随 机 过 程 基 础

李育强　　　编

2018年 3 月

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对相关内容作些必要的补充.

§1.1　概率论简述

为了方便叙述, 我们先回顾一些初等概率论课程的知识.

(A) 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.称由Ω中子集构成的集类$\mathfrak{F}$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) $A_k \in \mathfrak{F}, k = 1, 2, \cdots \Rightarrow \bigcup\limits_{k\geq 1} A_k \in \mathfrak{F}$.

称$\mathfrak{F}$中每个元素为随机事件, 它是样本空间的子集. 从应用角度看, $\mathfrak{F}$ 中包含的随机事件才是人们的兴趣所在. $\mathbb{P}$是$\mathfrak{F} \to [0,1]$上的函数, 用来度量$\mathfrak{F}$中随机事件发生的可能性. 我们称它为概率, 如果满足如下条件

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

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

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

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

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

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

$$\mathbb{P}(A|B) = \mathbb{P}(AB)/\mathbb{P}(B).$$

本质上, 随机事件下的条件概率可以看作是一个将样本空间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

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

(2) (乘法公式) 对任意两个随机事件A, B,

$$\mathbb{P}(AB) = \mathbb{P}(B)\mathbb{P}(A|B).$$

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

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

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

(B) 单个随机变量及相关刻画

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

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

通常简记随机变量为X, Y, Z. 称随机变量X, Y是几乎处处相等的(记作$X \overset{a.s.}{=} Y$), 如果$\mathbb{P}(X = Y) = 1$. 称函数

$$F(x) = \mathbb{P}(X \le x)$$

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

(B1) 离散型随机变量

若随机变量可能的取值是有限的或可用自然数的次序将所有取值一一排列出来(可列的), 那么该随机变量称为是离散的。设离散型随机变量X的可能取值为

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

我们称

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

为X的概率分布列(离散分布). 称

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

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

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

为X的方差.

下面我们列举几个简单的常用离散分布(变量).

(1) 两点分布：X只有两种取值$0, 1$, 分布列为(q, p)且$0 < p = 1 - q < 1$.

(2) 二项分布$b(n,p)$：$X\in\{0,1,\cdots,n\}$, 分布列$(p_0,p_1,\cdots,p_n)$满足

$$p_k=C_n^kp^kq^{n-k},\qquad k=0,1,\cdots,n,\quad p=1-q\in(0,1).$$

(3) 几何分布：$X\in\{1,2,\cdots,n,\cdots\}$, 分布列$(p_1,\cdots,p_n,\cdots)$满足

$$p_k=pq^{k-1},\qquad k\geq 0,\quad p=1-q\in(0,1).$$

(4) 泊松分布(强度为λ)：$X\in\{0,1,2,\cdots,n,\cdots,\}$, 分布列$(p_0,p_1,\cdots,p_n,\cdots)$满足

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

(B2) 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$f(t)$为是X的分布密度函数.

对任意$k\geq 1$, 称

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

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

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

为X的方差.

一些典型连续型分布(变量)如下.

(1) 均匀分布：X服从(a,b)上均匀分布, 若分布密度函数

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

(2) 指数分布：X服从$(0,\infty)$上参数为$\lambda>0$的指数分布, 若密度函数为

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

(3) Γ-分布：X服从$(0,\infty)$上形状参数$a>0$,尺度参数$\lambda>0$的$\Gamma(a,\lambda)$分布, 若密度函数

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

(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, 若密度函数

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

(C) 有限多个随机变量

同时考虑样本空间Ω(概率空间$(\Omega,\mathfrak{F},\mathbb{P})$)上随机变量$X_1,\cdots,X_n$. 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).

(C1) 联合分布、协方差

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

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

称该多元函数f为$(X_1,\cdots,X_n)$的联合密度函数. 此时也称单个随机变量X_i的分布函数

$$F_{X_i}(x)=\mathbb{P}(X_i\le x)$$

为该联合分布的边际(边缘)分布; 称X_i的分布密度函数$f_{X_i}(x)$(若存在)为边际(边缘)密度函数, 而且

$$\begin{aligned}F_i(x)&=\int_{-\infty}^{x}\mathrm{d}t_i\int_{-\infty}^{\infty}\mathrm{d}t_1\cdots\int_{-\infty}^{\infty}f(t_1,\cdots,t_i,\cdots,t_n)\mathrm{d}t_n,\\ f_i(x)&=\int_{-\infty}^{\infty}\mathrm{d}t_1\cdots\int_{-\infty}^{\infty}f(t_1,\cdots,x,\cdots t_n)\mathrm{d}t_n.\end{aligned}$$

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

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

那么X的联合分布列为

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

相应随机变量X_i的边际分布列

$$p_k^{(i)}=\mathbb{P}(X_i=x_{ik})=\sum_{r=1,r\ne i}^{n}\sum_{k_r=0}^{\infty}\mathbb{P}(X_1=x_{1k_1},\cdots,X_i=x_{ik},\cdots,X_n=x_{nk_n}).$$

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

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

任意两个随机变量之间的协方差为

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

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

n元正态分布的是一个重要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, 若它们的联合分布密度函数

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Big\{-\frac{1}{1-\rho^2}\Big(\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{x-\mu_1}{\sigma_1}\frac{y-\mu_2}{\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\Big)\Big\},$$

其中$\mu_1=\mathbb{E}(X_1),\mu_2=\mathbb{E}(X_2)$, $\sigma_1^2=\mathbf{Var}X_1,\sigma_2^2=\mathbf{Var}X_2$, $\rho=\mathrm{Cov}(X_1,X_2)/(\sigma_1\sigma_2)$为相关系数. 引进协方差矩阵

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

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

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

记$(\mathbf{x}-\mu)=(x-\mu_1,y-\mu_2)^T$, 那么利用矩阵表示二元正态分布密度函数可写成

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

一般地n元正态随机变量$(X_1,\cdots,X_n)$的密度函数定义为

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

其中D为正定矩阵, $\mathbf{x}-\mu=(x_1-\mu_1,\cdots,x_n-\mu_n)^T$, $\mu_1,\cdots,\mu_n$ 为常数. 此时对任意i, $\mathbb{E}(X_i)=\mu_i$, 而且D就是$X_1,\cdots,X_n$之间的协方差矩阵.

下面这个例子介绍了参数为$(p_1,p_2,\cdots,p_m)$的多元二项分布.

例1.1 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, $p_1+\cdots+p_m=1$. 连续独立地抛n次. 以X_i表示落入第i个罐子的可能硬币数. 求$(X_1,X_2,\cdots,X_m)$ 的联合分布.

解 以n_i表示第i个罐子落入的硬币数, 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

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

容易看出任意前k项($k<m$)的分布列为

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

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

(C2) 随机变量独立性

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

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

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

若$X_1,\cdots,X_n$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密度等于各自密度函数的乘积, 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1.2 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$1\leq j\leq k_i$, $1\leq i\leq n$,

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

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, 并且随机变量$Y_i:=f_i(\mathbf{X}_i)$, $i=1,\cdots,k$, 也相互独立.

(C3) 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$\mathcal{A}\in\mathbb{R}^n$的可逆函数, n元随机变量$Y:=(Y_1,\cdots,Y_n)=\Phi(X)$, 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

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

(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的联合密度函数.

例1.2 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. 若$(T_1,\cdots,T_k)$的联合密度函数为

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

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, Jacobi行列式$J_\Phi\equiv 1$且

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

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

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

由此可知W_1的分布密度为

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

类似可计算得W_i, $i\geq 2$, 的分布密度为

$$f_{W_i}(x_i)=\lambda\mathrm{e}^{-\lambda x_i}.$$

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. □

(C4) 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$X+Y$的分布函数$G(z)=\mathbb{P}(X+Y\le z)$可用卷积表示, $G=F_X*F_Y$, 即对任意$z\in\mathbb{R}$

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

特别, 若X,Y均是连续型随机变量, 那么$X+Y$也是连续型随机变量, 它的分布密度函数

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

若$X_1,X_2,\cdots,X_n$是独立同分布的随机变量, 共同分布为F, 那么对任意$2\le k\le n$, $X_1+X_2+\cdots+X_k$的分布函数

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

其中$F^{*(1)}=F$. 若F有密度函数f, 那么$X_1+X_2+\cdots+X_k$的分布密度函数

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

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

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

密度函数分别为

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

例1.3 (1)若$X_1,\ X_2,\cdots,X_n$是分别服从$\Gamma(a_1,\lambda),\ \Gamma(a_2,\lambda),\cdots,\Gamma(a_n,\lambda)$分布的独立随机变量, 那么

$$X_1+\cdots+X_n\sim\Gamma(a_1+\cdots+a_n,\lambda).$$

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

$$X_1+\cdots+X_n\sim N(\mu_1+\cdots+\mu_n,\sigma_1^2+\cdots+\sigma_n^2).$$

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$x>0$

$$\mathbb{P}(|X|\ge x)\le \frac{\mathbb{E}(|X|^p)}{x^p},\qquad p>0\quad (\text{切比雪夫不等式}).$$

(2) 若随机变量X,Y的二阶矩都存在, 那么

$$[\mathbb{E}(XY)]^2\le \mathbb{E}(X^2)\mathbb{E}(Y^2),\qquad (\text{Schwarz 不等式}).$$

练习题

1.1 设$X_1,X_2,\cdots,X_n$是服从几何分布的独立同分布随机变量,

$$\mathbb{P}(X_1=k)=q^{k-1}(1-q),\qquad k\ge 1.$$

求$X_1+X_2+\cdots+X_n$的分布列.

1.2 设X_1,X_2,X_3是服从$(0,1)$上均匀分布的独立随机变量,

(1) 分别用$X_{(1)},X_{(2)},X_{(3)}$表示其中最大, 其次和最小的随机变量. 求它们的联合密度函数.

(2) 求(X_1^2,X_2,X_2+X_3)的联合密度函数.

1.3 设二元随机变量$\mathbf{X}=(X_1,X_2)$, $\mathbf{Y}=(Y_1,Y_2)$独立. 证明X_1 与Y_1独立, 并举例说明X_1,X_2 可以不独立.

1.4 用Schwarz不等式证明若X的二阶矩存在, 那么$\mathbb{P}(|X|>0)\ge[\mathbb{E}(X)]^2/\mathbb{E}(X^2)$.

1.5 若对任意$k\ge 0$, $\{a_n(k),n\ge 1\}$为非负单调不降数列,且$\lim\limits_{n\to\infty}a_n(k)=a(k)$, 证明

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

1.6 利用习题1.5的结论证明

(1) 若离散随机变量X的取值为$0,1,2,\cdots$, 那么$\mathbb{E}(X)=\sum_{k=1}^{\infty}\mathbb{P}(X\ge k)$.

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

1.7* 利用习题1.5的结论证明(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$f(x,y)$是非负可积函数, 那么$\int_0^{\infty}\mathrm{d}x\int_0^{x}f(x,y)\mathrm{d}y=\int_0^{\infty}\mathrm{d}y\int_y^{\infty}f(x,y)\mathrm{d}x$.

1.8* 设非负随机变量X服从离散分布F, 称$S=\{x:\mathbb{P}(\{x\})>0\}$为分布$F$的支撑集. 若存在非负实数$d>0$使得$S\subset\{nd:n\ge 0\}$, 则称$X$服从格子点分布. 若存在互素的非负整数$n_k,k\ge 0$使得$S=\{n_kd,k\ge 0\}$, 则称$d$为$F$的步长. 证明格子点分布的步长总存在且唯一.

1.9* 设正整数$k_1,\cdots,k_s$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$c_1,\cdots,c_s$使得$nd=c_1k_1+c_2k_2+\cdots+c_sk_s$. ([提示] $d=r_1n_1+r_2n_2+\cdots+r_sn_s$. 设$r_1,\cdots,r_l\ge 0$, $r_{l+1},\cdots,r_s<0$. 记$q=-(r_{l+1}n_{l+1}+\cdots+r_sn_s)>0$. 可验证对任意$n\ge(q/d)^2$, 结论成立.)

§1.2 条件数学期望

对任意随机变量X, 不论是离散的还是连续的, 我们知道数学期望是刻画它的一个非常重要的数值特征量. 这一节我们将进一步介绍一个非常重要的概念——条件数学期望。

(A)随机事件下的条件分布与均值

任意给定的两个随机变量(X,Y), 利用初等概率中随机事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

$$F_{X|A}(x):=\mathbb{P}(X\le x|Y\in A)=\frac{\mathbb{P}(X\le x,Y\in A)}{\mathbb{P}(Y\in A)}.$$

注意到随机事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$Y\in A$条件下我们可以定义X的函数$g(X)$的数学期望为

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

其中$f(x,A)$为分布函数$F_{X|A}(x)$的密度函数.

例2.1 设(X,Y)服从$(0,1)\times(0,2)$上的均匀分布, 求$\mathbb{E}(X|Y>1)$.

解 令$A=(1,\infty)$. 当(X,Y)为连续随机变量时,

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

其中$f(x,y)$为(X,Y)的联合密度函数. 由于(X,Y)服从$(0,1)\times(0,2)$上的均匀分布,

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

代入计算得

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

从而

$$\mathbb{E}(X|Y>1)=\int_0^1 xf(x,A)\mathrm{d}x=1/2.$$ □

特别, 给定两元离散随机变量(X,Y), y是Y的一个可能值. 在随机事件$\{Y=y\}$条件下, X的(条件)分布列为

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

其中x表示X的可能取值. 在随机事件$\{Y=y\}$条件下, X的(条件)数学期望为

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

二元函数$\Phi(X,Y)$在随机事件$\{Y=y\}$条件下变成$\Phi(X,y)$, 从而(条件)数学期望为

$$\mathbb{E}(\Phi(X,Y)|Y=y)=\sum_x \Phi(x,y)\mathbb{P}(X=x|Y=y)=\frac{\sum_x \Phi(x,y)\mathbb{P}(X=x,Y=y)}{\mathbb{P}(Y=y)}.$$

例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

$$\mathbb{E}(X|X+X_1=n).$$

解 令$Y=X+X_1$, 则$Y\sim(n+m,p)$二项分布.

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

注意到$\sum_{k=0}^m C_{n-1}^k C_m^{m-k}=C_{n+m-1}^m$,

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

当(X,Y)是二元连续随机变量时,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)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)的联合密度函数. 从而在随机事件$\{Y=y\}$条件下, X的(条件)数学期望定义为

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

类似对二元函数$\Phi(X,Y)$在随机事件$\{Y=y\}$条件下的数学期望为

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

例2.3 对任意$y>0$, 求$\mathbb{E}(X|Y=y)$, 其中(X,Y)的联合密度函数为

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

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

(B)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$\omega\in\Omega$,
$$\omega\in\Omega\longrightarrow\mathbb{E}(X|Y=y)\qquad\text{其中 }y=Y(\omega).\tag{1.2.3}$$
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$\mathbb{E}(X|Y=y)$没有意义. 但此时若令
$$\mathcal{Y}=\begin{cases}\{y:\mathbb{P}(Y=y)>0\}, & Y\text{离散},\\ \{y:\quad f_Y(y)\neq 0\}, & Y\text{连续},\end{cases}$$
以及$\Omega_Y=\{\omega:\quad Y(\omega)\in\mathcal{Y}\}$, 那么映射在集合$\Omega_Y$ 上是明确、可行的.

定义2.1 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$\omega\in\Omega_Y$,
$$Z(\omega)=g(Y(\omega))=\mathbb{E}(X|Y=y)\qquad\text{其中 }y=Y(\omega),\tag{1.2.4}$$

则称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$y\in\mathcal{Y}$, $g(y)=\mathbb{E}(X|Y=y)$. 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
$$\bar{\Omega}_Y:=\Omega\setminus\Omega_Y=\{\omega:Y\in\{y;f_Y(y)=0\}\}.$$
因此
$$\mathbb{P}(\bar{\Omega}_y)=\mathbb{P}(Y\in\{y;f_Y(y)=0\})=\int_A f_Y(y)\mathrm{d}y=0,$$
其中$A=\{y;f_Y(y)=0\}$. 这表明由定义2.1给出的任意两个条件数学期望，除在一个零概率事件上可能不同外, 是完全一致的, 即Z_1与Z_2几乎处处相同. 一般而言, 零概率事件不影响讨论, 几乎处处相等的随机变量的差别可以忽略. 因此

注2.1 在几乎处处相等意义下, 定义2.1 中的条件数学期望是唯一的. 通常把该"唯一"的条件数学期望记作$\mathbb{E}(X|Y)$. 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2.4 给定样本空间$\Omega=(0,1)$以及合适的事件集和概率, 使得对任意$(a,b)\subset(0,1)$,

$$\mathbb{P}((a,b))=b-a.$$

设X,Y均是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$\mathbb{E}(X|Y)$.

解 按定义$\mathbb{P}(Y=0)=1/4$, $\mathbb{P}(Y=1)=0$且$\mathbb{P}(Y=2)=3/4$. 因此$\mathcal{Y}=\{0,2\}$. 从而

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

因此

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

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

例2.5 设X,Y为二维高斯随机变量, 联合分布密度函数如下, 求$\mathbb{E}(X|Y)$.

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

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$\mathcal{Y}=\mathbb{R}$, 且对任意$y\in\mathbb{R}$,

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

所以对任意ω, $\mathbb{E}(X|Y)(\omega)=g(Y(\omega))=Y(\omega)/2$. 即$\mathbb{E}(X|Y)=Y/2$. □

例2.6 (X,Y)的联合分布密度函数为$f(x,y)$. 设$a<b$ 且$\mathbb{P}(a\le Y\le b)>0$. 证明

$$\int_a^b \mathbb{E}(X|Y=y)f_Y(y)\mathrm{d}y=\mathbb{E}(X|a\le Y\le b)\mathbb{P}(a\le Y\le b),$$

其中$f_Y(y)$为Y的密度函数.

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

$$\begin{aligned}\mathbb{E}(X|a\le Y\le b) &= \int_{-\infty}^{\infty} x f_{X|A}(x)\mathrm{d}x=\frac{\int_{-\infty}^{\infty}\mathrm{d}x\int_a^b xf(x,y)\mathrm{d}y}{\int_{-\infty}^{\infty}\int_a^b f(x,y)\mathrm{d}x\mathrm{d}y}\\ &= \frac{\int_{-\infty}^{\infty}\mathrm{d}x\int_a^b xf(x,y)\mathrm{d}y}{\mathbb{P}(A)}.\end{aligned}$$

另一方面, 记$\mathbb{E}(X|Y)$为$g(Y)$, 那么

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

因此

$$\mathbb{P}(a\le Y\le b)\mathbb{E}(X|a\le Y\le b)=\int_a^b \mathbb{E}(X|Y=y)f_Y(y)\mathrm{d}y.$$ □

随机变量下的条件数学期望有如下重要性质.

性质2.1 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$\mathbb{E}(X|Y)=\mathbb{E}(X)$.

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

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

(4) $|\mathbb{E}(X|Y)|\le\mathbb{E}(|X||Y)$.

(5) $\mathbb{E}(X)=\mathbb{E}(\mathbb{E}(X|Y))$; (全概率公式).

(6) 对任意$\omega\in\Omega$, 记$y=Y(\omega)$, 那么

$$\mathbb{E}(\Phi(X,Y)|Y)(\omega)=\mathbb{E}(\Phi(X,y)|Y=y).$$

特别对任意随机变量$Z=f(X),W=g(Y)$

$$\mathbb{E}(f(X)g(Y)|Y)=g(Y)\mathbb{E}(f(X)|Y),\qquad \mathbb{E}(g(Y)|Y)=g(Y).$$

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$y=Y(\omega)$,

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

其中$f_X(x)$表示随机变量X的分布密度函数.

(5) 直接计算得

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

(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)的联合分布密度函数为

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

此时由随机变量下条件期望定义

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

作变量代换$x=\phi_y^{-1}(z)$, 那么

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

其他请读者自己完成, 参见练习2-4. □

例2.7 某矿工身陷在有三个门的矿井中, 选择第1个门的通道进行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

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

由全概率公式

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

由此可知$\mathbb{E}(X)=10$. □

(C) 随机变量下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2.3 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

$$\begin{aligned}\mathcal{Y} &= \{(y_1,\cdots,y_n);\ f_Y(y_1,\cdots,y_n)\neq 0\}, \\ \Omega_Y &= \{\omega;\ (Y_1(\omega),\cdots,Y_n(\omega))\in\mathcal{Y}\},\end{aligned}$$

其中$f_Y(y_1,\cdots,y_n)$为$(Y_1,\cdots,Y_n)$的边际密度. 若随机变量

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

满足: 对任意$\omega\in\Omega_Y$,

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

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$\mathbb{E}(X|Y_1,\cdots,Y_n)$.

此时条件数学期望仍遵循定义2.1后对条件数学期望概念的解释和说明, 并且对应的性质2.1仍然成立. 比如

(1) 若X与$Y_1,\cdots,Y_n$独立, 则$\mathbb{E}(X|Y_1,\cdots,Y_n)=\mathbb{E}(X)$;

(2) $\mathbb{E}(\sum\limits_{i=1}^{n} a_i X_i|Y_1,\cdots,Y_n)=\sum\limits_{i=1}^{n} a_i\mathbb{E}(X_i|Y_1,\cdots,Y_n)$, 其中$a_i$非随机.

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

$$=\mathbb{E}(\Phi(X,y_1,\cdots,y_n|Y_1=y_1,\cdots,Y_n=y_n);$$

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

作为全概率公式的拓展, 此时我们还有如下的所谓平滑性质.

性质2.2 对任意$n>m$, $\mathbb{E}\big(\mathbb{E}(h(X)|Y_1,\cdots,Y_n)\big|Y_1,\cdots,Y_m\big)=\mathbb{E}(h(X)|Y_1,\cdots,Y_m)$.

证明* 只考虑连续型随机变量的情形.

设$(X,Y_1,\cdots,Y_n)$的密度函数为$f(x,y_1,\cdots,y_n)$, 那么$(Y_1,\cdots,Y_n)$的密度函数

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

$(Y_1,\cdots,Y_m)$的密度函数$f_{Y_{1-m}}(y_1,\cdots,y_m)$为

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

记$g(Y_1,\cdots,Y_n)=\mathbb{E}(h(X)|Y_1,\cdots,Y_n)$, 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$(y_1,\cdots,y_m)\in\mathcal{Y}_{1-m}$,

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

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

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

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

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

由条件数学期望定义及其说明可知性质2.2成立. □

注2.2 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2.1 和性质2.2也成立.

例2.8 设X,Y为独立同分布随机变量, 求$\mathbb{E}(X|X+Y)$.

解 由于X,Y为独立同分布随机变量,

$$\mathbb{E}(Y|X+Y)=\mathbb{E}(Y|Y+X)=\mathbb{E}(X|X+Y).$$

因此

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

由此可得$\mathbb{E}(X|X+Y)=(X+Y)/2$. □

例2.9 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$n\geq 1$

$$\mathbb{E}(S_{n+1}|S_1,\cdots,S_n)=\mathbb{E}(S_{n+1}|S_n).$$

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$S_1,S_2,\cdots,S_n$独立, 因此由条件数学期望性质

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

例2.10 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$\mathbb{E}(N(t))$. 记第一个仪器的寿命为W, 显然, 若$W=w>t$, 那么$N(t)=1$, 而如果$W=w\leq t$且记从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

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

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

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

令$u(t)=\mathbb{E}(N(t))$, 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$u(0)=\mathbb{E}(N(0))=1$得$c=1$, 从而$\mathbb{E}(N(t))=u(t)=1+\lambda t$. □

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

$$\mathbb{P}(A|Y_1,\cdots,Y_n):=\mathbb{E}(\mathbf{1}_A|Y_1,\cdots,Y_n).$$

为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2.11 设二维连续随机变量的联合分布密度函数

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

求条件概率$\mathbb{P}(X<Y|Y)$.

解 对任意$y\geq 0$, 由定义

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

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

与通常的数学期望和方差的关系一样, 利用条件数学期望可定义条件方差:

$$\mathbf{Var}(X|Y)=\mathbb{E}\big((X-\mathbb{E}(X|Y))^2|Y\big)=\mathbb{E}(X^2|Y)-\big(\mathbb{E}(X|Y)\big)^2.$$

显然, 条件方差仍是随机变量. 且对任意$\omega\in\Omega$, 记$y=Y(\omega)$,

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

例2.12 接例2.3, 求$\mathbf{Var}(X|Y)$.

解 由例2.3易知$\mathbb{E}(X|Y)=(1+Y)/2$, 而且对任意$y>0$,

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

因此$\mathbb{E}(X^2|Y)=(1+Y)^2$, 进而$\mathbf{Var}(X|Y)=(1+Y)^2-(1+Y)^2/4=3(1+Y)^2/4$.

性质2.3 若X的二阶矩存在, 那么$\mathbf{Var}(X)=\mathbb{E}(\mathbf{Var}(X|Y))+\mathbf{Var}(\mathbb{E}(X|Y))$.
证明 由全概率公式可得

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

例2.13 设$\{X_t,t\in T\}$是一簇随机变量, N是取值在T上的随机变量. 若X_N是按函数复合方式定义的新的随机变量, 即对任意ω,

$$X_N(\omega):=X_{N(\omega)}(\omega),$$

则称X_N为复合随机变量. 现假设X_i, $i=1,2,\cdots$,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$\{X_i\}$ 独立且二阶矩也存在. 令$S_N=\sum_{k=1}^{N}X_k$. 求复合随机变量S_N 的均值与方差.
解 对任意$\omega\in\Omega$, 记$n=N(\omega)$. 由独立性和条件期望性质

$$\begin{aligned}\mathbb{E}(S_N|N)(\omega) &= \mathbb{E}(\sum_{k=1}^{N}X_i|N=n)=\mathbb{E}(\sum_{i=1}^{n}X_i|N=n)=n\mathbb{E}(X_1)\\ \mathbf{Var}(S_N|N)(\omega) &= \mathbb{E}((\sum_{k=1}^{N}X_i-n\mathbb{E}(X_1))^2|N=n)=\mathbb{E}\big(\big(\sum_{i=1}^{n}\big(X_i-\mathbb{E}(X_i)\big)\big)^2|N=n\big)\\ &= \mathbb{E}\big(\big(\sum_{i=1}^{n}\big(X_i-\mathbb{E}(X_i)\big)\big)^2\big)=n\mathbf{Var}(X_1).\end{aligned}$$

因此$\mathbb{E}(S_N|N)=N\mathbb{E}(X_1)$, $\mathbf{Var}(S_N|N)=N\mathbf{Var}(X_1)$. 从而

$$\begin{aligned}\mathbb{E}(S_N) &= \mathbb{E}(\mathbb{E}(S_N|N))=\mathbb{E}(N)\mathbb{E}(X_1);\\ \mathbf{Var}(S_n) &= \mathbb{E}(\mathbf{Var}(S_N|N))+\mathbf{Var}(\mathbb{E}(S_N|N))=\mathbb{E}(N)\mathbf{Var}(X_1)+[\mathbb{E}(X_1)]^2\mathbf{Var}(N).\end{aligned}$$ □

练习题

2.1 $X_i,i\geq 1$,是独立同分布的随机变量, 且$\mathbb{P}(X_1=1)=p$, $\mathbb{P}(X_1=0)=1-p$. 任取整数$1\leq k\leq n$, 对任意$1\leq i\leq n$, 求

$$\mathbb{P}(X_i=1|X_1+X_2+\cdots+X_n=k).$$

2.2 设二元随机变量(X,Y)的联合分布密度函数为$f(x,y)=6x^2y,\quad 0<x,y<1$. 求$\mathbb{E}(X|Y)$, $\mathbb{E}(X|X+Y)$.

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

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

2.4* 证明性质2.1中(2),(3),(4).
2.5 设X,Y是两有界随机变量且$\mathbb{E}(X|Y)=Y$, $\mathbb{E}(Y|X)=X$, 证明$\mathbb{P}(X=Y)=1$.

§1.3 收敛性与极限定理

在初等概率论中我们还简单地研究过无穷可列多随机变量的收敛问题. 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.

(A) 三类收敛的定义和极限定理

设$X_1,\cdots,X_n,\cdots$ 为概率空间$(\Omega,\mathfrak{F},\mathbb{P})$ 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$\varepsilon>0$,

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

(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

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

即至多存在一个例外集$A\in\Omega$使得$\mathbb{P}(A)=0$且对任意$\omega\in\Omega\setminus A$, 数列$X_n(\omega)\to X(\omega)$.

注3.1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$\psi_n(t)$点点收敛到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Laplace变换代替.

(2) 对几乎必然收敛，用$\varepsilon-N$ 语言可表述成$X_n\xrightarrow{a.s.}X$当且仅当存在一个集合$A\in\mathfrak{F}$满足$\mathbb{P}(A)=0$, 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$|X_n(\omega)-X(\omega)|<\varepsilon$.

利用这三种极限, 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3.1 (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$\mathbb{E}(|X_n|)<\infty$, 那么依概率有

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

定理3.2 (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, 即其均值和标准差分别为μ和σ, 那么对任意$x\in(-\infty,+\infty)$

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

即$\frac{\sum_{i=1}^n (X_i-\mu)}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$, 其中$N(0,1)$表示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. 为此我们对几乎必然收敛和依概率收敛做些必要的补充.

(B) Borel-Cantelli 引理

引理3.4 设$X, X_1, X_2, \cdots$为一列随机变量, 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 任取$\omega \in \{\lim_{n\to\infty} X_n = X\}$, 那么对任意$m > 0$, 存在$N(m,\omega)$, 当$n \geq N$时, $|X_n(\omega) - X(\omega)| < 1/m$, 即

$$\omega \in \{|X_n - X| < \frac{1}{m}\},$$

故

$$\omega \in \bigcap_{n\geq N(m,\omega)} \{|X_n - X| < \frac{1}{m}\} \subset \bigcup_{N=1}^{\infty} \bigcap_{n\geq N} \{|X_n - X| < \frac{1}{m}\}.$$

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$\lim_{n\to\infty} X_n(\omega) = X(\omega)$, 即$\omega \in \{\lim_{n\to\infty} X_n = X\}$. 因此引理结论成立. □

定理3.5 设$X, X_1, X_2, \cdots$为一列随机变量, 下列结论等价:

(1) $X_n \xrightarrow{a.s.} X$;

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

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

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

证明 (1)⇔(2): 引理3.4的自然推论.

(2) ⇒ (3): 显然成立,

(3) ⇒ (2): 注意到对任意$m > 0$, 存在$\varepsilon > 0$使得$1/m > \varepsilon$, 从而由(3)可知

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

因此

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

即(2)成立.

(3) ⇔ (4): 此时注意到

$$\bigcup_{n\geq N+1}\{|X_n - X| \geq \varepsilon\} \subset \bigcup_{n\geq N}\{|X_n - X| \geq \varepsilon\}.$$

由概率连续性,

$$\mathbb{P}\big(\bigcap_{N=1}^{\infty}\bigcup_{n\geq N}\{|X_n - X| \geq \varepsilon\}\big) = \lim_{N\to\infty}\mathbb{P}\big(\bigcup_{n\geq N}\{|X_n - X| \geq \varepsilon\}\big) = 0,$$

从而(3)与(4)等价. □

注3.2: 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

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

其中i.o是infinitely occur的缩写.

定理3.6(Borel-Cantelli引理) 设$A_1, A_2, \cdots$,是$(\Omega, \mathfrak{F}, \mathbb{P})$中一列随机事件, $p_k = \mathbb{P}(A_k)$,

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

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

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

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

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

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

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

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

注3.3 由定理3.5和定理3.6可知, 若对任意$\varepsilon>0$

$$\sum_{n=1}^{\infty}\mathbb{P}(|X_n-X|>\varepsilon)<\infty, \tag{1.3.1}$$

那么$X_n\overset{a.s.}{\to}X$. (1.3.1)也是我们常用的判断几乎必然收敛的充分条件.

性质3.7 若$X_n\overset{P}{\to}X$, 那么存在$\{X_n\}$的一个子列$\{X_{n_k}\}$使得

$$X_{n_k}\overset{a.s.}{\to}X.$$

证明 因为$X_n\overset{P}{\to}X$, 对任意$k\ge 1$, 由$\mathbb{P}(|X_n-X|>1/2^k)\to 0$可知存在$n_k$使得

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

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

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

即(1.3.1)对序列$\{X_{n_k}\}$成立, 从而推论成立. □

例3.1 设$X=\{X_k,k\ge 1\}$为独立同分布的随机变量, 且$X_k\sim b(1,p)$, 试用Borel-Cantelli引理证明

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k=p,\quad \text{a.s.} \tag{1.3.2}$$

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

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

直接计算可知

$$\mathbb{E}\Big(\big(\sum_{i=1}^{k}(X_i-p)\big)^4\Big)=kpq(p^3+q^3)+3k(k-1)p^2q^2. \tag{1.3.3}$$

因此$\sum\limits_{k=1}^{\infty}\mathbb{P}(A_k)<\infty$. 所以(1.3.2)成立. □

事实上例3.1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3.3 (强大数定律) 设$X_n, n\ge 1$, 为一列独立同分布的随机变量, 若$\mathbb{E}(|X_n|)<\infty$, 那么

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

(C) 柯西基本列

定义3.1 称随机变量序列$\{X_n\}$是几乎必然意义下的基本列或柯西(Cauchy) 列, 若至多存在一个零概率集A, 使得对任意$\omega\in\Omega\setminus A$ 以及任意$\varepsilon>0$, 存在一个$N=N(\varepsilon,\omega)$ 使得当$n,m>N$ 时$|X_n(\omega)-X_m(\omega)|<\varepsilon$.

显然, 由于收敛的数列当且仅当它是柯西数列我们容易推得

命题3.8 $\{X_n\}$几乎必然收敛当且仅当$\{X_n\}$是几乎必然收敛意义下的柯西列.

与定理3.5类似, 我们有如下结果. 证明请读者自己补充.

定理3.9 设$X_1, X_2, \cdots$为一列随机变量, 下列结论等价:

(1) $\{X_n\}$ 是几乎必然意义下的基本列

(2) 对任意$\varepsilon>0$,

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

(3) 对任意$\varepsilon>0$,

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

由定理3.9, 我们易得下面的推论.

推论3.10 $\{X_n\}$ 是几乎必然意义下的基本列当且仅当存在非负数列δ_k使得

$$\sum_{k=1}^{\infty}\delta_k<\infty \text{ 而且 } \sum_{k=1}^{\infty}\mathbb{P}(|X_{k+1}-X_k|\ge\delta_k)<\infty.$$

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

$$\begin{aligned}\{|X_{m+v}-X_m|\ge\varepsilon\}&\subset\{|X_{m+v}-X_m|\ge\sum_{k=m}^{m+v}\delta_k\}\\&\subset\bigcup_{k=m}^{v-1}\{|X_{k+1}-X_k|\ge\delta_k\}\subset\bigcup_{k=m}^{\infty}\{|X_{k+1}-X_k|\ge\delta_k\}.\end{aligned}$$

因此由定理3.9可知结论成立. □

定义3.2 称随机变量序列$\{X_n\}$是依概率收敛意义下的基本列或Cauchy列, 若对任意$\varepsilon>0$, $\delta>0$ 存在一个$N=N(\varepsilon,\delta)$ 使得当$n,m>N$ 时

$$\mathbb{P}(|X_n(\omega)-X_m(\omega)|\ge\varepsilon)\le\delta.$$

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3.11 依概率收敛的基本列一定存在一个子列使其为几乎必然收敛的基本列.

证明 设$\{X_n\}$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$m>n_k$

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

取子列$\{X_{n_k}\}$, 那么

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

由推论3.8可知$\{X_{n_k}\}$就是几乎必然意义下的基本列. □

记该基本列的极限为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$\{X_n\}$是概率收敛的. 因此$\{X_n\}$依概率收敛当且仅当它是概率收敛的基本列.

(D) 简单性质

性质3.12 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

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

(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$Y_n \overset{a.s.}{\to} Y$(或$Y_n \overset{P}{\to} Y$)且$\mathbb{P}(Y=0)=0$, 那么

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

其中当$Y_n=0$时$1/Y_n$可以自由定义.

证明 我们只证(2), 其他的参见习题3.4.

若$X_n \overset{a.s.}{\to} X$ 且$Y_n \overset{a.s.}{\to} Y$, 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$X_nY_n \overset{a.s.}{\to} XY$.

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

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

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

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

可得$\mathbb{P}(|X_n| \geq M) \leq \delta/4$. 进一步, 由概率收敛条件, 对任意$\varepsilon > 0$, 存在$N > N_1$, 使得当$n \geq N$时

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

因此当$n > N$时

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

这表明$X_nY_n \overset{P}{\to} XY$. □

注3.4 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. 在几乎处处收敛意义下, 利用数学分析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,

$$\{X_n(\omega) \to +\infty\} = \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{X_n \geq m\}. \tag{1.3.4}$$

对依概率收敛到可正概率取无穷的极限的定义与表示, 我们不再展开.

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概率公式.

引理3.13 设事件集$\{A_i, i \geq 1\}$是样本空间Ω的一个分割, 即A_i两两不交且$\cup_{i=1}^{\infty} A_i = \Omega$. 若$X$是非负随机变量或$X$的数学期望存在, 那么

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{E}(X\mathbf{1}_{A_k}). \qquad \text{(全概率公式)}$$

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,

$$\mathbb{E}(X) = \int_0^{\infty} x f_X(x)\mathrm{d}x = \lim_{M\to\infty}\int_0^{M} x f_X(x)\mathrm{d}x = \lim_{M\to\infty}\mathbb{E}(X\mathbf{1}_{\{X\leq M\}}).$$

由于$\{A_i, i \geq 1\}$是样本空间Ω的一个分割, 对任意$n \geq 1$, $\mathbf{1}_{\Omega} = \sum_{i=1}^{n}\mathbf{1}_{A_i} + \mathbf{1}_{\cup_{k=n+1}^{\infty}A_i}$,

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

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

$$\mathbb{E}(X\mathbf{1}_{\{X\leq M\}}\mathbf{1}_{\cup_{k=n+1}^{\infty}A_i}) \leq M\mathbb{P}(\cup_{k=n+1}^{\infty}A_i) \to 0.$$

因此

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

由此可得

$$\mathbb{E}(X) \leq \sum_{k=1}^{\infty}\mathbb{E}(X\mathbf{1}_{A_k}).$$

另一方面, 对任意$n \geq 1$,

$$\mathbb{E}(X) \geq \mathbb{E}(X\mathbf{1}_{\cup_{k=1}^{n}A_k}) = \sum_{k=1}^{n}\mathbb{E}(X\mathbf{1}_{A_k}).$$

令$n\to\infty$得

$$\mathbb{E}(X) \geq \sum_{k=1}^{\infty}\mathbb{E}(X\mathbf{1}_{A_k}).$$

综合两方面, 可知引理3.13成立. □

定理3.14(单调收敛定理)设X_n为非负随机变量且对任意$n \geq 1$, $X_n \leq X_{n+1}$几乎处处成立. 记$X = \lim_{n\to\infty} X_n$, a.s. 那么

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

证明 由条件可知, 存在一个概率为0的事件A,使得在$\bar{A}$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$X_n \overset{a.s.}{\to} X$. 若$\mathbb{P}(X = \infty) = p > 0$, 由(1.3.4)可知对任意$M > 0$,

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

这表明存在$m(M)$, 当$n > m$ 时$\mathbb{P}(X_n > M) > p/2$, 从而$\mathbb{E}(X_n) > Mp/2$,

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

若$\mathbb{P}(X = \infty) = 0$. 此时由数学期望全概率公式

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

记$a_n(k) = \mathbb{E}(X_n \mathbf{1}_{\{k-1\leq X<k\}})$, $a(k) = \mathbb{E}(X \mathbf{1}_{\{k-1\leq X<k\}})$. 对任意$k \geq 1$, 显然$a_n(k)$, $a(k)$非负,且$a_n(k)$随n增加单调不降. 注意到

$$\begin{aligned} |a_n(k) - a(k)| &\leq \mathbb{E}(|X_n - X|\mathbf{1}_{\{k-1\leq X<k\}}) \\ &= \mathbb{E}(|X_n - X|\mathbf{1}_{\{k-1\leq X<k\}\cap\{|X_n-X|>\epsilon\}}) \\ &\quad +\mathbb{E}(|X_n - X|\mathbf{1}_{\{k-1\leq X<k\}\cap\{|X_n-X|\leq\epsilon\}}) \\ &\leq k\mathbb{P}(|X_n - X| > \epsilon) + \epsilon \end{aligned}$$

由于$X_n \to X, a.s.$ 因此随$n \to \infty$, $k\mathbb{P}(|X_n - X| > \epsilon) \to 0$. 再由$\epsilon$的任意小性可知当$n \to \infty$时

$$a_n(k) \to a(k).$$

由数列级数收敛知识(见习题1.5)可得

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

定理得证. □

练习题

3.1 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$S_{N_n}/N_n \to X, \quad a.s.$

3.2 试构造一个随机序列$\{X_n\}$使得$X_n \xrightarrow{P} 0$ 但$\mathbb{P}(X_n \to 0) = 0$.

3.3 验证(1.3.3)成立.

3.4* 证明性质3.1中(1)和(3).

3.5* 证明定理3.9.

$$\left(\begin{array}{ll} \text{提示: 先证明} & \{X_n\text{是Cauchy列}\} = \bigcap_{m=1}^\infty \bigcup_{n=1}^\infty \bigcap_{v=1}^\infty \{|X_{n+v} - X_n| \le \frac{1}{m}\}, \\ \text{以及} & \bigcup_{v=1}^\infty \{|X_{n+v} - X_n| \ge \epsilon\} \subset \bigcap_{k=1}^n \bigcup_{v=1}^\infty \{|X_{k+v} - X_k| \ge \epsilon/2\}. \end{array} \right)$$

§1.4 特征函数、Laplace变换与母函数

特征函数, Laplace变换与母函数是我们刻画分布函数的重要工具.

(A) 特征函数

特征函数作为证明中心极限定理的主要工具, 在《概率论》课程中一般都有介绍. 为了后文需要, 我们先简要地回顾和概述特征函数的相关性质.

设随机变量X的分布函数为F. 我们称复值函数

$$\begin{aligned}\psi_X(t) &= \mathbb{E}(\mathrm{e}^{itX}) = \int_{-\infty}^{\infty} \mathrm{e}^{itx}\mathrm{d}F(x)\\ &= \int_{-\infty}^{\infty}\cos(tx)\mathrm{d}F(x) + i\int_{-\infty}^{\infty}\sin(tx)\mathrm{d}F(x)\\ &= \mathbb{E}(\cos(tX)) + i\mathbb{E}(\sin(tX)), \qquad t\in(-\infty,\infty)\ \ i\text{为虚数单位}.\end{aligned}$$

为X的特征函数.

由定义可知, 对任意$t\in\mathbb{R}$, $|\psi_X(t)|\le\psi_X(0)=1$, 而且作为函数$\psi_X(t)$在$\mathbb{R}$上一致连续.

特征函数与分布函数之间有如下关联:

定理4.1(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, 反之, 若特征函数不同则对应的分布函数也不同.

定理4.2(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$X_n \xrightarrow{d} X$的充要条件是对任意$t\in\mathbb{R}$, $\psi_{X_n}(t)$收敛而且$\lim\limits_{t\to 0}\lim\limits_{n\to\infty}\psi_{X_n}(t)=1$.

性质4.3 (1) $\overline{\psi_X(t)}=\psi_X(-t)=\psi_{-X}(t)$; (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4.4 若X的n阶绝对矩$\mathbb{E}(|X|^n)<\infty$, 那么对任意$k\le n$, $\mathbb{E}(X^k)=i^{-k}\psi^{(k)}(0)$.

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, 联合分布函数为$F(x_1,\cdots,x_n)$, 称n元复值函数

$$\begin{aligned}\psi_{(X_1,X_2,\cdots,X_n)}(t_1,\cdots,t_n) &= \mathbb{E}(\mathrm{e}^{i(t_1X_1+\cdots+t_nX_n)})\\ &= \int_{-\infty}^{\infty}\mathrm{e}^{i(t_1x_1+\cdots+t_nx_n)}\mathrm{d}F(x_1,\cdots,x_n)\end{aligned}$$

为$(X_1,\cdots,X_n)$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4.1'(唯一性定理) 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.

性质4.4 随机变量$X_1,\cdots,X_n$相互独立当且仅当对任意的$t_1,\cdots,t_n\in\mathbb{R}$

$$\psi_{(X_1,X_2,\cdots,X_n)}(t_1,\cdots,t_n)=\prod_{i=1}^{n}\psi_{X_i}(t_i).$$

(B) Laplace 变换

对于非负随机变量分布的刻画, 用所谓的Laplace变换会更方便些.

定义4.1 若X为非负随机变量, 称函数$L(\theta)=\mathbb{E}(\mathrm{e}^{-\theta X})$ 为X的Laplace变换, 其中变量$\theta\in[0,\ +\infty)$.

若记X的分布函数为F, 通常也把Laplace变换$L(\theta)$记作

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

此时也称$L(\theta)$是分布F的Laplace变换.

当$F(x)$是具有密度为$f(x)$的连续分布时,

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

当$F(x)$是具有分布列$p_i=\mathbb{P}(X=x_i)$, $i=1,2,\cdots,\infty$, 的离散分布时,

$$L(\theta)=\sum_{i=1}^{\infty}\mathrm{e}^{-\theta x_i}p_i,$$

特别若x_i取非负整数, 令$p_k=\mathbb{P}(X=k)$, 那么

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

定理4.5 (Laplace变换反演公式) 设F是$[0,\infty)$上的分布函数, $L(\theta)$是F的Laplace变换, 那么对F的任意连续点x, 有

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

证明 只证明$F(x)$为连续分布的情形. 设其密度函数为f.

