, \cdots, mb\}, \{mb + 1, \cdots, mb + r\}.$$

注意到事件$\{T \geq k\}$发生时, 在k 之前任意相邻的b 步中游动X_i符号不能完全相同——否则就提前到访了0 或b. 因此由X_i的独立同分布性可知

$$P(T \geq k) \leq (1 - p^b - q^b)^m,$$

从而

$$P(T < \infty) = \lim_{n\to\infty} P(T < n) = 1 - \lim_{n\to\infty} P(T \geq k) = 1,$$

而且

$$\begin{aligned} E(T) &= \sum_{k=1}^{\infty} P(T \geq k) = \sum_{m=0}^{\infty}\sum_{r=1}^{b} P(T \geq mb + r) \\ &\leq b(1 + \sum_{m=1}^{\infty}(1 - p^b - q^b)^m) = \frac{b}{p^b + q^b} < \infty. \end{aligned}$$

由$T < \infty$ a.s. 可知, $S_T = 0$或b, a.s. 进而

$$P(\tau_b < \tau_0) = P(S_T = b) \quad \text{以及} \quad E(S_T) = bP(S_T = b).$$

下面我们分两种情况讨论.

(A) $p \neq q$时. 令$Y_n = S_n - n(p - q)$, 显然$\{Y_n; n \geq 0\}$关于$\{S_n; n \geq 0\}$为鞅. 由

$$E(|Y_{n+1} - Y_n||S_0, \cdots, S_n) = E(|X_{n+1} - (p - q)||S_0, \cdots, S_n) \leq 1 + |p - q|,$$

以及推论7.2.6可得$E(Y_T) = E(Y_0)$, 即

$$E(S_T) = a + (p - q)E(T). \tag{7.2.4}$$

此外, 令

$$M_n = \left(\frac{q}{p}\right)^{S_n}.$$

则$\{M_n;n\geq 0\}$也是鞅, 而且对任意$n\geq 0$, $M_{T\wedge n}\leq (1\vee \frac{q}{p})^b$, 由定理7.2.5可知,

$$E(M_T)=E(M_0)=\left(\frac{q}{p}\right)^a.$$

即

$$P(S_T=0)+\left(\frac{q}{p}\right)^b P(S_T=b)=\left(\frac{q}{p}\right)^a. \tag{7.2.5}$$

将$P(S_T=0)=1-P(S_T=b)$代入(7.2.5)解得

$$P(S_T=b)=\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}=P(\tau_b<\tau_0).$$

从而

$$E(S_T)=bP(S_T=b)=b\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}.$$

将其代入(7.2.4)得

$$E(T)=\frac{b}{p-q}\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}-\frac{a}{p-q}.$$

(B) $p=q=1/2$时. 此时令

$$M_n=S_n^2-n.$$

容易验证$\{M_n;n\geq 0\}$为鞅, 而且$E(|M_T|)<b^2+E(T)<\infty$. 此外,

$$\lim_{n\to\infty}E(|M_n|1_{\{T\geq n\}})\leq \lim_{n\to\infty}E((b^2+T)1_{\{T\geq n\}})=0.$$

由定理7.2.7可得, $E(M_T)=E(M_0)=a^2$. 即

$$P(S_T=b)b^2-E(T)=a^2.$$

另一方面, 可以验证(7.2.4)在$q=p$时仍成立, 从而

$$E(S_T)=bP(S_T=b)=a.$$

联立上面两式解得$P(S_T=b)=a/b=P(\tau_b<\tau_0)$以及$E(T)=a(b-a)$. □

▶**例7.2.11.** *已知$S_0=x>0$, $S_n=x+\sum\limits_{k=1}^n X_k$, $n\geq 1$, 其中独立同分布的随机变量$X_i,i\geq 1$, 服从正态分布$N(\mu,\sigma^2)$. 令$A=\{$存在$n>0$使得$S_n\leq 0\}$, 估计$P(A)$的上界.*

解 令$\tau=\inf\{n>0;S_n\leq 0\}$. τ为停时且$A=\{\tau<\infty\}$. 由$X_i\sim N(\mu,\sigma^2)$可知

$$E(\mathrm{e}^{-\lambda X_i})=\mathrm{e}^{-\lambda\mu+\lambda^2\sigma^2/2}.$$

取$\theta=2\mu/\sigma^2$, 由例7.1.5(2)可知$\{\mathrm{e}^{-\theta S_n};n\geq 0\}$为鞅. 进而由有界停时定理可知

$$\mathrm{e}^{-\theta x}=E(\mathrm{e}^{-\theta S_0})=E(\mathrm{e}^{-\theta S_{n\wedge\tau}})\geq E(\mathrm{e}^{-\theta S_\tau}1_{\{\tau\leq n\}})\geq P(\tau\leq n),$$

对任意$n \geq 1$成立. 令$n \to \infty$得$P(A) \leq \mathrm{e}^{-\theta x} = \mathrm{e}^{-2\mu x/\sigma^2}$. □

▶**例7.2.12.** 以Z_n表示第n年开始时水库蓄水量, Y_{n+1}表示第n年的净入库流量(Y_{n+1}为负表示净出库流量), b表示水库的最大库容. 显然$Z_{n+1} = (Z_n + Y_{n+1})^+ \wedge b$. 以$a$表示水库最低警戒库容, 令$T = \min\{n \geq 0; Z_n \leq a\}$, 试求$E(T)$的下界估计, 其中我们假定$Z_0 = z \in (a, b)$以及

$$E(Y_{n+1}|Z_0, Y_1, \cdots, Y_n) \leq m, \quad E(\exp\{-\lambda Y_{n+1}\}|Z_0, Y_1, \cdots, Y_n) \leq 1.$$

解 由$P(T < \infty) < 1$可知$E(T) = \infty$. 为估计$E(T)$下界, 不妨设$P(T < \infty) = 1$. 对任意$z \geq 0$, 令

$$g(z) = \frac{1}{m}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1 - \mathrm{e}^{-\lambda(z-a)}}{\lambda} - (z-a)\Big),$$

以及$f(z) = \begin{cases} g(z) \vee 0, & 0 \leq z \leq b, \\ g(b); & z > b. \end{cases}$ 易知当$z \geq 0$时$f(z) \geq g(z)$, 而且对任意$n \geq 0$, $f(Z_{n+1}) = f(Z_n + Y_{n+1})$. 进一步地, 由假设条件可知, 对任意$n \geq 0$, 当$a \leq Z_n \leq b$时,

$$\begin{aligned} E(f(Z_{n+1})|Z_0, Y_1, \cdots, Y_n) &= E(f(Z_n + Y_{n+1})|Z_0, Y_1, \cdots, Y_n) \\ &\geq E(g(Z_n + Y_{n+1})|Z_0, Y_1, \cdots, Y_n) \\ &= \frac{1}{m}E\Big(\mathrm{e}^{\lambda(b-a)}\frac{1 - \mathrm{e}^{-\lambda(Z_n + Y_{n+1} - a)}}{\lambda} - (Z_n - a + Y_{n+1})\Big|Z_0, Y_1, \cdots, Y_n\Big) \\ &\geq \frac{1}{m}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1 - \mathrm{e}^{-\lambda(Z_n - a)}}{\lambda} - (Z_n - a)\Big) - 1 \geq f(Z_n) - 1, \end{aligned}$$

而且当$0 \leq Z_n < a$时, 由$f(Z_n) = 0, f(Z_{n+1}) \geq 0$可知上式仍然成立. 因此

$$\{f(Z_n) + n; n \geq 0\}$$

为下鞅. 在给定$Z_0 = z \in (a, b)$条件下, 由下鞅有界停时定理可得

$$f(z) = E(f(Z_0)) \leq E(f(Z_{T\wedge n}) + T \wedge n) = E(f(Z_{T\wedge n})) + E(T \wedge n). \tag{7.2.6}$$

由于f为有界连续函数, 存在常数$M > 0$使得$n \to \infty$时,

$$|E(f(Z_{T\wedge n})) - E(f(Z_T))| \leq E(|f(Z_{T\wedge n}) - f(Z_T)|) \leq MP(T > n) \to 0.$$

因此

$$\lim_{n\to\infty} E(f(Z_{T\wedge n})) = E(f(Z_T)) = 0.$$

此外由单调收敛定理(定理1.4.23), 我们还有

$$\lim_{n\to\infty} E(T \wedge n)) = E(T).$$

因此在(7.2.6)两边令$n \to \infty$可得$E(T) \geq f(z)$, 即

$$E(T) \geq \frac{1}{m}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1 - \mathrm{e}^{-\lambda(z-a)}}{\lambda} - (z-a)\Big), \quad z \in (a, b).$$ □

练习题7.2

1 试构造一个鞅$\{X_n; n \geq 0\}$和停时T使得$E(X_T) \neq E(X_0)$.