对任意$\theta>0$, 可以对L求k阶导数,

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

对任意给定的$x\geq 0$, 取$\theta=n$, 则

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

设$X_i,\ i\ \geq\ 1$为独立同分布的随机变量, 服从强度为y的泊松分布, 那么$\mathbb{E}(X_1)=\mathbb{V}ar(X_1)=y$且

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

服从强度为ny的泊松分布. 此时

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

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

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

另一方面, 当$y-x\le-\epsilon/4$时

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

因此

$$\begin{aligned}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL^{(k)}(n)-F(x)\Big|&\le\int_0^{(x-\epsilon)\vee0}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}-1\Big|f(y)\mathrm{d}y\\&\quad+\int_{x+\epsilon}^{\infty}\sum_{k=0}^{\lfloor nx\rfloor}\mathrm{e}^{-ny}\frac{(ny)^k}{k!}f(y)\mathrm{d}y+2\int_{(x-\epsilon)\vee0}^{x+\epsilon}f(y)\mathrm{d}y\\&\le\frac{1}{n}(\frac{x}{\epsilon^2}+\frac{1}{\epsilon})+2(F(x+\epsilon)-F(x-\epsilon)).\end{aligned}$$

令$n\to\infty$, 我们可得

$$\limsup_{n\to\infty}\Big|\sum_{k=0}^{\lfloor nx\rfloor}\frac{(-1)^k}{k!}n^kL^{(k)}(n)-F(x)\Big|\le2(F(x+\epsilon)-F(x-\epsilon)).$$

由于ϵ的任意性以及$\lim_{\epsilon\to0}F(x+\epsilon)-F(x-\epsilon)=0$, 我们可得

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

因此

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

由定理4.2 我们可以立即得到如下的Laplace变换的唯一性定理.

定理4.6 $[0,\infty)$上不同的概率分布有不同的Laplace变换.

对于Laplace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阅参考书目[10]和[1]中有关内容.

定理4.7 设非负随机变量X_n的Laplace变换为$L_{X_n}(\theta)$. 那么存在r.v.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$\lim\limits_{\theta\to 0}\lim\limits_{n\to\infty}L_{X_n}(\theta)=1$.

注 定理4.7可以推广为如下的所谓广义连续性定理：$X_n\xrightarrow{d}X$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Laplace定义容易验证如下性质成立：

性质4.8 (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$L_{X+Y}(\theta)=L_X(t)L_Y(\theta)$, $t\ge 0$; (4) 对任意正整数k, $\mathbb{E}(X^k)=(-1)^kL_X^{(k)}(0)$, 其中$L_X^{(k)}$表示$L_X$的$k$阶导数.

例4.1 设X服从$(0,1)$上均匀分布, 那么$L_X(\theta)=(1-\mathrm{e}^{-\theta})/\theta$.

例4.2 若X服从$(0,\infty)$上参数为$\lambda>0$的指数分布, 那么

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

例4.3 若X服从Γ分布, 密度函数

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

那么

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

例4.4 设X_n服从二项分布$b(n,\frac{1}{n})$, 那么对任意$\theta>0$, 当$n\to\infty$

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

注意$\mathrm{e}^{-(1-\mathrm{e}^{-\theta})}$是强度为1的泊松分布的Laplace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.

(C) 母函数

在研究只取有限或无限个非负整数值的随机变量时, 用所谓母函数来代替特征函数或Laplace变换较为方便.

定义4.2 设X为非负整数值的随机变量, 称函数$\phi_X(s)=\mathbb{E}(s^X)$ 为X的母函数, 其中变量$s\in(-1,1)$.

设X的分布列为

$$p_k=\mathbb{P}(X=k),\qquad k\geq 0,$$

那么

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

对照母函数和Laplace变换, 容易看出所谓母函数就是X为非负整数值的随机变量条件下的Laplace变换, 只不过将其中的$\mathrm{e}^{-\theta}$记作s. 因此母函数也拥有Laplace变换的性质, 我们不在赘述, 请读者自己补充. 下面我们主要介绍和解释母函数性质的一些变化.

性质4.9 若非负整数值随机变量X的n阶矩存在, 则母函数$\psi_X(s)$的k阶导数$\phi_X^{(k)}(s)$存在($|s|\leq 1$), 且X的n阶矩可由$\phi_X(s)$前n阶导数在$s=1$表示. 特别

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

证明 由数学分析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4.10(反演公式) 设非负整数值随机变量X的分布列为$p_k,k\geq 0$, 母函数为$\psi_X(s)$. 那么

$$p_k=\phi_X^{(k)}(0)/k!.$$

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4.5 设X服从几何分布, 对任意$k\geq 1$, $p_k=\mathbb{P}(X=k)=p(1-p)^{k-1}$. 那么

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

例4.6 设X_i, $i\geq 1$是独立同分布的随机变量, 服从$(0,1)$上均匀分布. Y与$X_k,k\geq 1$独立, 服从几何分布,

$$p_k=\mathbb{P}(Y=k)=pq^{k-1},\qquad k\geq 1,\quad p=1-q\in(0,1).$$

令$S=\sum_{i=1}^{Y}X_i$, 求S的Laplace变换$L_S(\theta)$.

解 由定义和条件数学期望性质

$$\begin{aligned}L_S(\theta)&=\mathbb{E}(\mathrm{e}^{-\theta S})=\mathbb{E}(\mathbb{E}(\mathrm{e}^{-\theta S}|Y))=\mathbb{E}(\mathbb{E}(\mathrm{e}^{-\theta\sum_{i=1}^{Y}X_i}|Y))\\&=\sum_{k=1}^{\infty}\mathbb{E}(\mathrm{e}^{-\theta\sum_{i=1}^{Y}X_i}|Y=k)p_k=\sum_{k=1}^{\infty}\mathbb{E}(\mathrm{e}^{-\theta\sum_{i=1}^{k}X_i}|Y=k)p_k.\end{aligned}$$

由于

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

因此

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

容易将母函数的定义推广到多元非负整数值的随机变量, 得到联合分布的母函数. 为简单, 下面仅介绍二元随机变量情形.

定义4.2 设X_1, X_2为非负整数值的随机变量, 联合分布列为

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

称$[-1,1]\times[-1,1]$上的二元函数

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

为二元随机变量(X_1,X_2)的母函数.

与定理4.10一样, 此时我们有反演公式

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

性质4.11 非负整数值随机变量X_1, X_2相互独立当且仅当对任意的$s,t\in[-1,1]$

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

练习题

4.1 在X是连续型非负随机变量条件下证明其Laplace变换$L_X(t)$在$[0,+\infty)$内是一致连续的.

4.2 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Laplace变换$L_X(t)$满足

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

4.3 设X_i, $i\geq 1$是独立同分布的随机变量, 服从参数为λ的指数分布. N是服从强度为θ的Poisson分布的随机变量且与$X_k, k\geq 1$独立. 令$S=\sum_{i=1}^{N}X_i$, 求S的Laplace变换$L_S(\theta)$.

4.4 证明性质4.9.

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

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

第二章　随机过程概述

本章我们将简要地解释随机过程的基本概念并介绍两类基础随机模型：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及其“逆”。它们在应用领域有着广泛的应用。

§2.1　随机过程的基本概念

随机过程简单而言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. 虽然在概率论中我们对无穷多随机变量的极限性质做了初步研究, 但这在应用中远远不够. 先看下面两个具体的例子.

例5.1 观察赌场里某个赌徒正在进行的赌博(比如抽扑克牌比大小). 若赌徒在输完筹码才会离开, 问要进行多少局，赌局才会结束？由于赌局的随机性，每一局后赌徒拥有的赌资是随机的. 要对这个问题做出估计，可以用随机变量$Y(n)$表示第n局后赌徒拥有的筹码, 那么所估计的赌局数为

$$k = \inf\{n; Y(n) \leq 0\}.$$

要合理估计k值, 一般要掌握随机变量族$\{Y(n), n \geq 0\}$的随机变化规律.

例5.2 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13元买入15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5.1 称$X=\{X(t);\ t\in T\}$为概率空间$(\Omega,\mathfrak{F},\mathbb{P})$ 上随机过程, 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, T 称为是指标集或参数集, 特别若指标集具有时间属性时, 也称之为时间参数.

按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5.1 对随机过程还可从其他角度进行分类. 比如, 若对所有$t\in T$, $X(t)\in R^d$其中$d>1$, 则也称X是向量值随机过程; 若$T\in R^n$ 的子集时(此时T往往对应空间或时空指标), 我们又称X为随机场. 为了简单, 本书不讨论随机场这种情形, 从现在起本书讨论的随机过程, 如无特别说明, 都是实值随机过程, 即$S\in\mathbb{R}$.

随机过程从函数角度看也是集合$T\times\Omega\to\mathbb{R}$的二元函数. 对给定的时刻$t,\omega$, $X(t,\omega)$是一个实数; 对给定的t和变化的ω, 随机过程$X(t,\cdot)$ 为随机变量; 对变化的t和给定的$\omega\in\Omega$, $X(\cdot,\omega)$是一个指标集T上的函数:

$$X(\cdot,\omega):T\to\mathbb{R}\ \ \text{对任意}t\in T, X(\cdot,\omega)(t)=X(t,\omega).$$

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).

对概率空间$(\Omega,\mathfrak{F},\mathbb{P})$上有限个随机变量$X_1,\cdots,X_m$, 我们一般用联合分布函数

$$F(x_1,x_2,\cdots,x_m)\triangleq\mathbb{P}(X_1\le x_1,X_2\le x_2,\cdots,X_m\le x_m),$$

其中$x_1,\cdots,x_m\in\mathbb{R}$ 刻画他们的统计规律. 对于随机过程, 我们可以借用这一工具.

定义5.2 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

$$F_{t_1,t_2,\cdots,t_k}(x_1,x_2,\cdots,x_k)$$

为随机过程X的一个k阶有限维分布, 称函数族

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

为随机过程X的有限维分布族.

注5.2 当S为离散时,有限维分布族可用如下有限维分布列代替

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

性质5.1 随机过程的有限维分布满足

(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. □

定理5.2*(Kolmogorov) 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, 则在样本空间

$$\Omega=S^T=\{(s_t;t\in T);s_t\in S\},$$

以及σ-代数

$$\begin{aligned}\mathfrak{F}=\sigma\Big\{&\{(s_t,t\in T);\ (s_{t_1},\cdots,s_{t_n})\in B\};\\ &B\subset\mathfrak{B}(S^n),\ t_1,\cdots,t_n\in T,n\geq 1\Big\}\end{aligned}$$

上存在唯一的概率$\mathbb{P}$, 使得$(\Omega,\mathfrak{F},\mathbb{P})$上随机过程$X=\{X(t),t\in T\}$, 其中

$$X(t,\omega)=\omega_t,\qquad\qquad \omega=(\omega_t)_{t\in T},$$

的有限维分布族恰好是$\{F_{t_1,t_2,\cdots,t_k}(x_1,x_2,\cdots,x_k);\ t_1,\cdots,t_k\in T,k\geq 1\}$.

定理的证明超出本课程范围, 在此省略. 由该定理可看出, 有限维分布族在一定意义下确定唯一的随机过程.

注5.3* 定理中$\mathfrak{B}(S^n)$表示由S^n中所有开集产生的σ-代数, 通常称之为Borel代数. 特别, 若 S^n为离散集, 那么$\mathfrak{B}(S^n)$就是由S^n中所有子集构成的集合类.

下面是一个解释如何利用Kolmogorov定理给出随机过程及对应概率空间在数学上的严格构造的简单例子.

例5.3* 从装有M个红球和N个白球的袋中重复做有放回的抽取, 若抽到红色记作0，白球记作1, 试建立严格的数学模型描述随机抽取的动态变化.

解 令$p=\frac{N}{N+M}$, $q=1-p$, a_i, $i\geq 1$, 取值为0或1. 显然该过程中任意抽取s次, 抽取结果为$a_1,\cdots,a_s$的概率为

$$p^{a_1+\cdots+a_s}q^{s-(a_1+\cdots+a_s)}.$$

由抽取过程我们可以得到如下的有限分布列族

$$\begin{gathered}\{P_{n_1,n_2,\cdots,n_s}(a_1,\cdots,a_s);\ n_1,n_2,\cdots,n_s\text{为互不相同的正整数},\\ a_i\in\{0,1\},i=1,2,\cdots,s,s\geq 1\}\end{gathered}$$

其中$P_{n_1,n_2,\cdots,n_s}(a_1,\cdots,a_s)=p^{a_1+\cdots+a_s}q^{s-(a_1+\cdots+a_s)}$. 容易验证, 该有限维分布列簇是对称的且具有相容性. 因此由Kolmogorov定理, 令

$$\begin{aligned}\Omega&=\{0,1\}^{\mathbf{N}}=\{(a_1,a_2,\cdots,);\ a_i=0,1,i=1,2,\cdots\},\\ \mathfrak{F}&=\sigma\Big\{\{(\omega_i,i\geq 1);\ \omega_{i_1}=a_{i_1},\cdots,\omega_{i_k}=a_{i_k}\};\\ &\qquad\qquad i_l\in T,a_{i_l}\in\{0,1\},l=1,\cdots,k,k\geq 1\Big\}\end{aligned}$$

存在唯一概率$\mathbb{P}$使得随机过程$X=\{X_n,n\ge 1\}$, 其中$X_n(\omega)=\omega_n$的有限维分布与已知观察一致. 此时Ω中随机事件

$$\{(\omega_n;\ n\ge 1);\ \omega_{i_k}=a_{i_k},k=1,2,\cdots,s\}$$

发生概率$P=p^{a_{i_1}+\cdots+a_{i_s}}q^{s-(a_{i_1}+\cdots+a_{i_s})}$. □

定义5.3 设$X=\{X_t;t\in T\}$, $Y=\{Y_t;t\in T\}$都是概率空间$(\Omega,\mathfrak{F},\mathbb{P})$上的随机过程, 称$Y$是$X$ 的一个版本(Version), 若X,Y有限维分布族相同; 称Y是X的一个修正(Modification), 若对任意$t\in T$, $\mathbb{P}(X_t=Y_t)=1$; 称X与Y不可区分(Indistinguish), 若$\mathbb{P}(X_t=Y_t,t\in T)=1$.

显然衡量两个过程是否一样，概念“不可区分”是最强的. 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. 这两个结论一般都逆不真, 但若X,Y 为离散参数随机过程, 那么”不可区分”与”修正“等价.

定义5.4 设$X=\{X_t;t\in T\}$, $Y=\{Y_s;s\in S\}$都是概率空间$(\Omega,\mathfrak{F},\mathbb{P})$上的随机过程, 若对任意$k,m\ge 1$, $t_1,\cdots,t_k\in T$, $s_1,\cdots,s_m\in S$, $(X_{t_1},\cdots,X_{t_k})$与$(Y_{s_1},\cdots,Y_{s_m})$ 独立, 则称X和Y是独立的.

例5.4 设W_i, $i\ge 1$是独立同分布的随机变量, S_n为W_i的部分和序列, N为一个固定常数. 对任意$k\ge 1$, 令$T_k=S_{N+k}-S_N$. 那么随机过程$\{T_k,k\ge 1\}$与$\{S_n,1\le n\le N\}$独立. 事实上对任意$m\ge 1$, 任取$k_1,\cdots,k_m\ge 1$, 由W_i, $i\ge 1$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5.5 设$X=\{X(t);t\in T\}$是定义在概率空间$(\Omega,\mathfrak{F},\mathbb{P})$上的随机过程, 分别称如下的函数

$$m_X(t)=\mathbb{E}(X(t)),\qquad D_X(t)=\mathbb{V}ar(X(t)),$$

$$R_X(s,t)=\mathbb{E}(X(s)X(t)),\qquad \mathbb{C}ov_X(s,t)=\mathbb{C}ov(X(s),X(t))$$

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.

例5.5 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$t, t+h, s, s+h \in T$, $X(t+h)-X(t)$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$X=\{X(t), t \geq 0\}$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$t \geq 0$，

$$m(t)=Ct+\mathbb{E}(X(0)), \quad D(t)=C_1 t+\mathbf{Var}(X(0)), \tag{2.1.1}$$

其中$C=\mathbb{E}(X(1))-\mathbb{E}(X(0))$, $C_1=\mathbf{Var}(X(1))-\mathbf{Var}(X(0))$.

如果T为非负整数, 独立增量过程就是一列相互独立随机变量的部分和；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.

例5.6 随机过程$X=\{X_n; n \geq 1\}$称为伯努利过程, 若$\{X_n; n \geq 1\}$ 为独立同分布(i.i.d)随机变量, 且

$$\mathbb{P}(X_n=1)=p, \qquad \mathbb{P}(X_n=0)=q=1-p, \quad 0<p<1, n \geq 1.$$

例5.3中构造的随机模型其实就是一个伯努利过程. 伯努利过程常用来描述一列独立重复实验, 通常将$\mathbb{P}(X_n=1)=p$看作第n次实验成功概率, $\mathbb{P}(X_n=0)=q$ 看作第n次实验失败的概率. 令$T_0=0$,

$$T_1=\inf\{n>0; X_n=1\},$$

其中规定$\inf\emptyset=+\infty$. 因此T_1是个可以取$+\infty$的随机变量, 而且由

$$\{T_1=n\}=\{X_n=1, X_1, \cdots, X_{n-1}\text{均不为1}\} \tag{2.1.2}$$

可知$T_1=n$是否发生由$X_1, \cdots, X_n$的取值确定或者说由时刻n记此前观察到的结果确定. 一般地, 我们称具有这种性质的随机变量为**随机时间**或**停时**.

进一步, 对任意$k>1$, 我们还可以定义$T_k=\inf\{n>T_{k-1}; X_n=1\}$. 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

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

这表明若$\sum\limits_{n\geq 1}|f(n)|<\infty$, 则对任意$\varepsilon>0$, 存在$m>0$使得

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

因此

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

例5.7 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, 若损坏后更新的费用为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, 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, 为此厂家需额外收取$c\alpha^{T_k}$ 元费用. 要保证长期使用则应多收取的费用为$\sum_{k=1}^{\infty} c\alpha^{T_k}$. 因此由(2.1.3), 预期应增收费用为

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

练习题

5.1 试解释对离散随机过程而言, “不可区分”与“修正”等价.
5.2 验证(2.1.1)中$m(t)$与$D(t)$的表达式.
5.3 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.

§2.2 直线上随机游动

定义6.1 设$X_n, n \ge 1$为独立同分布的随机变量, 随机变量X_0与$\{X_n, n \ge 1\}$独立. 令

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

称随机过程$W = \{W_n;\ n \ge 0\}$为(直线上)随机游动或1维随机游动. 称X_0的取值为随机游动W的初值或初始位置. 若随机游动W定义中X_0是整数值随机变量且

$$\mathbb{P}(X_n = 1) = p, \qquad \mathbb{P}(X_n = -1) = q = 1 - p,$$

则称随机过程$W = \{W_n;\ n \ge 0\}$为直线上的简单随机游动或(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.

注6.1 用初等概率论的眼光看, 随机游动W实质就是独立同分布随机变量的部分和. 但当我们把X_n形象地理解成困在一根细管中的蚂蚁随机走出的一步时, W就描述了这只蚂蚁在细管中随机走过的路程. 这也是我们称W为随机游动的一种形象解释. 为了叙述方便, 有时也形象地称X_n为第n步的步长.

注6.2 由定义可知简单随机游动W是平稳独立增量过程.

由大数定律我们知道$W_n/n \overset{a.s.}{\to} \mathbb{E}(X_1)$. 因此, 只要$p \ne q$, 简单随机游动$|W_n| \overset{a.s.}{\to} +\infty$, 从而对几乎所有的$\omega$, 以及任意常数$C$, $|W_n(\omega)| \le C$ 只能出现有限次, 换句话说, 存在某个$N(\omega)$, 在此之后，$|W_n(\omega)|$都比C大. 显然, 这种趋势性的认识对我们整体把握随机现象是非常有必要的, 也是非常基础的. 但这还不够, 因为我们还经常会对随机游动发展过程出现的某些现象产生兴趣. 比如, 第一次到达指定位置的时间是多少？它一定会回到指定位置吗？指定两个位置会先到达哪个？以及在给定的时间内它呆在给定区域的可能性等等. 这类问题不仅有趣而且跟实际应用密切相关, 也是我们用随机过程刻画随机现象时常要解决的问题. 下面我们就从随机游动这一模型出发, 围绕这类问题做些简单的介绍, 帮助大家初步了解随机过程所讨论的内容. 为了简便, 我们只讨论**整数值的随机游动**, 即$X_k, k \ge 0$, 都是整数值随机变量.

首先我们引进记号

$$p_n(x, y) := \mathbb{P}(W_n = y | W_0 = x).$$

$p_n(x, y)$表示随机游动在初值为x条件下, 时刻n时值为y的概率. 注意到X_i是独立同分布的随机变量, 对任意$k \ge 0$, 条件概率

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

进而

$$\mathbb{P}(W_{k+n}=y|W_k=x) = \frac{\mathbb{P}(\sum_{i=1}^n X_i = y-x, X_0=0)}{\mathbb{P}(X_0=0)} = p_n(0, y-x).$$

由此可知

$$\mathbb{P}(W_{k+n}=y|W_k=x) = \mathbb{P}(W_n=y|W_0=x) = p_n(x,y) = p_n(0, y-x). \quad (2.2.1)$$

因此$p_n(x,y)$不仅表示随机游动起点由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. 我们常把后者简记成$p_n(y-x)$. 特别, 若W的初值取定为0, 那么对任意x, $\mathbb{P}(W_n=x)=p_n(x)$.

为获得随机游动W的任意有限维分布, 任取

$$0=k_0<k_1<\cdots<k_n, \ \ r_0, r_1, \cdots, r_n \in \mathbf{Z}.$$

由X_i的独立同分步性质可得

$$\begin{aligned}
&\mathbb{P}(W_{k_0}=r_0, W_{k_1}=r_1, \cdots, W_{k_n}=r_n) \\
&= \mathbb{P}\Big(X_0=r_0, \sum_{i=k_0+1}^{k_1} X_i = r_1-r_0, \cdots, \sum_{i=k_{n-1}+1}^{k_n} X_i = r_n - r_{n-1}\Big) \\
&= \mathbb{P}(X_0=r_0)\mathbb{P}\Big(\sum_{i=k_0+1}^{k_1} X_i = r_1-r_0\Big)\cdots\mathbb{P}\Big(\sum_{i=k_{n-1}+1}^{k_n} X_i = r_n-r_{n-1}\Big) \\
&= \mathbb{P}(X_0=r_0)\mathbb{P}(W_{k_1-k_0}=r_1-r_0)\cdots\mathbb{P}(W_{k_n-k_{n-1}}=r_n-r_{n-1}) \\
&= \mathbb{P}(X_0=r_0)p_{k_1}(r_0,r_1)p_{k_2-k_1}(r_1,r_2)\cdots p_{k_n-k_{n-1}}(r_{n-1},r_n). \quad (2.2.2)
\end{aligned}$$

可见条件概率簇$\{p_n(x,y)\}$对我们认识随机游动的统计规律而言是基础的.

对简单随机游动, 我们可以非常方便地计算$\{p_n(x,y)\}$. 由于简单随机游动每次步长只能是+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

$$p_n(x,y) = \begin{cases} C_n^{(n+y-x)/2} p^{(n+y-x)/2} q^{(n+x-y)/2}, & |y-x| \le n\text{且同奇偶}; \\ 0, & \text{其他}. \end{cases} \quad (2.2.3)$$

例6.1 设W是(q,p)-简单随机游动. 对$s\in[0,1)$, 令$\phi(s)=\sum\limits_{n=0}^{\infty} p_n(0,0)s^n$. 求$\phi(s)$.

解 由(2.2.3)可得

$$p_n(0,0) = \begin{cases} 0, & n=2k-1, \\ C_{2k}^k p^k q^k, & n=2k. \end{cases}$$

因此

$$\begin{aligned}\phi(s) &= 1+\sum_{k=1}^{\infty} C_{2k}^{k} p^k q^k s^{2k} = 1+\sum_{k=1}^{\infty}\frac{(2k)!}{k!k!}(pqs^2)^k = 1+\sum_{k=1}^{\infty}\frac{(2k-1)!!}{k!}(2pqs^2)^k \\ &= 1+\sum_{k=1}^{\infty}\frac{-\frac{1}{2}(-\frac{1}{2}-1)\cdots(-\frac{1}{2}-(k-1))}{k!}(-4pqs^2)^k\end{aligned}$$

注意到函数$(1+x)^{-1/2}$的幂级数展开式为

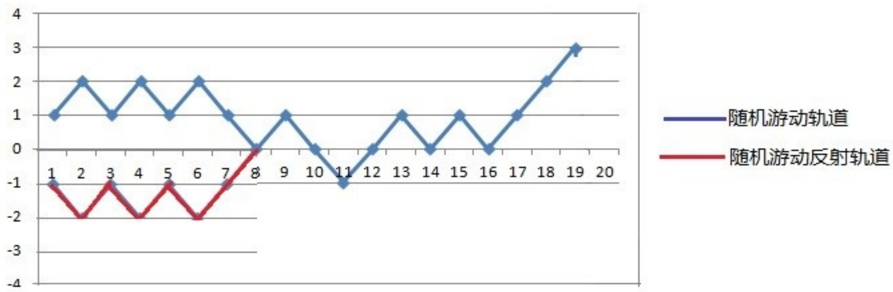
$$(1+x)^{-1/2} = 1+\sum_{k=1}^{\infty}\frac{-\frac{1}{2}(-\frac{1}{2}-1)\cdots(-\frac{1}{2}-(k-1))}{k!}x^k, \quad |x|<1.$$

因此所求函数$\phi(s)=(1-4pqs^2)^{-1/2}$. □

下面我们介绍更多的随机游动的其他性质和规律.

(A) 反射原理与对称原理

将**简单随机游动**W中其各位置点(n, W_n)在平面上绘点, 并按n的次序用直线依次相连, 我们称这样的折线为W的一根轨道。如图



引理6.1(反射原理) 设x, y为正整数, 对**简单随机游动**W而言, 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.

证明 令A表示从(n, x)到$(n+m, y)$并途中位置回到零点的所有轨道, 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. 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(如上图红线所示), 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6.2 若W是**简单对称随机游动**, 那么

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

证明 注意到W是对称随机游动, 即$p=q=1/2$时, 随机游动每一步向上或向下走的机会是对称的. 将W的任意两个可能点(n,x), $(n+m,y)$ 它们连接起来的轨道是等可能的, 而且与从(n,x)出发走过m步的任一轨道发生的概率相同. 注意到从(n,x)出发走过m步的轨道数为2^m, 由引理5.1得

$$\begin{aligned}\mathbb{P}(W_{n+m}=y,\min_{n<k<m}W_k\le 0\ \mid\ W_n=x)&=\frac{A\text{中轨道数}}{2^m}\\&=\frac{B\text{中轨道数}}{2^m}=\mathbb{P}(W_{n+m}=y|W_n=-x).\end{aligned}$$ □

例6.2 设W是简单对称随机游动, 初始值为0, 求概率

$$\mathbb{P}(W_1>0,W_2>0,\cdots,W_{2n}>0).$$

解 由于W是初值为0的简单对称随机游动,因此$W_1>0$意味着$W_1=1$, 而且W_{2n}只取偶数, 因此由全概率公式

$$\begin{aligned}\mathbb{P}(W_1>0,\cdots,W_{2n}>0)&=\sum_{k=1}^{\infty}\mathbb{P}(W_1>0,W_2>0,\cdots,W_{n-1}>0,W_{2n}=2k)\\&=\sum_{k=1}^{\infty}\mathbb{P}(W_1=1,\min_{1<m<2n}W_m>0,W_{2n}=2k)\end{aligned}$$

利用$\{\min_{1<m<2n}W_m>0\}$的对立事件$\{\min_{1<m<2n}W_m\le 0\}$得

$$\begin{aligned}&\mathbb{P}(W_1>0,\cdots,W_{2n}>0)\\&\quad=\sum_{k=1}^{\infty}[\mathbb{P}(W_1=1,W_{2n}=2k)-\mathbb{P}(W_1=1,\min_{1<m<2n}W_m\le 0,W_{2n}=2k)]\\&\quad=\sum_{k=1}^{\infty}\mathbb{P}(W_1=1)[\mathbb{P}(W_{2n}=2k|W_1=1)-\mathbb{P}(\min_{1<m<2n}W_m\le 0,W_{2n}=2k|W_1=1)].\end{aligned}$$

由命题6.2可知

$$\begin{aligned}\mathbb{P}(W_1>0,\cdots,W_{2n}>0)&=\mathbb{P}(W_1=1)\sum_{k=1}^{\infty}[p_{2n-1}(1,2k)-p_{2n-1}(-1,2k)]\\&=\mathbb{P}(W_1=1)\sum_{k=1}^{\infty}[p_{2n-1}(2k-1)-p_{2n-1}(2k+1)]=\frac{p_{2n-1}(1)}{2}=\frac{(2n-1)!}{n!(n-1)!4^n}.\ \square\end{aligned}$$

注6.3 直接计算可知

$$\frac{p_{2n-1}(1)}{2}=\frac{1}{2}p_{2n}(0)=\frac{1}{2}\mathbb{P}(W_{2n}=0),$$

因此对初值为0的简单对称随机游动而言一个有趣的现象是

$$\mathbb{P}(W_1\neq 0,W_2\neq 0,\cdots,W_{2n}\neq 0)=\mathbb{P}(W_{2n}=0).$$

引理6.3(对称原理) 若W是初值为0的**随机游动**, 那么对任意$n\geq 1$ 以及$b>a\geq 0$,

$$\mathbb{P}(W_1>0,\cdots,W_{n-1}>0,W_n\in[a,b])=\mathbb{P}(W_1<W_n,\cdots,W_{n-1}<W_n,W_n\in[a,b]).$$

证明 令$Y_1=X_n$, $Y_2=X_{n-1},\cdots,Y_n=X_1$ 以及对任意$k=1,\cdots,n$,

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

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

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

分布相同. 因此

$$\begin{aligned}&\mathbb{P}(W_1>0,\cdots,W_{n-1}>0,W_n\in[a,b])=\mathbb{P}(S_1>0,\cdots,S_{n-1}>0,S_n\in[a,b])\\&\quad=\mathbb{P}(W_n-W_{n-1}>0,\cdots,W_n-W_1>0,W_n\in[a,b])\\&\quad=\mathbb{P}(W_1<W_n,\cdots,W_{n-1}<W_n,W_n\in[a,b]).\end{aligned}$$ $\square$

注6.4 对称原理主要利用增量的独立同分布性质(平稳独立增量). 引理6.3中的结论表达式只是对称原理的一种表达式. 根据问题需要, 可以写出不同表示.

(B) 首达时及相关问题

对任意整数i, 令

$$\tau_i=\inf\{n\geq 1;\ W_n=i\}\text{ 以及 }\kappa_i=\inf\{n\geq 0;\ W_n=i\}.$$

直观看, 若$W_0=i$, τ_i表示W首次回到位置i的时间, 称为首回时; κ_i表示W 首次到到i的时间, 一般称其为首达时. 显然若$W_0\neq i$, $\tau_i=\kappa_i$. 由定义可知

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

注意, 若W是初值为0的简单随机游动, $i>0$, 那么

$$\{\tau_i=n\}=\{W_1<W_n,\cdots,W_{n-1}<W_n,W_n=i\}. \tag{2.2.4}$$

记

$$f_0(n)=\mathbb{P}(\tau_0=n|W_0=0),\qquad f_0=\sum_{n=1}^{\infty}f_0(n).$$

那么$f_0(n)$表示从0出发在n时刻首次回到0的概率, 而f_0表示从0出发在有限时间内回到0的概率.

命题6.4 设W是(q,p)简单随机游动, 那么

$$f_0(2k)=\begin{cases}2pq, & k=1,\\ \frac{(2k-3)!!}{2^k}\frac{(4pq)^k}{k!}, & k\geq 1,\end{cases} \text{以及 } f_0=\begin{cases}1, & p=q=1/2,\\ 1-|p-q|, & p\neq q.\end{cases}$$

证明 注意到

$$\begin{aligned}
p_{2k}(0,0)&=\mathbb{P}(W_{2k}=0|W_0=0)=\sum_{m=1}^{k}\mathbb{P}(W_{2k}=0,\tau_0=2m|W_0=0)\\
&=\sum_{m=1}^{k}\frac{\mathbb{P}(W_{2k}=0,\tau_0=2m,W_0=0)}{\mathbb{P}(\tau_0=2m,W_0=0)}\frac{\mathbb{P}(\tau_0=2m,W_0=0)}{\mathbb{P}(W_0=0)}\\
&=\sum_{m=1}^{k}\frac{\mathbb{P}(W_{2k}=0,W_{2m}=0,W_k\neq 0,1\leq k<2m,W_0=0)}{\mathbb{P}(W_{2m}=0,W_k\neq 0,1\leq k<2m,W_0=0)}\mathbb{P}(\tau_0=2m|W_0=0)\\
&=\sum_{m=1}^{k}\mathbb{P}\Big(\sum_{i=2m+1}^{2k}X_i=0\Big)\mathbb{P}(\tau_0=2m|W_0=0)=\sum_{m=1}^{k}p_{2k-2m}(0,0)f_0(2m)
\end{aligned}$$

再令

$$F_0(s)=\sum_{k=1}^{\infty}f_0(2k)s^{2k},\quad s\in[0,1).$$

利用例6.1中定义的$\phi(s)$,

$$\phi(s)=1+\sum_{k=1}^{\infty}\sum_{m=1}^{k}p_{2k-2m}(0,0)f_0(2m)s^{2k}=1+F_0(s)\phi(s).$$

由此可得

$$F_0(s)=1-(1-4pqs^2)^{1/2}=2pqs^2+\sum_{k=2}^{\infty}\frac{(2k-3)!!}{2^k}\frac{(4pq)^k}{k!}s^{2k}.$$

对比$F_0(s)$定义中系数函数可得命题中$f_0(2k)$的值.

另一方面由于W是简单随机游动, 从0出发只能在偶数步才可能回到0, 从而

$$f_0 = \sum_{k=1}^{\infty} f_0(2k).$$

注意到$f_0 \le 1$且$f_0(2k)$非负, 因此$\sum_{k=1}^{\infty} f_0(2k)$ 为正项收敛级数. 从而

$$f_0 = \lim_{s\uparrow 1} F_0(s) = 1 - \sqrt{1-4pq} = \begin{cases} 1, & p = q = 1/2, \\ 1 - |p-q|, & p \ne q. \end{cases} \qquad \square$$

下面我们将简单随机游动稍微扩展. 讨论正步长只能为1, 负步长可以是任意整数, 即

$$\mathbb{P}(X_1 = 1) > 0 \text{ 而且 } \sum_{k=-\infty}^{1} \mathbb{P}(X_1 = k) = 1,$$

的整数值随机游动. 这样的随机游动又常称为不带右跳的随机游动. 我们介绍这类随机游动在位置$i > 0$的首次到达时间的分布. 注意, 此时若W初值为0, (2.2.4)仍然成立.

我们先看一个排列组合方面的引理

引理6.5 对任意正整数m, 在满足$x_1 + \cdots + x_n = m$, 其中$x_i \le 1$, $i = 1, \cdots, n$, 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$(-2, -1, 0, 1)$和$(1, -1, 0, 1)$.

证明* 设$(u_1, u_2, \cdots, u_n)$是方程$x_1 + \cdots + x_n = m$的整数解. 由于$(u_1, u_2, \cdots, u_n)$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$S_1, \cdots, S_n$. 设其最大值为M并令$l = \min\{k, S_k = M\}$. 那么$M \ge m$, $u_l = 1$且对任意$l < i \le n$, $u_{l+1} + \cdots + u_i \le 0$. 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$j < m$, 序列块$(v_{k_j+1}, \cdots, v_{k_{j+1}})$一定满足

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

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

由此我们可以检验$(v_1, v_2, \cdots, v_n)$的循环排列中所有

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

满足引理要求. 其他的任何一个循环

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

都不满足引理要求, 因为此时存在一个j使得$k_j+1 < i \le k_{j+1}$, 该循环的前$n-i+k_j$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. □

定理6.6 设W是初值为0不带右跳的随机游动, 对任意正整数m,

$$\mathbb{P}(W_1 < W_n, \cdots, W_{n-1} < W_n, W_n = m) = \frac{m}{n}\mathbb{P}(W_n = m).$$

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

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

任取A中一个基本事件$u = (u_1, \cdots, u_n)$, 该基本事件的n个循环排列也是A中的基本事件, 记u及其循环排列事件构成的集合为Γ_u, 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, 因此上式只涉及至多可数个事件Γ_u的并. 另一方面, 由引理6.5, 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

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

注意到$B \subset A$, 我们可得

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

因此由全概率公式

$$\mathbb{P}(B)=\mathbb{P}(B\cap\bigcup_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}\Gamma_u)=\sum_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}\mathbb{P}(B\cap\Gamma_u)=\frac{m}{n}\sum_{\substack{\Gamma_u\in A\\ \text{且互不相交}}}\mathbb{P}(A\cap\Gamma_u)=\frac{m}{n}\mathbb{P}(A).$$

所以定理6.6成立. □

由定理6.6以及(2.2.4) 我们可得

推论6.7 设W是初值为0的(q,p)简单随机游动. 对任意$i\geq 1$,

$$\mathbb{P}(\tau_i=n)=\frac{i}{n}\mathbb{P}(W_n=i)=\begin{cases}\frac{i}{n}C_n^{(n+i)/2}p^{(n+i)/2}q^{(n-i)/2}, & \frac{n+i}{2}\leq n\text{且为正整数};\\ 0, & \text{其他}.\end{cases}$$

例6.3(选票问题), 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 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是独立同分布的, 那么

$$\mathbb{P}(X_i=1)=p=\frac{n}{n+m},\ \ \mathbb{P}(X_i=-1)=q=\frac{m}{n+m}.$$

以随机游动$W_k=X_1+\cdots+X_k$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

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

由对称原理及定理6.6

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

从而

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

□

(C) 随机游动应用举例

命题6.8 设W为(q,p)-简单随机游动, 记$\rho=p/q$. 任给两个状态(整数)u,v使得$u<0,v>0$, 从0出发, 在到达v之前先到达u的概率

$$\mathbb{P}(\kappa_u<\kappa_v|W_0=0)=\begin{cases}(1-\rho^v)/(1-\rho^{v-u}), & \rho\neq 1,\\ v/(v-u), & \rho=1.\end{cases}$$

证明 对任意整数$x \le 0 \le y$, 令$g(x,y) = \mathbb{P}(\kappa_x < \kappa_y | W_0 = 0)$, 易知

$$g(0,v) = 1; \qquad g(u,0) = 0; \qquad u < 0, v > 0. \tag{2.2.5}$$

由全概率公式可得

$$\mathbb{P}(\kappa_u < \kappa_v | W_0 = 0) = \mathbb{P}(\kappa_u < \kappa_v, W_1 = 1 | W_0 = 0) + \mathbb{P}(\kappa_u < \kappa_v, W_1 = -1 | W_0 = 0).$$

再由X_n的独立同分布性质可得

$$\begin{aligned}
&\mathbb{P}(\kappa_u < \kappa_v, W_1 = 1 | W_0 = 0) \\
&= \sum_{n > m = 1} \mathbb{P}\Big(W_m = u, W_n = v, W_k \notin \{u, v\}, \\
&\qquad\qquad 1 \le k < m, W_l \neq v, m+1 \le l < n, W_1 = 1 \Big| W_0 = 0\Big) \\
&= \sum_{n > m = 1} \mathbb{P}\Big(W_m = u-1, W_n = v-1, W_k \notin \{u-1, v-1\}, \\
&\qquad\qquad 1 \le k < m, W_l \neq v-1, m+1 \le l < n \Big| W_1 = 0\Big) \mathbb{P}(W_1 = 1 | W_0 = 0) \\
&= p \sum_{n > m = 0} \mathbb{P}\Big(W_m = u-1, W_n = v-1, W_k \notin \{u-1, v-1\}, \\
&\qquad\qquad 0 \le k < m, W_l \neq v-1, m+1 \le l < n \Big| W_0 = 0\Big) \\
&= pg(u-1, v-1)
\end{aligned}$$

同理可知

$$\mathbb{P}(\kappa_u < \kappa_v, W_1 = -1 | W_0 = 0) = qg(u+1, v+1).$$

因此

$$g(u,v) = pg(u-1, v-1) + qg(u+1, v+1).$$

由此可得

$$g(u+1, v+1) - g(u,v) = \rho(g(u,v) - g(u-1, v-1)), \quad u < 0, v > 0$$

反复使用该递推关系式并最后使用(2.2.5)得

$$\begin{aligned}
g(0, v+|u|) &- g(-1, v+|u|-1) \\
&= \rho^{|u|}(g(u,v) - g(u-1, v-1)) \\
&= \rho^{|u|+v-1} g(u-v+1, 1).
\end{aligned}$$

从而由

$$1 = g(0, v+|u|) = \begin{cases} \frac{1-\rho^{|u|+v}}{1-\rho} g(u-v+1, 1), & \rho \neq 1, \\ (|u|+v) g(u-v+1, 1), & \rho = 1; \end{cases}$$

可得

$$g(u,v)=\begin{cases}\frac{1-\rho^v}{1-\rho}g(u-v+1,1)=(1-\rho^v)/(1-\rho^{v-u}), & \rho\neq 1,\\ vg(u-v+1,1)=v/(v-u), & \rho=1.\end{cases}$$

命题证毕. □

下面我们介绍一个随机游动应用在随机试验上的简单例子.

例6.4 甲乙两种新药对某种疾病可能的治愈率分别假定为p_1, p_2. 为判定两种药物的优劣, 我们关心p_1, p_2间的大小关系. 为此给出如下的试验与判别准则: 每次试验治疗两病人, 一人接受甲药治疗, 一人接受乙药治疗. 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. 问这种试验判别的失误率?

分析 不妨假定对药物在不同患者的疗效是独立的. 试验停止标准为累积治愈人数差达到M人.

令$X_i=\begin{cases}1, & 第i对病人中接受\text{A}药治疗者被治愈,\\ 0, & 否则,\end{cases}$

$Y_i=\begin{cases}1, & 第i对病人中接受\text{B}药治疗者被治愈,\\ 0, & 否则.\end{cases}$

那么$\mathbb{P}(X_i=1,Y_i=0)=p_1(1-p_2)$, $\mathbb{P}(X_i=0,Y_i=1)=p_2(1-p_1)$. 由于试验关注的是治愈人数的差, 因此只需要关注引起治愈人数差值改变的试验, 即那些同时治愈或同时无效的试验可以不考虑.

解 令$U_i=X_i-Y_i$, 其中X_i,Y_i如上定义. 那么

$$\mathbb{P}(U_i=1)=\mathbb{P}(X_i=1,Y_i=0|X_i-Y_i\neq 0)=\frac{p_1(1-p_2)}{p_1(1-p_2)+p_2(1-p_1)}\triangleq p,$$

$$\mathbb{P}(U_i=-1)=\mathbb{P}(X_i=0,Y_i=1|X_i-Y_i\neq 0)=\frac{p_2(1-p_1)}{p_1(1-p_2)+p_2(1-p_1)}\triangleq q.$$

若$p_1>p_2$则$p>q$, 此时试验出现失误表现在(q,p)-随机游动

$$W_0=0,\qquad W_n=\sum_{i=1}^n U_i$$

在没有到达M之前先到达$-M$, 因此失误率

$$P=g(-M,M)=\frac{1-\rho^M}{1-\rho^{2M}},\quad \rho=p/q.$$

假设$p_1=0.6$, $p_2=0.5$, $M=10$. 此时$p=0.6$, $q=0.4$, $\rho=1.5$, 失误率

$$P=(1.5^{10}-1)/(1.5^{20}-1)=1/(1.5^{10}+1)\approx 0.017.$$ □

练习题 以下总设$W=\{W_n,n\geq 0\}$为(p,q)-随机游动。

6.1 证明对任意$m \geq 0$以及$i_0, \cdots, i_m \in \mathbf{Z}$,

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

6.2 证明$\mathbb{P}(\kappa_{u+r} < \kappa_{v+r} | W_0 = r) = \mathbb{P}(\kappa_u < \kappa_v | W_0 = 0)$.
6.3 设$u > 0$, 求$\mathbb{P}(\tau_u < \infty | W_0 = 0)$.
6.4 设赌徒在每一赌局中以概率p赢得一个单位财富, 以概率$q = 1 - p$输掉一个单位财富. 假定赌徒财富累积到N时自动离开而财富为0时则输光被迫离开. 若赌徒初始财富为M, $0 < M < N$, 求该赌徒输光的概率.
6.5 某娱乐公司为了吸引顾客推出一种特别活动. 活动规则如下：顾客可以参与一种赌博游戏, 输赢概率各50%. 公司为顾客第一局买单：即赢了的顾客获得一个筹码而输了的顾客则无需付出. 若顾客继续游戏, 那么按正常游戏规则执行, 即赢了的顾客获得一个筹码而输了的顾客付出一个筹码. 假设某个顾客参与了该活动但没有购买任何筹码, 问他至少能玩10局的机会是多少？若他运气够好, 玩了至少10局, 问在第10 局结束时他的筹码平均有多少？

§2.3 泊松(Poisson)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.