2 证明对任意鞅或非负下鞅$\{X_n; n \geq 0\}$及有限停时τ, $E(|X_\tau|) \leq \lim\limits_{n\to\infty} E(|X_n|)$.

3 以$S = \{S_n; n \geq 0\}$表示初值为0的简单(q,p)-随机游动, 且$q < p$. 对任意$k \in \mathbb{N}$, 令$\tau_k = \inf\{n \geq 1, S_n = k\}$. 利用例7.2.10的结果, (1) 当$k < 0$时, 求$P(\tau_k < \infty)$; (2) 当$k > 0$时, 求$E(\tau_k)$.

4 (接第二章第二节习题5) 观察每次治疗效果, 当一种药物的累积治愈人数超过另一种药物的累积治愈人数并达到预先给定的一个标准时, 试验就停止并判定该药比另一种药物疗效好. 试讨论这种试验判别的失误率.

7.3 鞅的不等式与收敛定理

作为本章以及本书的最后一节, 我们简要地介绍鞅的几个不等式以及一个关于鞅收敛的基本定理.

7.3.1 极大不等式与Doob不等式

★**定理7.3.1.** 设$\{X_n; n \geq 0\}$为上鞅, 则对任意$\lambda > 0$有

$$\lambda P(\max_{0\leq k\leq n} X_k \geq \lambda) \leq E(X_0) - E(X_n 1_{\{\max_{0\leq k\leq n} X_k < \lambda\}}); \tag{7.3.1}$$

$$\lambda P(\min_{0\leq k\leq n} X_k \leq -\lambda) \leq -E(X_n 1_{\{\min_{0\leq k\leq n} X_k \leq -\lambda\}}); \tag{7.3.2}$$

$$\lambda P(\max_{0\leq k\leq n} |X_k| \geq \lambda) \leq E(X_0) + 2E(X_n^-), \quad \text{其中} \ \ X_n^- = (-X_n) \vee 0. \tag{7.3.3}$$

证明 (1) 令$T = \min\{k; X_k \geq \lambda\} \wedge n$, 则$T$为有界停时. 由有界停时定理可得

$$\begin{aligned} E(X_0) \geq E(X_T) &= E(X_T 1_{\{\max_{0\leq k\leq n} X_k < \lambda\}}) + E(X_T 1_{\{\max_{0\leq k\leq n} X_k \geq \lambda\}}) \\ &\geq \lambda P(\max_{0\leq k\leq n} X_k \geq \lambda) + E(X_n 1_{\{\max_{0\leq k\leq n} X_k < \lambda\}}). \end{aligned}$$

移项后即得(7.3.1).

(2)令$S = \min\{k; X_k \leq -\lambda\} \wedge n$, 则$S$为有界停时, 且$S \leq n$. 由有界停时定理可得

$$\begin{aligned} E(X_n) \leq E(X_S) &= E(X_S 1_{\{\min_{0\leq k\leq n} X_k \leq -\lambda\}}) + E(X_S 1_{\{\min_{0\leq k\leq n} X_k > -\lambda\}}) \\ &\leq -\lambda P(\min_{0\leq k\leq n} X_k \leq -\lambda) + E(X_n 1_{\{\max_{0\leq k\leq n} X_k > -\lambda\}}). \end{aligned}$$

移项后即得(7.3.2).

(3) 由(7.3.1),(7.3.2)两项直接相加可得(7.3.3). □

♠**注记7.3.2.** 若$\{X_n; n \geq 0\}$为鞅或非负下鞅, 那么$\{|X_n|; n \geq 0\}$仍是下鞅, 进而$\{-|X_n|; n \geq 0\}$为上鞅. 此时由(*7.3.2*)可知

$$P(\max_{0\leq k\leq n} |X_k| \geq \lambda) \leq \frac{1}{\lambda} E(|X_n| 1_{\{\max_{0\leq k\leq n} |X_k| \geq \lambda\}}).$$

▼**引理7.3.3.** 若随机变量ζ, η满足$E(|\zeta|^p) < \infty, E(|\eta|^q) < \infty$, 其中$1/p + 1/q = 1$. 那么$E|\zeta\eta| < \infty$而且

$$E(|\zeta\eta|) \leq [E(|\zeta|^p)]^{1/p} [E(|\eta^q|)]^{1/q}. \tag{7.3.4}$$

证明 由$\ln x$为凹函数可知, 对任意$x, y \in (0, \infty)$, $\lambda \in [0, 1]$,

$$\ln(\lambda x + (1-\lambda)y) \geq \lambda \ln x + (1-\lambda) \ln y.$$

从而对任意$x, y \in (0, \infty)$,
$$x^{1/p}y^{1/q} \le \frac{x}{p} + \frac{y}{q}.$$
显然上式对$x, y \in [0, \infty)$也成立.

当$E(|\zeta|^p)E(|\eta|^q) = 0$时, ζ和η中至少有一个几乎处处为零. 此时
$$E(|\zeta\eta|) = 0 = [E(|\zeta|^p)]^{1/p}[E(|\eta^q|)]^{1/q}.$$
当$E(|\zeta|^p)E(|\eta|^q) \ne 0$时, 令$x = |\zeta|^p/E(|\zeta|^p)$, $y = |\eta|^q/E(|\eta|^q)$, 那么
$$\frac{|\zeta\eta|}{E(|\zeta|^p)^{1/p}(E(|\eta|^q))^{1/q}} \le \frac{1}{p}\frac{|\zeta|^p}{E(|\zeta|^p)} + \frac{1}{q}\frac{|\eta|^q}{E(|\eta|^q)}.$$
两边求期望得
$$E\Big(\frac{|\zeta\eta|}{(E(|\zeta|^p))^{1/p}(E(|\eta|^q))^{1/q}}\Big) \le \frac{1}{p}E\Big(\frac{|\zeta|^p}{E(|\zeta|^p)}\Big) + \frac{1}{q}E\Big(\frac{|\eta|^q}{E(|\eta|^q}\Big) = 1.$$
移项得所需结论. □

一般地, 我们称不等式(7.3.4)为Hölder不等式.

★定理7.3.4. 设$\{X_n; n \ge 0\}$为鞅或非负下鞅, $X_n^* = \max_{0\le k\le n} |X_k|$. 那么

(1) 对任意$\lambda > 0$及$r \ge 1$, $P(X_n^* \ge \lambda) \le \lambda^{-r}E(|X_n|^r)$. (极大不等式)

(2) 对任意$r > 1$, $E(|X_n^*|^r) \le (\frac{r}{r-1})^r E(|X_n|^r)$. (*Doob* 不等式)

证明 若$E(|X_n|^r) = \infty$, 则结论显然成立. 下设$E(|X_n|^r) < \infty$.

首先注意到, 不论$\{X_n; n \ge 0\}$本身是鞅还是非负下鞅, 由命题7.1.10可知$\{|X_n|^r, n \ge 0\}$ 为非负下鞅. 令$T = \min\{k, |X_k| \ge \lambda\}$以及$\tilde{T} = T \wedge n$. 显然停时$\tilde{T} \le n$. 由有界停时定理可得
$$E(|X_n|^r) \ge E(|X_{\tilde{T}}|^r) \ge \lambda^r P(|X_{\tilde{T}}| \ge \lambda) = \lambda^r P(T \le n) = \lambda^r P(X_n^* \ge \lambda).$$
由此立即可得
$$P(X_n^* \ge \lambda) \le \lambda^{-r}E(|X_n|^r).$$
进一步地, 若$r > 1$,
$$\begin{aligned}E(|X_n^*|^r) &= \int_0^\infty r\lambda^{r-1}P(|X_n^*| \ge \lambda)\mathrm{d}\lambda \le \int_0^\infty r\lambda^{r-2}E(|X_n|1_{\{|X_n^*|\ge\lambda\}})\mathrm{d}\lambda \\ &= E\Big(|X_n|\int_0^{X_n^*} r\lambda^{r-2}\mathrm{d}\lambda\Big) = \frac{r}{r-1}E(|X_n||X_n^*|^{r-1}),\end{aligned}$$
其中不等号源自注记7.3.2. 由Hölder不等式得
$$E(|X_n^*|^r) \le \frac{r}{r-1}(E(|X_n|^r))^{1/r}(E(|X_n^*|^r))^{1-1/r}.$$
由此可知Doob不等式成立. □