(A) 指数分布的有关性质

由初等概率论我们已知随机变量X服从参数为$\lambda > 0$的指数分布, 是指X是正的随机变量, 且分布函数$F(x) = 1 - \mathrm{e}^{-\lambda x}$, 其中$x \geq 0$. X的密度函数为

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

而且$\mathbb{E}(X) = 1/\lambda$, $\mathbf{Var}(X) = 1/\lambda^2$. 我们也常称参数$\lambda$为速率参数.

由初等概率论我们还知指数分布的以下几个性质.

(1) 对任意$n \geq 1$, 若$X_1, \cdots, X_n$是服从参数为λ的指数分布的独立变量, 那么$X_1 + X_2 + \cdots + X_n$服从$\Gamma(n, \lambda)$ 分布. 该分布的密度函数为

$$\gamma_n(x) = \begin{cases} \frac{\lambda^n x^{n-1} \mathrm{e}^{-\lambda x}}{(n-1)!}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

(2) 若$X_1, \cdots, X_n$是分别服从参数为λ_i的指数分布的独立随机变量, 那么$\min\limits_{1 \leq k \leq n} X_k$服从参数为$\lambda_1 + \cdots + \lambda_n$的指数分布.

例7.1 设$X_1, \cdots, X_n$是独立随机变量, 分别服从参数为$\lambda_1, \cdots, \lambda_n$的指数分布. 对任意$1 \leq i \leq n$以及$t > 0$, 求

$$\mathbb{P}(\min_{1 \leq k \leq n} X_k = X_i) \text{ 以及} \mathbb{P}(\min_{1 \leq k \leq n} X_k = X_i > t).$$

解 令$X = \min\limits_{k \neq i} X_k$, 则

$$\min_k X_k = \min\{X, X_i\}.$$

由指数分布性质, X服从速率为$\sum_{k \neq i} \lambda_i$的指数分布. 因此

$$\begin{aligned} \mathbb{P}(\min_{1 \leq k \leq n} X_k = X_i) &= \mathbb{P}(X_i < X) = \int_0^\infty \lambda_i \mathrm{e}^{-\lambda_i x} \mathbb{P}(X > x) \mathrm{d}x \\ &= \int_0^\infty \lambda_i \mathrm{e}^{-\sum_{i=1}^n \lambda_i x} \mathrm{d}x = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}. \\ \mathbb{P}(\min_{1 \leq k \leq n} X_k = X_i > t) &= \mathbb{P}(X_i > t, X_i < X) = \int_t^\infty \lambda_i \mathrm{e}^{-\sum_{i=1}^n \lambda_i x} \mathrm{d}x \\ &= \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n} \mathrm{e}^{-\sum_{i=1}^n \lambda_i t}. \end{aligned}$$

□

注意到

$$\mathbb{P}(\min_{1\le k\le n} X_k = X_i > t) = \mathbb{P}(\min_{1\le k\le n} X_k = X_i)\mathbb{P}(\min_{1\le k\le n} X_k > t).$$

这表明最小值的大小, 与最小值是由哪个随机变量达到这两个随机事件是独立的.

定义7.1 一个非负随机变量X称为是无记忆的, 若对任意$t, s \ge 0$,

$$\mathbb{P}(X > t + s | X > s) = \mathbb{P}(X > t).$$

性质7.1 若X是非负连续型随机变量. X服从指数分布当且仅当X是无记忆的.

证明 只要证充分性. 令$F(t)$为X的分布函数, $H(t) = \ln(1 - F(t))$. 由无记忆性可知对任意$t, s \ge 0$,

$$1 - F(s + t) = (1 - F(s))(1 - F(t)).$$

由此可得

$$H(s + t) = H(s) + H(t).$$

由于F右连续, 因而H右连续, 进而由上方程可知存在正常数λ使得$H(t) = -\lambda t$. 这表明

$$F(t) = 1 - \mathrm{e}^{-\lambda t},$$

即X服从指数分布. □

指数分布的无记忆性还能推广为

性质7.2 设$X_1, \cdots, X_n$是独立随机变量, 分别服从参数为$\lambda_1, \cdots, \lambda_n$的指数分布. 那么在条件$\min\limits_{1\le i\le n} X_i = X_k$ 发生时, $X_i - X_k, i \ne k$ 仍是独立随机变量且分别服从参数为λ_i, $i \ne k$的指数分布.

证明 仅以$k = 1$为例(其他情形类似). 对任意$t_i \ge 0, i = 2, \cdots, n$,

$$\begin{aligned}
&\mathbb{P}(X_2 - X_1 > t_2, \cdots, X_n - X_1 > t_n, \min_k X_k = X_1)\\
&\quad= \mathbb{P}(X_2 - X_1 > t_2, \cdots, X_n - X_1 > t_n)\\
&\quad= \int_0^\infty \lambda_1 \mathrm{e}^{-\lambda_1} x \mathrm{d}x \int_{x+t_2}^\infty \lambda_2 \mathrm{e}^{-\lambda_2 u_2} \mathrm{d}u_2 \cdots \int_{x+t_n}^\infty \lambda_n \mathrm{e}^{-\lambda_n u_n} \mathrm{d}u_n\\
&\quad= \int_0^\infty \lambda_1 \mathrm{e}^{-\sum_{i=1}^n \lambda_i x} \mathrm{d}x \mathrm{e}^{-\sum_{i=2}^\infty \lambda_i t_i} = \frac{\lambda_1}{\lambda_1 + \cdots + \lambda_n} \mathrm{e}^{-\sum_{i=2}^\infty \lambda_i t_i}.
\end{aligned}$$

由例7.1可知

$$\mathbb{P}(X_2 - X_1 > t_2, \cdots, X_n - X_1 > t_n | \min_k X_k = X_1) = \mathrm{e}^{-\sum_{i=2}^\infty \lambda_i t_i} = \prod_{i=2}^\infty \mathbb{P}(X_i > t_i).$$

这表明性质7.2成立. □

例7.2 有两台机器生产同一种产品, 生产产品所需时间分布服从速率为λ_1和λ_2的指数分布而且每件产品所用时间是独立的. 现有三件产品要生产, 先在每台机器上各

生产一件产品, 先生产完产品的机器紧接着生产第三件产品. 求第三件产品在最后生产出来的概率.

解 令$X_i, i = 1,2,3$, 分别表示第i件产品生产所需时间. X_1服从参数为λ_1的指数分布, X_2服从参数为λ_2的指数分布. 当$X_1 < X_2$时, 第三个产品在第一台机器上生产, X_3服从参数为λ_1的指数分布, 且与X_1, X_2独立. 当$X_1 > X_2$时, 第三个产品在第二台机器上生产, X_3服从参数为λ_2的指数分布, 且与X_1, X_2独立. 因此所求概率可以表示成

$$p = \mathbb{P}(X_3 + X_1 > X_2, X_1 < X_2) + \mathbb{P}(X_3 + X_2 > X_1, X_2 < X_1).$$

注意到

$$\begin{aligned}\mathbb{P}(X_3 + X_1 > X_2, X_1 < X_2) &= \mathbb{P}(X_3 + X_1 > X_2 | X_1 < X_2)\mathbb{P}(X_1 < X_2)\\ &= \mathbb{P}(X_3 > X_2 - X_1 | X_1 < X_2)\mathbb{P}(X_1 < X_2).\end{aligned}$$

由例7.1以及性质7.2可知

$$\mathbb{P}(X_3 > X_2 - X_1 | X_1 < X_2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}, \ \ \mathbb{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

因此

$$\mathbb{P}(X_3 + X_1 > X_2, X_1 < X_2) = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2}.$$

类似可同样得

$$\mathbb{P}(X_3 + X_2 > X_1, X_2 < X_1) = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2}.$$

因此所求概率$p = 2\lambda_1\lambda_2/(\lambda_1 + \lambda_2)^2$. □

(B) 计数过程

用数学模型刻画随机现象时, 一种简单而有效的工作就是记录所关心的随机事件发生的次数. 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7.1 非负整数值的随机过程$N = \{N(t); t \geq 0\}$称为计数过程或点过程, 若其样本函数为右连续的单调不降函数. 若还有$\mathbb{P}(N(t) - N(t-) \leq 1$对任意$t \geq 0) = 1$, 则称$N$ 是简单计数过程, 其中$N(t-)$表示N在t点的左极限.

直观地看简单计数过程要求在同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$k \geq 1$,

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

有意义, 其中$T_0 \equiv 0$并约定$\inf \emptyset = \infty$. 由定义可知,

(1) T_k就是第k次事件发生的随机时刻, 通常称之为第k次(随机事件)到达时间.

(2) $N(t)$是时刻t之前(包含t)发生的随机事件数, 第$N(t)$次事件必须发生在时刻t之前(含t), 因此$T_{N(t)} \le t$, 即对任意ω,

$$T_{N(t,\omega)}(\omega) \le t.$$

(3) 若$T_k < \infty$, 那么$N(T_k) \ge k$, 即对任意ω,

$$N(T_k(\omega), \omega) \ge k.$$

特别N 为简单计数过程时$N(T_k) = k$.

综上可得

$$N(t) = \sup\{k \ge 0; T_k \le t\}, \tag{2.3.1}$$

即对任意ω, $N(t,\omega) = \sup\{k; T_k(\omega) \le t\}$, 其中约定$\sup\emptyset = 0$.

与随机游动一样, 我们也可以从随机变量部分和序列出发认识计数过程.

设$S = \{S_k, k \ge 0\}$是非负随机变量$X_i, i \ge 0$, 的部分和序列. 对这样的部分和序列, 在应用和理论研究中经常碰到的一个问题是: 在已知部分和的条件下, 估计具体涉及的求和项数k或相关问题?

注意到, 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$S^{-1}: D = \{x_k;\ x_k = S_k, k \ge 0\} \mapsto \mathbb{N}$使得$S^{-1}(x_k) = k$. 进一步, 这个逆映射还可推广到非负实数$\bar{\mathbb{R}}_+$上, 使得对任意$x \in \bar{\mathbb{R}}_+$, 当$x_k \le x < x_{k+1}$时$S^{-1}(x) = k$, 即

$$S^{-1}(x) = \sup\{k \ge 0; S_k \le x\}. \tag{2.3.2}$$

如果X_i可以为0, 那么S_k作为k的函数不一定可逆, 但(2.3.2)定义的函数$S^{-1}(x)$仍有意义, 而且只要S_k是非负随机变量的部分和序列, $S^{-1}(x)$就是使部分和$S_k \le x$成立的最大求和项数. 通常, 我们称过程

$$S^{-1} = \{S^{-1}(t), t \ge 0\}$$

为部分和序列$\{S_k, k \ge 0\}$的逆过程.

对计数过程N而言, 令$W_k = T_k - T_{k-1}$. 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, 一般地我们称其为第k 个等待时间. 此时第k 个事件发生的时间

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

为序列$\{W_i, i \ge 1\}$的部分和. 由(2.3.1)可知, 计数过程N就是部分和序列$\{T_k, k \ge 0\}$的逆过程.

(C) 泊松过程及其刻画

下面我们研究一类特殊的计数过程, 在这个过程里, 我们要求部分和序列是独立同分布的指数随机变量之和. 我们称这样的计数过程为泊松(Poisson)过程.

定义7.2 称计数过程$N=\{N(t);\ t\geq 0\}$是强度为λ的泊松过程, 若N具有独立同分布的等待时间序列$\{W_k;\ k\geq 1\}$且等待时间分布服从速率为λ的指数分布.

例7.3 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. 以$N(t)$表示在$[0,t]$ 时间内更换下来的灯泡数. 若假定灯泡的寿命是独立同分布的且服从指数分布，那么N就是泊松过程.

以$W_k, k=1,2,\cdots$表示泊松过程的等待时间变量簇, 它们是独立且服从参数为λ的指数分布的随机变量列. 对任意$t\geq 0$, 泊松过程

$$N(t)=\sup\{k\geq 0;\sum_{n=1}^{k}W_n\leq t\},\quad t\geq 0,$$

其中规定$\sum_{n=1}^{0}=0$. 对任意$k\geq 1$, 第k次事件发生时刻$T_k=\sum_{i=1}^{k}W_i$服从$\Gamma(k,\lambda)$分布, 概率密度函数为$\gamma_k(x)$.

由于指数分布的随机变量只取正数, 而且$\{N(0)>0\}\subset\{W_1=0\}$, 因此

$$\mathbb{P}(N(0)=0)=1-\mathbb{P}(N(0)\geq 1)\geq 1-\mathbb{P}(W_1=0)=1,$$

而且

$$\begin{aligned}\mathbb{P}(N(t)-N(t-)\geq 2\text{对某个}t\geq 0\text{成立}) &= \mathbb{P}(\text{存在某个}k\text{使得}W_k=0)\\ &= \bigcup_{k=1}^{\infty}\mathbb{P}(W_k=0)=0.\end{aligned}$$

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.

命题7.3 强度为λ的泊松过程在任何一个时刻$t>0$都服从强度为λt的泊松分布. 因此

$$\mathbb{E}(N(t))=\lambda t,\quad \mathbb{V}ar(N(t))=\lambda t.$$

证明 对任意给定的$t>0$和非负整数k, 由(2.3.1)可知

$$\{N(t)=k\}=\{T_k\leq t<T_{k+1}\},\tag{2.3.3}$$

因此

$$\begin{aligned}\mathbb{P}\big(N(t)=k\big) &= \mathbb{P}(T_k\leq t<T_{k+1})=\mathbb{P}(T_k\leq t)-\mathbb{P}(T_{k+1}\leq t)\\ &= \int_0^t[\gamma_k(x)-\gamma_{k+1}(x)]\mathrm{d}x=\int_0^t\Big[\frac{\lambda^k x^{k-1}}{(k-1)!}-\frac{\lambda^{k+1}x^k}{k!}\Big]\mathrm{e}^{-\lambda x}\mathrm{d}x\end{aligned}$$

由分部积分可得

$$\mathbb{P}\big(N(t)=k\big)=\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}.$$

这表明$N(t)$服从强度为λt的泊松分布. □

定理7.4 强度为λ的泊松过程是一个独立增量过程且对任意$t, s \geq 0$, $N(t+s)-N(s)$服从参数为λt的泊松分布, 即对任意整数$k \geq 0$,

$$\mathbb{P}(N(s+t)-N(s)=k)=\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}. \tag{2.3.4}$$

证明* 只要证明对任意$n \geq 1$以及$0=t_0<t_1<\cdots<t_n$和非负整数$k_1, \cdots, k_n$,

$$\mathbb{P}\big(N(t_r)=\sum_{i=1}^{r} k_i, r=1,2,\cdots,n\big)=\mathrm{e}^{-\lambda t_n}\prod_{i=1}^{n}\frac{[\lambda(t_i-t_{i-1})]^{k_i}}{k_i!}. \tag{2.3.5}$$

为了表示简便, 记N的等待时间序列为$\{W_r\}$, 事件发生时刻为$\{T_r\}$. 对任意$r \geq 1$, 令$s_r=\sum_{i=1}^r k_i$以及

$$A_r=\{T_{s_1}\leq t_1, T_{s_1+1}>t_1, \cdots, T_{s_{r-1}}\leq t_{r-1}, T_{s_{r-1}+1}>t_{r-1}\},$$

$$B_r=\{T_{s_r}\leq t_r, T_{s_r+1}>t_r\},$$

$$C_r=\{T_{s_1}\leq t_1, T_{s_1+1}>t_1, \cdots, T_{s_{r-1}}\leq t_{r-1}\}.$$

利用等式(2.3.3), (2.3.5)可简写为

$$\mathbb{P}\big(A_{n+1}\big)=\mathrm{e}^{-\lambda t_n}\prod_{i=1}^{n}\frac{[\lambda(t_i-t_{i-1})]^{k_i}}{k_i!}. \tag{2.3.6}$$

当$n=1$时由命题7.3可知(2.3.6)显然成立. 下设$n=m-1$时成立, 即

$$\mathbb{P}\big(A_m\big)=\mathrm{e}^{-\lambda t_{m-1}}\prod_{i=1}^{m-1}\frac{[\lambda(t_i-t_{i-1})]^{k_i}}{k_i!}.$$

当$n=m$时, 注意到$A_{m+1}=A_m B_m$, 由归纳假设,

$$\begin{aligned}\mathbb{P}(A_{m+1}) &= \mathbb{P}(A_m)\mathbb{P}(B_m|A_m)\\ &= \mathbb{P}(B_m|A_m)\mathrm{e}^{-\lambda t_{m-1}}\prod_{i=1}^{m-1}\frac{[\lambda(t_i-t_{i-1})]^{k_i}}{k_i!}.\end{aligned} \tag{2.3.7}$$

下面我们分情况讨论.
(1)当$k_m=0$时, $s_m=s_{m-1}$, $B_m=\{T_{s_{m-1}}<t_m, T_{s_{m-1}+1}>t_m\}$.

$$\begin{aligned}\mathbb{P}(B_m|A_m) &= \frac{\mathbb{P}(A_m B_m)}{\mathbb{P}(A_m)}=\frac{\mathbb{P}(C_m, T_{s_{m-1}+1}>t_m)}{\mathbb{P}(C_m, T_{s_{m-1}+1}>t_{m-1})}\\ &= \frac{\mathbb{E}(\mathbf{1}_{C_m}\mathbb{P}(W_{s_{m-1}+1}>t_m-T_{s_{m-1}}|T_1,\cdots,T_{s_{m-1}}))}{\mathbb{E}(\mathbf{1}_{C_m}\mathbb{P}(W_{s_{m-1}+1}>t_{m-1}-T_{s_{m-1}}|T_1,\cdots,T_{s_{m-1}}))}\\ &= \frac{\mathbb{E}\big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_m-T_{s_{m-1}})}\big)}{\mathbb{E}\big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_{m-1}-T_{s_{m-1}})}\big)}=\mathrm{e}^{-\lambda(t_m-t_{m-1})}.\end{aligned} \tag{2.3.8}$$

(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

$$\mathbb{P}(B_m|A_m) = \frac{\mathbb{P}(A_mB_m)}{\mathbb{P}(A_m)} = \frac{\mathbb{P}(C_m, t_{m-1} < T_{s_m} < t_m, T_{s_m+1} > t_m)}{\mathbb{P}(C_m, T_{s_{m-1}+1} > t_{m-1})}.$$

由$\{T_k\}$的独立增量性, 以及对任意$k > l$,

$$T_k - T_l \sim \gamma_{k-l}(x) = \frac{\lambda^{k-l}}{(k-l-1)!}x^{k-l-1}\mathrm{e}^{-\lambda x}, \quad x > 0, \tag{2.3.9}$$

与(2.3.8)类似处理可得

$$\begin{aligned}
&\mathbb{P}(C_m, t_{m-1} < T_{s_m} < t_m, T_{s_m+1} > t_m) \\
&= \mathbb{E}\Big(\mathbf{1}_{C_m}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}} \lambda\mathrm{e}^{-\lambda s}\mathrm{d}s\int_{t_m-T_{s_{m-1}}-s}^{\infty}\lambda\mathrm{e}^{-\lambda u}\mathrm{d}u\Big) \\
&= \mathbb{E}\big(\mathbf{1}_{C_m}\lambda(t_m - t_{m-1})\mathrm{e}^{-\lambda(t_m - T_{s_{m-1}})}\big) \\
&= \lambda(t_m - t_{m-1})\mathrm{e}^{\lambda(t_m - t_{m-1})}\mathbb{E}\big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_{m-1} - T_{s_{m-1}})}\big)
\end{aligned}$$

因此$k_m = 1$时,

$$\begin{aligned}
\mathbb{P}(B_m|A_m) &= \frac{\lambda(t_m - t_{m-1})\mathrm{e}^{\lambda(t_m - t_{m-1})}\mathbb{E}\big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_{m-1} - T_{s_{m-1}})}\big)}{\mathbb{E}\big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_{m-1} - T_{s_{m-1}})}\big)} \\
&= \lambda(t_m - t_{m-1})\mathrm{e}^{\lambda(t_m - t_{m-1})}.
\end{aligned} \tag{2.3.10}$$

(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

$$\begin{aligned}
\mathbb{P}(A_mB_m) &= \mathbb{P}(C_m \cap \{T_{s_{m-1}+1} > t_{m-1}\} \cap B_m) \\
&= \mathbb{P}(C_m, t_m - T_{s_{m-1}} > T_{s_{m-1}+1} - T_{s_{m-1}} > t_{m-1} - T_{s_{m-1}}, \\
&\qquad T_{s_m} - T_{s_{m-1}+1} < t_m - T_{s_{m-1}+1}, T_{s_m+1} - T_{s_m} > t_m - T_{s_m}).
\end{aligned}$$

再次应用$\{T_k\}$的独立增量性以及(2.3.9)式可得, 上式等于

$$\begin{aligned}
&\mathbb{E}\Big(\mathbf{1}_{C_m}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\lambda\mathrm{e}^{-\lambda s}\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\gamma_{k_m-1}(u)\mathrm{d}u\int_{t_m-T_{s_{m-1}}-s-u}^{\infty}\lambda\mathrm{e}^{-\lambda v}\mathrm{d}v\Big) \\
&= \mathbb{E}\Big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_m - T_{s_{m-1}})}\int_{t_{m-1}-T_{s_{m-1}}}^{t_m-T_{s_{m-1}}}\mathrm{d}s\int_0^{t_m-s-T_{s_{m-1}}}\frac{\lambda^{k_m}u^{k_m-2}}{(k_m-2)!}\mathrm{d}u\Big) \\
&= \mathbb{E}\Big(\mathbf{1}_{C_m}\mathrm{e}^{-\lambda(t_m - T_{s_{m-1}})}\frac{\lambda^{k_m}(t_m - t_{m-1})^{k_m}}{k_m!}\Big).
\end{aligned}$$

因此$k_m > 1$时,

$$\mathbb{P}(B_m|A_m) = \frac{\mathbb{P}(A_mB_m)}{\mathbb{P}(A_m)} = \frac{\lambda^{k_m}(t_m - t_{m-1})^{k_m}}{k_m!}\mathrm{e}^{-\lambda(t_m - t_{m-1})}. \tag{2.3.11}$$

分别将(2.3.8), (2.3.10)和(2.3.11)代入(2.3.7)可知不论k_m取什么值, (2.3.6)对$n=m$也成立. 由归纳法原理, (2.3.6)对一切$n \geq 1$ 都成立, 因而定理得证. □

定理7.5 不恒为0的计数过程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.

记N的等待时间序列为$\{W_k;\ k \geq 1\}$, 令$T_k = \sum_{i=1}^k W_i,\ k \geq 1$. 为证明定理7.5，我们需要如下的引理.

引理7.6 $\{W_k;\ k \geq 1\}$ 为服从参数为λ的指数分布的独立同分布随机变量列当且仅当对任意$k \geq 1$及任意$0 < t_1 \leq t_2 \leq \cdots \leq t_n$,

$$\mathbb{P}(T_k \leq t_k, k=1,2,\cdots,n) = \int_0^{t_1} \mathrm{d}s_1 \int_{s_1}^{t_2} \mathrm{d}s_2 \cdots \int_{s_{n-1}}^{t_n} \lambda^n \mathrm{e}^{-\lambda s_n} \mathrm{d}s_n. \quad (2.3.12)$$

证明 由数学归纳法容易证明必要性，留作习题. 充分性由例1.2可得, 此略. □

引理7.7 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$k_0 < k_1 < \cdots < k_{n+1}$,

$$\begin{aligned}&\mathbb{P}(N(t_{n+1}) \geq k_{n+1} | N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \leq m \leq n)\\ &\qquad = \mathbb{P}(N(t_{n+1} - t_n) \geq k_{n+1} - k_n).\end{aligned}$$

证明 由于计数过程单调不降, 对任意$t \geq 0$, $N(t-) = \lim_{s \uparrow t} N(s)$. 因此,

$$\begin{aligned}&\mathbb{P}(N(t_{n+1}) \geq k_{n+1}, N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \leq m \leq n)\\ &= \lim_{\varepsilon \to 0} \mathbb{P}(N(t_{n+1}) \geq k_{n+1}, N(t_m) = k_m, N(t_m - \varepsilon) = k_m - 1, 0 \leq m \leq n).\end{aligned}$$

不妨设$0 < \varepsilon < \min_{1 \leq k \leq n}\{t_k - t_{k-1}\}$, 由平稳独立增量性可得

$$\begin{aligned}&\mathbb{P}(N(t_{n+1}) \geq k_{n+1}, N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \leq m \leq n)\\ &= \mathbb{P}(N(t_{n+1}) - N(t_n) \geq k_{n+1} - k_n)\\ &\qquad \times \lim_{\varepsilon \to 0} \mathbb{P}(N(t_m) = k_m, N(t_m - \varepsilon) = k_m - 1, 0 \leq m \leq n).\\ &= \mathbb{P}(N(t_{n+1} - t_n) \geq k_{n+1} - k_n) \mathbb{P}(N(t_m) = k_m, N(t_m-) = k_m - 1, 0 \leq m \leq n).\end{aligned}$$

由此可知引理成立. □

定理7.5的证明 必要性由定理7.4可得. 下证充分性. 为此, 我们需证明$W_k, k \geq 1$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7.6, 我们只需证明(2.3.12)对任意$n \geq 1$成立.

当$n=1$时, 对任意$t \geq 0$, 注意到

$$\mathbb{P}(T_1 \leq t) = \mathbb{P}(N(t) \geq 1) = 1 - \mathbb{P}(N(t) = 0)$$

令$f(t) = \mathbb{P}(N(t) = 0)$. 对任意$s, t > 0$, 由平稳独立增量性以及条件$N(0) = 0$可得

$$f(s+t) = \mathbb{P}(N(s+t) = 0) = \mathbb{P}(N(s+t) = 0, N(t) = 0) = f(s)f(t).$$

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$\mathbb{P}(N(t+s) - N(s) \ge 1) = 1 - f(t) = 1$, 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

$$\mathbb{P}(N(1) - N(s) \ge 2) = 1.$$

令$s \to 1$得

$$\mathbb{P}(N(1) - N(1-) \ge 2) = 1.$$

这与N为简单计数过程矛盾, 因此

$$\mathbb{P}(T_1 \le t) = \mathbb{P}(N(t) \ge 1) = \int_0^t \lambda \mathrm{e}^{-\lambda s} \mathrm{d}s, \tag{2.3.13}$$

即$n = 1$时(2.3.12)成立.

下设$n = m - 1$时(2.3.12)也成立, 那么$n = m$时,由归纳假设

$$\begin{aligned}\mathbb{P}(T_k \le t_k, 1 \le k \le m) = & \int_0^{t_1} \mathrm{d}s_1 \cdots \int_{s_{m-2}}^{t_{m-1}} \lambda^{m-1} \mathrm{e}^{-\lambda s_{m-1}} \\ & \times \mathbb{P}(T_m \le t_m | T_k = s_k, 1 \le k < m) \mathrm{d}s_{m-1}.\end{aligned} \tag{2.3.14}$$

由于N是初值为0的简单计数过程, 具有平稳独立增量, 由引理7.7和(2.3.13)可知

$$\begin{aligned}&\mathbb{P}(T_m \le t_m | T_k = s_k, 1 \le k < m) \\ &= \mathbb{P}(N(t_m) \ge m | N(s_k) = k, N(s_k-) = k-1, 1 \le k < m) \\ &= \mathbb{P}(N(t_m - s_{m-1}) \ge 1) \\ &= \int_0^{t_m - s_{m-1}} \lambda \mathrm{e}^{-\lambda s_m} \mathrm{d}s_m = \int_{s_{m-1}}^{t_m} \lambda \mathrm{e}^{-\lambda(s_m - s_{m-1})} \mathrm{d}s_m.\end{aligned}$$

将其代入(2.3.14)整理后即得(2.3.12). □

由定理7.4和7.5可知不恒为0的计数过程N是强度为λ的泊松过程当且仅当N是初值为0简单计数过程且具有平稳独立增量, 增量$N(t+s) - N(s)$服从参数为λt的泊松分布.

例7.4 假设某银行顾客按强度为λ的泊松过程N到达, 已知T时刻到达银行的顾客数是10. 问(1)再经过时间T, 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$c = \mathbb{E}(N(2T) - N(T))$. 由平稳独立增量性可知

$$c = \mathbb{E}(N(T)) = \lambda T.$$

(2)对任意t, 令$\tilde{N}(t)=N(t+T)-N(T)$. 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$t=\mathbb{E}(T_5)$. 由T_5服从$\Gamma(5,\lambda)$分布可知$t=5/\lambda$. □

(D) 到达时间的条件分布及其应用

定理7.8 设$\{T_n, n\geq 1\}$是强度为λ的泊松过程$N=\{N(t), t\geq 0\}$的到达时刻序列, 那么对任意$0\leq s_1<t_1<\cdots\leq s_n<t_n\leq t$,

$$\mathbb{P}(T_j\in(s_j,t_j), 1\leq j\leq n|N(t)=n)=\frac{n!}{t^n}\prod_{j=1}^{n}(t_j-s_j).$$

证明 任取$0=t_0\leq s_1<t_1\leq s_2<\cdots\leq s_n<t_n\leq t$,

$$\begin{aligned}&\mathbb{P}(T_j\in(s_j,t_j], 1\leq j\leq n|N(t)=n)\\&=\mathbb{P}\big(N(t_j)-N(s_j)=1, N(s_j)-N(t_{j-1})=0,\\&\qquad\qquad 1\leq j\leq n, N(t)-N(t_n)=0\big|N(t)=n\big)\\&=\prod_{j=1}^{n}[\lambda(t_j-s_j)\mathrm{e}^{-\lambda(t_j-t_{j-1})}]\mathrm{e}^{-\lambda(t-t_n)}/[\frac{(\lambda t)^n}{n!}\mathrm{e}^{-\lambda t}]=\frac{n!}{t^n}\prod_{j=1}^{n}(t_j-s_j).\end{aligned}$$

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

$$p(x_1,\cdots,x_n)=n!t^{-n},\quad 0<x_1<x_2<\cdots<x_n<t,$$

从而对任意$0\leq s_1<t_1<\cdots\leq s_n<t_n\leq t$, 概率

$$\mathbb{P}(X_{(j)}\in(s_j,t_j), 1\leq j\leq n)=\frac{n!}{t^n}\prod_{j=1}^{n}(t_j-s_j).$$

比较上式与定理7.8结论，我们可知在分布意义下

$$(T_1,T_2,\cdots,T_n|N(t)=n)\stackrel{d}{=}(X_{(1)},X_{(2)},\cdots,X_{(n)}),$$

即给定条件$N(t)=n$后, $T_1,\cdots,T_n$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7.5 (接例7.4)求第10位顾客到达的时间分布和平均时间.

解 由定理7.8, 在条件$N(T)=10$下, 第10个顾客到达时间T_{10}与10个独立的$(0,T]$上均匀分布随机变量的最大次序统计量相同. 因此T_{10}的密度函数为

$$f(x)=10\frac{x^9}{T^{10}}.$$

平均时间

$$\mathbb{E}(T_{10}) = \int_0^T xf(x)\mathrm{d}x = 10T\int_0^1 x^{10}\ dx = \frac{10T}{11}. \qquad \square$$

作为定理7.8的推论, 我们有如下重要结论.

推论7.9 设n元函数$f(x_1, x_2\cdots, x_n)$满足对于任意一个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

$$\mathbb{E}(f(T_1, T_2, \cdots, T_n)|N(t) = n) = \mathbb{E}(f(X_1, X_2, \cdots, X_n)),$$

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7.8可知

$$\mathbb{E}(f(T_1, T_2, \cdots, T_n)|N(t) = n) = \mathbb{E}(f(X_{(1)}, X_{(2)}, \cdots, X_{(n)}).$$

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

$$\mathbb{E}(f(X_{(1)}, X_{(2)}, \cdots, X_{(n)}) = \mathbb{E}(f(X_1, \cdots, X_n)),$$

进而结论成立. $\square$

例7.6 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

$$\begin{aligned}\bar{T} &= \mathbb{E}(T) = \mathbb{E}(\sum_{i=1}^{N(t)} (t - T_i)) = \mathbb{E}\Big(\mathbb{E}(\sum_{i=1}^{N(t)} (t - T_i)|N(t))\Big) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(\sum_{i=1}^{n} (t - T_i)|N(t) = n)\mathbb{P}(N(t) = n)\end{aligned}$$

显然改变T_i次序不会影响$\sum_{i=1}^n (t - T_i)$的结果, 由推论7.10,

$$\begin{aligned}\bar{T} &= \sum_{n=1}^{\infty} \mathbb{E}\Big(\sum_{i=1}^{n} (t - X_i)\Big)\mathbb{P}(N(t) = n) = \sum_{n=1}^{\infty} \Big(nt - \mathbb{E}(\sum_{i=1}^{n} X_i)\Big)\mathbb{P}(N(t) = n) \\ &= \sum_{n=1}^{\infty} \frac{nt}{2}\mathbb{P}(N(t) = n) = \frac{t}{2}\mathbb{E}(N(t)) = \frac{\lambda t^2}{2}. \qquad \square\end{aligned}$$

例7.7* 设泊松过程N的强度为λ, T_k为到达时刻, $k\geq 1$. 若$\int_0^\infty |f(t)|\mathrm{d}t<\infty$, 证明

$$\mathbb{E}(\sum_{n=1}^\infty f(T_n))=\lambda\int_0^\infty f(t)\mathrm{d}t.$$

证明 先设f非负. 此时由单调收敛定理(定理3.14)

$$\begin{aligned}\mathbb{E}(\sum_{n=1}^\infty f(T_n)) &= \sum_{n=1}^\infty \mathbb{E}(f(T_n))=\sum_{n=1}^\infty\int_0^\infty f(t)\frac{\lambda^n t^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t}\mathrm{d}t\\ &= \int_0^\infty f(t)\sum_{n=1}^\infty\frac{\lambda^n t^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda t}\mathrm{d}t=\lambda\int_0^\infty f(t)\mathrm{d}t.\end{aligned}$$

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

$$\mathbb{E}\big(\sum_{n=1}^\infty |f(T_n)|\big)<\infty,\quad \mathbb{E}\big(\sum_{n=1}^\infty f^+(T_n)\big)<\infty,\quad \mathbb{E}\big(\sum_{n=1}^\infty f^-(T_n)\big)<\infty.$$

这表明级数$\sum\limits_{n=1}^\infty f(T_n)$, $\sum\limits_{n=1}^\infty f^+(T_n)$ 和$\sum\limits_{n=1}^\infty f^-(T_n)$ 都几乎处处收敛. 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

$$\begin{aligned}\mathbb{E}\big(\sum_{n=1}^\infty f(T_n)\big) &= \mathbb{E}\big(\sum_{n=1}^\infty f^+(T_n)\big)-\mathbb{E}\big(\sum_{n=1}^\infty f^-(T_n)\big)\\ &= \lambda\int_0^\infty[f^+(t)-f^-(t)]\mathrm{d}t=\lambda\int_0^\infty f(t)\mathrm{d}t.\end{aligned}$$ □

(E) 稀疏过程

对于计数过程$N=\{N(t)\}$记录的随机事件, 有时我们需要对随机事件进行分类, 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, 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的稀疏过程. 计数过程N与它的稀疏过程N_i之间满足: 对任意t,

$$N(t)=\sum_{i=1}^m N_i(t).$$

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$p_1, p_2, \cdots, p_m$为发生概率的多元二项分布(参见例1.1), 而且与其它不相交区间发生的随机事件无关. 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

$$\begin{aligned}\mathbb{P}(N_i(t) - N_i(s) = k_i, 1 \le i \le m | N(t) - N(s) = k)\\ = \frac{k!}{k_1!k_2!\cdots k_m!} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m},\end{aligned} \tag{2.3.15}$$

而且对任意$0 = t_0 < t_1 < \cdots < t_n$, 以及非负整数$k_{ij}, 1 \le i \le m, 1 \le j \le n$,

$$\begin{aligned}&\mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j, 1 \le i \le m, 1 \le j \le n | N(t_l) - N(t_{l-1}), l = 1, \cdots, n)\\ &= \prod_{j=1}^n \mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j, 1 \le i \le m | N(t_l) - N(t_{l-1}), l = 1, \cdots, n)\\ &= \prod_{j=1}^n \mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j, 1 \le i \le m | N(t_j) - N(t_{j-1})).\end{aligned} \tag{2.3.16}$$

定理7.10 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(2.3.16) 可知

$$\begin{aligned}&\mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j, j = 1, \cdots, n)\\ &= \mathbb{E}[\mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j, j = 1, \cdots, n | N(t_l) - N(t_{l-1}), l = 1, \cdots, n)]\\ &= \mathbb{E}[\prod_{j=1}^n \mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j | N(t_j) - N(t_{j-1}))]\\ &= \prod_{j=1}^n \mathbb{E}[\mathbb{P}(N_i(t_j) - N_i(t_{j-1}) = k_j | N(t_j) - N(t_{j-1}))]\\ &= \prod_{j=1}^n \mathbb{P}(N_i(t_j) - N(t_{j-1}) = k_j).\end{aligned}$$

因此N_i具有独立增量. 注意到对任意$t > s \ge 0$以及非负整数k,

$$\begin{aligned}\mathbb{P}(N_i(t) - N_i(s) = k) &= \sum_{r=0}^\infty \mathbb{P}(N_i(t) - N_i(s) = k, N(t) - N(s) = k + r)\\ &= \sum_{r=0}^\infty \mathbb{P}(N_i(t) - N_i(s) = k | N(t) - N(s) = k + r) \mathbb{P}(N(t) - N(s) = k + r).\end{aligned}$$

由(2.3.15)得

$$\begin{aligned}\mathbb{P}(N_i(t)-N_i(s)=k) &= \sum_{r=0}^{\infty} C_{k+r}^k p_i^k (1-p_i)^r \frac{[\lambda(t-s)]^{k+r}}{(k+r)!} \mathrm{e}^{-\lambda(t-s)} \\ &= \frac{[p_i\lambda(t-s)]^k \mathrm{e}^{-\lambda(t-s)}}{k!} \sum_{r=0}^{\infty} \frac{[\lambda(1-p_i)(t-s)]^r}{r!} \\ &= \frac{[p_i\lambda(t-s)]^k \mathrm{e}^{-\lambda(t-s)}}{k!} \mathrm{e}^{\lambda(1-p_i)(t-s)} = \frac{[p_i\lambda(t-s)]^k}{k!} \mathrm{e}^{-\lambda p_i(t-s)}\end{aligned}$$

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7.5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(2.3.16)可得

$$\begin{aligned}&\mathbb{P}(N_i(t_j)-N_i(t_{j-1})=k_{i,j}, 1\leq i\leq m, 1\leq j\leq n) \\ &\quad= \mathbb{P}(N_i(t_j)-N_i(t_{j-1})=k_{i,j}, 1\leq i\leq m, 1\leq j\leq n|A)\mathbb{P}(A) \\ &\quad= \mathbb{P}(A)\prod_{j=1}^n \mathbb{P}(N_i(t_j)-N_i(t_{j-1})=k_{i,j}, 1\leq i\leq m|N(t_j)-N(t_{j-1})=k_j).\end{aligned}$$

再由(2.3.15)和N的独立增量泊松分布可知

$$\begin{aligned}&\mathbb{P}(N_i(t_j)-N_i(t_{j-1})=k_{i,j}, 1\leq i\leq m, 1\leq j\leq n) \\ &\quad= \prod_{j=1}^n \Big[\frac{k_j!}{k_{1,j}!k_{2,j}!\cdots k_{m,j}!} p_1^{k_{1,j}}\cdots p_m^{k_{m,j}} \frac{[\lambda(t_j-t_{j-1})]^{k_j}}{k_j!} \mathrm{e}^{-\lambda(t_j-t_{j-1})}\Big] \\ &\quad= \prod_{i=1}^m \Big[\prod_{j=1}^n \frac{[\lambda p_i(t_j-t_{j-1})]^{k_{i,j}}}{k_{i,j}!} \mathrm{e}^{-\lambda p_i(t_j-t_{j-1})}\Big] \\ &\quad= \prod_{i=1} \mathbb{P}(N_i(t_j)-N_i(t_{j-1})=k_{i,j}, 1\leq j\leq n).\end{aligned}$$

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(参见习题5-3)可知稀疏过程$N_i, i=1,\cdots,m$独立. □

例7.8 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10%为男性, 90%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)已知一小时内有4个男顾客到达条件下, 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$p=\mathbb{P}(N_1(0.5)\geq 1)=1-\mathbb{P}(N_1(0.5)=0)=1-\mathrm{e}^{-1}$.

(2)所求期望为$n=4+\mathbb{E}(N_2(1))=4+18=22$. □

(F) 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. 泊松过程的一种推广是所谓非齐次泊松过程, 它的定义如下：

定义7.3 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, 若她满足:

(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.

由定义可知对非齐次的泊松过程, 在任意时刻t, $N(t)$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

$$\mathbb{P}(T_1>t)=\mathbb{P}(N(t)=0)=\mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u},$$

从而密度函数为

$$f(t)=\lambda(t)\mathrm{e}^{-\int_0^t\lambda(u)\mathrm{d}u},\quad t\geq 0.$$

例7.9 已知某网店在不同时间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. 问在$11:00-12:00$没有顾客访问的概率？

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$0<t\leq 0.5$时间段内$N(t)$的强度为10, 在$0.5<t\leq 1$时间段内强度为6. 因此所求概率$p=\mathrm{e}^{-8}$. □

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7.10得:

定理7.11 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, 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, (2.3.16)仍然成立.

对任意$0\leq s<t$和非负整数n, 任取非负整数$n_1,\cdots,n_m$使得$n_1+\cdots+n_m=n$,

$$\begin{aligned}&\mathbb{P}(N_i(t)-N_i(s)=n_i,1\leq i\leq m|N(t)-N(s)=n)\\&\quad=\mathbb{E}(\mathbb{P}(N_i(t)-N_i(s)=n_i,1\leq i\leq m|T_1,\cdots,T_n)|N(t)-N(s)=n))\end{aligned}$$

其中$T_1,\cdots,T_n$表示N在(s,t)内n次事件发生的时刻,从而为(s,t)上均匀分布的次序统计量. 记条件概率

$$\mathbb{P}(N_i(t)-N_i(s)=n_i,1\leq i\leq m|T_1,\cdots,T_n)=f(T_1,\cdots,T_n).$$

由于在不同时间, 事件分类的概率不同，因此可以将不同时间发生的事件看作不同的随机实验. 从而该条件概率可看作在n个独立的不同随机实验中，观察到m种结

果出现的次数为$(n_1, n_2, \cdots, n_m)$ 的概率. 由于实验是独立的, 该概率与实验次序无关。这表明改变函数f中变量次序并不改变f 的值. 另一方面, 条件事件

$$(N_i(t) - N_i(s) = n_i, 1 \le i \le m | T_1, \cdots, T_n)$$

中共有$\frac{n!}{n_1!n_2!\cdots n_m!}$种不同组合, 每种组合发生的概率都可表示成如下形式:

$$p_1(T_{t_1}) \cdots p_1(T_{t_{n_1}}) p_2(T_{t_{n_1+1}}) \cdots p_2(T_{t_{n_1+n_2}}) \cdots p_m(T_{t_{n-n_m+1}}) \cdots p_m(T_{t_n}),$$

其中$(t_1, \cdots, t_n)$是一个n级排列. 因此由推论6.9得

$$\begin{aligned}
&\mathbb{P}(N_i(t) - N_i(s) = n_i, 1 \le i \le | N(t) - N(s) = n) = \mathbb{E}(f(X_1, \cdots, X_n)) \\
&= \sum_{\text{所有组合}} \mathbb{E}\Big(\prod_{i=1}^{n_1} p_1(X_{t_i}) \prod_{i=n_1+1}^{n_1+n_2} p_2(X_{t_i}) \cdots \prod_{i=n-n_m+1}^{n} p_m(X_{t_i}) \Big) \\
&= \frac{n!}{n_1!n_2!\cdots n_m!} \Big[\frac{1}{t-s} \int_s^t p_1(u) \mathrm{d}u \Big]^{n_1} \cdots \Big[\frac{1}{t-s} \int_s^t p_m(u) \mathrm{d}u \Big]^{n_m}, \qquad (2.3.17)
\end{aligned}$$

其中$X_1, \cdots, X_n$为n个独立的(s, t)均与分布随机变量.

将(2.3.17)替换(2.3.15), 重复定理7.10的证明可得定理结论. □

例7.10 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时间服从速率为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

$$p_t = \mathbb{E}(\mathbf{1}_{\{c-t \le U_t\}}(1 - \mathrm{e}^{-(U_t - c + t) \wedge d})) = \frac{1 - \mathrm{e}^{-d(1+\lambda)}}{1+\lambda} \mathrm{e}^{\lambda(t-c)}.$$

当$d + c > t > c$时, 顾客被监控到的概率

$$p_t = \mathbb{E}(1 - \mathrm{e}^{-U_t \wedge (d+c-t)}) = \frac{1 - \mathrm{e}^{(1+\lambda)(t-c-d)}}{1+\lambda}.$$

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

$$m = K \int_0^{d+c} p_t \mathrm{d}t = K \Big[\frac{(1 - \mathrm{e}^{-d(1+\lambda)})(1 - \mathrm{e}^{-\lambda c})}{\lambda(1+\lambda)} + \frac{d}{1+\lambda} - \frac{1 - \mathrm{e}^{-(1+\lambda)d}}{(1+\lambda)^2} \Big]. \qquad □$$

练习题 以下总设$N = \{N(t), t \ge 0\}$是强度为λ的泊松过程.