►例7.3.5. 设$X_n, n \ge 1$, 为独立同分布的随机变量且
$$E(X_1) = 0, \quad Var(X_1) = \sigma^2.$$

则$S_n=\sum_{k=1}^n X_k$为鞅，由极大不等式可得柯尔莫哥洛夫不等式:

$$P(\max_{0\le k\le n} S_k\ge\lambda)\le\frac{1}{\lambda^2}E(S_n^2)=\frac{n\sigma^2}{\lambda^2},\qquad \lambda>0.$$

此外，由Doob不等式可得

$$E(|\max_{0\le k\le n} S_k|^2)\le 4E(S_n^2)=4n\sigma^2.\qquad\square$$

7.3.2 Doob收敛定理

对任意随机序列$X=\{X_n;n\ge 0\}$以及闭区间$[a,b]$，令

$$T_0=\inf\{n\ge 0;X_n\le a\},\qquad T_1=\inf\{n>T_0;X_n\ge b\};$$

$$T_{2j}=\inf\{n>T_{2j-1};X_n\le a\},\quad T_{2j+1}=\inf\{n>T_{2j};X_n\ge b\},\ j=1,2,\cdots.$$

显然$\{T_k;k\ge 1\}$为一列关于X的单调不降的停时.

♣**定义7.3.6.** 称$k=\sup\{j;T_{2j-1}\le n\}$ 为序列$\{X_0,\cdots,X_n\}$ 在时刻n 之前上穿区间$[a,b]$ 的次数，记作$U_a^b[X,n]$.

直观地看，$U_a^b[X,n]$就是在n之前，按时间先后顺序，X的取值从区间$[a,b]$的下方升到上方这种现象发生的次数. 按定义,

$$\{U_a^b[X,n]=j\}=\{T_{2j-1}\le n<T_{2j+1}\}.$$

★**定理7.3.7.** 设$\{X_n;n\ge 0\}$为上鞅，则对任意$n\ge 1$,

$$E(U_a^b[X,n])\le\frac{1}{b-a}E[(X_n-a)^-].$$

证明 对任意$k\ge 0$，由有界停时定理可得

$$\begin{aligned}0&\ge E(X_{T_{2k+1}\wedge n}-X_{T_{2k}\wedge n})\\&=E\Big((X_{T_{2k+1}\wedge n}-X_{T_{2k}\wedge n})(1_{\{T_{2k}\le n<T_{2k+1}\}}+1_{\{n\ge T_{2k+1}\}})\Big)\\&\ge E((X_n-a)1_{\{T_{2k}\le n<T_{2k+1}\}})+(b-a)P(T_{2k+1}\le n)\\&\ge E((X_n-a)1_{\{T_{2k}\le n<T_{2k+1}\}})+(b-a)P(U_a^b(x,n)\ge k+1).\end{aligned}$$

从而

$$\begin{aligned}P(U_a^b(X,n)\ge k+1)&\le\frac{1}{b-a}E(-(X_n-a)1_{\{T_{2k}\le n<T_{2k+1}\}})\\&\le\frac{1}{b-a}E((X_n-a)^-1_{\{T_{2k}\le n<T_{2k+1}\}}).\end{aligned}$$

两边关于k求和得

$$E(U_a^b(X,n))\le\frac{1}{b-a}E((X_n-a)^-1_{\{T_0\le n\}})\le\frac{1}{b-a}E((X_n-a)^-).\qquad\square$$

★定理7.3.8. (*Doob*收敛定理) 设$\{X_n;n\geq 0\}$为(上, 下)鞅, 若

$$M=\sup_n E(|X_n|)<\infty.$$

则存在随机变量X_∞使得当$n\to\infty$时$X_n\to X_\infty$ *a.s.* 而且

$$E(|X_\infty|)\leq\liminf_{n\to\infty}E(|X_n|).$$

证明 由于下鞅的负过程是上鞅, 我们只讨论上鞅的情形. 令$\mathbb{Q}$表示有理数全体, 设$a,b\in\mathbb{Q}$, $a<b$. 令$U_a^b(X)$为$\{X_n;n\geq 0\}$上穿$[a,b]$的次数, 即

$$U_a^b(X)=\lim_{n\to\infty}U_a^b(X,n).$$

由上穿不等式可得

$$E(U_a^b(X))\leq\frac{1}{b-a}\sup_n E((X_n-a)^-)\leq\frac{1}{b-a}(a+\sup_n E|X_n|)<\infty.$$