7.1 设随机变量X_1, X_2, X_3独立且分别服从参数为$\lambda_1, \lambda_2, \lambda_3$的指数分布. 求

(1)$\mathbb{P}(X_1 < X_2 < X_3)$; (2)$\mathbb{P}(\max\limits_{1\le k\le 3} X_k \le t)$; (3)$\mathbb{P}(X_1 < X_2 < X_3, \max\limits_{1\le k\le 3} X_k \le t)$.

7.2* 若例7.4中泊松过程N的强度λ未知, 试给出c和t的估计.

7.3 对任意$0 < s < t$, $n \ge 1$, 证明

$$\mathbb{P}(N(s) = k|N(t) = n) = C_n^k \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}, \quad k = 0, 1, \cdots, n$$

7.4 设随机变量T与N独立且对任意$t > 0$, $\mathbb{P}(T > t) = \mathrm{e}^{-\mu t}$. 证明

$$\mathbb{P}(N(T) = k) = \frac{\mu}{\lambda + \mu}\left(\frac{\lambda}{\lambda + \mu}\right)^k, \quad k = 0, 1, 2, \cdots.$$

7.5 对任意$k \ge 0$, 令T_k表示第k次随机事件发生的时间序列. 对任意$t > x > 0$, 求(1) $\mathbb{P}(t - T_k > x|N(t) = k)$, (2) $\mathbb{P}(t - T_{N(t)} > x)$.

7.6 对任意$k \ge 1$, 令W_k表示第k次事件的等待时间, 对任意$x > 0$, 求(1) $\mathbb{P}(W_{k+1} > x|N(t) = k)$, (2) $\mathbb{P}(W_{N(t)+1} > x)$.

7.7 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)已知最后两名观众在放映前5 分钟内到达, 问他们是放映前两分钟到达的概率？(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.

7.8* 证明引理7.6的必要性.

第三章　离散时间马氏链

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, 但在现实问题中并不常有. 人们更容易观察或更能近似观察到一种所谓的条件独立性. 对这种现象建模可以使用所谓的马氏过程. 本章介绍其中一类简单的模型——离散时间离散状态的马氏过程.

§3.1　马氏链的定义与举例

(A) 条件独立与马氏链

定义8.1 称随机事件A,B在事件C(要求$\mathbb{P}(C)>0$), 发生条件下独立, 若

$$\mathbb{P}(AB|C)=\mathbb{P}(A|C)\mathbb{P}(B|C).$$

简称A,B条件独立.

对条件独立, 如下性质成立

性质8.1 设A,B,C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) $\mathbb{P}(A|BC)=\mathbb{P}(A|C)$;

(3) $\mathbb{P}(B|AC)=\mathbb{P}(B|C)$.

证明 $(1)\Rightarrow(2),(3)$: 由$\mathbb{P}(AB|C)=\mathbb{P}(A|C)\mathbb{P}(B|C)$可得

$$\mathbb{P}(ABC)=\mathbb{P}(A|C)\mathbb{P}(B|C)\mathbb{P}(C)=\mathbb{P}(AC)\mathbb{P}(B|C)=\mathbb{P}(A|C)\mathbb{P}(BC)$$

从而

$$\mathbb{P}(A|BC)=\mathbb{P}(A|C),\qquad \mathbb{P}(B|AC)=\mathbb{P}(B|C).$$

$(2)\Rightarrow(1)$: 由$\mathbb{P}(A|BC)=\mathbb{P}(A|C)$可知

$$\mathbb{P}(ABC)=\mathbb{P}(A|C)\mathbb{P}(BC)$$

两边除以$\mathbb{P}(C)$得$\mathbb{P}(AB|C)=\mathbb{P}(A|C)\mathbb{P}(B|C)$.

类似可得$(3)\Rightarrow(1)$. 证毕. □

将条件独立的假设用到随机过程上就可定义所谓的马氏性(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8.1 假定例4.1中的赌徒在某个时刻t(现在)手中筹码为x, 经验告诉我们, 他在此后(未来)的输赢取决于此刻的筹码数而与t 之前(过去)的筹码及输赢无关.

例8.2 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: 从罐中任意抽出一个球, 如果是红球就放入一个白球, 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.

称具有这种马氏性的随机过程为马氏过程. 本课程只介绍其中具有离散参数和离散状态的情形, 我们称这种马氏过程为离散时间马氏链, 简称为马氏链. 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. 为方便，以下我们总设状态空间S 为整数集. 马氏链定义如下：

定义8.2 设$X = \{X_n; n \geq 0\}$是概率空间$(\Omega, \mathfrak{F}, \mathbb{P})$上随机过程, 状态空间$S \subset \mathbf{Z}$, 称$X = \{X_n; n \geq 0\}$为离散时间马氏链, 如果对任意$n \geq 1$, $i_0, \cdots, i_{n-1}, i, j \in S$,

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i). \qquad (3.1.1)$$

例8.3(存储模型) 设$\xi_i, i = 1, 2, \cdots$为独立同分布的非负整数值随机变量, 令

$$X_0 = K, \qquad X_n = \begin{cases} 0 \vee (X_{n-1} - \xi_n), & k < X_{n-1} \leq K, \\ 0 \vee (K - \xi_n), & X_{n-1} \leq k, \end{cases} \quad n \geq 1,$$

其中$k < K$为给定的两个正整数. 那么$X = \{X_n, n \geq 0\}$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$n \geq 1$ 以及$i_1, \cdots, i_{n+1} \in S$,

$$\begin{aligned} &\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_1 = i_1, X_0 = K) \\ &\quad = \begin{cases} \mathbb{P}((K - \xi_{n+1}) \vee 0 = i_{n+1} | X_n = i_n, \cdots, X_0 = K), & i_n \leq k, \\ \mathbb{P}((i_n - \xi_{n+1}) \vee 0 = i_{n+1} | X_n = i_n, \cdots, X_0 = K), & k < i_n \leq K, \end{cases} \end{aligned}$$

对任意$i \geq 1$, 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$\{\xi_i, i \geq 1\}$的独立性

$$\begin{aligned} &\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_1 = i_1, X_0 = K) \\ &\quad = \begin{cases} \mathbb{P}((K - \xi_{n+1}) \vee 0 = i_{n+1}), & i_n \leq k, \\ \mathbb{P}((i_n - \xi_{n+1}) \vee 0 = i_{n+1}), & k < i_n \leq K, \end{cases} \\ &\quad = \begin{cases} \mathbb{P}((K - \xi_{n+1}) \vee 0 = i_{n+1} | X_n = i_n), & i_n \leq k, \\ \mathbb{P}((i_n - \xi_{n+1}) \vee 0 = i_{n+1} | X_n = i_n), & k < i_n \leq K, \end{cases} \\ &\quad = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n). \end{aligned}$$

所以X为离散时间马氏链. □

一般称该马氏链为离散存储模型, 其中ξ_i表示第i个周期的市场需求量, X_i表示第i个周期末的商品仓储量, K为最大库存量, k为补货策略边界值, 即若在每个周期

末若商品仓储量大于k则不补货, 若不大于k, 则在下一个周期开始时补足货物到K. 在物流管理问题中常关心的问题是确定合适的k,K值使得仓储成本最小.

注8.2 需要指出的是在马氏过程定义中要求确切知道现在n时刻的状态. 若现在的状态不是唯一已知的, 那么即使是马氏链, 条件独立性也可能不成立. 例如, 由习题6.1可知(q,p)-随机游动W是马氏链, 若$\mathbb{P}(X_0=0)=\mathbb{P}(X_0=2)=1/2$ 且$p\neq q$, 则

$$p=\mathbb{P}(W_2=2|W_1\in\{1,3\},W_0=0)\neq\mathbb{P}(W_2=2|W_1\in\{1,3\})=\frac{p+pq}{1+p}.$$

(B) 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8.2 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$i_1,\cdots,i_{n+1}\in S$, 都有

$$\mathbb{P}(X_t=i_{n+1}|X_{t_1}=i_1,X_{t_2}=i_2,\cdots,X_{t_n}=i_n)=\mathbb{P}(X_t=i_{n+1}|X_{t_n}=i_n).\quad(3.1.2)$$

证明* 只需要证明必要性. 首先注意到由全概率公式和(3.1.1), 对任意$n\geq k\geq 0$

$$\begin{aligned}
&\mathbb{P}(X_{n+1}=i_{n+1}|X_k=i_k,\cdots,X_{n-1}=i_{n-1},X_n=i_n)\\
&=\frac{\mathbb{P}(X_{n+1}=i_{n+1},X_n=i_n,\cdots X_k=i_k)}{\mathbb{P}(X_n=i_n,\cdots X_k=i_k)}.\\
&=\sum_{j_0,\cdots,j_{k-1}\in S}\frac{\mathbb{P}(X_{n+1}=i_{n+1},X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{\mathbb{P}(X_n=i_n,\cdots X_k=i_k)}\\
&=\sum_{j_0,\cdots,j_{k-1}\in S}\Big[\frac{\mathbb{P}(X_{n+1}=i_{n+1},X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{\mathbb{P}(X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}\\
&\qquad\times\frac{\mathbb{P}(X_n=i_n,\cdots X_k=i_k,X_{k-1}=j_{k-1},\cdots,X_0=j_0)}{\mathbb{P}(X_n=i_n,\cdots X_k=i_k)}\Big]\\
&=\mathbb{P}(X_{n+1}=i_{n+1}|X_n=i_n)
\end{aligned}\quad(3.1.3)$$

下面我们证明对任意$m\geq 1$, $n\geq 0$,

$$\begin{aligned}
&\mathbb{P}(X_{n+1}=i_{n+1},\cdots,X_{n+m}=i_{n+m}|X_k=i_k,X_{k+1}=i_{k+1}\cdots,X_n=i_n)\\
&\qquad=\mathbb{P}(X_{n+1}=i_{n+1},\cdots,X_{n+m}=i_{n+m}|X_n=i_n),\quad 0\leq k\leq n.
\end{aligned}\quad(3.1.4)$$

当$m=1$时(3.1.4)显然对任意$n\geq 0$以及$0\leq k\leq n$成立. 设$m\leq l-1$时(3.1.4)显然对任意$n\geq 0$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

$$\mathbb{P}(D|ABC)=\mathbb{P}(D|C)=\mathbb{P}(D|BC),\quad \mathbb{P}(C|AB)=\mathbb{P}(C|B).$$

因此

$$\begin{aligned}\text{l.h.s of (3.1.4)} &= \mathbb{P}(CD|AB) = \mathbb{P}(ABCD)/\mathbb{P}(AB)\\ &= \mathbb{P}(D|ABC)\mathbb{P}(C|BA) = \mathbb{P}(D|C)\mathbb{P}(C|B) = \mathbb{P}(D|CB)\mathbb{P}(C|B)\\ &= \mathbb{P}(DC|B) = \text{r.h.s of (3.1.4)}.\end{aligned}$$

由数学归纳法可知(3.1.4)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

$$\begin{aligned}&\mathbb{P}(X_t = i_{n+1}|X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)\\ &\quad = \frac{\mathbb{P}(X_t = i_{n+1}X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)}{\mathbb{P}(X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)}\\ &\quad = \sum_{j_s \in S, 0 \leq s < t}^{s \neq t_1, \cdots, t_n} \frac{\mathbb{P}(X_t = i_{n+1}, X_{t_k} = i_k, X_s = j_s, 0 \leq s < t, s \neq t_k, 1 \leq k \leq n)}{\mathbb{P}(X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)}\end{aligned}$$

再由(3.1.4)得

$$\begin{aligned}&\mathbb{P}(X_t = i_{n+1}|X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)\\ &\quad = \sum_{j_s \in S, t_n < s < t} \mathbb{P}(X_t = i_{n+1}, X_s = j_s, t_n < s < t|X_{t_n} = i_n)\\ &\qquad \times \sum_{j_s \in S, 0 \leq s < t_n}^{s \neq t_1, \cdots, t_n} \frac{\mathbb{P}(X_{t_k} = i_k, X_s = j_s, 0 \leq s < t_n, s \neq t_k, 1 \leq k \leq n)}{\mathbb{P}(X_{t_1} = i_1, X_{t_2} = i_2, \cdots, X_{t_n} = i_n)}\\ &\quad = \mathbb{P}(X_t = i_{n+1}|X_{t_n} = i_n)\end{aligned}$$

因此X是马氏链. □

进一步, 由(3.1.4)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$0 \leq t_1 < \cdots < t_m < n$, 都有

$$\mathbb{P}(AB|X_n = i_n) = \mathbb{P}(A|X_n = i_n)\mathbb{P}(B|X_n = i_n). \tag{3.1.5}$$

具体论证过程请读者完成(参见习题8.5).

(C) 转移概率矩阵与C-K方程

由(3.1.2)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

$$\mathbb{P}(X_{n+m} = j|X_n = i).$$

定义8.3 称条件概率$\mathbb{P}(X_{n+m}=j|X_n=i)$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

$$\mathbb{P}(X_{n+m}=j|X_n=i)=\mathbb{P}(X_m=j|X_0=i)$$

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$p_{i,j}^{(m)}$, 称之为从状态i到状态j 的m步转移概率. 特别简记$p_{i,j}^{(1)}$ 为$p_{i,j}$, 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8.4 称马氏链X是时齐的, 如果X的转移概率是平稳的.

定理8.3 设$X=\{X_n;n\geq 0\}$是时齐马氏链, 则对任意$n,m\geq 1$及$i,j\in S$

$$p_{i,j}^{(n+m)}=\sum_{k\in S}p_{i,k}^{(n)}p_{k,j}^{(m)}. \tag{3.1.6}$$

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(3.1.6). 由全概率公式

$$p_{i,j}^{(n+m)}=\mathbb{P}(X_{n+m}=j|X_0=i)=\sum_{k\in S}\mathbb{P}(X_{n+m}=j,X_n=k|X_0=i).$$

再由条件概率公式与马氏性(3.1.2)

$$\begin{aligned}\mathbb{P}(X_{n+m}=j,X_n=k|X_0=i)&=\frac{\mathbb{P}(X_{n+m}=j,X_n=k,X_0=i)}{\mathbb{P}(X_0=i)}\\&=\frac{\mathbb{P}(X_{n+m}=j,X_n=k,X_0=i)}{\mathbb{P}(X_n=k,X_0=i)}\frac{\mathbb{P}(X_n=k,X_0=i)}{\mathbb{P}(X_0=i)}\\&=\mathbb{P}(X_{n+m}=j|X_n=k)\mathbb{P}(X_n=k|X_0=i)=p_{k,j}^{(m)}p_{i,k}^{(n)}\end{aligned}$$

由此可知(3.1.6)成立. □

通常称(3.1.6)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.

本课程以后若无特别说明, 所涉及马氏链X都是时齐的.

定义8.5 将马氏链$X=\{X_n;n\geq 0\}$所有的转移概率排成矩阵形式$\mathbf{P}=(p_{i,j})_{i,j\in S}$, 称此矩阵为$X$的转移概率矩阵; 称$\mathbf{P}^{(n)}=(p_{i,j}^{(n)})_{i,j\in S}$ 为X的n阶(步)转移概率矩阵.

注8.4 转移概率矩阵满足:

(1)$p_{i,j}\geq 0$, $i,j\in S$, (2)$\sum_{j\in S}p_{i,j}=1$, $i\in S$.

一般称满足这两条件的矩阵为随机矩阵.

例8.4 由习题6.1以及(2.2.1)可知直线上的(q,p)随机游动是一个时齐马氏链, 转移概率矩阵为

$$\mathbf{P}=\begin{pmatrix} \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \\ \cdots & 0 & p & 0 & 0 & 0 & \cdots \\ \cdots & q & 0 & p & 0 & 0 & \cdots \\ \cdots & 0 & q & 0 & p & 0 & \cdots \\ \cdots & 0 & 0 & q & 0 & p & \cdots \\ \cdots & 0 & 0 & 0 & q & 0 & \cdots \\ \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \end{pmatrix}.$$

当S只有有限个状态时, 比如$S=\{1,2,\cdots,N\}$, 由C-K方程, 对任意$n,m>0$

$$\begin{aligned}\mathbf{P}^{(n+m)} &= \big(p_{i,j}^{(n+m)}\big)_N = \Big(\sum_{k=1}^{N} p_{i,k}^{(n)} p_{k,j}^{(m)}\Big)_{N\times N} \\ &= \big(p_{i,j}^{(n)}\big)_N \big(p_{i,j}^{(m)}\big)_N = \mathbf{P}^{(n)}\mathbf{P}^{(m)} = \mathbf{P}^{n+m}.\end{aligned}$$

类似的记法也可用到S有无穷多个状态的情形. 因此我们总有

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)}\mathbf{P}^{(m)} = \mathbf{P}^{n+m}.$$

显然, 对状态有限的马氏链, 要了解其n步转移概率矩阵, 可以通过线性代数中学过的矩阵运算实现.

例8.5 通常我们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

$$\begin{gathered}\mathbb{P}(X_{n+1}=1|X_n=1)=0.8,\quad \mathbb{P}(X_{n+1}=2|X_n=1)=0.1,\\ \mathbb{P}(X_{n+1}=3|X_n=1)=0.1,\quad \mathbb{P}(X_{n+1}=1|X_n=2)=0.3,\\ \mathbb{P}(X_{n+1}=2|X_n=2)=0.4,\quad \mathbb{P}(X_{n+1}=3|X_n=2)=0.3,\\ \mathbb{P}(X_{n+1}=1|X_n=3)=0.05,\quad \mathbb{P}(X_{n+1}=2|X_n=3)=0.1,\\ \mathbb{P}(X_{n+1}=3|X_n=3)=0.85.\end{gathered}$$

那么$\{X_n\}$构成一个马氏链, 转移概率矩阵

$$\mathbf{P}=\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.05 & 0.1 & 0.85 \end{pmatrix}.$$

由高等代数知识可知, $\mathbf{P}$有三个不同特征值1,0.3和0.75, 对应特征向量可取为

$$(1,1,1)',\ (1,-6,1)' \text{ 和 } (-26,-6,19)'.$$

因此

$$\mathbf{P}=\begin{pmatrix}1&1&-26\\1&-6&-6\\1&1&19\end{pmatrix}\begin{pmatrix}1&0&0\\0&0.3&0\\0&0&0.75\end{pmatrix}\left[\frac{1}{315}\begin{pmatrix}108&45&162\\25&-45&20\\-7&0&7\end{pmatrix}\right].$$

从而对任意$n\geq 1$,

$$\mathbf{P}^{(n)}=\mathbf{P}^n=\begin{pmatrix}1&1&-26\\1&-6&-6\\1&1&19\end{pmatrix}\begin{pmatrix}1&0&0\\0&0.3^n&0\\0&0&0.75^n\end{pmatrix}\left[\frac{1}{315}\begin{pmatrix}108&45&162\\25&-45&20\\-7&0&7\end{pmatrix}\right].$$

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,

$$\mathbf{P}^{(5)}=\begin{pmatrix}0.480&0.143&0.377\\0.373&0.145&0.482\\0.243&0.142&0.615\end{pmatrix}.$$

由此可知

$$p_{2,3}^{(5)}=\mathbb{P}(X_8=3|X_0=2)=0.482.$$

在该模型下, 祖父为社会中层的家庭到他曾孙辈有48.2%的概率处于社会下层. □

(D) 有限维分布表示

定理8.4 设$\mathbf{P}=(p_{i,j})_{i,j\in S}$是马氏链$X$的转移概率矩阵, 且设$\mathbb{P}(X_0=i)=p_i$, 那么对任意$0\leq t_1<t_2<\cdots<t_n$以及$i_1,\cdots,i_n\in S$,

$$\mathbb{P}(X_{t_1}=x_1,X_{i_2}=x_2,\cdots,X_{i_n}=x_n)=\sum_{i\in S}p_ip_{i,i_1}^{(t_1)}p_{i_1,i_2}^{(t_2-t_1)}\cdots p_{i_{n-1},i_n}^{(t_n-t_{n-1})}.$$

证明 由条件概率公式与马氏性

$$\begin{aligned}&\mathbb{P}(X_{t_1}=i_1,\cdots,X_{t_n}=i_n)\\&=\mathbb{P}(X_{t_n}=i_n|X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)\mathbb{P}(X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)\\&=p_{i_{n-1},i_n}^{(t_n-t_{n-1})}\mathbb{P}(X_{t_{n-1}}=i_{n-1},\cdots,X_{t_1}=i_1)=\cdots\\&=p_{i_{n-1},i_n}^{(t_n-t_{n-1})}\cdots p_{i_1,i_2}^{(t_2-t_1)}\sum_{i\in S}\mathbb{P}(X_{t_1}=i_1,X_0=i)\\&=\sum_{i\in S}p_ip_{i,i_1}^{(t_1)}p_{i_1,i_2}^{(t_2-t_1)}\cdots p_{i_{n-1},i_n}^{(t_n-t_{n-1})}.\end{aligned}$$

□

定理8.4表明马氏链的有限维分布由初始分布以及转移矩阵唯一确定. 进而, 对任意给定的随机矩阵$\mathbf{P}$以及初始分布$\pi=\{p_i,i\in S\}$, 由Kolmogorov 定理可知存在一个马氏链X使其转移概率矩阵恰为$\mathbf{P}$.

例8.6(排队模型) 设$\xi_i, i \geq 1$为独立同分布的非负整数值随机变量, 另设X_0也为一非负整数值随机变量且与$\{\xi_i, i \geq 1\}$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$X = \{X_n; n \geq 0\}$为离散时间马氏链. 若记$p_k = \mathbb{P}(\xi_1 = k)$, 求$X$的转移概率矩阵. 进一步若设$u_k = \mathbb{P}(X_0 = k)$, 求$\mathbb{P}(X_0 = 1, X_1 = 2, X_3 = 4)$.

证明 显然X的状态空间$S = \{0, 1 \cdots\}$. 对任意$n \geq 1$ 以及非负整数$i_0, \cdots, i_{n+1}$,

$$\begin{aligned} &\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_1 = i_1, X_0 = i_0) \\ &\quad = \begin{cases} \mathbb{P}(i_n - 1 + \xi_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_0 = i_0), & i_n > 0, \\ \mathbb{P}(\xi_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_0 = i_0), & i_n = 0, \end{cases} \end{aligned}$$

与例8.3类似的分析可知$\{X_n = i_n, \cdots, X_1 = i_1, X_0 = i_0\}$由随机变量$X_0, \xi_1, \cdots, \xi_n$确定. 因此由条件中独立性假设,

$$\begin{aligned} &\mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \cdots, X_1 = i_1, X_0 = i_0) \\ &\quad = \begin{cases} \mathbb{P}(i_n - 1 + \xi_{n+1} = i_{n+1}), & i_n > 0, \\ \mathbb{P}(\xi_{n+1} = i_{n+1}), & i_n = 0, \end{cases} \\ &\quad = \begin{cases} \mathbb{P}(i_n - 1 + \xi_{n+1} = i_{n+1} | X_n = i_n), & i_n > 0, \\ \mathbb{P}(\xi_{n+1} = i_{n+1} | X_n = i_n), & i_n = 0, \end{cases} \\ &\quad = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n). \end{aligned}$$

所以X为离散时间马氏链. 直接计算可知

$$\begin{aligned} p_{i,j} = \mathbb{P}(X_1 = j | X_0 = i) &= \begin{cases} \mathbb{P}(\xi_1 = j + 1 - i), & i \neq 0 \\ \mathbb{P}(\xi_1 = j), & i = 0 \end{cases} \\ &= \begin{cases} p_{j+1-i}, & i \neq 0 \text{且} j + 1 - i \geq 0, \\ 0; & i \neq 0 \text{且} j + 1 - i < 0, \\ p_j; & i = 0. \end{cases} \end{aligned}$$

因此X的转移概率矩阵为

$$\mathbf{P} = \begin{array}{c} \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & \cdots \\ p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & \cdots \\ 0 & p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ 0 & 0 & p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & 0 & 0 & p_0 & p_1 & p_2 & \cdots \\ 0 & 0 & 0 & 0 & p_0 & p_1 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots \end{pmatrix} \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \cdot \end{array}$$

由定理8.4, 有限维分布

$$\begin{aligned}\mathbb{P}(X_0=1,X_1=2,X_3=4)&=u_1p_{1,2}p_{2,4}^{(2)}\\&=u_1p_2(p_{2,1}p_{1,4}+p_{2,2}p_{2,4}+p_{2,3}p_{3,4}+p_{2,4}p_{4,4}+p_{2,5}p_{5,4})\\&=u_1p_2(p_0p_4+p_1p_3+p_2p_2+p_3p_1+p_4p_0)\\&=u_1p_2(2p_0p_4+2p_1p_3+p_2p_2).\end{aligned}$$ □

练习题

8.1 在仓储模型中设$\mathbb{P}(\xi_1=0)=0.1$,$\mathbb{P}(\xi_1=1)=0.2$, $\mathbb{P}(\xi_1=2)=0.3$, $\mathbb{P}(\xi_1=3)=0.2$, $\mathbb{P}(\xi_1=4)=0.2$, $k=2$, $K=4$. 求该模型的转移概率矩阵.

8.2 设马氏链的状态空间$S=\{0,1,2\}$, 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$n$阶转移矩阵$\mathbf{P}^{(n)}$.

8.3 设$N=\{N(t),t\geq 0\}$为泊松过程. 任取$t_0>0$. 证明$Y=\{N(nt_0),n\geq 0\}$是马氏链.

8.4 (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$\{X_n\}$就是个马氏链. 试写出该模型的转移概率矩阵.

8.5 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$S_m\subset S$, $m\geq 0$,

$$\mathbb{P}\left(\bigcap_{m=n+1}^{\infty}\{X_m\in S_m\}\Big|X_n=i,\bigcap_{m=0}^{n-1}\{X_m\in S_m\}\right)=\mathbb{P}\left(\bigcap_{m=1}^{\infty}\{X_m\in S_m\}\Big|X_0=i\right).$$

8.6* 试证明(3.1.5).

§3.2 状态分类

以下总设马氏链$X=\{X_n;\ n\ge 0\}$的转移概率矩阵$\mathbf{P}=(p_{i,j})_{i,j\in S}$. 为方便, 分别记$\mathbb{E}(\cdot|X_0=i)$和$\mathbb{P}(\cdot|X_0=i)$ 为$\mathbb{E}_i(\cdot)$和$\mathbb{P}_i(\cdot)$.

(A) 互通、本质与不可约

定义9.1 如果存在$n\ge 0$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$n\ge 0$, $p_{i,j}^{(n)}=0$, 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$i\leftrightarrow j$.

由于$p_{i,j}^{(0)}=\delta(i,j)=\begin{cases}1, & i=j\\ 0, & i\ne j\end{cases}$, 因此对任意状态$i$, 它与自己总是互通的.

命题9.1 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,

$$p_{i,k}^{(n+r)}\ge p_{i,j}^{(n)}p_{j,k}^{(r)}>0;\ \ p_{k,i}^{(l+m)}\ge p_{k,j}^{(l)}p_{j,i}^{(m)}>0.$$

因此$i\leftrightarrow k$. □

定义9.2 记$C(i)=\{k\in S;k\leftrightarrow i\}$, 称$C(i)$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

$$C(i)\cap C(j)\ne\emptyset\Leftrightarrow C(i)=C(j).$$

特别, 若$p_{ii}=1$, 则称i为吸收的, 此时$C(i)=\{i\}$. 反过来, 由$C(i)=\{i\}$不能推出i是吸收的.

例9.1 若马氏链X的转移概率矩阵为

$$\begin{array}{cc}\mathbf{P}=\begin{pmatrix}1/3 & 2/3 & 0 & 0 & 0\\ 0 & 1/2 & 1/2 & 0 & 0\\ 3/4 & 1/4 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 1 & 0\end{pmatrix} & \begin{array}{c}0\\1\\2\\3\\4\end{array}\end{array}$$

那么由$0\to 1$, $1\to 2$, $2\to 0$可知$\{0,1,2\}\subset C(0)$. 再由$0\not\to 3$, $0\not\to 4$可知$\{0,1,2\}=C(0)$. 同样可知$C(3)=\{3,4\}$. 因此X的状态空间包含两个互通类$C(0)$, $C(3)$, 此时X不是不可约的马氏链. □

定义9.3 称状态i是本质的, 若对任意$j\in S$, 由$i\to j$可得$j\to i$.

显然吸收态是本质的, 进一步容易看出

性质9.2 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. □

定义9.4 若马氏链X的所有状态都是互通的, 即对任意$i \in S$, $S = C(i)$, 则称X是不可约的(irreducible).

不可约链的状态都是本质的.

(B) 周期性

定义9.5 设i是马氏链X的一个状态, 若正整数集$\{n \geq 1;\ p_{i,i}^{(n)} > 0\}$非空, 其中元素的最大公因数$d_i$称为是状态$i$的周期. 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.

由周期定义可知, 若$d_i \dagger n$, 则$p_{i,i}^{(n)} = 0$.

命题9.3 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$p_{i,j}^{(n)} > 0$, $p_{j,i}^{(m)} > 0$. 则由C-K方程,

$$p_{i,i}^{(n+m)} \geq p_{i,j}^{(n)} p_{j,i}^{(m)} > 0, \quad p_{j,j}^{(n+m)} \geq p_{j,i}^{(m)} p_{i,j}^{(n)} > 0.$$

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9.6 若马氏链X的每一个状态都有相同的周期d, 称X是d周期马氏链. 特别, 若$d = 1$, 则称X是非周期的.

例9.2 (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$i, j \in$ **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$\mathbf{Z} = C(0)$, 从而W是不可约的. 对任意$i \in \mathbf{Z}$, 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9.3 考虑例7.6中定义的排队模型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$j > i$, 由$p_{j,i}^{(j-i)} = p_0^{j-i} > 0$可知$j \to i$.

反过来, 当$i \neq 0$时

$$p_{i,j}^{((k-1)(j-i))} \geq p_{i,i+(k-1)(j-i)}^{(j-i)} p_{i+(k-1)(j-i),j}^{((k-2)(j-i))} \geq p_k^{j-i} p_0^{(k-2)(j-i)} > 0,$$

即$i \to j$; 而当$i = 0$时, 由$p_k > 0$可知$0 \to k$以及$k \to j$可知$0 \to j$.

因此, 对任意$i,j\in S$, $i\leftrightarrow j$, 从而X是不可约的. 注意到$p_{0,0}=p_0>0$, 状态0的周期为$d_0=1$. 由命题9.2可知X的所有状态周期均为1, 因此X是非周期的. □

定理9.4 设状态i的周期为d, 则存在正整数N使得对任意$n\ge N$, $p_{i,i}^{(nd)}>0$.

证明 由周期定义, 存在$1\le n_1,\cdots,n_s$使得$p_{i,i}^{(n_k)}>0$, $k=1,\cdots,s$, 而且$n_1,\cdots,n_s$的最大公因数为d. 由习题1.9, 存在正整数N, 对任意$n\ge N$, 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

$$p_{i,i}^{(nd)}\ge\prod_{k=1}^{s}(p_{i,i}^{(n_k)})^{r_k}>0.$$

这表明结论成立. □

由此易得如下两个推论.

推论9.5 记状态i的周期为d_i, 若$p_{j,i}^{(m)}>0$, 那么存在$N>0$使得$p_{j,i}^{(m+nd_i)}>0$, $n\ge N$.

推论9.6 若状态i是非周期的, 那么存在$N>0$使得对任意$n>N$, $p_{i,i}^{(n)}>0$.

(C) 常返与非常返

定义9.7 对任意状态$i\in S$, 称

$$\tau_i=\inf\{n\ge1;X_n=i\}$$

为状态i的首次回访时间, 其中规定$\inf\emptyset=+\infty$.

注9.1 τ_i是一个可能取值$+\infty$的随机变量, 由全概率公式可知

$$\{\tau_i\le n\}=\bigcup_{k=1}^{n}\{\tau_i=k\}=\bigcup_{k=1}^{n}\{X_k=i,X_{k-1}\ne i,\cdots,X_1\ne i\},$$

因此对任意$n\ge1$, 事件$\{\tau_i\le n\}$是否发生只与X在n及n之前的状态有关而与n之后的状态无关.

对任意$i,j\in S$, 令

$$f_{i,j}^{(n)}=\mathbb{P}_i(\tau_j=n)=\mathbb{P}(X_n=j,X_{n-1}\ne j,\cdots,X_1\ne j|X_0=i), \tag{3.2.1}$$

$$f_{i,j}=\mathbb{P}_i(\tau_j<\infty)=\sum_{n=1}^{\infty}f_{i,j}^{(n)}. \tag{3.2.2}$$

由定义可知, 当$j\ne i$时, $f_{i,j}^{(n)}$表示从i出发后第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. 显然, 对任意i,j, 若$i\not\to j$, 则$f_{i,j}=0$. 此外由

$$f_{i,j}=\sum_{n=1}^{\infty}\mathbb{P}(X_n=j,X_{n-1}\ne j,\cdots,X_1\ne j|X_0=i)=\mathbb{P}_i(\bigcup_{k=1}^{\infty}\{X_k=j\}),$$

可得

$$0 \le f_{i,j}^{(n)} \le p_{i,j}^{(n)} \le f_{i,j} \le 1. \tag{3.2.3}$$

定义9.8 称状态i是常返的(recurrent), 如果$f_{i,i} = 1$; 否则称为非常返的或暂留的(滑过的)(transient).

性质9.7 若i是常返的, $i \to j$, 那么$f_{j,i} = 1$.

证明 定义

$$e_{i,j}^{(n)} = \mathbb{P}_i(X_n = j, X_v \ne i, v = 1, \cdots, n-1). \tag{3.2.4}$$

它表示从i出发, 中途不仅过i而在第n步到达j的概率(禁忌概率). 因为$i \to j$, 存在$k > 0$, 使得$e_{i,j}^{(k)} > 0$. 若$f_{j,i} < 1$, 那么

$$\begin{aligned}
&\mathbb{P}_i(\tau_i = \infty) \ge \mathbb{P}_i(X_k = j, X_n \ne i, 1 \le n \ne k)\\
&\quad = \mathbb{P}_i(X_k = j, X_n \ne i, 1 \le n < k)\mathbb{P}(X_v \ne i, v > k | X_k = j, X_n \ne i, 1 \le n < k)\\
&\quad = e_{ij}^{(k)}\mathbb{P}(X_n \ne i, n \ge 1 | X_0 = j) = e_{ij}^{(k)}(1 - f_{j,i}) > 0.
\end{aligned}$$

此与i常返矛盾, 因此$f_{j,i} = 1$. □

由性质9.7以及(3.2.3)可知, 若i常返, 那么i是本质的.

定理9.8 状态i常返当且仅当$\sum\limits_{n=0}^{\infty} p_{i,i}^{(n)} = \infty$. 而在非常返时,

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} = \frac{1}{1 - f_{i,i}}.$$

证明 令$F_{i,j}(s) = \sum\limits_{n=1}^{\infty} s^n f_{i,j}^{(n)}$, $P_{i,j}(s) = \sum\limits_{n=0}^{\infty} s^n p_{i,j}^{(n)}$, $s \in (-1, 1)$. 对任意$i, j \in S$,

$$\begin{aligned}
p_{i,j}^{(n)} &= \sum_{k=1}^{n} \mathbb{P}_i(\tau_j = k, X_n = j) = \sum_{k=1}^{n} \mathbb{P}_i(X_k = j, X_v \ne j, 1 \le v < k, X_n = j)\\
&= \sum_{k=1}^{n} \Big[\mathbb{P}(X_n = j | X_k = j, X_v \ne j, 1 \le v < k, X_0 = i)\\
&\qquad\qquad \times \mathbb{P}_i(X_k = j, X_v \ne j, 1 \le v < k)\Big]\\
&= \sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)}.
\end{aligned} \tag{3.2.5}$$

因此

$$\begin{aligned}
P_{i,i}(s) &= 1 + \sum_{n=1}^{\infty} s^n p_{i,i}^{(n)} = 1 + \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{i,i}^{(k)} p_{i,i}^{(n-k)}\\
&= 1 + \sum_{k=1}^{\infty} s^k f_{i,i}^{(k)} \sum_{r=0}^{\infty} s^r p_{i,i}^{(r)} = 1 + F_{i,i}(s) P_{i,i}(s)
\end{aligned}$$

由此可得

$$P_{i,i}(s)=\frac{1}{1-F_{i,i}(s)}. \tag{3.2.6}$$

注意到$F_{i,i}(s)$与$P_{i,i}(s)$均为非负系数的幂级数函数, 由习题1.5可知

$$f_{i,i}=F_{i,i}(1-),\quad \sum_{n=0}^{\infty}p_{i,i}^{(n)}=P_{i,i}(1-).$$

因此$f_{i,i}=1\Leftrightarrow\sum\limits_{n=0}^{\infty}p_{i,i}^{(n)}=\infty$, 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. □

由定理9.8容易推出

推论9.9 设状态i是常返的, 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9.9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9.2表明(q,p)-简单随机游动是不可约的, 再由命题6.4可知, (q,p)-简单随机游动的状态0是常返的当且仅当$p=q=1/2$. 因此(q,p)-简单随机游动是不可约常返的当且仅当$p=q=1/2$.

例9.4 (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. 此外我们还有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我们可如下计算其n步转移概率

$$p_{(0,0,0),(0,0,0)}^{(n)} = \begin{cases} 0, & n = 2k-1, \\ \sum\limits_{r+l+s=k} \frac{(2k)!}{r!r!l!l!s!s!} \frac{1}{6^{2k}}, & n = 2k. \end{cases}$$

因此

$$\begin{aligned} \sum_{n=0}^{\infty} p_{(0,0,0),(0,0,0)}^{(n)} &= \sum_{k=0}^{\infty} \sum_{r+l+s=k} \frac{(2k)!}{r!r!l!l!s!s!} \frac{1}{6^{2k}} \\ &= \sum_{k=0}^{\infty} \frac{1}{6^{2k}} \frac{(2k)!}{k!k!} \sum_{r=0}^{k} \sum_{l=0}^{k-r} \Big[\frac{k!}{r!l!(k-r-l)!}\Big]^2 \\ &\le \sum_{k=0}^{\infty} \frac{1}{6^{2k}} \frac{(2k)!}{k!k!} \max_{0\le r\le m\le k} \Big\{\frac{k!}{r!(m-r)!(k-m)!}\Big\} \sum_{r=0}^{k} \sum_{l=0}^{k-r} \frac{k!}{r!l!(k-r-l)!} \end{aligned}$$

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,

$$\frac{k!}{r!(m-r)!(k-m)!} = C_k^m C_m^r.$$

注意到对任意正整数$m \ge r$, 组合数

$$C_m^r \le C_m^{\lfloor m/2 \rfloor},$$

其中$\lfloor \cdot \rfloor$表示取整运算, 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

$$\sum_{n=0}^{\infty} p_{(0,0,0),(0,0,0)}^{(n)} \le \sum_{k=0}^{\infty} \frac{3^k}{6^{2k}} \frac{(2k)!}{k!(k-m_k)!\lfloor m_k/2 \rfloor!(m_k - \lfloor m_k/2 \rfloor)!}$$

注意到$m_k \approx 2k/3$, 由k充分大时的Stirling公式 $k! \approx \sqrt{2\pi k}(k/\mathrm{e})^k$, 近似计算后得

$$\frac{(2k)!}{k!(k-m_k)!\lfloor m_k/2\rfloor!(m_k-\lfloor m_k/2\rfloor)!} \approx \frac{3\sqrt{3}}{2\pi\sqrt{\pi}}\frac{2^{2k}3^k}{k^{3/2}}.$$

因此存在常数$M>0$使得

$$\sum_{n=0}^{\infty} p^{(n)}_{(0,0,0),(0,0,0)} \le M\sum_{k=1}^{\infty}\frac{1}{k^{3/2}} < \infty.$$

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, 参见习题9.6. □

命题9.10 令$g_{i,i}(m)=\mathbb{P}_i(\text{至少有}m\text{个}n\ge 1\text{使得}X_n=i)$, $g_{i,i}=\mathbb{P}_i(\{X_n=i\},\ \text{i.o.})$, 那么$g_{i,i}(m)=f^m_{i,i}$. 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

$$\begin{aligned} g_{i,i}(m+1) &= \mathbb{P}_i(\text{至少有}m+1\text{个}n\ge 1\text{使得}X_n=i)\\ &= \sum_{k=1}^{\infty}\mathbb{P}_i(\tau_i=k, \text{至少有}m\text{个}n\ge k+1\text{使得}X_n=i)\\ &= \sum_{k=1}^{\infty}\mathbb{P}_i(\text{至少有}m\text{个}n\ge k+1\text{使得}X_n=i|\tau_i=k)\mathbb{P}_i(\tau_i=k) \end{aligned}$$

由时齐马氏性(参见习题8.5)

$$\begin{aligned} g_{i,i}(m+1) &= \sum_{k=1}^{\infty}\mathbb{P}(\text{至少有}m\text{个}n\ge 1\text{使得}X_n=i|X_0=i)\mathbb{P}_i(\tau_i=k)\\ &= g_{i,i}(m)\sum_{k=1}^{\infty}f^{(k)}_{i,i} = g_{i,i}(m)f_{i,i}. \end{aligned}$$

由此递归关系并注意到$g_{i,i}(1)=f_{i,i}$可得, 对任意$m\ge 1$,

$$g_{i,i}(m)=f^m_{i,i}.$$

令$m\to\infty$, 由概率连续性可知$g_{i,i}=\lim\limits_{m\to\infty}f^m_{i,i}$, 因此后一结论自然成立. □

由命题9.10可知, 状态i是常返的, 那么从i出发几乎必然可以无穷次地回访i, 这正是“常返”的字面含义; 若i是非常返(暂留)的, 那么由$\mathbb{P}_i(\{X_n=i\},\ \text{i.o.})=0$可知, 对几乎所有的$\omega$, 都存在一个时刻$N(\omega)$使得在$N(\omega)$后$X$不再访问$i$.

定理9.11 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$Y=\{Y_n, n\ge 0\}$ 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$n \geq 1$, $k_0 \in S, k_1, \cdots, k_{n-1} \in S \setminus \{i\}$, $m \geq 1$, $0 \leq n_1 < \cdots < n_m$, $j_1, \cdots, j_m \in S$. 令$k_n = i$,

$$\begin{aligned}
&\mathbb{P}\Big(\bigcap_{l=0}^{n-1}\{X_l = k_l\} \cap \{\tau_i = n\} \cap \bigcap_{r=1}^{m}\{Y_{n_r} = j_r\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{l=0}^{n}\{X_l = k_l\} \cap \bigcap_{r=1}^{m}\{X_{n+n_r} = j_r\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{l=0}^{n-1}\{X_l = k_l\} \cap \{X_n = i\}\Big)\mathbb{P}\Big(\bigcap_{r=1}^{m}\{X_{n+n_r} = j_r\}|X_n = i\Big) \\
&= \mathbb{P}\Big(\bigcap_{l=0}^{n-1}\{X_l = k_l\} \cap \{X_n = i\}\Big)\mathbb{P}_i\Big(\bigcap_{r=1}^{m}\{X_{n_r} = j_r\}\Big)
\end{aligned}$$

两边对全部的$n \geq 1$，$k_0 \in S$以及$k_1, \cdots, k_{n-1} \in S \setminus \{i\}$求和得

$$\mathbb{P}\Big(\bigcap_{r=1}^{m}\{Y_{n_r} = j_r\}\Big) = \mathbb{P}_i\Big(\bigcap_{r=1}^{m}\{X_{n_r} = j_r\}\Big)$$

进而

$$\begin{aligned}
&\mathbb{P}\Big(\bigcap_{l=0}^{n-1}\{X_l = k_l\} \cap \{\tau_i = n\} \cap \bigcap_{r=1}^{m}\{Y_{n_r} = j_r\}\Big) \\
&= \mathbb{P}\Big(\bigcap_{l=0}^{n-1}\{X_l = k_l\} \cap \{X_n = i\}\Big)\mathbb{P}_i\Big(\bigcap_{r=1}^{m}\{Y_{n_r} = j_r\}\Big).
\end{aligned}$$

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9.11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 除特别说明外, 以下总设X为马氏链, 转移概率矩阵为$(p_{i,j})_{i,j \in S}$.

9.1 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$d|n-m$.
9.2 若状态i非常返, 那么对任意$j \in S$, $\lim_{n\to\infty} p_{j,i}^{(n)} = 0$.
9.3 若j为吸收状态, 那么$\lim\limits_{n\to\infty} p_{i,j}^{(n)} = f_{i,j}$.
9.4 若$C(i)$,$C(j)$是两不相同的本质互通类, 那么$f_{i,j} = 0$.
9.5 元素有限的本质类必是常返类.
9.6 证明平面上简单对称随机游动是常返的.
9.7* 若状态j的周期为d, 那么集合$\{n \geq 1;\ f_{j,j}^{(n)} > 0\}$的最大公因子也为$d$.
[提示] 由$p_{j,j}^{(n)} = 0 \Rightarrow f_{j,j}^{(n)} = 0$可知$d$是集合$A = \{n \geq 1 : f_{j,j}^{(n)} > 0\}$的公因子. 另一方面令$n_1 = min\{n \geq 1, p_{j,j}^{(n)} > 0\}$, $n_m = min\{n > n_{m-1}, p_{j,j}^{(n)} > 0$ 且 $(n_1, \cdots, n_{m-1}) \nmid n_m\}$, $1 \leq m$. 由j的周

期为d可知这样的n_m只有有限个(记为k). 记$d_m=(n_1,\cdots,n_m)$, $1\le m\le k$. 那么$d_k=d$, 而且对任意$0<r<n_m$, 若$p_{j,j}^{(r)}>0$, 则$d_{m-1}|r$, 进而由$d_{m-1}\not| n_m \Rightarrow d_{m-1}\not|(n_m-r)$ 可知$f_{j,j}^{(n_m-r)}=0$. 因此$0<p_{j,j}^{(n_m)}=\sum_{r=0}^{n_m-1}p_{j,j}^{(r)}f_{j,j}^{(n_m-r)}=f_{j,j}^{(n_m)}$ 表明$n_1,\cdots,n_k\in A$从而A的最大公因子不大于d. 综合两方面可知结论成立.