于是$U_a^b(X)<\infty$, *a.s.* 令

$$W_{a,b}=\{\liminf_{n\to\infty}X_n<a,\ \ \limsup_{n\to\infty}X_n=b\},$$

$$W=\bigcup_{a,b\in\mathbb{Q},a<b}W_{a,b}.$$

由$W_{a,b}\subset\{U_a^b(X)=\infty\}$可知$P(W_{a,b})=0$, 从而$P(W)=0$. 显然对任意$\omega\notin W$,

$$\liminf_{n\to\infty}X_n(\omega)=\limsup_{n\to\infty}X_n(\omega),$$

从而$\lim\limits_{n\to\infty}X_n(\omega)$存在, 记其为$X_\infty(\omega)$. 若$\omega\in W$则补充定义$X_\infty(\omega)=0$. 因此

$$X_n\to X_\infty,\ \ a.s.$$

令$p=P(|X_\infty|=\infty)$. 由

$$\lim_{n\to\infty}P(|X_n|\geq c)=P(|X_\infty|\geq c)\geq p,$$

可知: 当$c=3M/p$时, 存在N, 对任意$n>N$,

$$P(|X_n|\geq c)\geq p/2.$$

若$p>0$, 则

$$E(|X_n|)\geq\frac{3M}{2}.$$

这与条件$M=\sup_n E(|X_n|)$矛盾, 因此$p=0$, 即$P(|X_\infty|<\infty)=1$.

另一方面, 对任意$c>0$, $n>0$, $\varepsilon>0$,

$$\begin{aligned}E(|X_\infty|1_{\{|X_\infty|\leq c\}})&=E(|X_\infty|1_{\{|X_\infty|\leq c,|X_n|\leq c\}})+E(|X_\infty|1_{\{|X_\infty|\leq c,|X_n|>c\}})\\&\leq E(|X_\infty|1_{\{|X_\infty|\leq c,|X_n|\leq c\}})+E(|X_n|1_{\{|X_\infty|\leq c,|X_n|>c\}})\\&\leq E(|X_\infty-X_n|1_{\{|X_\infty|\leq c,|X_n|\leq c\}})+E(|X_n|1_{\{|X_\infty|\leq c\}})\\&\leq\varepsilon P(|X_\infty|\leq c,|X_n|\leq c)+2cP(|X_n-X_\infty|>\varepsilon)+E(|X_n|)\\&\leq\varepsilon+2cP(|X_n-X_\infty|>\varepsilon)+E(|X_n|).\end{aligned}$$

先令$n\to\infty$再令$\varepsilon\to 0$得

$$E(|X_\infty|1_{\{|X_\infty|\le c\}})\le\liminf_{n\to\infty}E(|X_n|).$$

再令$c\to\infty$, 我们就得到定理后一个结论. □

▶**例7.3.9.** 设非负随机过程$X=\{X_n;\ n\ge 0\}$满足

$$E(X_{n+1}|X_0,\cdots,X_n)=2X_n+1,\qquad n\ge 0.$$

证明当$n\to\infty$时$X_n/2^n$几乎必然收敛.

证明 由条件容易验证, 对任意$n\ge 0$,

$$E\Big(\frac{X_{n+1}}{2^{n+1}}\Big|X_0,\cdots,X_n\Big)=\frac{X_n}{2^n}+\frac{1}{2^n}\ge\frac{X_n}{2^n}.$$

由此可知$\{X_n/2^n;n\ge 0\}$是下鞅, 而且

$$\begin{aligned}E\Big(\frac{X_{n+1}}{2^{n+1}}\Big)&=E\Big(E\Big(\frac{X_{n+1}}{2^{n+1}}\Big|X_0,\cdots,X_n\Big)\Big)=E\Big(\frac{X_n}{2^n}+\frac{1}{2^n}\Big)\\&=E(X_1/2)+\frac{1}{2}+\cdots+\frac{1}{2^n}\le E(X_1/2)+1<\infty.\end{aligned}$$

由定理7.3.8可知$n\to\infty$时$X_n/2^n$几乎必然收敛. □

▶**例7.3.10.** 设X是时齐的不可约常返马氏链, $\mathbf{P}=(p_{ij})_{i,j\in S}$为其转移概率矩阵. 若$y(i)$, $i\in S$, 为方程组

$$u(i)=\sum_{j\in S}p_{ij}u(j),\qquad i\in S,$$

的有界解, 试证明$y(i),i\in S$, 恒为常数.