9.8 设$X=\{X(n);n\ge 0\}$为马氏链, 令$\tau_i^{(0)}=0$, $\tau_i^{(k)}=\inf\{n>\tau_i^{(k-1)};X(n)=i\}$,

$$W_i^{(k)}=\tau_i^{(k)}-\tau_i^{(k-1)},\quad k\ge 1.$$

证明$W_i^{(k)},k\ge 1$,相互独立, 而且$W_i^{(k)},k\ge 2$,有相同分布.

§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. 这一节我们将通过寻找恰当的计算方法, 进一步讨论这类随机变量的分布、性质与应用. 为叙述方便, 若无特别声明, 本节总设马氏链$X=\{X_n;\ n\geq 0\}$的转移概率矩阵$\mathbf{P}=(p_{i,j})_{i,j\in S}$.

(A) 访问概率与分布

首先我们将单个状态i的回访时间τ_i推广到状态空间S中任一子集A上. 令

$$\tau_A=\inf\{n\geq 1, X_n\in A\}.$$

τ_A表示首次进入(回到)集合A的时间. 对任意正整数$n\geq 1$, 令

$$f_{i,A}^{(n)}=\mathbb{P}_i(\tau_A=n), \qquad f_{i,A}=\sum_{n=1}^{\infty} f_{i,A}^{(n)},$$

以及对任意$u\in[0,1]$, 定义

$$F_{i,A}(u)=\sum_{n=1}^{\infty} f_{i,A}^{(n)}u^n.$$

那么$f_{i,A}$就是从i出发在有限时间内访问(回访)A的概率. 由于级数$\sum_{n=1}^{\infty} f_{i,A}^{(n)}\leq 1$, $f_{i,A}=F_{i,A}(1)$. 当$f_{i,A}=1$时, 序列$\{f_{i,A}^{(n)}\}$就是随机变量τ_A的分布列, $F_{i,A}(u)$就是对应的矩母函数; 当$f_{i,A}<1$ 时, 从i出发在有限时间内到达A的事件可能不发生, 但由幂级数与系数关系可知序列$\{f_{i,A}^{(n)}\}$与函数$F_{i,A}(u)$ 仍是一一对应的.

定理10.1 对任意给定的$A\subset S$以及$u\in[0,1]$, $\{F_{i,A}(u); i\in S\}$是线性方程组

$$z_i=\sum_{k\notin A} up_{i,k}z_k+u\sum_{j\in A} p_{i,j}, \qquad i\in S, \tag{3.3.1}$$

的最小非负解. 即$\{F_{i,A}(u); i\in S\}$本身是方程组(3.3.1)的非负解而且对该方程组的任意非负解 $\{g_i(u); i\in S\}$, $F_{i,A}(u)\leq g_i(u)$ 对一切$i\in S$ 都成立.

证明 记$A(i,n,u)=\sum_{m=1}^{n} f_{i,A}^{(m)}u^m$, 则$\lim\limits_{n\to\infty} A(i,n,u)=F_{i,A}(u)$. 由

$$\begin{aligned}A(i,n+1,u) &= \sum_{m=1}^{n+1} f_{i,A}^{(m)}u^m=f_{i,A}^{(1)}u+\sum_{m=2}^{n+1} f_{i,A}^{(m)}u^m\\ &= \sum_{j\in A} p_{i,j}u+\sum_{m=2}^{n+1}\sum_{k\notin A}\mathbb{P}_i(X_m\in A, X_1=k, X_v\notin A, 1<v<m)u^m\end{aligned}$$

由马氏性和时齐性可知

$$\begin{aligned}&\mathbb{P}_i(X_m\in A, X_1=k, X_v\notin A, 1<v<m)\\ &\quad=\mathbb{P}_i(X_1=k)\mathbb{P}(X_m\in A, X_v\notin A, 2<v<m|X_1=k)=p_{i,k}f_{k,A}^{(m-1)}.\end{aligned}$$

因此

$$\begin{aligned}A(i,n+1,u) &= u\sum_{j\in A}p_{i,j}+\sum_{k\notin A}\sum_{m=2}^{n+1}up_{i,k}f_{k,A}^{(m-1)}u^{m-1}\\ &= u\sum_{j\in A}p_{i,j}+u\sum_{k\notin A}p_{i,k}A(k,n,u).\end{aligned}$$

两边令$n\to\infty$得(参见习题3-1)

$$F_{i,A}(u)=u\sum_{j\in A}p_{i,j}+u\sum_{k\notin A}p_{i,k}F_{k,A}(u).$$

因此$\{F_{i,A}(u),i\in S\}$确实为方程组(3.3.1)的非负解.

任取(3.3.1)的一组非负解$\{g_i(u),i\in S\}$, 由

$$g_i(u)=u\sum_{j\in A}p_{i,j}+u\sum_{k\notin A}p_{i,k}g_k(u)$$

可知

$$g_i(u)\ge A(i,1,u)=u\sum_{j\in A}p_{i,j}$$

对任意$i\in S$成立. 设对任意$i\in S$, $g_i(u)\ge A(i,k,u)$对 $k=n-1$成立, 那么由

$$A(i,n,u)=\sum_{j\in A}up_{i,j}+\sum_{k\notin A}up_{i,k}A(k,n-1,u)\le\sum_{j\in A}up_{i,j}+\sum_{k\notin A}up_{i,k}g_k(u)=g_i(u)$$

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

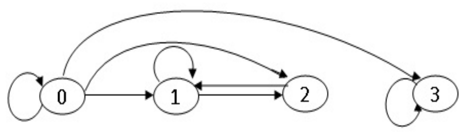
$$g_i(u)\ge A(i,n,u).$$

令$n\to\infty$ 得$g_i(u)\ge F_{i,A}(u)$对所有$i\in S$成立. 因此$F_{i,A}(u)$为(3.3.1)的最小非负解.□

定理10.1为我们提供了一种通过求解代数方程来获得τ_A分布的方法. 特别, 若X是有限状态马氏链, 那么由定理10.1, 我们可以通过求解含有限个变量的线性方程组获得τ_A的分布信息.

例10.1 设马氏链X的转移概率矩阵为$\begin{pmatrix}1/4 & 1/4 & 1/4 & 1/4\\ 0 & 1/3 & 2/3 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$. 求$F_{i,1}(u)$以及$\mathbb{P}_0(\tau_1=n)$, 其中$i=0,1,2,3$, n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9.1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$f_{0,1}^{(1)} = 1/4, \quad f_{0,1}^{(k)} = 5/4^k, k \ge 2.$ □

例10.2设$W=\{W_n\}$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9.1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

$$F_{0,1}(u) = up + uqF_{-1,1}(u). \tag{3.3.2}$$

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

$$\begin{aligned} f_{i,i+j}^{(n)} &= \mathbb{P}(\tau_{i+j} = n | W_0 = i) = \mathbb{P}(W_n = i+j, W_k < i+j, 0<k<n | W_0 = i) \\ &= \mathbb{P}(W_n - W_0 = j, W_k - W_0 < j, 0<k<n) \\ &= \mathbb{P}(W_n = j, W_k < j, 0<k<n | W_0 = 0) = f_{0,j}^{(n)}. \end{aligned}$$

特别

$$\begin{aligned} f_{0,2}^{(n)} &= \mathbb{P}(W_n = 2, W_k < 2, 0<k<n | W_0 = 0) \\ &= \sum_{j=1}^{n-1} \mathbb{P}(W_n = 2, W_k < 2, j<k<n, W_j = 1, W_l < 1, 0<l<j | W_0 = 0) \end{aligned}$$

由马氏性

$$\begin{aligned}&\mathbb{P}(W_n=2,W_k<2,j<k<n,W_j=1,W_l<1,0<l<j|W_0=0)\\&\quad=\mathbb{P}(W_j=1,W_l<1,0<l<j|W_0=0)\mathbb{P}(W_n=2,W_k<2,j<k<n|W_j=1)\end{aligned}$$

注意到W还是时齐的,

$$\begin{aligned}&\mathbb{P}(W_n=2,W_k<2,j<k<n|W_j=1)\\&\quad=\mathbb{P}(W_{n-j}=2,W_k<2,0<k<n-j|W_0=1)=f_{1,2}^{(n-j)}.\end{aligned}$$

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

$$\begin{aligned}F_{-1,1}(u)\;=\;F_{0,2}(u)&=\sum_{n=2}^{\infty}f_{0,2}^{(n)}u^n=\sum_{n=2}^{\infty}\sum_{j=1}^{n-1}f_{0,1}^{(j)}u^jf_{0,1}^{(n-j)}u^{n-j}\\&=\sum_{j=1}^{\infty}f_{0,1}^{(j)}u^j\sum_{k=1}^{\infty}f_{0,1}^{(k)}u^k=\big[F_{0,1}(u)\big]^2\end{aligned}$$

将其带入(3.3.2), 注意到$F_{0,1}(u)$为最小非负解, 解得

$$F_{0,1}(u)=\frac{2pu}{1+\sqrt{1-4pqu^2}}.\qquad\square$$

由定理10.1容易得到如下关于访问概率的常用结果.

推论10.2 对任意给定的$j\in S$, $\{f_{i,j};i\in S\}$是线性方程组

$$z_i=\sum_{k\neq j}p_{i,k}z_k+p_{i,j},\qquad i\in S,\tag{3.3.3}$$

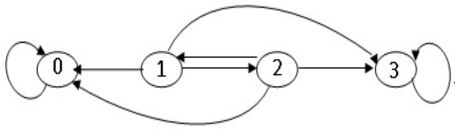
的最小非负解.

例10.3 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). 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 由转移概率矩阵我们可以画出状态一步转移图如下：



所求比例为概率f_{10}. 由于$3 \not\to 0$, $f_{3,0}=0$. 由推论9.2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$f_{10}=7/8$, 即最终付清得比例为87.5%. □

方程组(3.3.1)与(3.3.3)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10.3 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(3.3.3)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9.9和性质9.7可知对任意$i,j\in S$, $f_{i,j}=1$. 从而方程组(3.3.1)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(3.3.3)的最小非负解, 从而由推论10.2可知$f_{j,j}=1$, 即j为常返的, 再由推论9.9 可知不可约链X为常返的. □

例10.4 设马氏链X的状态空间为非负整数, 转移概率矩阵为

$$\mathbf{P}=\begin{array}{cc} \begin{pmatrix} r_0 & p_0 & 0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{matrix} \end{array}$$

其中$p_0>0$, $r_0+p_0=1$, $q_i+r_i+p_i=1$, $p_i,q_i>0$, $i\geq 1$. 通常我们称X为(广义)带反射壁的随机游动. 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. 取$j=0$, 考察如下方程组

$$\begin{cases} z_0=\sum_{k\neq 0}p_{0,k}z_k+p_{0,0}=p_0z_1+r_0 \\ z_1=\sum_{k\neq 0}p_{1,k}z_k+p_{1,0}=r_1z_1+p_1z_2+q_1 \\ z_n=\sum_{k\neq 0}p_{z,k}z_k+p_{n,0}=q_nz_{n-1}+r_nz_n+p_nz_{n+1}, \quad n\geq 2. \end{cases} \tag{3.3.4}$$

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(3.3.4)的非负解必满足$z_0=p_0z_1+r_0$, 而且对任意$n\geq 1$,

$$z_{n+1}=z_1+(z_1-1)\sum_{k=1}^{n}\frac{q_1\cdots q_k}{p_1\cdots p_k}.$$

因此(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10.3, 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

$$0<z_0=p_0z_1+r_0<1\quad 而且\quad 0<z_1+A(z_1-1)<z_n\leq z_1<1$$

对所有$n\geq 1$都成立, 即(3.3.3)存在非负解$\{z_i,i\geq 0\}$使得$z_1<1$. 此时由定理9.3可知X是非常返的. □

(B) 平均访问时间

对任意$i\in S$和$A\subset S$ 我们还可以定义

$$m_{i,A}=\sum_{n=1}^{\infty}nf_{i,A}^{(n)}. \tag{3.3.5}$$

当$f_{i,A}=1$, 即$\mathbb{P}_i(\tau_A<\infty)=1$时,

$$m_{i,A}=\sum_{n=1}n\mathbb{P}_i(\tau_A=n)=\mathbb{E}_i(\tau_A)$$

表示从i出发首次到达状态j的平均时间. 特别, 若i 是常返的, $m_{i,i}$表示首次回到i的平均时间, 称为平均回访时间.

注10.1 当$f_{i,A}<1$时上面的概率解释不成立, 但$m_{i,A}$仍在一定条件下刻画了τ_A的"均值"信息.

注意到

$$m_{i,A}=\frac{\mathrm{d}F_{i,A}(u)}{\mathrm{d}u}\Big|_{u=1}.$$

因此若解方程(3.3.1)得到函数$F_{i,A}(u)$, 我们可以通过求导运算方便地算出$m_{i,A}$.

例10.5 对例10.1中X, 求$m_{i,1}$, 其中$i=0,1,2,3$.

解 对例10.1中结果$F_{i,1}(u)$求导, 并在导函数中令$u=1$可得

$$m_{0,1}=11/9,\ \ m_{1,1}=5/3,\ \ m_{2,1}=1,\ \ m_{3,1}=0.$$ □

通过求解方程组(3.3.1)得到函数$F_{i,A}(u)$有时候并不方便, 下面的定理给出了通过解方程直接得到$m_{i,A}$的方法.

定理10.4 对任意给定的$A\subset S$, $\{m_{i,A};i\in S\}$是线性方程组

$$z_i=\sum_{k\notin A}p_{i,k}z_k+f_{i,A},\qquad i\in S \tag{3.3.6}$$

的最小非负解.

证明 记$b_{i,A}(k)=\sum_{m=1}^{n} m f_{i,A}^{(m)}$, 则$n\to\infty$时$b_{i,A}(n)\uparrow m_{i,A}$. 由

$$\begin{aligned}
b_{i,A}(n+1) &= \sum_{m=1}^{n+1} m f_{i,A}^{(m)} = \sum_{m=1}^{n} m f_{i,A}^{(m+1)} + \sum_{m=1}^{n+1} f_{i,A}^{(m)} \\
&= \sum_{m=1}^{n+1} f_{i,A}^{(m)} + \sum_{m=1}^{n}\sum_{k\notin A} m\mathbb{P}_i(X_{m+1}\in A, X_1=k, X_v\notin A, 2\le v\le m) \\
&= \sum_{m=1}^{n+1} f_{i,A}^{(m)} + \sum_{k\notin A} p_{i,k}\sum_{m=1}^{n} m f_{k,A}^{(m)} = \sum_{k\notin A} p_{i,k} b_{k,A}(n) + \sum_{m=1}^{n+1} f_{i,A}^{(m)}.
\end{aligned}$$

两边令$n\to\infty$得

$$m_{i,A} = f_{i,A} + \sum_{k\neq j} p_{i,k} m_{k,A}.$$

因此$\{m_{i,A}\}_{i\in S}$确实为方程组(3.3.6)的非负解. 任取(3.3.6)的一组非负解$\{u_i\}_{i\in S}$, 由

$$u_i = f_{i,A} + \sum_{k\notin A} p_{i,k} u_k$$

可知$u_i \ge b_{i,A}(1) = f_{i,A}^{(1)}$对任意$i\in S$成立.

设对任意$i\in S$, $u_i\ge b_{i,A}(k)$对$k=n-1$成立, 那么由

$$b_{i,A}(n) = \sum_{m=1}^{n+1} f_{i,A}^{(m)} + \sum_{k\notin A} p_{i,k} b_{k,A}(n-1) \le f_{i,j} + \sum_{k\notin A} p_{i,k} u_k = u_i$$

可知对任意$i\in S$, $u_i\ge b_{i,A}(k)$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都有

$$u_i \ge b_{i,A}(n).$$

令$n\to\infty$ 得$u_i\ge m_{i,A}$对所有$i\in S$成立, 因此$m_{i,A}$为(3.3.6)的最小非负解. □

例10.6 设有$n+1, n\ge 1$,个选手依次两两进行比赛, 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$\{X_k\}$表示在第k次比赛结束时比赛胜者连赢场次, $X_0=0$. 容易看出$X=\{X_k,\ k\ge 0\}$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$f_{i,j}=1, i,j\ge 1$. 比赛结束的条件是$X_k=n$, 比赛结束的时间为

$$\tau_n=\inf\{k\ge 1, X_k=n\}$$

从而平均比赛次数

$$T=\mathbb{E}_0(\tau_n)=m_{0,n}.$$

显然$m_{0,n}=m_{1,n}+1$, 由定理9.4可知$m_{i,n}, 1\le i<n$,是方程组

$$\begin{cases}z_i=z_1/2+z_{i+1}/2+1, \quad 1\le i<n-1,\\ z_{n-1}=z_1/2+1.\end{cases}$$

的最小非负解. 由此可得

$$\begin{cases}z_i=2z_{i-1}-z_1-2, \qquad 2\le i\le n-1\\ z_{n-1}=z_1/2+1.\end{cases}$$

因此对任意$2\le i\le n-1$,

$$\begin{aligned}z_i &= 2^2z_{i-2}-2z_1-2^2-z_1-2\\ &= 2^{i-1}z_1-(2^{i-2}+\cdots+2+1)z_1-(2^{i-1}+\cdots+2)\\ &= z_1-(2^i-2)\end{aligned}$$

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10.7 设X是带反射壁的(q,p)-随机游动, 即在例9.4中$r_0=0$, $p_0=1$; 而对一切$i\ge 1$, $r_i=0$, $q_i=q$, $p_i=p$. 若$q>p$, 求$m_{n,0}$, $n\ge 0$.

解 由例10.4可知$q>p$时, X常返, 对$n\ge 0$, $f_{n,0}=1$. $\{m_{n,0}\}_{n\ge 0}$为下列方程组的最小非负解:

$$\begin{cases}z_0=\sum_{k\ne 0}p_{0,k}z_k+1=z_1+1\\ z_1=\sum_{k\ne 0}p_{1,k}z_k+1=pz_2+1\\ z_n=\sum_{k\ne 0}p_{n,k}z_k+1=qz_{n-1}+pz_{n+1}+1, \quad n\ge 2\end{cases} \tag{3.3.7}$$

由此可得

$$\begin{cases}z_0=\sum_{k\ne 0}p_{0,k}z_k+1=z_1+1\\ z_2-z_1=rz_1-\frac{1}{p}\\ z_{n+1}-z_n=r(z_n-z_{n-1})-\frac{1}{p}, \quad n\ge 2\end{cases}$$

其中$r=q/p$. 所以对$n\geq 1$

$$z_{n+1}-z_n=r^nz_1-\frac{1}{p}[1+r+\cdots+r^{n-1}]=r^nz_1-\frac{1}{p}\frac{r^n-1}{r-1},$$

进而

$$z_{n+1}=z_1\sum_{k=0}^{n}r^k-\frac{1}{p(r-1)}(\sum_{k=0}^{n}r^k-n-1)=z_1\sum_{k=0}^{n}r^k-\frac{1}{q-p}(\sum_{k=0}^{n}r^k-n-1).$$

上式两边除以$\sum_{k=0}^{n}r^k$得,

$$\frac{z_{n+1}}{\sum_{k=0}^{n}r^k}=z_1-\frac{1}{q-p}\left(1-\frac{n+1}{\sum_{k=0}^{n}r^k}\right),n\geq 1.$$

由$r=q/p>1$可知当$n\to\infty$时

$$\sum_{k=0}^{n}r^k\to\infty \text{ 而且 } n/\sum_{k=0}^{n}r^k\to 0.$$

因此作为(3.3.7)的非负解$z_n,n\geq 0$, 必有

$$z_1\geq\frac{1}{q-p},$$

从而

$$z_n\geq\frac{n}{q-p},\qquad n\geq 1.$$

直接验证可知

$$z_0=1+\frac{1}{q-p},\ z_n=n/(q-p),\ n\geq 1$$

为方程(3.3.7)的一组非负解从而为最小非负解. 因此由定理10.4

$$m_{0,0}=1+\frac{1}{q-p}=\frac{2q}{q-p};\ \ m_{n,0}=\frac{n}{q-p},n\geq 1.$$ □

值得注意的是, 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10.8 设X为(q,p)-随机游动, $q>p$, 求$m_{a,0}$其中$a>0$.

解: 将0的转移概率$p_{0,k}$重新定义为$p_{0,k}=\begin{cases}1, & k=1\\ 0 & k\neq 1\end{cases}$. 此时从$a>0$出发的随机游动即为带反射壁的$(q,p)$-随机游动. 因此由例10.7可知$m_{a,0}=a/(q-p)$. □

练习题

10.1 设X的转移概率矩阵$\mathbf{P}=\begin{pmatrix}0&1/2&1/2&0\\1/2&0&0&1/2\\0&0&1/3&2/3\\0&0&1/2&1/2\end{pmatrix}$, 求$(f_{i,j})$与$(m_{i,j})$.

10.2 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?

10.3 若马氏链X只有三个状态且$(m_{i,j})=\begin{pmatrix}2&3&3\\1&4&4\\1&4&4\end{pmatrix}$, 求$X$的转移概率矩阵$\mathbf{P}$.

10.4 设X是带反射壁的(q,p)-随机游动, $n>0$, (1) 若$q<p$, 求$f_{n,0}$, (2) 若$q>p$, 求$m_{0,n}$.

10.5 (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, 把它换成别的颜色后放回. 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?

§3.4 平稳分布与转移概率极限

上一节我们介绍了首访时间(回访周期)的“均值”$m_{i,j}$. 这是一个非常重要的概念, 不仅体现在应用上, 还是刻画马氏链性质的重要指标.

(A) 正常返, 零常返与转移概率极限

首先我们利用回访周期的平均值, 把常返状态进一步区分为.

定义11.1 设$i\in S$常返, 称i是正常返的, 若$m_{i,i}<\infty$; 称i是零常返的, 若$m_{i,i}=+\infty$.

直观而言状态i是正常返的意味着从i出发能“较大概率”地在短时间内回访自己, 而零常返则意味着虽然从i出发一定会回访自己，但从回访时间平均看, 可以说“遥不可及”.

定理11.1 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11.2 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$\{b_n;\ n\geq 0\}$为收敛数列, 那么

$$\lim_{n\to\infty}\frac{\sum_{k=0}^{n}a_kb_{n-k}}{\sum_{k=0}^{n}a_k}=\lim_{n\to\infty}b_n.$$

证明 由所设条件, 对任意给定$N\geq 1$, 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$n\geq N$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

$$\Big|\sum_{k=0}^{n}a_k(b_{n-k}-b)\Big|\leq\varepsilon\sum_{k=0}^{n-N}a_k+B\sum_{k=n-N+1}^{n}a_k$$

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$+\infty$, 则对一切$M>0$, 存在N, 使得当$n\geq N$时, $b_n\geq M$. 于是

$$\sum_{k=0}^{n}a_kb_{n-k}\geq M\sum_{k=0}^{n-N}a_k+\big(\min_{0\leq n\leq N}b_n\big)\sum_{k=n-N+1}^{n}a_k.$$

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11.1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

$$\{X_n = j\} = \bigcup_{k=0}^{n-1} \{X_n = j, X_k = i, X_v \neq i, v = k+1, \cdots, n-1\}.$$

从而, 由马氏性及(3.2.4)

$$\begin{aligned} p_{i,j}^{(n)} = \mathbb{P}_i(X_n = j) &= \mathbb{P}_i(\bigcup_{k=0}^{n-1} \{X_n = j, X_k = i, X_v \neq i, v = k+1, \cdots, n-1\}) \\ &= \sum_{k=0}^{n-1} p_{i,i}^{(k)} \mathbb{P}_i(X_n = j, X_v \neq i, v = k+1, \cdots, n-1 | X_k = i) \\ &= \sum_{k=0}^{n-1} p_{i,i}^{(k)} e_{ij}^{(n-k)}, \end{aligned} \tag{3.4.1}$$

进而

$$\begin{aligned} n &= \sum_{j \in S} \sum_{m=1}^{n} p_{i,j}^{(m)} = \sum_{j \in S} \sum_{m=1}^{n} \sum_{k=0}^{m-1} p_{i,i}^{(k)} e_{ij}^{(m-k)} \\ &= \sum_{j \in S} \sum_{k=0}^{n-1} p_{i,i}^{(k)} \sum_{m=1}^{n-k} e_{ij}^{(m)} = \sum_{k=0}^{n-1} p_{i,i}^{(k)} \sum_{m=1}^{n-k} \sum_{j \in S} e_{ij}^{(m)} \end{aligned}$$

注意到

$$\begin{aligned} \sum_{j \in S} e_{ij}^{(m)} &= \mathbb{P}_i(\bigcup_{j \in S} \{X_m = j, X_v \neq i, v = 1, \cdots, m\}) \\ &= \mathbb{P}_i(X_v \neq i, v = 1, \cdots, m-1) = \mathbb{P}_i(\tau_i \geq m). \end{aligned}$$

因此

$$n = \sum_{k=0}^{n-1} p_{i,i}^{(k)} \sum_{m=1}^{n-k} \mathbb{P}_i(\tau_i \geq m)$$

令$b_n = \sum\limits_{m=1}^{n} \mathbb{P}_i(\tau_i \geq m)$, 则$n = \sum\limits_{k=0}^{n-1} p_{i,i}^{(k)} b_{n-k}$. 若$i$正常返, $\lim\limits_{n\to\infty} b_n = m_{ii} < \infty$; 若$i$零常返或非常返, $\lim\limits_{n\to\infty} b_n = \infty$. 由引理11.2可知定理成立. □

由定理11.1, 容易得到如下结果, 证明简单, 此略.

推论11.3 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11.4 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11.4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11.1 直线上(q,p)-随机游动W当$p = q = 1/2$时是零常返的.
解 已知$p = q = 1/2$时W常返. 此时对$n = 2k, 2k+1$, 当$k \geq 1$充分大时由striling公式可知

$$\sum_{m=0}^{n} p_{0,0}^{(m)} = 1 + \sum_{m=1}^{k} \frac{C_{2m}^m}{4^m} \approx \sum_{m=1}^{k} \frac{1}{\sqrt{\pi m}}.$$

因此

$$\sum_{m=0}^{n} p_{0,0}^{(m)} \approx \sqrt{\frac{2n}{\pi}}.$$

这表明$\frac{1}{n}\sum\limits_{m=0}^{n} p_{0,0}^{(m)} \to 0$, 从而状态0为零常返的. 由$W$不可约可知$W$零常返. □

定理11.5 对任意$i, j \in S$以及$u \in [0,1]$,

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} p_{i,j}^{(k)} u^k}{\sum_{k=0}^{n} p_{j,j}^{(k)} u^k} = F_{i,j}(u).$$

证明 $u = 0$时结果显然成立. 下设$u \in (0,1]$. 由(3.2.5)可得对任意$n \geq 1$,

$$\sum_{k=1}^{n} p_{i,j}^{(k)} u^k = \sum_{k=1}^{n}\sum_{r=1}^{k} f_{ij}^{(r)} p_{j,j}^{(k-r)} u^k = \sum_{r=1}^{n} f_{ij}^{(r)} u^r \sum_{k=0}^{n-r} p_{j,j}^{(k)} u^k = \sum_{k=0}^{n-1} p_{j,j}^{(k)} u^k \sum_{r=1}^{n-k} f_{ij}^{(r)} u^r.$$

令

$$a_n = p_{jj}^{(n)} u^n, \quad b_n = \sum_{r=1}^{n} f_{ij}^{(r)} u^r.$$

那么$b_n \to F_{i,j}(u)$, 而且不论j是否常返引理11.2条件成立. 由此可知定理成立. □

推论11.6 对任意$i, j \in S$, $\frac{1}{n}\sum_{k=1}^{n} p_{i,j}^{(k)}$极限存在, 而且

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_{i,j}^{(k)} = \begin{cases} f_{ij}/m_{jj}, & j\text{正常返}, \\ 0, & j\text{为其他状态}. \end{cases}$$

证明 由定理11.1与定理11.5(令$u=1$)立刻可得. □

注11.1 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, $\lim\limits_{n\to\infty} p_{i,j}^{(n)}=0$. 后一结论是定理11.5的推论, 前两结论我们将在第五章第二节给出解释和证明.

对任意$i,j\in S$, 令$P_{i,j}(s)=\sum_{n=1}^{\infty} p_{i,j}^{(n)}s^n$. 由定理11.5和(3.2.6)可得

推论11.7 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$u=1$时

$$\sum_{n=1}^{\infty} p_{i,j}^{(n)}=P_{i,j}(1)=\begin{cases} f_{i,j}/(1-f_{j,j}), & f_{j,j}<1;\\ 0, & f_{i,j}=0\text{且}f_{j,j}=1;\\ \infty, & f_{i,j}>0\text{且}f_{j,j}=1.\end{cases}$$

(B) 平稳分布

定义11.2 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$i\in S$

$$u_i=\sum_{k\in S} u_k p_{k,i}. \tag{3.4.2}$$

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.

显然, 若$\mathbf{u}$为X的不变测度, 对任意$n\geq 1$, 由C-K方程, $u_i=\sum_{k\in S} u_k p_{k,i}^{(n)}$. 这表明若$X$以平稳分布$(\pi_i, i\in S)$为初始分布, 那么$X$在每一个时刻的分布都与初始分布相同. 这也是我们称其为“平稳分布”的原因所在.

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11.2 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. □

不是所有的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11.8 设X为不可约马氏链, X是正常返的当且仅当X存在平稳分布$(\pi_i, i\in S)$. X的平稳分布唯一而且

$$\pi_i = 1/m_{i,i} > 0, \quad i \in S.$$

证明 先证必要性. 设X是正常返的, 那么对任意$i,j\in S$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_{j,j}^{(k)} = \frac{1}{n}\sum_{k=1}^{n} p_{i,j}^{(k)} = 1/m_{j,j} > 0.$$

注意到

$$\frac{1}{n}\sum_{k=1}^{n} p_{j,j}^{(k)} = \frac{1}{n}\sum_{k=1}^{n}\sum_{i\in S} p_{j,i}^{(k-1)} p_{i,j} = \sum_{i\in S}\Big[\frac{1}{n}\sum_{k=1}^{n} p_{j,i}^{(k-1)}\Big] p_{i,j}.$$

两边令$n\to\infty$,

$$\frac{1}{m_{j,j}} \ge \lim_{N\to\infty}\sum_{i\in A_N}\frac{p_{i,j}}{m_{ii}} = \sum_{i\in S}\frac{p_{i,j}}{m_{i,i}}.$$

其中A_N表示S中元素个数为N的有限子集且

$$A_N \subset A_{N+1} \subset \cdots, \qquad \bigcup_N A_n = S.$$

另一方面, 由

$$1 = \sum_{i \in S} \Big[\frac{1}{n} \sum_{k=1}^{n} p_{j,i}^{(k)} \Big]$$

可知

$$1 \geq \lim_{N \to \infty} \sum_{i \in A_N} \frac{1}{m_{i,i}} = \sum_{i \in S} \frac{1}{m_{i,i}}.$$

因此

$$1 \geq \sum_{j \in S} \frac{1}{m_{j,j}} \geq \sum_{j \in S} \sum_{i \in S} \frac{p_{i,j}}{m_{i,i}} \geq \lim_{N \to \infty} \sum_{i \in A_N} \Big[\lim_{M \to \infty} \sum_{j \in A_M} \frac{p_{i,j}}{m_{i,i}} \Big] = \sum_{i \in S} \frac{1}{m_{i,i}}.$$

这表明对任意$j \in S$,

$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{p_{i,j}}{m_{i,i}},$$

即$(\frac{1}{m_{j,j}}, j \in S)$为$X$的不变测度. 进一步可得

$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{1}{m_{i,i}} \Big[\frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)} \Big],$$

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,

$$\frac{1}{m_{j,j}} = \sum_{i \in S} \frac{1}{m_{i,i} m_{j,j}}. \tag{3.4.3}$$

这表明$\sum\limits_{i \in S} 1/m_{i,i} = 1$, 即$(1/m_{j,j}, j \in S)$为$X$ 的平稳分布.

再证充分性. 设$(\pi_j, j \in S)$是X的任一平稳分布. 对任意$j \in S$,

$$\pi_j = \sum_{i \in S} \pi_i p_{i,j} = \sum_{i \in S} \pi_i \Big[\frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)} \Big]. \tag{3.4.4}$$

由推论10.6

$$\pi_j = \sum_{i \in S} \pi_i \Big[\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} p_{i,j}^{(k)} \Big] = \begin{cases} [\sum\limits_{i \in S} \pi_i f_{i,j}]/m_{j,j}, & j\text{正常返} \\ 0, & \text{其他}. \end{cases}$$

另一方面由$\sum\limits_{j\in S}\pi_j=1$可知存在$j_0\in S$使得$\pi_{j_0}>0$. 由于$j_0\to j$, 存在$N$使得$p_{j_0,j}^{(N)}>0$. 因此

$$\pi_j=\sum_{i\in S}\pi_i p_{i,j}=\sum_{i\in S}\pi_i p_{i,j}^{(N)}\ge \pi_{j_0}p_{j_0,j}^{(N)}>0.$$

因此j正常返, 又由于X是不可约的, 因此X是正常返链, 而且由于对任意$i,j\in S$, $f_{i,j}=1$, 因此$0<\pi_j=1/m_{j,j}$, 这表明平稳分布必唯一. □

对不可约的马氏链而言, 由定理11.8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在任何时刻的分布都保持不变. 更有意思的是:

定理11.9 若$X=\{X_n\}$是不可约正常返马氏链, 记其平稳分布为$\pi=(\pi_i,i\in S)$, 其中$\pi_i=1/m_{ii}$. 那么对任意$i,j\in S$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}p_{i,j}^{(k)}=\pi_j. \tag{3.4.5}$$

进一步, 若X还是非周期的, 任意给定X的一个初始分布$\mu=(\mu_i,i\in S)$, 那么随着$n\to\infty$, X_n的分布收敛到π, 即对任意$j\in S$, $\mathbb{P}(X_n=j)\to\pi_j$.

证明 由于X是不可约常返的, 由性质8.7可知, 对任意$i,j\in S$, $f_{i,j}=1$. 由推论10.6可知前半部分结论成立.

进一步, 若X还是非周期的, 由注10.1可知, 对任意的$i,j\in S$,

$$p_{i,j}^{(n)}\to\frac{1}{m_{jj}}=\pi_j.$$

因此当X的初始分布为μ时,

$$\begin{aligned}\mathbb{P}(X_n=j)&=\sum_{i\in S}\mathbb{P}(X_n=j,X_0=i)\\&=\sum_{i\in S}\mathbb{P}(X_0=i)\mathbb{P}(X_n=j|X_0=i)=\sum_{i\in S}\mu_i p_{i,j}^{(n)}.\end{aligned}$$

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$n\to\infty$时, 我们可得

$$\lim_{n\to\infty}\mathbb{P}(X_n=j)=\lim_{n\to\infty}\sum_{i\in S}\mu_i p_{i,j}^{(n)}=\sum_{i\in S}(\mu_i\lim_{n\to\infty}p_{i,j}^{(n)})=\sum_{i\in S}\mu_i\pi_j=\pi_j.$$

即定理后半部分也成立. □

例11.3 (例8.5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$\mathbf{Q}^T\pi=\pi'$, 即

$$\begin{cases}0.8\pi_1+0.3\pi_2+0.05\pi_3=\pi_1\\0.1\pi_1+0.4\pi_2+0.1\pi_3=\pi_2\\\pi_1+\pi_2+\pi_3=1\end{cases}$$

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. □

此外, 注意到$N_n(j)=\sum_{k=1}^n\mathbf{1}_{\{X_k=i\}}$表示过程$X$在时刻$n$之前(含$n$)到达状态$j$的次数.

$$\mathbb{E}(N_n(j))=\mathbb{E}(\sum_{k=1}^n\mathbf{1}_{\{X_k=i\}})=\sum_{k=1}^n\mathbb{P}(X_k=i)$$

表示X在时刻n之前(含n)到达状态j平均次数, $\mathbb{E}(N_n(j))/n$则表示到达状态j的平均比率. 由(3.4.5), 对X的任意初始分布$\mu=(\mu_i,i\in S)$,

$$\frac{\mathbb{E}(N_n(j))}{n}=\frac{1}{n}\sum_{k=1}^n\mathbb{P}(X_k=i)=\frac{1}{n}\sum_{k=1}^n\sum_{i\in S}\mu_ip_{i,j}^{(k)}=\sum_{i\in S}\Big[\mu_i\Big(\frac{1}{n}\sum_{k=1}^np_{i,j}^{(k)}\Big)\Big].$$

令$n\to\infty$得

$$\frac{\mathbb{E}(N_n(j))}{n}\to\pi_j.$$

这表明平稳分布还刻画了马氏链在长时间内访问各状态的比率.

例11.4 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. 问一个顾客长期消费后在各个品牌上的消费比例?

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$\mathbf{Q}^T\pi=\pi'$, 即

$$\begin{cases}0.8\pi_1+0.2\pi_2+0.1\pi_3=\pi_1\\0.1\pi_1+0.6\pi_2+0.1\pi_3=\pi_2\\\pi_1+\pi_2+\pi_3=1\end{cases}$$

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. □

由定理11.8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11.10 不可约马氏链X是正常返的当且仅当方程组

$$u_i=\sum_{k\in S}u_kp_{k,i}$$

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$\{u_i/\sum_{i\in S}u_i\}$为平稳分布, 从而推论成立. □

例11.5 讨论例10.4中广义带反射壁随机游动X的正常返性.

解 由例10.4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

$$z_i=\sum_{k=0}^{\infty}z_kp_{k,i},\quad i\geq 0.$$

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$n\geq 0$, $q_{n+1}z_{n+1}=p_nz_n$, 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11.9, 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

11.1 证明任意不可约有限状态马氏链必是正常返的.

11.2 设不可约链X存在状态j使得方程组

$$z_i=\sum_{k\neq j}p_{i,k}z_k+f_{ij},\qquad i\neq j$$

的一组有限解$\{u_i, i \neq j\}$满足
$$\sum_{k\neq j} p_{j,k}u_k < \infty.$$
若j 是常返的, 证明j 是正常返的.
11.3 设X是一个不可约非常返的马氏链链, 令$l_j = \sup\{n \geq 0, X_n = j, X_k \neq j, k \geq n\}$, 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$l_j$的分布列和均值.([提示] 均值为$m_{ii}/(1 - f_{ii})$)
11.4 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.
11.5 设X的转移概率满足：对任意$i, j \in S = \{0, 1, 2, \cdots\}$,
$$p_{i,j} = \begin{cases} \lambda_i + (1 - \lambda_i)u_i, & i = j, \\ (1 - \lambda_i)u_j, & i \neq j, \end{cases}$$
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$\sum\limits_{i=0}^{\infty} u_i/(1 - \lambda_i) < \infty$.

11.6 设X为不可约正常返马氏链, 利用等式(3.4.1)及推论10.10证明, 对任意$i, j \in S$, $\pi_j = \pi_i e_{ij}$, 其中$\{\pi_i; i \in S\}$ 为X的平稳分布, $e_{ij} = \sum\limits_{n=1}^{\infty} e_{ij}^{(n)}$.
11.7 证明不可约正常返马氏链的不变测度μ必是有限测度(即$\sum\limits_{i\in S} \mu(i) < \infty$).

第四章　马氏链的一些简单应用

马氏链在实际问题中有着广泛应用. 这一章我们试着用马氏链及相关理论建立一些应用模型，解决一些应用问题.

§4.1　存储与排队模型

(A) 存储模型

先看一个马氏链在商品存储决策中简单应用的例子. 为此我们先给出一个重要的极限定理(遍历性定理).

定理12.1 记不可约正常返马氏链$X=\{X_n;\ n\geq 0\}$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

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}g(X_k)=\sum_{j\in S}g(j)\pi_j:=\mathbb{E}_\pi(g),\qquad a.s.$$

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9.11易知$\{W_i(k),k\geq 2\}$为i.i.d序列, 且$\mathbb{E}(W_i(2))=m_{i,i}<\infty$. 由强大数定律

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}W_i(k)=\lim_{n\to\infty}\frac{1}{n}\tau_i(n)=m_{i,i},\qquad a.s.$$

再令

$$Y_k=\sum_{n=\tau_i(k)+1}^{\tau_i(k+1)}g(X_m).$$

仍由定理9.11可知$Y_k,k\geq 1$, 为独立同分布随机变量. 注意到

$$\begin{aligned}\mathbb{E}(Y_1)&=\mathbb{E}(\sum_{m=\tau_i(k)+1}^{\tau_i(k+1)}g(X_m))=\mathbb{E}_i(\sum_{m=1}^{\tau_i(1)}g(X_m))=\mathbb{E}_i(\sum_{m=1}^{\tau_i}g(X_m))\\&=g(i)+\mathbb{E}_i(\sum_{m=1}^{\tau_i-1}g(X_m))=g(i)+\sum_{n=1}^{\infty}\mathbb{P}_i(\tau_i=n)\mathbb{E}_i(\sum_{m=1}^{n-1}g(X_m)|\tau_i=n)\\&=g(i)+\sum_{n=1}^{\infty}\mathbb{P}_i(\tau_i=n)\sum_{m=1}^{n-1}\sum_{j\neq i}g(j)\mathbb{P}_i(X_m=j|\tau_i=n)\\&=g(i)+\sum_{n=1}^{\infty}\sum_{m=1}^{n-1}\sum_{j\neq i}g(j)\mathbb{P}_i(X_m=j,\tau_i=n)\end{aligned}$$

由于是非负项求和, 我们可以交换n,m的求和次序, 得

$$\begin{aligned}\mathbb{E}(Y_1) &= g(i)+\sum_{m=1}^{\infty}\sum_{j\neq i}g(j)\mathbb{P}_i(X_m=j,\tau_i\geq m+1)\\ &= g(i)+\sum_{m=1}^{\infty}\sum_{j\neq i}g(j)e_{i,j}^{(m)}=g(i)+\sum_{j\neq i}g(j)e_{i,j}.\end{aligned}$$

由习题11.5及定理11.7可知

$$\mathbb{E}(Y_1)=\sum_{j\in S}g(j)\pi_j/\pi_i=m_{i,i}\mathbb{E}_\pi(g)<\infty.$$

由强大数定律

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}Y_k=m_{i,i}\mathbb{E}_\pi(g),\qquad a.s. \tag{4.1.1}$$

记$N_n=\sum_{k=1}^{n}\mathbf{1}_{\{X_k=i\}}$. 显然$N_n$单调非降,记极限$N$. 对任意$m>0$,

$$\begin{aligned}\mathbb{P}(N\geq m) &= \mathbb{P}(\sum_{k=1}^{\infty}\mathbf{1}_{\{X_k=i\}}\geq m)\\ &= \sum_{r=1}^{\infty}\mathbb{P}(\tau_i(1)=r,\sum_{k=r+1}^{\infty}\mathbf{1}_{\{X_k=i\}}\geq m-1).\end{aligned}$$

再由定理9.11蕴含的独立性可得

$$\begin{aligned}\mathbb{P}(N\geq m) &= \sum_{r=1}^{\infty}\mathbb{P}(\tau_i(1)=r)\mathbb{P}_i(\sum_{k=1}^{\infty}\mathbf{1}_{\{X_k=i\}}\geq m-1)\\ &= \sum_{r=1}^{\infty}\mathbb{P}(\tau_i(1)=r)\mathbb{P}_i(N\geq m-1)\\ &= \mathbb{P}(\tau_i(1)<\infty)\mathbb{P}_i(N\geq m-1).\end{aligned}$$

注意到i是常返的,

$$\mathbb{P}(N\geq m)=\mathbb{P}_i(N\geq m-1)=\cdots=\mathbb{P}_i(N\geq 1)=1.$$

这表明$N_n\to\infty$, a.s. 由习题3.2可知

$$\tau_i(N_n)/N_n\to m_{ii}\quad a.s.\ \text{以及}\ \frac{1}{N_n}\sum_{k=1}^{N_n}Y_k=m_{ii}\mathbb{E}_\pi(g),\quad a.s.$$

注意到

$$\tau_i(N_n)\leq n\leq\tau_i(N_n+1).$$

由
$$\frac{\tau_i(N_n)}{N_n} \le \frac{n}{N_n} < \frac{\tau_i(N_n+1)}{N_n},$$
以及
$$\sum_{k=0}^{N_n-1} Y_k = \sum_{m=1}^{\tau_i(N_n)} g(X_m) \le \sum_{m=1}^{n} g(X_m) \le \sum_{m=1}^{\tau_i(N_n+1)} g(X_m) = \sum_{k=1}^{N_n} Y_k,$$
进而
$$\frac{N_n-1}{n}\Big[\frac{1}{N_n-1}\sum_{k=0}^{N_n-1} Y_k\Big] \le \frac{1}{n}\sum_{m=1}^{n} g(X_m) \le \frac{N_n}{n}\Big[\frac{1}{N_n}\sum_{k=1}^{N_n} Y_k\Big],$$
可知结论成立. □

在定理12.1假设前提下, 进一步我们可得如下推论.

推论12.2 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$\sum\limits_{i\in S} g(i)\pi_i < \infty$, 那么推论12.2就是定理12.1 的直接应用. 若$\sum\limits_{i\in S} g(i)\pi_i = \infty$, 那么对任意$N$, 令$g_N(i) = i \wedge N$. 对$g_N$ 直接应用定理11.1 得
$$\liminf_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} X_k \ge \limsup_{N\to\infty}\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} g_N(X_k) = \limsup_{N\to\infty}\mathbb{E}_\pi g_N(i) = \sum_{i\in S} g(i)\pi_i = \infty.$$
因此推论12.2成立. □

例12.1 考虑例8.1中的(k,K)商品存储模型X，其中$K=3$, 表示需求的随机变量ξ满足
$$p_0 = \mathbb{P}(\xi=0) = 0.3, p_1 = \mathbb{P}(\xi=1) = 0.4, p_2 = \mathbb{P}(\xi=2) = 0.2, p_3 = \mathbb{P}(\xi=3) = 0.1$$
若每销售一件商品可以获利12元，每件商品存储一个单位时间需花费2元，试确定最优的k使得从长期看平均收益最大.

解 对任意给定的$k\in\{0,1,2\}$, 根据进货策略，在第n个单位时间获得的净利润r_n为
$$r_n = \begin{cases} 12(3-X_n) - 2X_n = 36 - 14X_n, & X_{n-1} \le k \\ 12(X_{n-1}-X_n) - 2X_n = 12X_{n-1} - 14X_n, & k < X_{n-1} \le 3 \end{cases} =: g(X_{n-1}, X_n).$$
由假设,X的转移概率矩阵$(p_{i,j})_4$满足
$$p_{i,j} = \begin{cases} \mathbb{P}((3-\xi)\vee 0 = j) = \mathbb{P}(\xi = 3-j) = p_{3-j}, & i \le k, \\ \mathbb{P}((i-\xi)\vee 0 = j) = p_{i-j}, & k < i \le 3, 0 < j \le i \\ \mathbb{P}((i-\xi)\vee 0 = j) = p_i + \cdots + p_3, & k < i \le 3, j = 0 \\ \mathbb{P}((i-\xi)\vee 0 = j) = 0, & k < i \le 3, i < j \le 3 \end{cases}$$

显然X是有限状态不可约马氏链, 从而正常返，记其平稳分布为

$$\pi=(\pi_0,\pi_1,\pi_2,\pi_3).$$

由习题12.3及定理12.1可知平均利润

$$\begin{aligned}\frac{1}{n}\sum_{m=1}^{n} r_m=\frac{1}{n}\sum_{m=1}^{n} g(X_{m-1},X_m)\to r(k)&=\sum_{i=0}^{3}\sum_{j=0}^{3} g(i,j)\pi_i\, p_{i,j}\\ &=\sum_{i=0}^{k}\sum_{j=0}^{3}(36-14j)\pi_i p_{3-j}+\sum_{i=k+1}^{3}\sum_{j=0}^{3}(12i-14j)\pi_i p_{i,j}\\ &=9.4\sum_{i=0}^{k}\pi_i+12\sum_{i=k+1}^{3} i\pi_i-14\sum_{i=k+1}^{3}\Big(\pi_i\sum_{j=1}^{i} jp_{i-j}\Big).\end{aligned}$$

直接计算可知, 当$k=2$时

$$r(2)=9.4\sum_{i=0}^{2}\pi_i+36\pi_3-14\pi_3\sum_{j=1}^{3} jp_{3-j}=9.4\sum_{i=0}^{3}\pi_i=9.4.$$

当$k=1$时由

$$(p_{i,j})=\begin{pmatrix}0.1&0.2&0.4&0.3\\0.1&0.2&0.4&0.3\\0.3&0.4&0.3&0\\0.1&0.2&0.4&0.3\end{pmatrix}$$

可求得平稳分布

$$\pi=(19/110,30/110,40/110,21/110).$$

此时平均收益极限

$$\begin{aligned}r(1)&=9.4(\pi_0+\pi_1)+12(2\pi_2+3\pi_3)-14(\pi_2+1.9\pi_3)\\&=9.4+0.6\pi_2\approx 9.62.\end{aligned}$$

当$k=0$时由

$$(p_{i,j})=\begin{pmatrix}0.1&0.2&0.4&0.3\\0.7&0.3&0&0\\0.3&0.4&0.3&0\\0.1&0.2&0.4&0.3\end{pmatrix}$$

可求得平稳分布

$$\pi=(343/1070,300/1070,280/1070,147/1070).$$

此时平均收益极限

$$\begin{aligned}r(0)&=9.4\pi_0+12(\pi_1+2\pi_2+3\pi_3)-14(0.3\pi_1+\pi_2+1.9\pi_3)\\&=9.4-1.6\pi_1+0.6\pi_2\approx 9.11.\end{aligned}$$

因此$(1,3)$策略是长期平均利润准则下的最优策略. □

(B) 排队模型

我们现在考虑马氏链在服务系统中应用的例子. 在介绍这个例子之前我们需要两个判断马氏链是否(正)常返的工具.

定理12.3 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$i\neq j$,

$$\sum_{k\in S}p_{i,k}y_k\leq y_i, \tag{4.1.2}$$

而且$\lim\limits_{i\to+\infty}y_i=+\infty$, 那么$X$为常返的.

证明 不失一般性可设对任意$i\in S$, $y_i>0$, 令$\tilde{p}_{i,s}=\begin{cases}p_{i,s}, & i\neq j,\\ \delta_{j,s}, & i=j\end{cases}$, Y是转移概率矩阵为$\tilde{\mathbf{P}}=(\tilde{p}_{i,j})$的时齐马氏链. 则对$Y$而言,

(1)$\tilde{f}_{ij}=\mathbb{P}_i(\tau_j(Y)<\infty)=\mathbb{P}_i(\tau_j(X)<\infty)=f_{i,j}$.

(2) j为吸收态. 对任意$i\in S$, 若$i\neq j$, 由$i\to j$可知i非常返, 注11.1表明对任意$k\in S$, $\tilde{p}_{k,i}^{(n)}\to 0$; 若$i=j$, 由习题9.3可知$\tilde{p}_{k,j}^{(n)}\to\tilde{f}_{kj}$.

对任给的$i\in S$, $i\neq j$, 迭代(4.1.2)可得,

$$\sum_{k\in S}\tilde{p}_{i,k}^{(n)}y_k\leq y_i.$$

对任意$\varepsilon>0$, 由$y_n\to\infty$可知, 存在充分大的$N>j$, 使得

$$\inf_{n>N}y_n\geq y_i/\varepsilon.$$

由

$$\inf_{n>N}y_n\sum_{k>N}\tilde{p}_{i,k}^{(n)}\leq\sum_{k>N}\tilde{p}_{i,k}^{(n)}y_k\leq y_i$$

可得

$$\sum_{k>N}\tilde{p}_{i,k}^{(n)}<\varepsilon.$$

因此

$$\sum_{0\leq k\leq N}\tilde{p}_{i,k}^{(n)}>1-\varepsilon$$

令$n\to\infty$并结合上面的观察(2)可知$\tilde{f}_{i,j}\geq 1-\varepsilon$. 再令$\varepsilon\to 0$得$\tilde{f}_{i,j}=1$. 由上面的观察(1)及定理10.1 可知$f_{jj}=1$. 因此$j$是常返的，进而不可约链$X$是常返的. □

定理12.4 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

$$\begin{cases}\sum\limits_{k\in S}p_{i,k}z_k\leq z_i-1, & i\neq j,\\ \sum\limits_{k\in S}p_{j,k}z_k<\infty.\end{cases} \tag{4.1.3}$$

证明"必要性"：令$z_j=0$以及$z_i=m_{ij}$, $i\neq j$. 由定理10.3.

$$\sum_{k\in S}p_{j,k}z_k=\sum_{k\neq j}p_{j,k}m_{k,j}<m_{jj}<\infty$$

而对任意$i\neq j$,

$$\sum_{k\in S}p_{i,k}z_k=\sum_{k\neq j}p_{i,k}m_{k,j}=m_{ij}-f_{ij}=m_{ij}-1=z_i-1.$$

即存在(任)一个状态j以及非负数列$\{z_i, i\in S\}$使得(4.1.3)成立.

"充分性"：记$\delta=\sum_{k\in S}p_{j,k}z_k$, 则$0<\delta<\infty$. 任取$i\neq j$, 对任意$n\geq 2$, 迭代(4.1.3) 可得,

$$\begin{aligned}
0\leq \sum_k p_{i,k}^{(n)}z_k &= \sum_{k\in S}\sum_{l\in S}p_{i,l}^{(n-1)}p_{l,k}z_k=\sum_{l\in S}p_{i,l}^{(n-1)}\sum_{k\in S}p_{l,k}z_k\\
&= \sum_{l\neq j}p_{i,l}^{(n-1)}(z_l-1)+p_{i,j}^{(n-1)}\delta=\sum_{l\neq j}p_{i,l}^{(n-1)}z_l-1+p_{i,j}^{(n-1)}(1+\delta)\\
&\leq \sum_{l\in S}p_{i,l}^{(n-1)}z_l-1+p_{i,j}^{(n-1)}(1+\delta)\leq\cdots\\
&\leq z_i-n+\sum_{k=1}^{n-1}p_{i,j}^{(k)}(1+\delta).
\end{aligned}$$

这表明

$$\sum_{k=1}^{n-1}p_{i,j}^{(k)}(1+\delta)\geq n-z_i.$$

两边除n并令$n\to\infty$得

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}p_{i,j}^{(k)}\geq\frac{1}{1+\delta}>0.$$

由推论11.6可知j为正常返状态, 从而X是正常返的. □

例12.2 观察服务系统中的某一个窗口, 问平均会有多少人在排队？窗口没人的时间占多大比例？

要回答这样的实际问题, 我们通常需要做些假设. 设顾客相继随机到达, 一般认为达到的时间间隔构成一列独立同分布的随机变量序列, 顾客到达后通常按"先来先服务"的原则接受服务台的服务, 假设为每个顾客的服务时间也是独立同分布的随机变量且与顾客到达的时间间隔相互独立, 称满足这些假定的服务系统为$G/G/s$系统, 其中第一个G表示顾客到达的时间间隔分布, 第二个G表示服务时间分布, s为服务台(窗口)的数量.

为了简单, 现假设顾客到达时间间隔分布服从参数为λ的指数分布, 服务时间服从密度函数为g的连续分布, 只考虑一个窗口. 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$X=\{X_n;\ n\geq 0\}$为例6.2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$\{\xi_n;\ n\geq 1\}$ 是独立同分布的随机变量, 分布列为

$$p_k=\mathbb{P}(\xi_n=k)=\mathbb{E}(\mathbb{P}(\xi_n=k|U_n))=\int_0^\infty\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}g(t)\mathrm{d}t,$$

而且

$$X_n=\begin{cases}X_{n-1}-1+\xi_n, & X_{n-1}>0,\\ \xi_n, & X_{n-1}=0.\end{cases}$$

此时X的转移概率$p_{i,j}$满足条件：对任意$j\geq 0$, $p_{0,j}=p_j$. 而$i>0$时，对任意$j\geq i-1$, $p_{i,j}=p_{j-i+1}$, 对任意$0\leq j<i-1$, $p_{i,j}=0$. 一般地, 我们称X为$M(\lambda)/G/1$排队系统的嵌入链, 所求问题可以通过计算如下统计量获得有用信息

$$(1)\ \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n X_k,\qquad\qquad (2)\ \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{X_k=0\}}.$$

下面我们总设$0<p_0$, $p_0+p_1<1$, $\rho=\sum_{k=1}^\infty kp_k=\mathbb{E}(\xi_1)<\infty$. 对任意$s\in[0,1]$, 令

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

表示服务期间到达顾客分布的矩母函数. 显然$\rho=A'(1)$而且由习题12.2可知, $\rho>1$时, 方程$A(s)=s$在$(0,1)$内存在唯一解. 我们首先指出

命题12.5 若$\rho>1$, X非常返; 若$\rho=1$, X零常返; 若$\rho<1$, X 正常返.

证明 当$\rho>1$时, 记$A(s)=s$在$(0,1)$中的唯一解为u, 那么对任意$i\geq 0$, 令$z_i=u^i$, 那么

$$\begin{aligned}z_i&=u^i=u^{i-1}A(u)=\sum_{k=0}^\infty p_k u^{k+i-1}\\&=\sum_{k=i-1}^\infty p_{k-i+1}z_k=\sum_{k=i-1}^\infty p_{i,k}z_k=\sum_{k\neq 0}p_{i,k}z_k+p_{i,0},\quad i\neq 0\end{aligned}$$

由于$\{f_{i,0}\}_{i\neq 0}$是方程组

$$z_i=\sum_{k\neq 0}p_{i,k}z_k+p_{i,0},\quad i\neq 0$$

的最小非负解, 因此$f_{i,0}<1$. 这表明状态0非常返, 从而X非常返.

当$\rho\leq 1$时, 对任意$i\geq 1$, 令$y_i=i$,

$$\begin{aligned}y_i=i-1+1&\geq i-1+\rho=i-1+\sum_{k=0}^\infty kp_k=\sum_{k=0}^\infty(k+i-1)p_k\\&=\sum_{k=i-1}^\infty p_{k+1-i}k=\sum_{k\in S}p_{i,k}y_k.\end{aligned}$$

由定理12.3可知此时X是常返的.

当$\rho < 1$时, 由

$$i - 1 + \rho = \sum_{k=0}^{\infty} p_{i,k} k, \quad i \geq 1$$

可得

$$\frac{i}{1-\rho} - 1 = \sum_{k=0}^{\infty} p_{1,k} \frac{k}{1-\rho}$$

且

$$\sum_{k=0}^{\infty} p_{0,k} \frac{k}{1-\rho} = \frac{\sum_{k=0}^{\infty} k p_k}{1-\rho} = \frac{1}{1-\rho} < \infty.$$

由定理12.4可知, 此时X正常返.

当$\rho = 1$时, 若X正常返，则记其不变测度为$\pi = \{\pi_k\}_{k\geq 0}$. 对任意$s \in [0,1]$, 令

$$\pi(s) = \sum_{k\geq 0} \pi_k s^k.$$

那么由

$$\pi_i = \sum_{k\in S} \pi_k p_{k,i} = \pi_0 p_i + \sum_{k=1}^{i+1} \pi_k p_{i+1-k}$$

可知对任意$0 < s < 1$,

$$\begin{aligned}\pi(s) &= \sum_{i\geq 0} \pi_i s^i = \pi_0 \sum_{i\geq 0} p_i s^i + \sum_{k=1}^{\infty}\sum_{m=0}^{\infty} \pi_k p_m s^{m+k-1} \\ &= \pi_0 A(s) + \frac{1}{s}(\pi(s) - \pi_0) A(s).\end{aligned}$$

由此可得

$$\pi(s) = \frac{\pi_0 (s-1) A(s)}{s - A(s)} \tag{4.1.4}$$

以及

$$\frac{A(s)-1}{s-1} = 1 - \pi_0 \frac{A(s)}{\pi(s)}.$$

令$s \to 1-$得

$$\rho = A'(1) = 1 - \pi_0 \tag{4.1.5}$$

因此$\rho = 1$意味着$\pi_0 = 0$从而状态0非正常返, 这与X正常返假设矛盾. □

注意到X正常返条件下(4.1.4)是平稳分布的矩母函数，结合(4.1.5)可知

$$\pi(s)=(1-\rho)\frac{(1-s)A(s)}{s-A(s)},$$

从而

$$\pi'(s)=(1-\rho)\Big\{\frac{(s-1)A'(s)}{s-A(s)}+\frac{A(s)[s-A(s)-(1-A'(s))(s-1)]}{(s-A(s))^2}\Big\}$$

由推论12.2, 推论11.6, 当$\rho<1$时

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbb{P}(X_k=0)=\pi_0=1-\rho,$$

进一步若要求$A''(1-)=\mathbb{E}(\xi_1^2)-\mathbb{E}(\xi_1)<\infty$, 由$s\to 1$时$s-A(s)\sim(1-\rho)(s-1)$可得

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k=\sum_{k=0}^{\infty}k\pi_k=\pi'(1-)=\rho+\frac{A''(1-)}{2(1-\rho)}.$$

练习题

12.1 在例11.1中若令$p_0=p_1=p_2=p_3=0.25$, 其他条件不变, 求最佳策略$(k,3)$.
12.2 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.
12.3* 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$\{\pi_ip_{i,j}\}_{(i,j)\in\mathcal{M}}$.
12.4* 设X是状态空间$S=\{0,1\}$的两状态时齐马氏链, 转移概率矩阵

$$\mathbb{P}=\begin{pmatrix} p & 1-p\\ 1-q & q\end{pmatrix},$$

其中$0<p,q<1$. 若已知$X_0,\cdots,X_n$的一组观察值$x_0,\cdots,x_n$, 其中依时间先后次序相邻的两个观察值为$00,01,10,11$的分别有$n_{00},n_{01},n_{10},n_{11}$对.
(1) 证明p, q的极大似然估计分别是

$$\hat{p}_n=\frac{n_{00}}{n_{00}+n_{01}},\quad \hat{q}_n=\frac{n_{11}}{n_{10}+n_{11}}.$$

(2) 证明$\lim_{n\to\infty}\hat{p}_n=p,a.s.$

§4.2 可逆马氏链与蒙特卡罗模拟*

一个用Markov链描述的系统, 常常可以利用它的长时间后的稳定分布, 作为其参数的稳定设计的依据. 在前面我们已经看到, 在很宽的条件下, 不变分布就是这个稳定分布. 而求不变分布需要解一个方程, 一般来说需要很大的工作量. 所幸的是, 在实际中有很多问题满足一种对称性条件,即所谓可逆性条件. 对于满足这种条件的Markov 链, 有一种求不变分布的简便方法. 在本节中, 我们将给出这方面的理论和应用.

(A) 可逆马氏链

定义13.1 以$\mathbf{P} = (p_{i,j})_{i,j\in S}$ 为转移矩阵的Markov链X, 称为可逆的, 如果存在概率分布$\mu = \{\mu_i : i \in S\}$使得对任意$i, j \in S$,

$$\mu_i p_{i,j} = \mu_j p_{j,i}. \tag{4.2.1}$$

μ称为X的可逆初分布, 简称可逆分布. 物理学中, 条件(4.2.1)也被称为细致平衡条件.

由此定义可以看出, 若不可约马氏链X可逆, 那么可逆分布μ必惟一而且由

$$\sum_i \mu_i p_{i,j} = \sum_i \mu_j p_{j,i} = \mu_j.$$

可知μ是X的平稳分布. 这表明可逆不可约链的平稳分布可以利用方程(4.2.1)求出.

例13.1 考察状态$S = \{0, 1, \cdots, M\}$, 转移概率为

$$\begin{aligned} 0 < p_{i,i+1} &= \alpha_i = 1 - p_{i,i-1} < 1, i = 1, \cdots, M-1, \\ 0 < p_{0,1} &= \alpha_0 = 1 - p_{0,0}, p_{M,M} = \alpha_M = 1 - p_{M,M-1} < 1, \end{aligned}$$

的马氏链X, 求X的平稳分布.

解 容易看出X是不可约的. 令$\nu_0 = c$, 其中c为待定常数. 对任意$M > i \geq 0$,

$$\nu_i p_{i,i+1} = \nu_{i+1} p_{i+1,i},$$

即对任意$M > i \geq 0$

$$\nu_i \alpha_i = \nu_{i+1}(1 - \alpha_{i+1}).$$

由此可得, 对任意$M > i \geq 0$

$$\nu_{i+1} = \prod_{k=1}^{i+1} \frac{\alpha_{k-1}}{1-\alpha_k} \nu_0.$$

再令$\sum_{i=0}^M \nu_i = 1$可得

$$c = \frac{1}{1 + \sum_{i=1}^M \prod_{k=1}^i \frac{\alpha_{k-1}}{1-\alpha_k}}.$$

此时$\nu = (\nu_0, \nu_1, \cdots)$为$X$的可逆分布, 进而也是$X$的平稳分布. □

命题13.1 记可逆不可约马氏链X的可逆分布为$\mu = \{\mu_i : i \in S\}$. 那么对任意$i \in S$, $\mu_i > 0$, 而且对任意$j \in S$, $p_{i,j} = 0$当且仅当$p_{j,i} = 0$.

证明 由(4.2.1)可知对任意$n \geq 1$,

$$\mu_i p_{i,j}^{(n)} = \mu_j p_{j,i}^{(n)}. \tag{4.2.2}$$

由于μ为分布以及X不可约, 存在$i_0 \in S$使得$\mu_{i_0} > 0$ 以及对任意$i_0 \neq i \in S$, 存在$n \geq 0$使得$p_{i_0,i}^{(n)} > 0$. 将其代入上式可知

$$\mu_i p_{i,i_0}^{(n)} > 0.$$

这表明$\mu_i > 0$. 再由(4.2.1)可知对任意$j \in S$, $p_{i,j} = 0$当且仅当$p_{j,i} = 0$. □

由定理10.8可知可逆不可约马氏链一定是正常返的; 反之不真.

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13.2可知

$$p_{j,i_n} \cdots p_{i_2,i_1} p_{i_1,i_0} > 0.$$

在利用(4.2.1)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$j \in S$

$$\mu_j = \frac{v(i_0, j)}{\sum_{j \in S} v(i_0, j)}. \tag{4.2.3}$$

下面定理为我们判断不可约正常返马氏链是否是可逆链提供了一个更直观的方法.

定理13.2 对于只要$p_{i,j}=0$就有$p_{j,i}=0$的不可约正常返马氏链X, 它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. 即对于一切状态$i,i_1,i_2,\cdots,i_n$有

$$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_n,i}=p_{i,i_n}p_{i_n,i_{n-1}}\cdots p_{i_1,i}. \tag{4.2.4}$$

证明 前面分析已给出了必要性,下证充分性. 任取$i,j\in S$使得$i\neq j$且$p_{i,j}>0$. 对任意$n\geq 1$, 由(4.2.4) 可知,

$$p_{i,i_1}p_{i_1,i_2}\cdots p_{i_{n-1},j}p_{j,i}=p_{i,j}p_{j,i_{n-1}}\cdots p_{i_2,i_1}p_{i_1,i}.$$

关于所有$i_1,\cdots,i_{n-1}$求和得

$$p_{i,j}^{(n)}p_{j,i}=p_{i,j}p_{j,i}^{(n)}.$$

进一步关于n求和并求平均得

$$\lim_{k\to\infty}\frac{1}{k}\sum_{n=1}^{k}p_{i,j}^{(n)}p_{j,i}=p_{i,j}\lim_{k\to\infty}\frac{1}{k}\sum_{n=1}^{k}p_{j,i}^{(n)}.$$

由于X不可约且正常返, 由推论10.6以及定理10.8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. □

定理13.3 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$\sum_{i\in S}\pi_i=1$. 一定存在S上的可逆不可约马氏链X, 使得π是X的平稳分布.

证明 任取一个S上不可约的马氏链Y使其转移概率矩阵为$\mathbf{Q}=(q_{i,j})_{i,j\in S}$满足

$$q_{i,j}=0\Leftrightarrow q_{j,i}=0.$$

对任意$i\neq j\in S$, 令

$$\alpha_{i,j}=\min\{\frac{\pi_j q_{j,i}}{\pi_i q_{i,j}},1\},\qquad p_{i,j}=q_{i,j}\alpha_{i,j},$$

以及

$$p_{i,i}=q_{i,i}+\sum_{j\neq i}q_{i,j}(1-\alpha_{i,j}).$$

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$\alpha_{j,i} = 1$ 且

$$\pi_i p_{i,j} = \pi_j q_{j,i} = \pi_j q_{j,i} \alpha_{j,i} = \pi_j p_{j,i},$$

而$\alpha_{i,j} = 1$时, 若$\pi_i q_{i,j} = \pi_j q_{j,i}$ 则$\alpha(j,i) = 1$,且

$$\pi_i p_{i,j} = \pi_i q_{i,j} = \pi_j q_{j,i} = \pi_j q_{i,j};$$

若$\pi_i q_{i,j} > \pi_j q_{j,i}$ 则$\alpha(j,i) < 1$,且

$$\pi_i p_{i,j} = \pi_i q_{i,j} = \pi_j q_{j,i} \alpha(j,i) = \pi_j q_{i,j}.$$

总之, 此时π为X的可逆分布, 从而为X的平稳分布. □

注13.1 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11.9可知π还是X的极限分布, 即X_n的分布收敛到π.

注13.2 常称定理13.3证明中构造转移矩阵$\mathbf{P}$进而得到可逆马氏链的方法为Hastings-Metropolis 算法, 其中称转移矩阵$\mathbf{Q} = (q_{i,j})_{i,j \in S}$为预选矩阵.

(B) 马氏链蒙特卡罗(MCMC)方法

在许多复杂的统计问题中,有时很难直接对各种统计量进行理论分析或计算, 需要借助随机模拟对问题进行讨论. 但有时直接模拟也会存在一些问题. 比如, 我们常常需要生成满足特定随机分布的随机数, 但当给定随机分布比较复杂, 特别当给定随机分布是高维分布且各分量相依时, 直接对随机分布取样得到随机数就难以实现; 再比如, 假定X是一个随机变量, 当我们直接计算诸如$\mathbb{E}(h(X))$这样的算式有困难时, 一种自然的模拟是借助大数定律, 通过生成一列与X同分布的随机数$x_1, x_2, \cdots, x_n$, 通过计算$\sum_{i=1}^n h(x_i)/n$来近似, 但当生成X的随机数有困难时, 这种方案也难以实现.

马氏链蒙特卡罗(MCMC)方法是一种能解决上述诸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11.9)和遍历性定理(定理12.1和推论12.2), 通过模拟马氏链的轨道(实现), 近似生成特定分布的随机数或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.

(B1) 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. 因此Metropolis采样对模拟或近似给定分布π的理论有效性可由定理13.3得到. 在对过程X具体采样时, 转移矩阵$\mathbf{P}$可以如下方式实现, 具体证明请读者自己完成.

(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) 若$\frac{\pi_j q_{j,i}}{\pi_i q_{i,j}} \geq 1$, 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$U[0,1]$的随机数U, 如果$U \leq \frac{\pi_j q_{j,i}}{\pi_i q_{i,j}}$, 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13.2 试用Metropolis采样法生成所有n级排列所组成集合$\mathfrak{F}$上的分布π使得对任意对任意$v \in \mathfrak{F}$, $\pi_v = cT(v)$, 其中$T(v)$表示排列v中数n的位置.
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n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

$$q_{v,u} = \begin{cases} 0, & u \notin N(v) \\ 1/|N(v)| = 1/n, & u \in N(v), \end{cases}$$

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n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

$$q(v,u) = 0 \Leftrightarrow q(u,v) = 0.$$

由于要得到的分布π 满足对任意$v \in \mathfrak{F}$, $\pi_v = c$为常数. 因此

$$\alpha(v,u) = \min\left(\frac{\pi_u q_{u,v}}{\pi_v q_{v,u}}, 1\right) = \min\left(\frac{T(u)}{T(v)}, 1\right).$$

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采样法得到π的一个近似分布, 基本流程如下

(s1) 给X_0任意赋予一个n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$n-1$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, 我们可以编写计算机程序模拟分布π, 请感兴趣的读者自己完成.

(B2) Gibbs 采样

Gibbs采样是MCMC模拟中使用的最为广泛的一种马氏链轨道采样方法. 它本质上仍是一种特殊的Hastings-Metropolis算法. Gibbs采样下的MCMC模拟步骤与基本假设如下:

设$X=(X_1,X_2,\cdots,X_n)$是一个多维离散随机向量. 记其所有可能的取值集合为S. 假设

(h1) X的分布满足：对任意$x=(x_1,x_2,\cdots,x_n)\in S$,

$$\mathbb{P}(X=x)=cg(x),$$

其中c为常数.

(h2) 对任意$1\le i\le n$ 以及任意x_j, 其中$1\le j\le n$ 且$j\neq i$, 条件分布

$$\mathbb{P}(X_i=\cdot|X_j=x_j,j\neq i)$$

存在且已知.

注13.4 由概率分布的归一性可知$c=[\sum_{x\in S}g(x)]^{-1}$, 但在很多时候$g(x)$的求和计算不容易实现, 此时我们可以认为$c$是未知的.

注13.5 上述条件(h2)表明在已知了X的任意$n-1$个分量取值后, 我们就可以一定的分布确定X的取值.

在以上条件假设下, Gibbs采样法模拟生成的向量值马氏链$Y=\{Y_n,n\ge 0\}$按如下的转移方式生成样本:

(s1) 假设目前的状态是$Y_k=x=(x_1,x_2,\cdots,x_n)$;

(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$\mathbb{P}(X_i=\cdot|X_j=x_j,j\neq i)$生成$x_i$的随机数, 比如$x_i=y$;

(s4) 将Y_{k+1}的值取成$y=(x_1,\cdots,x_{i-1},y,x_{i+1},\cdots,x_n)$并回到(s1).

记如上方式生成的马氏链Y的一步转移概率为$q(x,y)$, 那么

(q1) 当x,y至少有两个分量不同时$q(x,y)=0$;

(q2) 当x,y只有一个分量不同时, 假设不同的分量下标为i,

$$q(x,y)=\frac{1}{n}\mathbb{P}(X_i=y_i|X_j=x_j,j\neq i)=\frac{cg(y)}{n\mathbb{P}(X_j=x_j,j\neq i)};$$

(q3) 当$x=y$时

$$q(x,x)=1-\sum_{y\neq x}q(x,y).$$

容易检验此时X是一个不可约马氏链, $q(x,y)=0\Leftrightarrow q(y,x)=0$ 而且$q(x,x)>0$. 进一步, 此时Hastings-Metropolis算法中的函数

$$\alpha(x,y)=\min\left(\frac{cg(y)q(y,x)}{cg(x)q(x,y)},1\right)=\min\left(\frac{cg(y)cg(x)}{cg(x)cg(y)},1\right)=1,$$

这表明Gibbs抽样法得到的马氏链Y也是按Hastings-Metropolis算法构造的非周期不可约可逆马氏链并以X的分布为平稳分布, 因此当模拟的马氏链Y运行足够多的步骤后, 即k足够大后, Y_k的值就可以看作X 分布的一个取样点.

例13.3 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$x=(x_1,x_2,\cdots,x_{50})$,

$$\mathbb{P}(X=x)=c\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{49}x_ix_{i+1}}.$$

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n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$n-1$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, 我们也可以编写计算机程序模拟分布π, 请感兴趣的读者自己完成.

最后我们指出的是MCMC方法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. 对这些内容感兴趣的读者请参阅有关参考文献, 本书不再详述.

§4.3 分支过程

这一节我们介绍一类特殊的马氏模型——分支过程. 这类过程源于家族姓氏消亡的统计调查, 现在生物, 经济, 金融以及社会学研究等领域有着广泛应用. 一类理想的分支过程定义如下：

定义14.1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, 递归定义随机过程$\{X_n, n\geq 0\}$ 如下

$$X_{n+1}=\sum_{k=1}^{X_n} Y_{n+1,k}, \quad n\geq 0,$$

其中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. 称随机过程$\{X_n, n\geq 0\}$为G-W(Galton-Watson)分支过程.

若以X_n表示第n代的个体数，$Y_{n+1,k}$则表示第n代中第k个个体的后代数. 直观看, G-W分支过程理想地假设了家族中每个个体产生后代的行为是一样的且彼此无关. 容易证明此时$X=\{X_n; n\geq 0\}$是一个时齐马氏链, 转移概率

$$p_{i,j}=\mathbb{P}(Y_{1,1}+\cdots+Y_{1,i}=j).$$

由定义可知$p_{0,0}=1$, 即状态0为吸收态. 记$p_k=\mathbb{P}(Y=k)$. 为避免平凡, 以下总设$p_1<1$. 若$p_0>0$, 则对任意i, $p_{i,0}=p_0^i>0$，这表明$i\to 0$. 假定$X_0=1$,

$$\tau_0=\inf\{n\geq 1, X_n=0\}$$

那么所求概率

$$q=\mathbb{P}_1(\tau_0<\infty)=f_{10}=\lim_{n\to\infty} p_{1,0}^{(n)}.$$

注意到若初始时刻有k个个体, 要使姓氏消亡, 意味着每个个体及其后代的姓氏都得消亡, 又由于每个个体及其后代的表现是相互独立的, 因此

$$f_{k,0}=\mathbb{P}_k(\tau_0<\infty)=f_{10}^k=q^k, \qquad p_{k,0}^{(n)}=\mathbb{P}_k(X_n=0)=(p_{1,0}^{(n)})^k.$$

命题14.1 若每个个体产生后代的平均值不超过1, 即$m=\mathbb{E}(Y)\leq 1$, 那么X_n必然消亡, 即$q=1$.

证明 由定理10.1可知

$$q=f_{1,0}=\sum_{k=1}^{\infty} p_{1,k}f_{k,0}+p_{1,0}=\sum_{k=1}^{\infty} p_k q^k+p_0.$$

令$A(s)=\sum_{k=0}^{\infty} p_k s^k$. 那么$q$满足方程$q=A(q)$. 由于$0\leq q\leq 1$, 而由习题14.2可知该方程在$[0,1)$内无解. 直接验算可知$q=1$. □

命题14.2 若每个个体产生后代的平均值超过1, 即$m=\mathbb{E}(Y)>1$, 那么X_n以正概率不消亡, 即$q<1$.

证明 由命题12.6可知q是$A(s)=s$在区间$[0,1]$上一个解. 注意到0是吸收状态, 对任意$n\ge 2$

$$p_{1,0}^{(n)}=\sum_{k=1}^{\infty}p_{1,k}p_{k,0}^{(n-1)}+p_{1,0}=\sum_{k=1}^{\infty}p_k(p_{1,0}^{(n-1)})^k+p_0=A(p_{1,0}^{(n-1)}),$$

其中$p_{1,0}^{(1)}=p_0$. 对方程$s=A(s)$的任意非负解u以及任意n, 都满足

$$\begin{aligned}u &= A^n(u)=A^{n-1}(A(u))\\ &\ge A^{n-1}(A(p_0))=A^{n-1}(p_{1,0}^2)\\ &\ge \cdots\ge A(p_{1,0}^{(n)})=p_{1,0}^{(n+1)}.\end{aligned}$$

令$n\to\infty$, 得$q=f_{1,0}\le u$. 这表明q是$s=A(s)$的最小非负解. 由习题14.2可知当$m>1$时, $A(s)=s$在$[0,1)$内有唯一非负解, 从而$0<q<1$. □

对GW模型而言, 0是吸收态而所有非零状态都是非常返的(参见习题14.1), 因此若X不灭绝则意味着X_n要趋向正无穷, 直观而言就是人口会爆炸. 这种两极化现象显然与人们的直观感知不一致. 一种对GW 模型简单的修正是引进所谓的“移民”, 得到如下的带移民GW过程.

定义14.2设$\{Y_{n,k},n\ge 1,k\ge 1\}$为一族独立同分布的非负整数值的随机变量, $\{I_n,n\ge 1\}$也是一族独立同分布的非负整数值随机变量且与$\{Y_{n,k},n\ge 1,k\ge 1\}$ 独立, 称随机过程$\{X_n,n\ge 0\}$ 为带移民的G-W过程, 如果

$$X_{n+1}=\sum_{k=1}^{X_n}Y_{n+1,k}+I_{n+1},\quad n\ge 0,$$

其中X_0是给定的非负整数值随机变量且与$\{Y_{n,k},n\ge 1,k\ge 1\}$与$\{I_n,n\ge 1\}$独立.

直观看, 在带移民的G-W过程模型中I_n表示第n代移入的人口数而且移入数量与n之前的的人口数无关, 与第n代新生的人口也无关. 由独立性假设, 容易验证此时$X=\{X_n,n\ge 0\}$仍为时齐马氏链. 记$p_k=\mathbb{P}(Y=k)$, $q_k=\mathbb{P}(I_1=k)$. 为简单, 我们约定$p_0>0$, $q_k>0$对所有$k\ge 0$ 成立. 对任意$i\ge 0$, 由

$$p_{i,0}=\mathbb{P}(Y_{11}+\cdots+Y_{1i}+I_1=0)=p_0^iq_0>0,\quad p_{0,i}=\mathbb{P}(I_1=i)=q_i>0$$

可知X是不可约非周期的.

记$m=\mathbb{E}(Y)$, $\sigma^2=\mathrm{Var}(Y)$, $\lambda=\mathbb{E}(I_1)$, $\beta^2=\mathrm{Var}(I_1)$. 由习题14.3可知

命题14.3 对带移民的G-W分支过程X, 如下的条件矩公式成立.

(1) $\mathbb{E}(X_n|X_{n-1})=mX_{n-1}+\lambda$;

(2) $\mathbb{E}(X_n^2|X_{n-1})=(mX_{n-1}+\lambda)^2+\sigma^2X_{n-1}+\beta^2$.

命题14.4 若$m<1$, $0<\lambda<\infty$, 那么X正常返.

证明 若X不是正常返的, 那么由注11.1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

$$\mathbb{P}_1(X_n \geq M) = 1 - \sum_{k=0}^{M-1} p_{1,k}^{(n)} > 1/2,$$

从而$\mathbb{E}_1(X_n) > 1 + \frac{\lambda}{1-m}$. 但另一方面, 由命题14.3中(1)可知

$$\begin{aligned}\mathbb{E}_1(X_n) &= \mathbb{E}_1(\mathbb{E}(X_n|X_{n-1})) = m\mathbb{E}_1(X_{n-1}) + \lambda \\ &= \cdots = \lambda(1 + \cdots + m^{n-2}) + m^{n-1} \leq \frac{\lambda}{1-m} + 1.\end{aligned}$$

矛盾表明X一定正常返. $\square$

下面我们在σ^2和β^2有限条件下考虑m, λ的参数估计.

总设$m < 1$, $0 < \lambda < \infty$. 记此时的平稳分布为$\pi = \{\pi_i\}_{i \geq 0}$. 注意$X_{n-1}$已知条件下, X_n 的最佳估计是条件期望

$$\mathbb{E}(X_n|X_{n-1}) = mX_{n-1} + \lambda.$$

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的最小二乘估计量为

$$\hat{m}_n = \frac{n\sum_{k=1}^n X_{k-1}X_k - \sum_{k=1}^n X_{k-1}\sum_{k=1}^n X_k}{n\sum_{k=1}^n X_{k-1}^2 - (\sum_{k=1}^n X_{k-1})^2},$$

以及

$$\hat{\lambda}_n = \frac{1}{n}(\sum_{k=1}^n X_k - \hat{m}_n \sum_{k=1}^n X_{k-1}).$$

由推论12.2以及习题12.3可知

$$\begin{aligned}\hat{m}_n &= \frac{\frac{\sum_{k=1}^n X_{k-1}X_k}{n} - \frac{\sum_{k=1}^n X_{k-1}}{n}\frac{\sum_{k=1}^n X_k}{n}}{\frac{\sum_{k=1}^n X_{k-1}^2}{n} - (\frac{\sum_{k=1}^n X_{k-1}}{n})^2} \\ &\to \frac{\sum_{i=0}^\infty i\pi_i \sum_{j=0}^\infty jp_{i,j} - (\sum_{i=0}^\infty i\pi_i)^2}{\sum_{i=0}^\infty i^2\pi_i - (\sum_{i=0}^\infty i\pi_i)^2} \\ &= \frac{\sum_{i=0}^\infty i(mi+\lambda)\pi_i - (\sum_{i=0}^\infty i\pi_i)^2}{\sum_{i=0}^\infty i^2\pi_i - (\sum_{i=0}^\infty i\pi_i)^2}, \quad a.s. \end{aligned} \tag{4.3.1}$$

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14.3(1),

$$\sum_{i=0}^{\infty} i\pi_i = \mathbb{E}_\pi(X_0) = \mathbb{E}_\pi(X_1) = \mathbb{E}_\pi(\mathbb{E}_\pi(X_1|X_0)) = \mathbb{E}_\pi(mX_0+\lambda)$$

因此

$$\sum_{i=0}^{\infty} i\pi_i = \frac{\lambda}{1-m}.$$

同样由命题14.3(2),

$$\begin{aligned}\sum_{i=0}^{\infty} i^2\pi_i &= \mathbb{E}_\pi(X_0^2) = \mathbb{E}_\pi(X_1^2) = \mathbb{E}_\pi(\mathbb{E}_\pi(X_1^2|X_0)) \\ &= m^2\mathbb{E}_\pi(X_0^2) + (2\lambda m+\sigma^2)\mathbb{E}_\pi(X_0) + \lambda^2 + \beta^2\end{aligned}$$

可得

$$\sum_{i=0}^{\infty} i^2\pi_i = (\frac{\sigma^2\lambda}{1-m}+\beta^2)\frac{1}{1-m^2} + \frac{\lambda^2}{(1-m)^2}.$$

因此

$$\begin{aligned}\sum_{i=0}^{\infty} i^2\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2 &= (\frac{\sigma^2\lambda}{1-m}+\beta^2)\frac{1}{1-m^2}, \\ \sum_{i=0}^{\infty} i(mi+\lambda)\pi_i - (\sum_{i=0}^{\infty} i\pi_i)^2 &= (\frac{\sigma^2\lambda}{1-m}+\beta^2)\frac{m}{1-m^2}.\end{aligned}$$

将以上结果代入(4.3.1)得

$$\hat{m}_n \to m, \quad a.s.$$

进而由定理12.1得

$$\hat{\lambda}_n \to \frac{\lambda}{1-m} - m\frac{\lambda}{1-m} = \lambda, \qquad a.s.$$

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.

练习题

14.1 不论p_0是否为零, 证明Galton-Watson模型所有非零状态都是非常返的.

14.2 设ξ为非负整数值随机变量, $\mathbb{P}(\xi=1)<1$, $A(s)$为ξ的矩母函数. 证明若$\mathbb{E}(\xi)>1$且$\mathbb{P}(\xi=0)>0$, 那么$A(s)=s$在$[0,1)$ 内有唯一解; 若$\mathbb{E}(\xi)\le 1$, 那么$A(s)=s$在$[0,1)$内无解.

14.3 证明命题14.3.

§4.4 隐马尔科夫链

本节我们简要介绍离散时间离散状态的隐马尔科夫模型.

(A) 隐马尔科夫链的定义与分布性质

定义15.1 设$X_n, Y_n, n \geq 0$, 是分别取值于集合S和W上的离散随机变量. 称$Z = \{Z_n = (X_n, Y_n), n \geq 0\}$为隐马尔科夫链, 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$n \geq 0$

$$\mathbb{P}(Z_{n+1} = (i_{n+1}, j_{n+1}) | Z_k = (i_k, j_k), 0 \leq k \leq n) = p_{i_n, i_{n+1}} q_{i_{n+1}, j_{n+1}}, \qquad (4.4.1)$$

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.

由定义可知所谓隐马氏链其实是一种具有特定转移概率机制的向量值时齐马氏链, 其转移概率

$$p_{(i_n, j_n),(i_{n+1}, j_{n+1})} = p_{i_n, i_{n+1}} q_{i_{n+1}, j_{n+1}}.$$

容易看出, 此时$X = \{X_n, n \geq 0\}$是一个以$\mathbf{P} = (p_{i,j})$为转移概率矩阵的时齐马氏链. 过程$Y = \{Y_n, n \geq 0\}$本身可以不是马氏的, 但在任意时刻n, 它是一个依赖于同期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. 在实际应用时, 模型Z中往往只有过程Y才是可观测、可记录的, 而马氏链X是隐含的、不可观测的. 这就是我们称Z 为隐马尔科夫链的主要原因.

隐马氏链在实际建模时会时常碰见.

例15.1 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$X = \{X_n\}$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$Z_n = (X_n, Y_n)$就是一个刻画股票价格波动率的隐马氏链模型.

例15.2 评估一台机器的生产状态. 将机器状态X_n分为好(状态1)和差(状态2)两种, 要求依据机器生产产品的质量情况Y_n实时评判机器的生产状态. 假定产品的质量分布只依赖于机器当前生产状态, 而机器状态的转变构成一个马氏链, 那么$Z_n = (X_n, Y_n)$就构成了一个隐马尔科夫链.

例15.3 设X_n是取值于状态空间$S = \{1, 2, \cdots, L\}$的Markov 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, $\{w_n\}$是独立同分布的随机干扰,只取有限个值,且与随机过程$\{X_n\}$独立. 那么(X_n, Y_n)就是一个隐Markov 模型.

以下总设$Z_n = (X_n, Y_n)$是一个隐马尔科夫链, 若已知X_0的初始分布$\pi = (\pi_i, i \in S)$, 那么由马氏链的有限维分布表示可知, 对任意$n \geq 0$ 以及$i_0, i_1, \cdots, i_n \in S$,

$j_0, j_1, \cdots, j_n \in W$,

$$\mathbb{P}(Z_k = (i_k, j_k), 0 \le k \le n) = \pi_{i_0} q_{i_0, j_0} \prod_{k=1}^{n} p_{i_{k-1}, i_k} q_{i_k, j_k}. \tag{4.4.2}$$

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$(\pi, \mathbf{P}, \mathbf{Q})$为隐马尔科夫链$Z$的参数组.