证明 令$Y_n=y(X_n)$, 由马氏性可知, 对任意$n\ge 0$,

$$\begin{aligned}E(Y_{n+1}|X_0,\cdots,X_n)&=E(y(X_{n+1})|X_0,\cdots,X_n)=E(y(X_{n+1})|X_n)\\&=\sum_{j\in S}p_{X_n,j}y(j)=y(X_n)=Y_n.\end{aligned}$$

因此$\{Y_n;n\ge 0\}$为鞅且有界. 由定理7.3.8可知存在随机变量U使得

$$Y_n=y(X_n)\to U,\quad a.s.$$

注意到X_n 是不可约常返马氏链, 对任意$i,j\in S$,

$$P(X_n=i,\text{i.o.})=P(X_n=j,\text{i.o.})=1.$$

因此存在$\omega\in\Omega$使得$Y_n(\omega)\to U(\omega)$ 且存在子列n'_k, n''_k

$$X_{n'_k}(\omega)=i,\qquad X_{n''_k}(\omega)=j.$$

这意味着

$$y(i)=Y_{n'_k}(\omega)\to U(\omega),\ \text{且}\ y(j)=Y_{n''_k}(\omega)\to U(\omega),$$

即$y(i)=y(j)$. 因i,j的任意性可知, $y(i),i\in S$, 为常值函数. □

练习题7.3

1 设$\{X_n; n \geq 0\}$为鞅, $E(X_n) = 0$且$E(X_n^2) < \infty$, $n \geq 0$, 证明

$$P(\max_{0 \leq k \leq n} X_k \geq \lambda) \leq \frac{E(X_n^2)}{E(X_n^2) + \lambda^2}.$$

2 令$S_0 = 0$, $S_n = X_1 + \cdots + X_n$, $n \geq 1$, 其中随机变量X_n满足条件: 存在$\lambda > 0$, 使得对任意$n \geq 1$,

$$E(\exp(\lambda X_{n+1}) | X_1, \cdots, X_n) \leq 1.$$

证明

$$P(\max_{n \geq 0}(x + S_n) > l) \leq \mathrm{e}^{-\lambda(l-x)}, \quad x \leq l.$$

3 设$\{X_k; k \geq 1\}$为独立随机变量序列且存在$t > 0$使得对任意$k \geq 1$,

$$\psi_k(t) = E(\exp(tX_k)) < \infty.$$

若$n \to \infty$时,

$$\Psi_n(t) = \prod_{k=1}^{n} \psi_k(t) \to \Psi(t) \in (0, \infty),$$

证明$S_n = \sum\limits_{k=1}^{n} X_k$随$n \to \infty$几乎必然收敛.

参考文献

[1] William, Feller. An introduction to probability theory and its applications (II). New York: John Wiley & Sons, 1971.

[2] Ross, Sheldom M. Introduction to Probability Models, Tenth Edition. Singapore: Elsevier, 2010.

[3] Karlin, Samuel and Tailor, Howard M. A First Course in Stochastic Processes. New York: Academic Press, 1975.

[4] Karatzas, Ioannis and Shreve, Steven E. Brownian Motion and Stochastic Calculus. New York: Springer-Verlag, 1988.

[5] Durrett, Rick. Probability: Theory and Examples. New York: Springer-Verlag, 2010.

[6] Durrett, Rick. Essentials of Stochastic Processes. New York: Springer-Verlag, 2012.

[7] Athreya, K. B. and Ney, P.E. Branching processes. New York: Springer-Verlag, 1972.

[8] Marcus, M.B and Rosen, J. Markov processes, Gaussion processes and local times. Cambridge: Cambridge Univerisity Press. 2006

[9] 钱敏平, 龚光鲁. 随机过程论(第二版). 北京: 北京大学出版社, 1997.

[10] 方兆本, 廖柏其. 随机过程(第二版). 北京: 科学出版社, 2004.

[11] 何声武. 随机过程引论. 北京: 高等教育出版社, 1999.

[12] 严士健, 刘秀芳. 测度与概率. 北京: 北京师范大学出版社, 2003.