由(4.4.1), (4.4.2)和全概率公式, 条件概率公式可得

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = i_{n+1} | X_k = i_k, Y_k = j_k, 0 \le k \le n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_k = i_k, 0 \le k \le n) \\
&\quad = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) = p_{i_n, i_{n+1}}, \qquad (4.4.3) \\
&\mathbb{P}(Y_{n+1} = j_{n+1} | X_{n+1} = i_{n+1}, X_k = i_k, Y_k = j_k, 0 \le k \le n) \\
&\quad = \mathbb{P}(Y_{n+1} = j_{n+1} | X_{n+1} = i_{n+1}, X_k = i_k, 0 \le k \le n) \\
&\quad = \mathbb{P}(Y_{n+1} = j_{n+1} | X_{n+1} = i_{n+1}, Y_k = j_k, 0 \le k \le n) \\
&\quad = \mathbb{P}(Y_{n+1} = j_{n+1} | X_{n+1} = i_{n+1}) = q_{i_{n+1}, j_{n+1}}. \qquad (4.4.4)
\end{aligned}$$

进一步, 若给定的$j_0, j_1, \cdots, j_n \in W$, 对任意$0 \le m \le n$, 令

$$F_m(i) = \mathbb{P}(X_m = i, Y_k = j_k, 0 \le k \le m), \qquad i \in S.$$

那么由(4.4.2)和全概率公式

$$\begin{aligned}
F_0(i) &= \pi_i q_{i, j_0} \\
F_m(i) &= \sum_{i_k \in S, 0 \le k < m} \pi_{i_0} q_{i_0, j_0} \Big[\prod_{k=1}^{m-1} p_{i_{k-1}, i_k} q_{i_k, j_k} \Big] p_{i_{m-1}, i} q_{i, j_m} \\
&= q_{i, j_m} \sum_{i_{m-1} \in S} p_{i_{m-1}, i} F_{m-1}(i_{m-1}), \qquad m \ge 1. \qquad (4.4.5)
\end{aligned}$$

对任意$0 \le m < n$, 令

$$B_m(i) = \mathbb{P}(Y_n = j_n, \cdots, Y_{m+1} = j_{m+1} | X_m = i),$$

那么由(4.4.4), 条件概率公式和全概率公式

$$\begin{aligned}
B_n(i) &\equiv 1, \\
B_m(i) &= \sum_{i_k \in S, m < k \le n} p_{i, i_{m+1}} q_{i_{m+1}, j_{m+1}} \Big[\prod_{k=m+2}^{n} p_{i_{k-1}, i_k} q_{i_k, j_k} \Big] \\
&= \sum_{i_{m+1} \in S} p_{i, i_{m+1}} q_{i_{m+1}, j_{m+1}} B_{m+1}(i_{m+1}), \qquad m \le n-1. \qquad (4.4.6)
\end{aligned}$$

基于公式(4.4.5)和(4.4.6), 我们可以通过递推的方法得到$F_m(i)$和$B_m(i)$.

由$F_m(i)$和$B_m(i)$的定义可得

$$\begin{aligned}\mathbb{P}(Y_k=j_k,0\le k\le n) &= \sum_{i\in S}\mathbb{P}(X_n=i,Y_k=j_k,0\le k\le n)=\sum_{i\in S}F_n(i) \quad (4.4.7)\\ &= \sum_{i\in S}\mathbb{P}(X_0=i,Y_k=j_k,0\le k\le n)\\ &= \sum_{i\in S}\pi_i q_{i,j_0}B_0(i). \quad (4.4.8)\end{aligned}$$

更一般地, 对任意$0\le m\le n$,

$$\begin{aligned}\mathbb{P}(Y_k=j_k,0\le k\le n) &= \sum_{i\in S}\mathbb{P}(X_m=i,Y_k=j_k,0\le k\le n)\\ &= \sum_{i\in S}\mathbb{P}(X_m=i,Y_k=j_k,0\le k\le m)\\ &\quad\times\mathbb{P}(Y_k=j_k,m<k\le n|X_m=i,Y_k=j_k,0\le k\le m)\\ &= \sum_{i\in S}F_m(i)\mathbb{P}(Y_k=j_k,m<k\le n|X_m=i)=\sum_{i\in S}F_m(i)B_m(i). \quad (4.4.9)\end{aligned}$$

因此

$$\mathbb{P}(X_m=i|Y_k=j_k,0\le k\le n)=\frac{F_m(i)B_m(i)}{\sum_{i\in S}F_m(i)B_m(i)}. \quad (4.4.10)$$

$$\begin{aligned}&\mathbb{P}(Y_{n+1}=j|Y_k=j_k,0\le k\le n)\\ &\quad=\sum_{i,i_{n+1}\in S}\mathbb{P}(Y_{n+1}=jX_{n+1}=i_{n+1},X_n=i|Y_k=j_k,0\le k\le n)\\ &\quad=\sum_{i,i_{n+1}\in S}\frac{p_{i,i_{n+1}}q_{i_{n+1},j}F_n(i)}{\sum_{i\in S}F_n(i)}. \quad (4.4.11)\end{aligned}$$

例15.4 在例15.1中, 为简单, 我们假设股票价格波动率分$-1,0,1$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, 股票价格波动率分$-1,0,1$ 的概率分别为$0.2,0.3,0.5$, 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

$$p_{-1,1}=1/4,\ p_{-1,-1}=3/4,\ p_{1,1}=4/5,\ p_{1,-1}=1/5.$$

若已知$\mathbb{P}(X_0=1)=4/5$以及3天的价格波动率$Y_0=1,Y_1=0,Y_2=-1$. 试问(1)第三个观察日股票在状态1的概率, (2)估计第四个观察日股票价格波动最可能的层次.

解 (1) 给定$j_0=1, j_1=0, j_2=-1$, 套用公式(4.4.5), 计算可知

$$\begin{aligned}
F_0(1)=\pi_1 q_{1,1}=0.8\times 0.5=0.4,\ \ F_0(-1)=\pi_{-1}q_{-1,1}=0.2\times 0.2=0.04;\\
F_1(1)=q_{1,0}[p_{-1,1}F_0(-1)+p_{1,1}F_0(1)]=0.099,\\
F_1(-1)=q_{-1,0}[p_{-1,-1}F_0(-1)+p_{1,-1}F_0(1)]=0.022;\\
F_2(1)=q_{1,-1}[p_{-1,1}F_1(-1)+p_{1,1}F_1(1)]=0.01684,\\
F_2(-1)=q_{-1,-1}[p_{-1,-1}F_1(-1)+p_{1,-1}F_1(1)]=0.02178.
\end{aligned}$$

因此由(4.4.10)(取$m=n=2$),

$$\mathbb{P}(X_2=1|Y_0=1,Y_1=0,Y_2=-1)=\frac{0.01684}{0.01684+0.02178}=\frac{842}{1931}\approx 0.4355.$$

(2) 为了计算方便, 我们令

$$\begin{aligned}
&a(1)=p_{1,1}q_{1,1}+p_{1,-1}q_{-1,1}=0.44,\ a(-1)=p_{-1,1}q_{1,1}+p_{-1,-1}q_{-1,1}=0.275;\\
&b(1)=p_{1,1}q_{1,0}+p_{1,-1}q_{-1,0}=0.28,\ b(-1)=p_{-1,1}q_{1,0}+p_{-1,-1}q_{-1,0}=0.225;\\
&c(1)=p_{1,1}q_{1,-1}+p_{1,-1}q_{-1,-1}=0.28,\ c(-1)=p_{-1,1}q_{1,-1}+p_{-1,-1}q_{-1,-1}=0.5.
\end{aligned}$$

由(4.4.11)可得(取$n=2$)

$$\begin{aligned}
\mathbb{P}(Y_3=1|Y_0=1,Y_1=0,Y_2=-1)&=\frac{a(1)F_2(1)+a(-1)F_2(-1)}{F_2(1)+F_2(-1)}\approx 0.3469;\\
\mathbb{P}(Y_3=0|Y_0=1,Y_1=0,Y_2=-1)&=\frac{b(1)F_2(1)+b(-1)F_2(-1)}{F_2(1)+F_2(-1)}\approx 0.2490;\\
\mathbb{P}(Y_3=-1|Y_0=1,Y_1=0,Y_2=-1)&=\frac{c(1)F_2(1)+c(-1)F_2(-1)}{F_2(1)+F_2(-1)}\approx 0.4041.
\end{aligned}$$

因此第四个观察日股票价格波动最可能的层次为-1.

(B) 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$\{Y_k,k\le n\}$, 需要估计$X_m$, $m\le n$的最佳值. 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的最佳值是在给定观察序列条件下X_m以最大概率取到的值.

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(4.4.10)

$$\mathbb{P}(X_m=i|Y_k=j_k,0\le k\le n)=\frac{F_m(i)B_m(i)}{\sum_{i\in S}F_m(i)B_m(i)}$$

可知, 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

$$\begin{aligned}
&\max_{i_k\in S,0\le k\le n}\mathbb{P}(X_k=i_kY_k=j_k,0\le k\le n)\\
&\quad=\max_{i_k\in S,0\le k\le n}[p_{i_{n-1},i_n}q_{i_n,j_n}\mathbb{P}(X_k=i_kY_k=j_k,0\le k<n)]\\
&\quad=\max_{i_n\in S}\left\{q_{i_n,j_n}\max_{i_k\in S,0\le k<n}\left[p_{i_{n-1},i_n}\mathbb{P}(X_k=i_kY_k=j_k,0\le k<n)\right]\right\},
\end{aligned}\tag{4.4.12}$$

可知, 对任意$0\le m\le n$, 若令

$$\begin{aligned}
M_m(i)&=\max_{i_k\in S,0\le k<m}\mathbb{P}(X_m=i,X_k=i_k,Y_k=j_k,0\le k<m)\\
&=\max_{i_k\in S,0\le k<m}\left[p_{i_{m-1},i}\mathbb{P}(X_k=i_kY_k=j_k,0\le k<m)\right],
\end{aligned}$$

则由马氏性可得

$$\begin{aligned}
M_m(i)&=\max_{i_k\in S,0\le k<m}\left[p_{i_{m-1},i}q_{i_{m-1},j_{m-1}}p_{i_{m-2},i_{m-1}}\mathbb{P}(X_k=i_kY_k=j_k,0\le k<m-1)\right]\\
&=\max_{i_{m-1}\in S}[p_{i_{m-1},i}q_{i_{m-1},j_{m-1}}M_{m-1}(i_{m-1})],
\end{aligned}\tag{4.4.13}$$

并且由(4.4.12)可得

$$\max_{i_k\in S,0\le k\le n}\mathbb{P}(X_k=i_kY_k=j_k,0\le k\le n)=\max_{i\in S}[q_{i,j_n}M_n(i)].\tag{4.4.14}$$

注意到对任意$i\in S$, $M_0(i)=\pi_i$其中π是X的初始分布, 利用(4.4.13), 由递归计算可得任意的$M_m(i)$, 进而由上式可知n及n之前所有时刻的X的最佳值与Y的观测值的联合概率.

为了得到获得n及n之前所有时刻的X的最佳值, 由(4.4.14)和(4.4.13)进一步可知n及n之前所有时刻的X的最佳值$i_0,i_1,\cdots,i_n$中

(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;

(2) 对任意$1\le m\le n$, 给定i_m, i_{m-1}就是使得函数

$$p_{i,i_m}q_{i,j_{m-1}}M_{m-1}(i),\qquad i\in S,$$

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15.5 (接例15.4) 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$M_0(1)=0.8, M_0(-1)=0.2$,

$$M_1(1)=\max_{i=-1,1} p_{i,1}q_{i,j_0}M_0(i)=\max_{i=-1,1} p_{i,1}q_{i,1}M_0(i)=0.32, \quad I_0(1)=1$$

类似可知

$$M_1(-1)=\max_{i=-1,1} p_{i,-1}q_{i,j_0}M_0(i)=0.08, \quad I_0(-1)=1.$$

$$M_2(1)=\max_{i=-1,1} p_{i,1}q_{i,j_1}M_1(i)=0.0768, \quad I_1(1)=1.$$

$$M_2(-1)=\max_{i=-1,1} p_{i,-1}q_{i,j_1}M_1(i)=0.0192, \quad I_1(-1)=1.$$

由于$\max\limits_{i=-1,1} q_{i,j_2}M_2(i)=\max\limits_{i=-1,1} q_{i,-1}M_2(i)=0.01536$在$i=1$时取到最大值. 因此前三个观察日该股票最可能的状态序列是$i_2=1, i_1=I_1(1)=1, i_0=I_0(1)=1$. □

(C) 隐Markov模型的参数估计*

在应用中, 对建立的隐Markov 模型, 我们还常需要从一段观测序列$\{Y_k, k\le n\}$出发, 估计模型参数$\theta=(\pi,\mathbf{P},\mathbf{Q})$. 这类问题又称为学习问题, 属于参数估计范畴.

若假设$\{X_k, k\ge 1\}$的样本也已知, 此时作为马氏链$Z=\{Z_k, k\ge 0\}$是完全已知的. 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, 由习题8.8(2) 以及大数定律可以证明, 只要i常返就可以用

$$\hat{p}_{i,j}=\frac{A_{i,j}}{A_i}$$

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

$$\hat{q}_{i,l}=\frac{B_{i,l}}{A_i}$$

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往往是不可测量的.

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的最大似然估计$\hat{\theta}$满足

$$\mathbb{P}(Y_k=l_k, 0\le k\le N|\hat{\theta})=\max_{\theta}\mathbb{P}(Y_k=l_k, 0\le k\le N|\theta),$$

其中$\mathbb{P}(\cdot|\theta)$表示给定参数θ下的概率.

一般而言, 直接求解该问题非常困难. 为此,可采取的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$\mathbb{P}(Y_k=l_k, 0\le k\le N|\theta)$的取值, 并将递推算法的最终结果作为参数$\theta$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

$$\begin{aligned} T(\theta|\theta_n) = & \sum_{i_k\in S, 0\le k\le N} \Big[\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n) \\ & \times \ln(\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta))\Big]. \end{aligned}$$

(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,

$$\begin{aligned} 0 \le & \ T(\theta_{n+1}|\theta_n) - T(\theta_n|\theta_n) \\ = & \sum_{i_k\in S, 0\le k\le N} \Big[\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n) \\ & \times \ln\Big(\frac{\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_{n+1})}{\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n)}\Big)\Big] \end{aligned}$$

因为对任意$x \in (0, \infty)$, $\ln x < x - 1$, 由上式可得

$$\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_{n+1}) \ge \mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n).$$

这表明上述算法得到的参数θ_n确实能使条件概率$\mathbb{P}(Y_k = j_k, 0 \le k \le N|\theta_n)$随着算法次数的增加而优化(增加).

将函数$T(\theta|\theta_n)$展开得

$$\begin{aligned} T(\theta|\theta_n) = & \sum_{\substack{i_k\in S \\ 0\le k\le N}} \Big[\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n) \ln(\pi_{i_0} q_{i_0,l_0} \prod_{n=1}^N p_{i_{n-1},i_n} q_{i_n,l_n})\Big] \\ = & \sum_{i_k\in S, 0\le k\le N} \Big[\mathbb{P}(X_k = i_k, Y_k = l_k, 0 \le k \le N|\theta_n) \\ & \times (\ln \pi_{i_0} + \sum_{n=0}^N \ln q_{i_n,l_n} + \sum_{n=1}^N \ln p_{i_{n-1},i_n})\Big] \\ = & \sum_{i_0\in S} \mathbb{P}(X_0 = i_0, Y_k = l_k, 0 \le k \le N|\theta_n) \ln \pi_{i_0} \\ & + \sum_{i_n\in S} \sum_{n=0}^N \mathbb{P}(X_n = i_n, Y_k = l_k, 0 \le k \le N|\theta_n) \ln q_{i_n,l_n} \\ & + \sum_{i_{n-1},i_n\in S} \sum_{n=1}^N \mathbb{P}(X_{n-1} = i_{n-1}, X_n = i_n, Y_k = l_k, 0 \le k \le N|\theta_n) \ln p_{i_{n-1},i_n}. \end{aligned}$$

或重写为

$$\begin{aligned}T(\theta|\theta_n) =& \sum_{i\in S}\mathbb{P}(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)\ln\pi_i\\ &+\sum_{\substack{i\in S\\ l\in W}}\sum_{n=0}^{N}\mathbb{P}(X_n=i,Y_k=l_k,0\le k\le N|\theta_n)\mathbf{1}_{\{Y_n=l\}}\ln q_{i,l}\\ &+\sum_{i,j\in S}\sum_{n=1}^{N}\mathbb{P}(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)\ln p_{i,j}.\end{aligned}$$

由**习题15.1**可知, 当$\theta=(\pi^*,\mathbf{P}^*,\mathbf{Q}^*)$时$T(\theta|\theta_n)$取到最大值, 其中, 对任意$i,j\in S,l\in W$, $\pi^*=(\pi_i^*),\mathbf{P}^*=(p_{i,j}^*),\mathbf{Q}^*=(q_{i,l}^*)$满足

$$\pi_i^* = \frac{\mathbb{P}(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)}{\sum_{i\in S}\mathbb{P}(X_0=i,Y_k=l_k,0\le k\le N|\theta_n)}, \tag{4.4.15}$$

$$p_{i,j}^* = \frac{\sum_{n=1}^{N}\mathbb{P}(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)}{\sum_{i,j\in S}\sum_{n=1}^{N}\mathbb{P}(X_{n-1}=i,X_n=j,Y_k=l_k,0\le k\le N|\theta_n)}, \tag{4.4.16}$$

$$q_{i,l}^* = \frac{\mathbb{P}(X_n=i,Y_k=l_k,0\le k\le N|\theta_n)\mathbf{1}_{\{Y_n=l\}}}{\sum_{\substack{i\in S\\ l\in W}}\sum_{n=0}^{N}\mathbb{P}(X_n=i,Y_k=l_k,0\le k\le N|\theta_n)\mathbf{1}_{\{Y_n=l\}}}. \tag{4.4.17}$$

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(4.4.15)-(4.4.17)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最大似然估计的统计方法. 隐Markov 模型参数的估计, 是EM 算法的一种典型运用.

对隐Markov 模型, 人们在应用时还会关心许多其他的问题, 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, 属于分类范畴. 对于这些问题, 我们不展开, 对此感兴趣的读者可以参阅相关资料, 进一步学习和了解.

练习题

15.1 设$z_i>0$ $(i\in T)$. 证明

$$\sum_{i\in T}z_i\ln\left(\frac{z_i}{\sum_{i\in T}z_i}\right)=\sup_{\substack{x_i>0,i\in T\\ \sum_i x_i=1}}\sum_{i\in T}z_i\ln x_i.$$

15.2 试寻找一些具体问题, 用隐Markov 模型建模并做出分析和预测.

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Poisson过程的推广; 一般而言, 不具有马氏性, 但对某些特定的时刻却有“从头再来” 的特征.

§5.1　更新过程与更新方程

给定概率空间$(\Omega, \mathfrak{F}, \mathbb{P})$及其上计数过程$N = \{N(t); t \geq 0\}$, 对任意整数$k \geq 1$, 令$T_k$ 计数过程为第k次(随机事件)发生时间, W_k为第k个等待时间, 即

$$T_k = \inf\{t \geq 0, N(t) \geq k\} \qquad \text{以及} \qquad W_k = T_k - T_{k-1},$$

其中$T_0 = 0$. 计数过程N与到达时刻序列$\{T_k;\ k \geq 1\}$满足如下关系

$$N(t) = \sup\{k;\ T_k \leq t\} = \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq t\}},$$

而且对任意$k \geq 0$以及$t \geq 0$

$$\{N(t) \geq k\} = \{T_k \leq t\}, \tag{5.1.1}$$

$$\{N(t) = k\} = \{T_k \leq t < T_{k+1}\}, \tag{5.1.2}$$

$$N(t) + 1 = \inf\{k, T_k > t\}. \tag{5.1.3}$$

(A) 更新过程与更新性

定义16.1 若计数过程$N = \{N(t);\ t \geq 0\}$具有独立同分布的等待时间序列, 即存在独立同分布的非负随机变量序列$\{W_k;\ k \geq 1\}$, 使得

$$N(t) = \sup\{k:\ \sum_{i=1}^{k} W_i \leq t\}. \tag{5.1.4}$$

则称N为更新过程. 称W_k的分布F为等待时间分布.

显然泊松过程是一个更新过程. 因此更新过程可看作泊松过程的推广. 此外, 为了避免$N(t) \equiv \infty$这种平凡情形, 总设$F(0) < 1$.

对任意$t \geq 0$以及给定的到达时刻T_n

$$\begin{aligned}(N(t+T_n) - n|T_n) &= \Big(\sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq t+T_n\}} - n\Big|T_n\Big)\\ &= \Big(\sum_{k=n+1}^{\infty} \mathbf{1}_{\{T_k - T_n \leq t\}}\Big|T_n\Big) = \Big(\sum_{k=n+1}^{\infty} \mathbf{1}_{\{\sum_{i=n+1}^{k} W_i \leq t\}}\Big|\sum_{i=1}^{n} W_i\Big).\end{aligned}$$

注意到W_i为独立同分布的随机变量, 因此

$$\big\{N(t+T_n)-n;\ t\geq 0|T_n\big\} \overset{f.d}{=} \Big\{\sum_{k=1}^{\infty}\mathbf{1}_{\{\sum_{i=1}^k W_i\leq t\}};\ t\geq 0\Big\} = \{N(t);\ t\geq 0\}.$$

这表明从恰好发生第n事件的时刻T_n往后看, 更新过程增量的随机变化情况与过程从零时刻开始的观测结果一致. 直观地说，在T_n之后，过程$N(t)$从“新”开始. 此后我们也常称T_n 为(第n次)更新时刻，称

$$\{N(t+T_n)-n;\ t\geq 0|T_n\} \overset{f.d.}{=} \{N(t);\ t\geq 0\}, \qquad n\geq 1$$

为更新性(更一般地, 我们也称其为再生性). 特别, 若$F(0)=0$, 对任意$n\geq 1$,

$$\{N(t+T_n)-N(T_n);\ t\geq 0|T_n\} \overset{f.d.}{=} \{N(t);\ t\geq 0\},$$

从而此时更新过程初值总应为0.

例16.1 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$N(t)=N_{[t]}$. 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：

$$\mathbb{P}(W_1=k)=\mathbb{P}(X_1=0,\cdots,X_{k-1}=0,X_k=1)=(1-p)^{k-1}p.$$

例16.2 某型商品专卖店采取(s,S)-式进货仓储策略, 即商店商品数不超过s 时将商品补足到S个, 假设开始时商店有S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, 那么$C(t)$就是一个更新过程, 等待时间W分布为

$$\mathbb{P}(W>t)=\mathbb{P}(\sum_{i=1}^{N(t)}\xi_i<S-s).$$

比如, 若还假设$\{\xi_i\}$与$N(t)$独立, $s=2,\ \ S=4$, 并设

$$\mathbb{P}(\xi_i=1)=p, \qquad \mathbb{P}(\xi_i=2)=q=1-p.$$

那么我们可得

$$\begin{aligned}\mathbb{P}(W>t) &= \mathbb{P}(\{(0,t]\text{内没有顾客}\}\cup\{(0,t]\text{内只有1个顾客且只买了1 件商品}\})\\ &= \mathrm{e}^{-\lambda t}+p\lambda t\mathrm{e}^{-\lambda t}.\end{aligned}$$

注意到$T_k=\sum_{i=1}^k W_i$为独立随机变量之和, 由强大数定律可知

$$T_n/n\to\mu=\mathbb{E}(W_1), \quad \text{a.s.}$$

因此$T_n\to\infty$, a.s. 此外

$$\mathbb{P}(T_n\leq t)=F^{*n}(t),$$

其中$F^{*0}(t)=\begin{cases}1, & t\geq 0\\ 0, & t<0\end{cases}$, 进而$N(t)$的分布列为

$$\mathbb{P}(N(t)=k)=\mathbb{P}(T_k\leq t)-\mathbb{P}(T_{k+1}\leq t)=F^{*k}(t)-F^{*(k+1)}(t),$$

并且对一切$t\geq 0$, $n\in\mathbf{N}$,

$$\mathbb{P}(N(t)<\infty)=\lim_{n\to\infty}\mathbb{P}(N(t)<n)=\lim_{n\to\infty}\mathbb{P}(T_n>t)=1,$$

$$\lim_{t\to\infty}\mathbb{P}(N(t)<n)=\lim_{t\to\infty}\mathbb{P}(T_n>t)=0.$$

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16.1 更新过程$N=\{N(t);\ t\geq 0\}$的任意阶矩存在, 即对任意$t\geq 0$, $k\geq 1$,

$$\mathbb{E}[(N(t))^k]<+\infty,$$

特别, $\qquad m(t):=\mathbb{E}(N(t))=\sum\limits_{n=1}^{\infty}F^{*n}(t).$

证明 对任意固定的$t\geq 0$,

$$\begin{aligned}\mathbb{E}[(N(t))^k] &= \sum_{n=1}^{\infty}n^k\mathbb{P}(N(t)=n)\leq\sum_{n=1}^{\infty}n^k\mathbb{P}(N(t)\geq n)\\ &=\sum_{n=1}^{\infty}n^k\mathbb{P}(T_n\leq t)=\sum_{n=1}^{\infty}n^kF^{*n}(t).\end{aligned}$$

由$\lim\limits_{n\to\infty}\mathbb{P}(T_n\leq t)=0$可知, 存在正整数$m$使得

$$a=F^{*m}(t)<1.$$

又因为对一切$n\geq 0$,

$$F^{*(m+n)}(t)=\int_0^t F^{*m}(t-s)\mathrm{d}F^{*n}(s)\leq F^{*m}(t)F^{*n}(t).$$

所以对任意$l=0,1,2\cdots$, $r=0,1,2,\cdots,m-1$,

$$F^{*(lm+r)}(t)\leq(F^{*m}(t))^l=a^l.$$

因此

$$\begin{aligned}\mathbb{E}[(N(t))^k] &\leq \sum_{n=1}^{\infty}n^kF^{*n}(t)=\sum_{l=0}^{\infty}\sum_{r=0}^{m-1}(lm+r)^kF^{*(lm+r)}(t)\\ &\leq m^{k+1}\sum_{l=0}^{\infty}(l+1)^ka^l<+\infty.\end{aligned}\tag{5.1.5}$$

此时

$$\begin{aligned} m(t)=\mathbb{E}(N(t)) &= \sum_{n=1}^{\infty} n\mathbb{P}(N(t)=n)=\sum_{n=1}^{\infty} n(F^{*n}(t)-F^{*(n+1)}(t)) \\ &= F(t)+\sum_{n=2}^{\infty}[n-(n-1)]F^{*n}(t)=\sum_{n=1}^{\infty} F^{*n}(t). \end{aligned} \quad \square$$

特别, 若N是泊松过程, 容易检验

$$m(t)=\sum_{n=1}^{\infty} F^{*n}(t)=\sum_{n=1}^{\infty}\int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!}\mathrm{e}^{-\lambda s}\mathrm{d}s=\int_0^t \lambda\mathrm{e}^{-\lambda s}\sum_{n=0}^{\infty}\frac{\lambda^n s^{n-1}}{(n-1)!}\mathrm{d}s=\lambda t,$$

这与命题7.3相关结果一致.

(B) 更新函数与更新方程

定义16.2 称$m(t)=\mathbb{E}(N(t))$为更新过程$N=\{N(t);\ t\geq 0\}$的更新函数. 当N的等待时间分布为F时, 也称m是分布F的更新函数.

注16.1 为了叙述方便, 若无特别说明, 本章此后更新过程的等待时间分布F都是离散格子点分布(参见例1.8) 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, 此时由习题16.6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12.1证明告诉我们级数$\sum_{n=1}^{\infty} F^{*n}(t)$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

$$\begin{aligned} \sum_{n=1}^{\infty} f_n(t) &= \Big[f(t)+\sum_{n=2}^{\infty}\int_0^t f_{n-1}(t-s)f(s)\mathrm{d}s\Big] \\ &\leq \sup_{0\leq s\leq t} f(s)[1+\sum_{n=1}^{\infty}\int_0^t f_n(s)\mathrm{d}s] \\ &= \sup_{0\leq s\leq t} f(s)[1+\sum_{n=1}^{\infty} F^{*n}(t)]<\infty \end{aligned}$$

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16.3 设$b(t)$是$[0,\infty)$上局部有界函数, 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16.2 一般地, 方程(3.2.2)中右边的积分$\int_0^t B(t-s)\mathrm{d}F(s)$是所谓的R-S(Riemann-Stieltjes)积分. 对此积分本书不展开介绍. 在注16.1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是离散型分布时, 该积分是和式

$$\sum_{0\le x_k\le t} m(t-x_k)p_k,$$

其中x_k为F支撑集(见例1.8)中的元素; 当F是密度函数为f的连续分布时, 该积分为牛顿- 莱布尼茨积分

$$\int_0^t m(t-s)f(s)\mathrm{d}s.$$

此后, 对任意给定的函数ϕ, 有时为表示简单我们会不区分F是连续还是离散分布而用积分$\int_0^\cdot \phi(s)\mathrm{d}F(s)$表示对应的求和或牛顿-莱布尼茨积分.

定理16.2 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 由全概率公式

$$m(t) = \mathbb{E}(N(t)) = \mathbb{E}(\mathbb{E}(N(t)|T_1)) = \int_0^\infty \mathbb{E}(N(t)|T_1 = s)\mathrm{d}F(s).$$

当$s = T_1 > t$时, $N(t) = 0$; 当$s = T_1 < t$时, 由更新性,

$$N(t) = N(t - T_1 + T_1) - 1 + 1 \overset{d}{=} 1 + N(t - T_1) = 1 + N(t-s).$$

因此 $\mathbb{E}(N(t)|T_1 = s) = \begin{cases} 0, & s > t, \\ 1 + \mathbb{E}(N(t-s)) = 1 + m(t-s), & s < t. \end{cases}$ 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16.3 通常称利用全概率公式和更新性建立更新方程的方法为更新技巧或更新方法.

例16.3 设更新过程$N(t)$的等待时间分布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$m(t)=\mathrm{e}^t-1$. □

利用更新函数我们可以给出更新方程唯一局部有界解的显示方式, 证明省略。

定理16.3 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16.4 (1) 与注16.2解释的一样, 方程(5.1.8)中的积分$\int_0^t b(t-s)\mathrm{d}m(s)$是R-S积分. 在注16.1条件下, 当$F$是具有局部有界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, $m(s)$是阶梯函数, 跳跃点在kd, 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,

$$\mathrm{d}m(t)\,=\,\mathrm{d}\big(\sum_{n=1}^{\infty}F^{*n}(t)\big)\approx\sum_{n=1}^{\infty}\mathbb{P}(t<T_n\le t+\mathrm{d}t).$$

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16.2与定理16.3, 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16.4 对任意$t,s\ge 0$, 求$\mathbb{E}(N^2(t))$.

解 记$K(t)=\mathbb{E}(N^2(t))$. 由全概率公式

$$K(t)=\mathbb{E}(\mathbb{E}(N^2(t)|T_1))=\int_0^{\infty}\mathbb{E}(N^2(t)|T_1=u)\mathrm{d}F(u).$$

由更新性

$$\mathbb{E}(N^2(t)|T_1=u)=\begin{cases}0, & u>t\\ \mathbb{E}((1+N(t-u))^2), & u<t.\end{cases}$$

即 $$\mathbb{E}(N^2(t)|T_1=u)=\begin{cases}0, & u>t\\ 1+2m(t-u)+K(t-u), & u<t.\end{cases}$$

所以$K(t)$满足方程

$$\begin{aligned}K(t) &= \int_0^t (1+2m(t-u)+K(t-u))\mathrm{d}F(u)\\ &= F(t)+2\int_0^t m(t-u)\mathrm{d}F(u)+\int_0^t K(t-u)\mathrm{d}F(u)\\ &= 2m(t)-F(t)+\int_0^t K(t-u)\mathrm{d}F(u).\end{aligned}$$

由定理16.1易知$K(t)$是局部有界函数. 由定理16.3以及(5.1.9)可得

$$\begin{aligned}K(t) &= 2m(t)-F(t)+\int_0^t (2m(t-s)-F(t-s))\mathrm{d}m(s)\\ &= m(t)+2\int_0^t m(t-s)\mathrm{d}m(s).\end{aligned}$$ □

例16.5 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16.3, 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为等待时间的更新过程是强度为2的泊松过程, 由命题7.3可知$m(s)=2s$. 因此

$$g(t)=t+\int_0^t (t-s)2\mathrm{d}s=t^2+t.$$ □

(C) 更新过程的几个统计量

定义16.4 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, $0 \le A(t) \le t$, $t \ge 0$ 且对任意$x \ge 0$, $0 \le y \le t$,

$$\begin{aligned}\{R(t) > x, A(t) \ge y\} &= \{(t-y, t+x]\text{中}N\text{无跳跃}\} \\ &= \{R(t-y) > x+y\} = \{A(t+x) \ge x+y\}.\end{aligned}$$

取$y=0$得

$$\{R(t) > x\} = \{A(t+x) \ge x\}.$$

定理16.4 设N的等待时间分布F均值有限. 令$\bar{F} = 1 - F(t)$, 则对任意$x, t \ge 0$,

$$\mathbb{P}(R(t) > x) = \bar{F}(t+x) + \int_0^t \bar{F}(t-s+x)\mathrm{d}m(s), \tag{5.1.10}$$

$$\mathbb{E}(R(t)) = \int_t^\infty \bar{F}(t)\mathrm{d}t + \int_0^t \int_{t-s}^\infty \bar{F}(u)\mathrm{d}u\mathrm{d}m(s), \tag{5.1.11}$$

证明 注意到对任意$x \ge 0$,

$$\mathbb{P}(R(t) > x | T_1 = s) = \begin{cases} 1, & s > t+x; \\ 0, & t < s \le t+x; \\ \mathbb{P}(R(t-s) > x), & s \le t. \end{cases}$$

利用更新技巧得

$$\begin{aligned}\mathbb{P}(R(t) > x) &= \mathbb{E}(\mathbb{P}(R(t) > x | T_1)) = \int_0^\infty \mathbb{P}(R(t) > x | T_1 = s)\mathrm{d}F(s) \\ &= \int_{t+x}^\infty \mathrm{d}F(s) + \int_0^t \mathbb{P}(R(t-s) > x)\mathrm{d}F(s) \\ &= \bar{F}(t+x) + \int_0^t \mathbb{P}(R(t-s) > x)\mathrm{d}F(s).\end{aligned}$$

由定理16.3可知(5.1.10)成立. 再注意到对任意$t \ge 0$

$$\mathbb{E}(R(t) | T_1 = s) = \begin{cases} s-t, & s > t; \\ \mathbb{E}(R(t-s)), & s \le t. \end{cases}$$

利用更新技巧

$$\begin{aligned}\mathbb{E}(R(t)) &= \mathbb{E}(\mathbb{E}(R(t)|T_1)) = \int_0^\infty \mathbb{E}(R(t)|T_1 = s)\mathrm{d}F(s) \\ &= \int_t^\infty (u-t)\mathrm{d}F(u) + \int_0^t \mathbb{E}(R(t-s))\mathrm{d}F(s) \\ &= \int_t^\infty \bar{F}(u)\mathrm{d}u + \int_0^t \mathbb{E}(R(t-s))\mathrm{d}F(s). \end{aligned} \tag{5.1.12}$$

此外由(5.1.10)可得

$$\mathbb{P}(R(t)>x)\le \bar{F}(x)+\bar{F}(x)m(t).$$

因此

$$\mathbb{E}(R(t))=\int_0^\infty \mathbb{P}(R(t)>x)\mathrm{d}x\le\int_0^\infty \bar{F}(x)\mathrm{d}x(1+m(t))=\mu(1+m(t))$$

是局部有界函数. 将定理16.3应用到方程(5.1.12)即得式(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$\mathbb{P}(L(t)>x)=1$; 若$0\le s=A(t)\le x$则意味者t时刻的总寿命一定大于s, 此时由更新性

$$\mathbb{P}(L(t)>x|A(t)=s)=\mathbb{P}(L(s)>x|A(s)=s)=\mathbb{P}(W>x|W>s)\ge\mathbb{P}(W>x),$$

其中W表示更新过程的任意一个等待时间. 因此不论t时刻的年龄如何, 总有

$$\mathbb{P}(L(t)>x|A(t))\ge\mathbb{P}(W>x),$$

从而

$$\mathbb{P}(L(t)>x)\ge\mathbb{P}(W>x). \tag{5.1.13}$$

这表明我们在某个固定t时刻观察到的等待时间会比通常“理论上”的等待时间长些——对于泊松过程我们可以给出更清楚的计算(参见习题7.6和16.4). 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：由于长的等待时间比短的等待时间更可能包含给定的时刻t, 因此t 时刻观察到的等待时间会更长.

练习题

16.1 设更新过程N的等待时间分布F的密度函数为$f(t)=t\mathrm{e}^{-t}$, $t>0$,

(1) 对任意$0<t_1<t_2$以及正整数n,m, 求$\mathbb{P}(N(t_1)=n,N(t_2)=m)$.

(2) 对任意$t>0$, 求更新函数$m(t)$.

16.2 求例16.3中$m(t)$其中$t\in(1,2]$.
16.3 求积分方程$g(t)=t+\int_0^t g(t-s)3\mathrm{e}^{-2s}\mathrm{d}s$ 的局部有界解.
16.4 求$\mathbb{P}(A(t)\ge x)$, $\mathbb{E}(A(t))$以及$\mathbb{P}(L(t)\ge x)$, $\mathbb{E}(L(t))$.
16.5 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.
16.6 设N是更新过程, 等待时间分布期望为μ, 证明$\mathbb{E}(T_{N(t)+1})=\mu(1+m(t))$.
16.7* 设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.

(A) 更新过程的大数定律与中心极限定理

定理17.1 对更新过程$N=\{N(t);\ t\ge 0\}$

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{\mu},\quad a.s.$$

其中$\mu=\infty$时$1/\mu$ 为0.

证明 由T_n的定义可知$T_{N(t)}\le t<T_{N(t)+1}$, 从而

$$\frac{T_{N(t)}}{N(t)}\le\frac{t}{N(t)}\le\frac{T_{N(t)+1}}{N(t)}=\frac{T_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}.\tag{5.2.1}$$

注意到$N(t)\to\infty$, a.s.以及$T_n/n\to\mu$, a.s. 由习题3.2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17.1 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次为此，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发生后从新开始, 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. 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; 同样以连续失败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13.1，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. 由习题2.3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

$$\mu=\frac{(1-p^k)(1-q^k)}{pq(p^{k-1}+q^{k-1}-p^{k-1}q^{k-1})}.$$ □

定理17.2 设更新过程$N=\{N(t);\ t\geq 0\}$满足条件

$$\mathbb{E}(W_1)=\mu<\infty, \mathbb{V}ar(W_1)=\sigma^2<\infty,$$

则对一切实数y,

$$\lim_{t\to\infty}\mathbb{P}\Big(\frac{N(t)-t/\mu}{\sigma\sqrt{t/\mu^3}}\leq y\Big)=\Phi(y),$$

其中$\Phi(y)$是标准正态分布函数.

证明 令$r_t=t/\mu+y\sigma\sqrt{t/\mu^3}$, $\hat{r}_t=[r_t]+1$, 则

$$\begin{aligned}\mathbb{P}\Big(\frac{N(t)-t/\mu}{\sigma\sqrt{t/\mu^3}}\leq y\Big) &= \mathbb{P}(N(t)\leq [r_t])=\mathbb{P}(N(t)<\hat{r}_t)\\ &= \mathbb{P}(T_{\hat{r}_t}>t)=\mathbb{P}\Big(\frac{T_{\hat{r}_t}-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}>\frac{t-\hat{r}_t\mu}{\sigma\sqrt{\hat{r}_t}}\Big).\end{aligned}$$

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. □

例17.2 一台机器相继不间断地加工某一种零件, 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17.2, 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

$$0.95=\mathbb{P}(N(6000)>x)=\mathbb{P}\Big(\frac{N(6000)-3000}{\sqrt{6000/24}}>\frac{x-3000}{5\sqrt{10}}\Big)\approx 1-\Phi\big(\frac{x-3000}{5\sqrt{10}}\big).$$

由此可知$x\approx 3000-1.64\times 5\sqrt{10}\approx 2974$. □

(B) 更新定理

下面的所谓关键更新定理是我们研究更新方程局部有界解极限的重要工具.

定理17.3 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. $h(t)$是$[0,\infty)$上非负可积函数. B 是更新方程的局部有界解,

$$B(t)=h(t)+\int_0^t B(t-s)\mathrm{d}F(s)=h(t)+\int_0^t h(t-s)\mathrm{d}m(s). \tag{5.2.2}$$

(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具有局部有界密度函数且存在$M > 0$使得$h(t)$在$[M,\infty)$ 上单调不增, 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$\nu(kd) = F(kd) - F((k-1)d) = \mathbb{P}(X = kd)$. 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

$$a_n = b_n + \sum_{k=0}^{n} a_{n-k}u_k. \tag{5.2.3}$$

注意, 由u_k的定义及条件可知, 存在$m_1, m_2, \cdots, m_r$使得

$$u_{m_1}, \cdots, u_{m_r} > 0\text{且}m_1, \cdots, m_r\text{互素}.$$

由习题1.9, 存在正整数J, 对任意正整数$j \ge J$ 存在非负整数$l_1, \cdots, l_r$, 使得

$$j = l_1m_1 + l_2m_2 + \cdots + l_rm_r. \tag{5.2.4}$$

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的假设条件可知

$$\sum_{k=0}^{\infty} |b_k| = \sum_{k=0}^{\infty} |h(c+kd)| < \infty.$$

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$M=\sup\limits_{n\geq 0}|a_n|$.

第二步, 证明$\limsup\limits_{n\to\infty} a_n \leq \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu<\infty, \\ 0, & \mu=\infty. \end{cases}$

为此记$l=\limsup\limits_{n\to\infty} a_n$. 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$\sum\limits_{k=0}^{\infty}u_k=1$ 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$\lim\limits_{n_i\to\infty}a_{n_i-m_k}=l$. 进一步反复使用以上方法可得对任意非负整数$l_1,\cdots,l_k$,

$$\lim_{n_i\to\infty}a_{n_i-\sum_{k=1}^{r}l_km_k}=l.$$

由(5.2.4)可知, 对任意$j\geq J$, $\lim\limits_{n_i\to\infty}a_{n_i-j}=l$. 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$U_n=\sum\limits_{k=n+1}^{\infty}u_k$, 由关系式$u_k=U_{k-1}-U_k$以及$U_{-1}=1$, (5.2.3)可改写为

$$U_0a_n+U_1a_{n-1}+\cdots+U_na_0=U_0a_{n-1}+U_1a_{n-2}+\cdots+U_{n-1}a_0+b_n,\quad n\geq 1.$$

记$A_n=\sum\limits_{k=0}^{n}U_ka_{n-k}$, 那么$A_0=b_0$且$A_n=A_{n-1}+b_n$, 进而$A_n=\sum\limits_{k=0}^{n}b_k$. 因此对任意$N$,

$$U_0a_{n_i'}+\cdots+U_Na_{n_i'-N}\leq A_{n_i'}\leq\sum_{k=0}^{\infty}|b_k|.$$

先令$n_i' \to \infty$再令$N \to \infty$并注意到$\sum\limits_{k=0}^{\infty} U_k = \sum\limits_{k=1}^{\infty} ku_k$得

$$l \leq \frac{\sum\limits_{k=0}^{\infty} b_k}{\sum\limits_{k=1}^{\infty} ku_k} = \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

第三步, 证明$\liminf\limits_{n\to\infty} a_n \geq \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$

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$k \geq 0$,

$$a_{n_i-k} \to \underline{l}.$$

记$g(N) = \sum\limits_{k=N+1}^{\infty} U_k$, 由$\mu = \sum\limits_{k=0}^{\infty} U_k < \infty$ 可知$g(N) \to 0$. 注意到

$$\sum_{k=0}^{n_i} b_k = \sum_{k=0}^{n_i} U_k a_{n_i-k} \leq \sum_{k=0}^{N} U_k a_{n_i-k} + g(N)M.$$

先令$n_i \to \infty$再令$N \to \infty$得

$$\underline{l} \geq \frac{\sum\limits_{k=0}^{\infty} b_n}{\sum\limits_{k=1}^{\infty} ku_k} = \begin{cases} d\sum\limits_{n=0}^{\infty} h(c+nd)/\mu, & \mu < \infty, \\ 0, & \mu = \infty. \end{cases}$$

综合第二、三两步的结果可知(1)成立.

(2) 此时证明更复杂, 此略 □

注17.1 事实上(2)对一切非格子点分布都成立.

注17.2 在(1)证明中我们得到了如下结果: 若数列$\{b_n\}$和非负数列$\{a_n\}$, $\{c_n\}$满足(1) $\sum\limits_{n=0}^{\infty} |b_n| < \infty$, (2) $\sum\limits_{n=0}^{\infty} c_n = 1$ 且数集$\{n \geq 1, c_n > 0\}$的最大公因子为1, (3) 对任意$n \geq 0$, $a_n = b_n + \sum\limits_{k=0}^{n} a_{n-k}c_k$. 那么

$$\lim_{n\to\infty} a_n = \frac{\sum\limits_{n=0}^{\infty} b_n}{\sum\limits_{n=1}^{\infty} nc_n},$$

其中商式分母为$+\infty$时取值为0.

例17.3 设X是以$P=(p_{i,j})_{i,j\in S}$为转移概率的马氏链, 若j非周期常返, 证明

$$\lim_{n\to\infty} p_{i,j}^{(n)} = \begin{cases} 0, & j\text{是零常返}, \\ f_{ij}/m_{jj}, & j\text{是正常返}. \end{cases}$$

证明 注意到对任意$n\geq 1$,

$$\begin{aligned} p_{i,j}^{(n)} &= \mathbb{P}_i(X_n=j)=\mathbb{P}_i(X_n=j,X_{n-1}=j)+\mathbb{P}_i(X_n=j,X_{n-1}\neq j) \\ &= \sum_{k=1}^{n}\mathbb{P}_i(X_n=j,X_k=j,X_v\neq j,v=k+1,\cdots,n-1) \\ &\qquad +\mathbb{P}_i(X_n=j,X_v\neq j,v=1,\cdots,n-1) \\ &= f_{ij}^{(n)}+\sum_{k=1}^{n-1}p_{i,j}^{(k)}f_{jj}^{(n-k)} = f_{ij}^{(n)}+\sum_{k=1}^{n-1}p_{i,j}^{(n-k)}f_{jj}^{(k)} \end{aligned}$$

对$n\geq 1$, 令$a_n=p_{i,j}^{(n)}$, $b_n=f_{ij}^{(n)}$, $c_n=f_{jj}^{(n)}$, 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(参见习题9.7). 容易验证注记17.1中其他两条件对$\{a_n\}$, $\{b_n\}$也成立, 因此

$$\lim_{n\to\infty}p_{i,j}^{(n)}=\lim_{n\to\infty}a_n=\frac{\sum\limits_{n=0}^{\infty}b_n}{\sum\limits_{n=1}^{\infty}nc_n}=\frac{\sum\limits_{n=0}^{\infty}f_{ij}^{(n)}}{\sum\limits_{n=1}^{\infty}nf_{jj}^{(n)}}=\begin{cases} 0, & j\text{是0常返}, \\ f_{ij}/m_{jj}, & j\text{是正常返}. \end{cases} \qquad \square$$

推论17.4(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$h(s)=\mathbf{1}_{[0,a]}(s)$, 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,

$$\lim_{N\to\infty}\frac{m(N)}{N}=\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}(m(k)-m(k-1))=\frac{1}{\mu_F}.$$

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是格子点分布的时, 利用习题17.7的结论, 同样可证明推论成立. 细节请读者自己补充. $\square$

例17.4 设更新过程N的等待时间分布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$\lim\limits_{t\to\infty}\mathbb{E}(L(t))$.

解 由习题16.3可知,

$$\mathbb{E}(L(t)) = \int_t^\infty s\mathrm{d}F(s) + \int_0^t \int_{t-s}^\infty u\mathrm{d}F(u)\mathrm{d}m(s).$$

令$h(t) = \int_t^\infty s\mathrm{d}F(s)$. 显然$h(t)$是单调不增的. 而且由于$F$均值$\mu$与方差$\sigma^2$都有限,

$$\int_0^\infty h(t)\mathrm{d}t = \int_0^\infty \int_t^\infty u\mathrm{d}F(u)\mathrm{d}t = \int_0^\infty u^2\mathrm{d}F(u) = \sigma^2 + \mu^2 < \infty.$$

因此由关键更新定理可得

$$\lim_{t\to\infty}\mathbb{E}(L(t)) = \frac{\int_0^\infty h(t)\mathrm{d}t}{\mu} = \mu + \frac{\sigma^2}{\mu}. \qquad \square$$

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等待时间的极限. 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或检验悖论现象.

例17.5 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，那么服务台空闲、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$\mathbb{E}(Y_i) = 1/\mu$; 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, $\mathbb{E}(U_i) = 1/\lambda$. 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. 更新系统的一个等待时长为$W_i = Y_i + U_i$, 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$\lim\limits_{t\to\infty}\mathbb{P}(A_t)$.

解 如上定义Y_i, U_i, W_i. 记W_i的分布为F. 则F是均值为$1/\mu + 1/\lambda$的连续分布. 那么

$$\begin{aligned} f(t) := \mathbb{P}(A_t) &= \mathbb{E}(\mathbb{P}(A_t|T_1)) = \int_0^\infty \mathbb{P}(A_t|T_1 = s)\mathrm{d}F(s) \\ &= \int_t^\infty \mathbb{P}(U_1 > t|W_1 = s)\mathrm{d}F(s) + \int_0^t \mathbb{P}(A_{t-s})\mathrm{d}F(s) \end{aligned}$$

注意到$\int_t^\infty \mathbb{P}(U_1 > t|W_1 = s)\mathrm{d}F(s) = \mathbb{P}(U_1 > t, W_1 > t) = \mathbb{P}(U_1 > t)$, 因此

$$f(t) = \mathbb{P}(U_1 > t) + \int_0^t f(t-s)\mathrm{d}F(s).$$

由关键更新定理,

$$\lim_{t\to\infty} f(t) = \frac{\int_0^\infty \mathbb{P}(U_1 > t)\mathrm{d}t}{1/\lambda + 1/\mu} = \frac{\mu}{\lambda + \mu}. \qquad \square$$

练习题

17.1 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.

17.2 设N是等待时间分布为连续分布的更新过程, 平均等待时间为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.

17.3 令$L(t)$表示时刻t更新过程的总寿命,若平均等待时间$\mu<\infty$, 证明$\lim\limits_{t\to\infty} L(t)/t=0$ 几乎必然成立.

17.4 设更新过程N的等待时间分布F连续且均值μ与方差σ^2都有限, 证明$\lim\limits_{t\to\infty}(m(t)-t/\mu)=(\sigma^2-\mu^2)/2\mu^2$.

17.5 设更新过程N的等待时间分布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$\lim\limits_{t\to\infty}\mathbb{E}(A(t))$.

17.6 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$\lim\limits_{t\to\infty} g(t)$. ([提示] 注意关键更新定理条件)

17.7* 设F是步长为d的格子点分布, 数学期望为μ, $m(t)$是以F为等待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, 本节简要介绍有偿更新过程与可终止更新过程.

(A) 有偿更新过程

定义18.1 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是以$\{W_n\}$为等待时间序列的更新过程, 对任意$t \geq 0$, 令

$$X(t) = \sum_{n=1}^{N(t)} Y_n,$$

称$X = \{X(t), t \geq 0\}$为有偿更新过程. 若N为Poisson过程, $\{Y_n, n \geq 1\}$与$\{W_n, n \geq 1\}$独立, 则称X为复合泊松过程.

定理18.1 设X为有偿更新过程, $\mathbb{E}(W_1) < \infty$, $\mathbb{E}|Y_1| < \infty$, 则

$$\lim_{t\to\infty} \frac{X(t)}{t} = \frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}, \quad \text{a.s.} \ \text{而且} \lim_{t\to\infty} \frac{\mathbb{E}(X(t))}{t} = \frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}.$$

证明 由强大数定理及$N(t) \to \infty$, a.s. 可知

$$\frac{X(t)}{t} = \frac{\sum_{k=1}^{N(t)} Y_k}{N(t)} \frac{N(t)}{t} \to \frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}, \quad a.s.$$

定理前半部分极限成立. 对后半部分极限, 注意到

$$\begin{aligned}
\mathbb{E}\Big(\sum_{k=1}^{N(t)+1} Y_k\Big) &= \sum_{k=1}^{\infty} \mathbb{E}\Big(\sum_{i=1}^{k} Y_i \mathbf{1}_{\{N(t)+1=k\}}\Big) = \sum_{k=1}^{\infty}\sum_{i=1}^{k} \mathbb{E}\Big(Y_i \mathbf{1}_{\{\sum\limits_{n=1}^{k-1} W_n \leq t, \sum\limits_{n=1}^{k} W_n > t\}}\Big) \\
&= \sum_{i=1}^{\infty}\sum_{k=i}^{\infty} \mathbb{E}\Big(Y_i \mathbf{1}_{\{\sum\limits_{n=1}^{k-1} W_n \leq t, \sum\limits_{n=1}^{k} W_n > t\}}\Big) \\
&= \sum_{i=1}^{\infty} \mathbb{E}\Big(Y_i - Y_i \sum_{k=1}^{i-1} \mathbf{1}_{\{\sum\limits_{n=1}^{k-1} W_n \leq t, \sum\limits_{n=1}^{k} W_n > t\}}\Big).
\end{aligned}$$

由$\{(W_n, Y_n), n \geq 1\}$的独立同分布性可得

$$\begin{aligned}
\mathbb{E}\Big(\sum_{k=1}^{N(t)+1} Y_k\Big) &= \sum_{i=1}^{\infty} \mathbb{E}(Y_i)\Big(1 - \sum_{k=1}^{i-1} \mathbb{P}(\sum_{n=1}^{k-1} W_n \leq t, \sum_{n=1}^{k} W_n > t)\Big) \\
&= \sum_{i=1}^{\infty} \mathbb{E}(Y_i)(1 - \mathbb{P}(N(t)+1 \leq i-1)) = \sum_{i=1}^{\infty} \mathbb{E}(Y_i)\mathbb{P}(N(t)+1 \geq i) \\
&= \mathbb{E}(Y_1)\sum_{i=1}^{\infty} \mathbb{P}(N(t)+1 \geq i) = \mathbb{E}(Y_1)(1 + m(t)). \qquad (5.3.1)
\end{aligned}$$

进而

$$\lim_{t\to\infty}\frac{1}{t}\mathbb{E}\Big(\sum_{i=1}^{N(t)+1}Y_i\Big)=\lim_{t\to\infty}\frac{g(t)}{t}=\mathbb{E}(Y_1)\lim_{t\to\infty}\frac{1+m(t)}{t}=\frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}.$$

另一方面,

$$\begin{aligned}\mathbb{E}(Y_{N(t)+1}) &= \int_0^\infty \mathbb{E}(Y_{N(t)+1}|T_1=s)\mathrm{d}F(s)\\ &= \int_t^\infty \mathbb{E}(Y_1|W_1=s)\mathrm{d}F(s)+\int_0^t \mathbb{E}(Y_{N(t-s)+1})\mathrm{d}F(s).\end{aligned}$$

记$B(t)=\mathbb{E}(Y_{N(t)+1})$, $b(t)=\int_t^\infty \mathbb{E}(Y_1|W_1=s)\mathrm{d}F(s)$, 则

$$B(t)=b(t)+\int_0^t B(t-s)\mathrm{d}F(s).$$

由$|b(t)|\le\int_0^\infty \mathbb{E}(|Y_1||W_1=s)\mathrm{d}F(s)=\mathbb{E}(|Y_1|)<\infty$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, $|b(t)|\le\varepsilon$. 所以

$$\begin{aligned}0\le\frac{1}{t}\int_0^t|b(t-s)|\mathrm{d}m(s) &= \frac{1}{t}\Big[\int_0^{t-T}+\int_{t-T}^t\Big]|b(t-s)|\mathrm{d}m(s)\\ &\le \varepsilon\frac{m(t-T)}{t}+\mathbb{E}(|Y_1|)\frac{m(t)-m(t-T)}{t}.\end{aligned}$$

先令$t\to\infty$再令$\varepsilon\to 0$, 由推论17.4可得

$$\lim_{t\to\infty}\frac{B(t)}{t}=0.$$

最后注意到$\mathbb{E}(X(t))=\mathbb{E}\Big(\sum\limits_{i=1}^{N(t)+1}Y_i\Big)-\mathbb{E}(Y_{N(t)+1})$, 综合以上结果可得

$$\lim_{t\to\infty}\frac{\mathbb{E}(X(t))}{t}=\frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}.$$ □

例18.1 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 可用有偿更新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,

$$Y_n=C\mathbf{1}_{\{W_n=T\}}+D\mathbf{1}_{\{W_n<T\}}$$

为第n次手机更新的费用, 那么在时刻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

$$\lim_{t\to\infty}\frac{X(t)}{t}=\frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}=\frac{C+(D-C)H(T-)}{\int_0^T \bar{H}(x)dx},$$

其中$H(T-)=\mathbb{P}(W_n<T)$. □

例18.2 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. 问长时间后顾客到达即可接受服务的平均概率.

解 可用有偿更新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

$$T(t)=\sum_{k=1}^{N(t)} Y_i+(t-T_{N(t)}-X_{N(t)+1})\mathbf{1}_{\{t>T_{N(t)}+X_{N(t)+1}\}}$$

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

$$\lim_{t\to\infty}\frac{T(t)}{t}=\frac{\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}=\frac{1/\lambda}{\mathbb{E}(X_1)+1/\lambda}.$$

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G分布, 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

$$\mathbb{E}(U)=\frac{1}{\mu},\quad \mathbb{E}(V)=\int_0^\infty \lambda u\mathrm{d}G(u)=\frac{\lambda}{\mu}.$$

由于服务顾客是独立同分布的, $\mathbb{E}(X_1|U,V)=U+V\mathbb{E}(X_1)$, 进而

$$\mathbb{E}(X_1)=\mathbb{E}(U)+\mathbb{E}(X_1)\mathbb{E}(V)=\frac{1}{\mu}+\mathbb{E}(X_1)\frac{\lambda}{\mu}.$$

由此可得$\mathbb{E}(X_1) = 1/(\mu - \lambda)$. 所以

$$\lim_{t\to\infty} \frac{T(t)}{t} = 1 - \frac{\lambda}{\mu}. \qquad \square$$

例18.3 设保险公司理赔事件发生服从强度为λ的Poisson过程, 每次理赔金额Y_i服从参数为μ的指数分布且相互独立. 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). 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

$$W(t) = x + ct - \sum_{i=1}^{N(t)} Y_i.$$

公司永不破产的概率

$$B(x) = \mathbb{P}(W(t) \ge 0, t \ge 0) = \mathbb{P}(x + ct - \sum_{i=1}^{N(t)} Y_i \ge 0, t \ge 0),$$

并约定$B(x) = 0, x < 0$. 那么由全概率公式

$$B(x) = \int \mathbb{P}(W(t) \ge 0, t \ge 0 | W_1 = s, Y_1 = y) \mathrm{d}H(s, y),$$

其中$H(s,y)$为(W_1, Y_1)的联合分布. 显然

$$\mathbb{P}(W(t) \ge 0, t \ge 0 | W_1 = s, Y_1 = y) = \begin{cases} 0, & y > x + cs, \\ B(x + cs - y), & y \le x + cs. \end{cases}$$

由假设$\mathrm{d}H(s,y) = \lambda e^{-\lambda s} \mu \mathrm{e}^{-\mu y} \mathrm{d}s\mathrm{d}y$, 因此

$$\begin{aligned} B(x) &= \int_0^\infty \lambda e^{-\lambda s} \mathrm{d}s \int_0^{x+cs} B(x + cs - y) \mu \mathrm{e}^{-\mu y} \mathrm{d}y \\ &= \int_x^\infty \frac{\lambda}{c} e^{-\frac{\lambda}{c}(u-x)} \mathrm{d}u \int_0^u B(u - y) \mu \mathrm{e}^{-\mu y} \mathrm{d}y. \end{aligned}$$

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

$$B(x)=1-\frac{\lambda}{c\mu}e^{(\frac{\lambda}{c}-\mu)x}.$$ □

(B)可终止更新过程

定义18.2 若更新过程等待时间分布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16.1一样我们仍要求等待时间分布F在$[0,\infty)$ 内或者有局部有界的导函数或者是阶梯函数且跳跃点包含在形如$\{nd,n\geq 0\}$的集合中. 前者类似连续分布，后者类似格子点分布.

例18.4 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$N(t)$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$m(t)=\mathbb{E}(N(t))$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, 在某次事件发生后会以正概率不再有事件发生. 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18.2 设$p=F(\infty)$, 即$1-p$是F在$+\infty$的概率. 对以F为等待时间分布的可终止更新过程N而言,

$$\mathbb{P}(\xi\leq t)=(1-p)(1+m(t)),\quad \mathbb{P}(N(\infty)=k)=p^k(1-p),\ \ k\geq 0,$$

以及

$$\mathbb{E}(\xi)=\frac{1}{1-p}\int_0^{\infty}(p-F(t))\mathrm{d}t,\quad m(\infty)=\mathbb{E}(N(\infty))=\frac{p}{1-p}.$$

证明 记N的第n个等待时间为W_n. 显然

$$\{N(\infty)=k\}=\{W_1<\infty,\cdots,W_k<\infty,W_{k+1}=\infty\}.$$

因此$\mathbb{P}(N(\infty)=k)=p^k(1-p)$, 进而$\mathbb{E}(N(\infty))=p/(1-p)$.

注意到$W_1=\infty$时$\xi=0$, 因此对任意$t\geq 0$, 由更新技巧可得

$$\begin{aligned}\mathbb{P}(\xi\leq t)&=\mathbb{E}(\mathbb{P}(\xi\leq t|W_1))=1-p+\int_0^{\infty}\mathbb{P}(\xi\leq t|W_1=s)\mathrm{d}F(s)\\&=1-p+\int_0^{t}\mathbb{P}(\xi\leq t-s)\mathrm{d}F(s).\end{aligned}$$

因此$\mathbb{P}(\xi\leq t)=(1-p)(1+m(t))$. 同理

$$\mathbb{E}(\xi)=\mathbb{E}(\mathbb{E}(\xi|W_1))=\int_0^{\infty}(s+\mathbb{E}(\xi))\mathrm{d}F(s)=\int_0^{\infty}s\mathrm{d}F(s)+p\mathbb{E}(\xi).$$

因此

$$\mathbb{E}(\xi)=\frac{\int_0^{\infty}s\mathrm{d}F(s)}{1-p}=\frac{1}{1-p}\int_0^{\infty}(p-F(t))\mathrm{d}t.\qquad\square$$

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3.14)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18.3 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为等待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

$$\lim_{t\to\infty} \mathrm{e}^{\lambda t}[m(\infty) - m(t)] = \frac{1}{\int_0^\infty \lambda s \mathrm{e}^{\lambda s} \mathrm{d}F(s)}. \qquad \square$$

例18.5 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$\tilde{W}_i$为等待时间间隔的可终止更新过程, 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

$$t = \mathbb{E}(\xi) + \tau = \frac{1}{1 - F(\tau)} \int_0^\tau (F(\tau) - F(t)) \mathrm{d}t + \tau. \qquad \square$$

练习题

18.1 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.

18.2 3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.

18.3 设N是等待时间分布为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$\lim\limits_{t\to\infty} \int_0^t A(s)\mathrm{d}s/t$.

18.4* 有偿更新过程X满足$\mathbb{E}(W_i) < \infty, \mathbb{E}|Y_i| < \infty$, 其中$W_i$ 服从连续分布. 令$m(t) = \mathbb{E}(X(t))$, 证明对任意$a > 0$

$$\lim_{t\to\infty} (m(t+a) - m(t)) = \frac{a\mathbb{E}(Y_1)}{\mathbb{E}(W_1)}.$$

18.5* 已知$X = \{X_n, n \geq 0\}$为状态空间$S = \{0, 1\}$的时齐马氏链, 初值为0, 转移概率矩阵

$$\mathbf{P} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}, \quad 0 < p, q < 1.$$

设$Y_n, n \geq 1$为一族均值为μ的独立同分布随机变量且与马氏链X独立. 令

$$W_k = \begin{cases} Y_k, & X_k = 0; \\ 0, & X_k = 1. \end{cases}$$

证明

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n W_k = \frac{\mu(1-q)}{2-p-q}.$$

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, 布朗运动已成为运动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

(A) 标准布朗运动

利用正态分布(随机变量), 我们定义布朗运动如下:

定义19.1 称实值随机过程$B = \{B(t);\ t \geq 0\}$ 为Wiener过程或布朗运动，若

(a) $B(0) = 0$ 且B具有连续轨道;

(b) B是独立增量过程, 即对任意$0 \leq t_1 < t_2 < \cdots < t_n$, $n \geq 2$, 增量

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

相互独立;

(c) 正态性, 对任意$0 \leq s < t$, $B(t) - B(s) \sim N(0, \sigma^2(t-s))$, 其中$\sigma^2$ 为正常数.

若$B_x(t) = x + B(t)$, 称$B_x = \{B_x(t);\ t \geq 0\}$为从$x$出发的布朗运动.

若布朗运动B满足$\sigma^2 = 1$, 则称其为标准布朗运动. 显然若B是方差为σ^2的布朗运动, 那么B/σ为标准布朗运动.

显然, 由布朗运动定义(b),(c)可知B具有平稳独立增量.

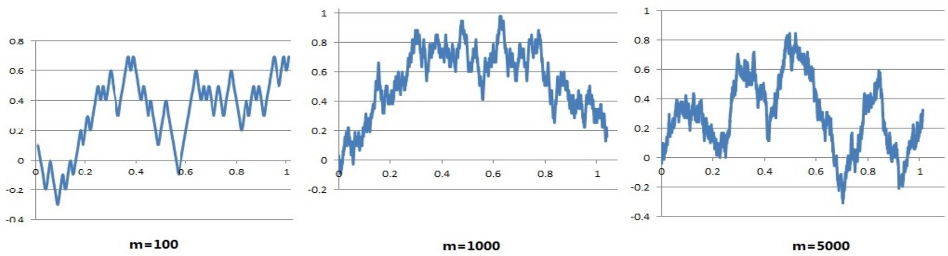
在一定意义上, 布朗运动B可以通过简单对称随机游动在时间和空间上采取一定方式连续化后得到. 由第一章第四节可知简单对称随机游动$S(n) = \sum\limits_{k=1}^{n} X_k$记录了每单位时间向左或向右等可能地跳一个单位的粒子在时刻n的轨迹变化. 令

$$S_m(t) = \sum_{k=1}^{[mt]} \frac{1}{\sqrt{m}} X_k.$$

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

$$S_m(t) = \frac{1}{\sqrt{m}} \sum_{k=1}^{[mt]} X_k = \frac{\sqrt{[mt]}}{\sqrt{m}} \frac{1}{\sqrt{[mt]}} \sum_{k=1}^{[mt]} X_k \to \sqrt{t} N(0,1) = N(0,t).$$

更严格的数学分析表明, 随机过程$S_m(t)$随着$m \to \infty$会收敛到标准布朗运动B. 下面分别是$m = 100, 1000, 5000$时, $S_m(t)$的轨道局部($t \in [0,1]$). 当m比较大时, $S_m(t)$可以近似看作布朗运动.



性质19.1 布朗运动B的均值函数$m(t)\equiv 0$; 方差函数$Var(B(t))=\sigma^2 t$; 协方差函数

$$R(s,t)=\mathbb{E}(B(s)B(t))=(s\wedge t)\sigma^2.$$

证明: 只需验证协方差函数. 不妨设$s<t$,

$$R(s,t)=\mathbb{E}(B(s)B(t))=\mathbb{E}(B(s)(B(t)-B(s))+\mathbb{E}(B^2(s))=\sigma^2 s. \qquad \square$$

例19.1 设$0<t_1<t_2<t_3$, 求$\mathbb{E}(B(t_1)B(t_2)B(t_3))$.
解 由独立增量与正态性可知

$$\begin{aligned}\mathbb{E}(B(t_1)B(t_2)B(t_3)) &= \mathbb{E}\big(B(t_1)\big(B(t_1)+B(t_2)-B(t_1)\big)\big(B(t_2)+B(t_3)-B(t_2)\big)\big)\\ &= \mathbb{E}\big(B(t_1)B(t_1)B(t_2)\big)+\mathbb{E}\big(B(t_1)\big(B(t_2)-B(t_1)\big)B(t_2)\big)\\ &= \mathbb{E}\big(B^3(t_1)\big)+\mathbb{E}\big(B(t_1)\big(B(t_2)-B(t_1)\big)^2\big)=0. \qquad \square\end{aligned}$$

定理19.2 设$B=\{B(t);\ t\geq 0\}$为标准布朗运动,那么下述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$\{X(t), 0 \le t \le u\} = \{B(u) - B(u-t), 0 \le t \le u\}$为标准布朗运动. $\square$

注19.1 (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$\{tB(1/t);\ t \ge 0\}$(其中$t = 0$时取值为0)也是标准布朗运动.

(B) 标准布朗运动的有限维分布与条件分布

下面我们研究标准布朗运动的有限维分布. 为此, 总设$B = \{B(t);\ t \ge 0\}$为标准布朗运动并记

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

即对任意$a \in R$, $t > 0$, $f(x-a,t)$是正态分布$N(a,t)$的密度函数. 容易验证

(1) 对任意$a, b \in R, s, t > 0$,

$$f(x-a,s)f(x-b,t) = f(a-b,s+t)f(x - \frac{ta+sb}{s+t}, \frac{st}{s+t}). \tag{6.1.1}$$

(2)对任意$x \in R$,

$$\int_x^\infty f(y-a,t)dy = \int_{(x-a)/\sqrt{t}}^\infty f(u,1)du = \Phi((a-x)/\sqrt{t})$$

其中Φ为标准正态分布函数.

定理19.3 对任意$n \ge 1$, $0 \le t_1 < t_2 < \cdots < t_n$, $(B(t_1), B(t_2), \cdots, B(t_n))$是均值为0, 协方差矩阵为

$$\mathbb{D} = (t_i \wedge t_j)_{i,j=1,\cdots n} = \begin{pmatrix} t_1 & t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & t_2 & \cdots & t_2 \\ t_1 & t_2 & t_3 & \cdots & t_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_1 & t_2 & t_3 & \cdots & t_n \end{pmatrix}$$

的n维正态随机变量. 因此$(B(t_1), \cdots, B(t_n))$的联合密度函数为

$$\rho(x_1, \cdots, x_n; t_1, \cdots, t_n) = f(x_1, t_1) \prod_{i=2}^n f(x_i - x_{i-1}, t_i - t_{i-1}).$$

证明 对任意$0 \le t_1 < t_2 < \cdots < t_n, \lambda_1, \cdots, \lambda_n \in \mathbb{R}$ 由布朗运动的平稳正态独立增量性质可知

$$\begin{aligned} \mathbb{E}\Big(\exp\Big\{\mathrm{i}\sum_{k=1}^n \lambda_k B(t_k)\Big\}\Big) &= \mathbb{E}\Big(\exp\Big\{\mathrm{i}\sum_{k=1}^n (B(t_k) - B(t_{k-1})) \sum_{j=k}^n \lambda_j\Big\}\Big) \\ &= \prod_{k=1}^n \mathbb{E}\Big(\exp\Big\{\mathrm{i}(B(t_k) - B(t_{k-1})) \sum_{j=k}^n \lambda_j\Big\}\Big) \\ &= \exp\Big\{-\sum_{k=1}^n \frac{t_k - t_{k-1}}{2}\Big(\sum_{j=k}^n \lambda_j\Big)^2\Big\}. \end{aligned}$$

直接验证可知

$$\begin{aligned}\sum_{k=1}^{n}\frac{t_k-t_{k-1}}{2}\Big(\sum_{j=k}^{n}\lambda_j\Big)^2 &= \frac{1}{2}\Big[\sum_{k=1}^{n}t_k\Big(\sum_{j=k}^{n}\lambda_j\Big)^2-\sum_{k=1}^{n-1}t_k\Big(\sum_{j=k+1}^{n}\lambda_j\Big)^2\Big]\\ &= \frac{1}{2}\Big[\sum_{k=1}^{n}2t_k\lambda_k\sum_{j=k+1}^{n}\lambda_j+\sum_{k=1}^{n}t_k\lambda_k^2\Big]\\ &= \frac{(\lambda_1,\cdots,\lambda_n)\mathbb{D}(\lambda_1,\cdots,\lambda_n)^T}{2}.\end{aligned}$$

因此$(B(t_1),\cdots,B(t_n))$服从均值为0协方差矩阵为$\mathbb{D}$的n维正态分布, 联合密度函数

$$\rho(x_1,\cdots,x_n;t_1,\cdots,t_n)=\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbb{D}|}}\mathrm{e}^{-\frac{(x_1,\cdots,x_n)\mathbb{D}^{-1}(x_1,\cdots,x_n)^T}{2}}.$$

由矩阵求逆计算可知$\mathbb{D}^{-1}$可表示为

$$\begin{pmatrix}1&-1&\cdots&0&0\\0&1&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&-1\\0&0&\cdots&0&1\end{pmatrix}\begin{pmatrix}\frac{1}{t_1}&0&\cdots&0&0\\0&\frac{1}{t_2-t_1}&\cdots&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&\frac{1}{t_{n-1}-t_{n-2}}&0\\0&0&\cdots&0&\frac{1}{t_n-t_{n-1}}\end{pmatrix}\begin{pmatrix}1&0&\cdots&0&0\\-1&1&\cdots&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&1&0\\0&0&\cdots&-1&1\end{pmatrix}$$

令$x_0=0$, $t_0=0$, 将$\mathbb{D}^{-1}$的上述表达式代入联合密度函数得

$$\begin{aligned}\rho(x_1,\cdots,x_n;t_1,\cdots,t_n) &= \frac{1}{\prod_{i=1}^{n}\sqrt{2\pi(t_i-t_{i-1})}}\mathrm{e}^{-\sum_{i=1}^{n}\frac{(x_i-x_{i-1})^2}{2(t_i-t_{i-1})}}\\ &= f(x_1,t_1)\prod_{i=2}^{n}f(x_i-x_{i-1},t_i-t_{i-1}).\end{aligned}$$ □

推论19.4 当$B(t_k)=a$时, $(B(t_1),\cdots,B(t_{k-1}),B(t_{k+1}),\cdots,B(t_n))$的密度函数为

$$\begin{aligned}&\varrho(x_1,\cdots,x_{k-1},x_{k+1},\cdots,x_n;t_1,\cdots t_{k-1},t_{k+1},\cdots,t_n|a,t_k)\\ &=\rho(x_1,\cdots,x_{k-1},a,x_{k+1},\cdots,x_n;t_1,\cdots,t_n)/f(a,t_k).\end{aligned}$$

特别 $\varrho(x_2,\cdots,x_n;t_2,\cdots,t_n|a,t_1)=f(x_2-a,t_2-t_1)\prod_{i=3}^{n}f(x_i-x_{i-1},t_i-t_{i-1})$,

$$\varrho(x_1,\cdots,x_{n-1};\cdots,t_{n-1}|a,t_n)=\frac{f(a-x_{n-1},t_n-t_{n-1})}{f(a,t_n)}\prod_{i=1}^{n-1}f(x_i-x_{i-1},t_i-t_{i-1}).$$

证明 由条件概率密度公式, 我们只需证明

$$\int_{-\infty}^{\infty}dy_1\cdots\int_{-\infty}^{\infty}\rho(y_1,\cdots,y_{k-1},a,y_{k+1},\cdots,y_n;t_1,\cdots,t_n)dy_n=f(a,t_k).\quad(6.1.2)$$

事实上由

$$\begin{aligned}\text{l.h.s of (6.1.2)} &= \int_{-\infty}^{\infty} f(y_1,t_1)dy_1 \int_{-\infty}^{\infty} f(y_2-y_1,t_2-t_1)dy_2 \\ &\times\cdots\int_{-\infty}^{\infty} f(y_{k-1}-y_{k-2},t_{k-1}-t_{k-2})f(a-y_{k-1},t_k-t_{k-1})dy_{k-1} \\ &\times\int_{-\infty}^{\infty} f(y_{k+1}-a,t_{k+1}-t_k)dy_{k+1}\cdots\int_{-\infty}^{\infty} f(y_n-y_{n-1},t_n-t_{n-1})dy_n\end{aligned}$$

可知

$$\begin{aligned}&\text{l.h.s of (6.1.2)} = \int_{-\infty}^{\infty} f(y_1,t_1)dy_1\cdots\int_{-\infty}^{\infty} f(y_{k-2}-y_{k-3},t_{k-2}-t_{k-3})dy_{k-2} \\ &\quad\times\int_{-\infty}^{\infty} f(y_{k-1}-y_{k-2},t_{k-1}-t_{k-2})f(a-y_{k-1},t_k-t_{k-1})dy_{k-1} \\ &=\int_{-\infty}^{\infty} f(y_1,t_1)dy_1\cdots\int_{-\infty}^{\infty} f(y_{k-2}-y_{k-3},t_{k-2}-t_{k-3})f(a-y_{k-2},t_k-t_{k-2})dy_{k-2} \\ &=\cdots=\int_{-\infty}^{\infty} f(y_1,t_1)f(a-y_1,t_k-t_1)dy_1 = \text{r.h.s. of (6.1.2)},\end{aligned}$$

其中第二个等号用了公式(6.1.1). □

例19.2 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

$$p=\mathbb{P}(B(1/2)>0|B(1)=1)$$

由于在$B(1)=1$条件下, $B(1/2)\sim N(1/2,1/4)$, 因此

$$p=\mathbb{P}(N(1/2,1/4)>0)=\mathbb{P}(N(0,1)>-1)=\Phi(1)\approx 0.8413,$$

其中Φ为标准正态分布函数. □

在给定两个时刻取值的条件下, 我们有如下的结论, 证明省略.

推论19.5 $B(s)=x, B(t)=y$时, 任取$s\le t_1<t_2<\cdots<t_n\le t$, $(B(t_1),\cdots,B(t_n))$是均值为

$$\Big(x+\frac{y-x}{t-s}(t_i-s)\Big)_{1\le i\le n},$$

协方差矩阵

$$\mathbb{D}=\Big(\frac{(t-t_{i\vee j})(t_{i\wedge j}-s)}{t-s}\Big)_{i,j=1,\cdots,n}$$

的n维正态随机变量.

(C) 马氏性与轨道性质

定理19.6 任取$s,t>0$, $0\le t_1<t_2<\cdots<t_n<t$以及$x_1,x_2,\cdots,x_n,x,y\in R$,

$$\mathbb{P}(B(s+t)\le y|B(t)=x,B(t_n)\le x_n,\cdots,B(t_1)\le x_1)=\mathbb{P}(B(t+s)\le y|B(t)=x)$$

证明 利用条件分布计算公式可得

$$\begin{aligned}&\mathbb{P}(B(s+t)\le y|B(t)=x,B(t_n)\le x_n,\cdots,B(t_1)\le x_1)\\&\quad=\frac{\mathbb{P}(B(s+t)\le y,B(t_n)\le x_n,\cdots,B(t_1)\le x_1|B(t)=x)}{\mathbb{P}(B(t_n)\le x_n,\cdots,B(t_1)\le x_1|B(t)=x)}.\end{aligned}$$

由推论19.4

$$\begin{aligned}&\mathbb{P}(B(s+t)\le y,B(t_n)\le x_n,\cdots,B(t_1)\le x_1|B(t)=x)\\&\quad=\int_{-\infty}^{x_1}f(y_1,t_1)\mathrm{d}y_1\cdots\int_{-\infty}^{x_n}f(y_n-y_{n-1},t_n-t_{n-1})\\&\qquad\qquad\times f(x-y_n,t-t_n)\mathrm{d}y_n\int_{-\infty}^{y}f(u-x,s)\mathrm{d}u,\end{aligned}$$

以及

$$\begin{aligned}&\mathbb{P}(B(t_n)\le x_n,\cdots,B(t_1)\le x_1|B(t)=x)\\&\quad=\int_{-\infty}^{x_1}f(y_1,t_1)\mathrm{d}y_1\cdots\int_{-\infty}^{x_n}f(y_n-y_{n-1},t_n-t_{n-1})f(x-y_{n-1},t-t_n)\mathrm{d}y_n.\end{aligned}$$

因此

$$\begin{aligned}&\mathbb{P}(B(s+t)\le y|B(t)=x,B(t_n)\le x_n,\cdots,B(t_1)\le x_1)\\&\quad=\int_{-\infty}^{y}f(u-x,s)\mathrm{d}u=\mathbb{P}(B(t+s)\le y|B(t)=x),\end{aligned}\tag{6.1.3}$$

其中最后一个等式由推论19.4推出. □

注19.1 定理19.6的结论形式恰是所谓连续时间连续状态随机过程马氏性的表示, 因此布朗运动也是连续时间马氏过程的一个具体例子. 记$B(t)=x$ 条件下$B(t+s)$的概率密度函数为$p(t,s+t,x,y)$, 由(6.1.3) 可知$p(t,s+t,x,y)=f(y-x,s)$. 在马氏过程中常称$p(t,s+t,x,y)$ 为从t 时刻状态x转移到$s+t$时刻状态y的转移概率密度(函数). 更多关于连续时间连续状态马氏过程的一般性介绍, 本书从略.

例19.3 在$\{B(1)=0,B(3)=u\}$条件下, 求事件$\{B(2)>u,B(4)>u\}$发生的概率.

解 由定理19.6, 所求概率为

$$\begin{aligned}&\mathbb{P}(B(2)>u,B(4)>u|B(3)=u,B(1)=0)\\&\quad=\mathbb{P}(B(2)>u|B(3)=u,B(1)=0)\mathbb{P}(B(4)>u|B(3)=u,B(2)>u,B(1)=0)\\&\quad=\mathbb{P}(B(2)>u|B(3)=u,B(1)=0)\mathbb{P}(B(4)>u|B(3)=u).\end{aligned}$$

由推论19.5可知(对应地取$s=1,t=3,t_1=2,\ x=0,\ y=u$), $\{B(3)=u,B(1)=0\}$条件下$B(2)\sim N(u/2,1/2)$. 所以

$$\begin{aligned}\mathbb{P}(B(2)>u|B(3)=u,B(1)=0) &= \mathbb{P}(N(u/2,1/2)>u)\\ &= \mathbb{P}(N(0,1)>u/\sqrt{2})=1-\Phi(u/\sqrt{2}).\end{aligned}$$

由推论19.4或布朗运动独立增量性可知$\{B(3)=u\}$条件下, $B(4)\sim N(u,1)$. 因此

$$\mathbb{P}(B(4)>u|B(3)=u)=1/2.$$

因此所求概率$p=(1-\Phi(u/\sqrt{2}))/2$. □

定理19.7 标准布朗运动$B=\{B(t);\ t\geq 0\}$的几乎所有轨道处处不可微.

证明 只考虑$t\in[0,1]$的情形. 任取轨道ω, 若$B(\cdot,\omega)$在s可微, 则存在$\delta\in(0,1/2)$与整数$l\geq 1$使得当$|t-s|<\delta$时

$$|B(t,\omega)-B(s,\omega)|<l|t-s|.$$

对任意$n>4/\delta$, 令$i=[ns]+1\leq n+1$, 对$j=i\pm 1,i\pm 2,i\pm 3$都应有

$$\begin{aligned}&|B(j/n,\omega)-B((j-1)/n,\omega)|\\ &\quad\leq |B(j/n,\omega)-B(s,\omega)|+|B((j-1)/n,\omega)-B(s,\omega)|\\ &\quad\leq l|j/n-s|+l|(j-1)/n-s|\leq 7l/n. \qquad (6.1.4)\end{aligned}$$

令$A^i_{n,l}=\{\omega;$对给定的 l,i (6.1.4) 对 $j=i+1,i+2,i+3$,或对 $j=i-1,i-2,i-3$成立$\}$. 显然随着l的增加, $A^i_{n,l}$ 单调不降. 特别取$\delta=1/m,\ m=2,3,\cdots$, 则

$$\begin{aligned}\Omega_1 &= \{\omega;\text{存在}s\in[0,1]\text{使得}B(\cdot,\omega)\text{在}s\text{点可微}\}\\ &\subset \bigcup_{l\geq 1,m\geq 2}\bigcap_{n\geq 4m}\bigcup_{i\leq n+1}A^i_{n,l}.\end{aligned}$$

但是当$n\to\infty$时

$$\begin{aligned}\mathbb{P}(\bigcup_{i\leq n+1}A^i_{n,l}) &\leq (n+1)\max_{1\leq i\leq n+1}\mathbb{P}(A^i_{n,l})\\ &\leq 2(n+1)\Big[\mathbb{P}(|B(j/n)-B((j-1)/n)|\leq 7l/n)\Big]^3\\ &= 2(n+1)\Big[\int_{-7l/n}^{7l/n}\frac{1}{\sqrt{2\pi/n}}\mathrm{e}^{-ny^2/2}dy\Big]^3\\ &\leq 2(n+1)(14l/\sqrt{2\pi n})^3\to 0.\end{aligned}$$

因此$\mathbb{P}(\Omega_1)\leq\lim\limits_{l\to\infty}\lim\limits_{n\to\infty}\mathbb{P}(\bigcup\limits_{i\leq n+1}A^i_{n,l})=0$. 即对几乎必然的$\omega$, $B(\cdot,\omega)$处处不可微. □

注意到$B(\cdot,\omega)$是连续的, 布朗运动轨道为我们提供了连续但处处不可微的函数实例.

练习题 以下总设$B=\{B(t),t\geq 0\}$为标准布朗运动.

19.1 对任意$0\leq s<t$, 求$B(s)+B(t)$的分布密度函数.
19.2 已知$B(1)=x$, 对任意$0\leq s_1<s_2\leq s_3<s_4\leq 1$, 证明

(1) $\mathbb{E}((B(s_4)-B(s_3))(B(s_2)-B(s_1)))=(s_4-s_3)(s_2-s_1)(x^2-1)$

(2) $\mathbb{E}[((B(s_4)-B(s_1))^2]=(s_4-s_1)+(s_4-s_1)^2(x^2-1)$

19.3 在例15.2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.
19.4 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$0\leq t<s$, $\mathbb{E}(Y(s)|B(t))=Y(t)$以及$\mathbb{E}(R(s)|B(t))=R(t)$.

§6.2 反射原理与极值分布

对任意$a \in R$, 定义

$$\tau_a = \inf\{t \geq 0;\ B(t) = a\} = \begin{cases} \inf\{t \geq 0;\ B(t) \geq a\}, & a \geq 0; \\ \inf\{t \geq 0;\ B(t) \leq a\}, & a < 0. \end{cases}$$

称τ_a为B首次到达状态a的时刻,简称为a的首达时.

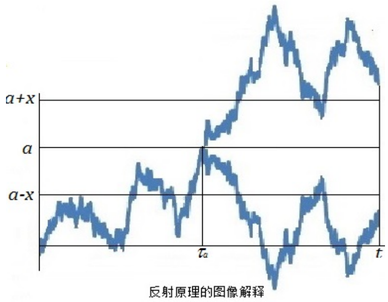
(A) 反射原理

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20.1(反射原理) 对任意$a \in R$及$x \geq 0$,

$$\mathbb{P}(B(t) \geq a + x, \tau_a \leq t) = \mathbb{P}(B(t) < a - x, \tau_a \leq t). \tag{6.2.1}$$

反射原理的直观图形解释如下,



反射原理的图像解释

对该直观解释不严格的数学表示如下: 设τ_a的分布函数为$F(s)$, 那么

$$\begin{aligned}
&\mathbb{P}(B(t) \geq a + x, \tau_a \leq t) = \mathbb{P}(B(t) > a + x, \tau_a \leq t) \\
&\quad = \int_0^t \mathbb{P}(B(t) > a + x | \tau_a = s) \mathrm{d}F(s) \\
&\quad = \int_0^t \mathbb{P}(B(t) > a + x | B(s) = a, B(u) \neq a, 0 \leq u < s) \mathrm{d}F(s) \quad (\text{轨道连续}) \\
&\quad = \int_0^t \mathbb{P}(B(t) > a + x | B(s) = a) \mathrm{d}F(s) \\
&\quad = \int_0^t \int_{a+x}^{\infty} f(y - a, t - s) \mathrm{d}F(s) = \int_0^t \int_{-\infty}^{a-x} f(y - a, t - s) \mathrm{d}F(s) \quad (\text{对称性}) \\
&\quad = \int_0^t \mathbb{P}(B(t) < a - x | B(s) = a) \mathrm{d}F(s) = \mathbb{P}(B(t) < a - x, \tau_a \leq t).
\end{aligned}$$

推论20.2 对任意$a \in R$, $\mathbb{P}(B(t) \geq a | \tau_a \leq t) = 1/2$.

(B) 极值分布

定理20.3 记$B^*(t)=\max\limits_{0\le u\le t}B(u)$. 对任意$z\ge 0$, $x\le z$,

$$\mathbb{P}(B^*(t)\ge z, B(t)<x)=\mathbb{P}(B(t)>2z-x)=\int_{2z-x}^{\infty}f(y,t)\mathrm{d}y.$$

因此$(B^*(t),B(t))$的联合密度函数为

$$h(z,x)=2f'(2z-x,t)=\frac{2(2z-x)}{\sqrt{2\pi t^3}}\mathrm{e}^{-\frac{(2z-x)^2}{2t}},\quad x\le z, z\ge 0.$$

证明 由轨道连续性及$B(0)=0$可知, 对任意$z\ge 0$,

$$\{B^*(t)\ge z\}=\{\tau_z\le t\} \tag{6.2.2}$$

从而由反射原理(6.2.1)可得

$$\begin{aligned}&\mathbb{P}(B^*(t)\ge z, B(t)<x)=\mathbb{P}(\tau_z\le t, B(t)<x)\\&\quad=\mathbb{P}(B(t)<z-(z-x),\tau_z\le t)=\mathbb{P}(B(t)>z+(z-x),\tau_z\le t)\\&\quad=\mathbb{P}(B(t)>2z-x)=\int_{2z-x}^{\infty}f(y,t)\mathrm{d}y.\end{aligned}$$

上式两边关于z,x求二阶混合偏导即得联合密度函数$h(z,x)$. □

例20.1 求$\mathbb{P}(B^*(t)>x|B(t)=B^*(t))$.

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

$$\mathbb{P}(B^*(t)>x|B(t)=B^*(t))=\int_x^\infty g(z)\mathrm{d}z=\mathrm{e}^{-\frac{x^2}{2t}}.$$ □

推论20.4 对任意$z\ge 0$,

$$\mathbb{P}(B^*(t)\ge z)=2\mathbb{P}(B(t)\ge z)=\mathbb{P}(|B(t)|\ge z),$$

因此$B^*(t)$与$|B(t)|$ 同分布, $B^*(t)$的密度函数

$$g(z)=\frac{2}{\sqrt{2\pi t}}\mathrm{e}^{-z^2/2t}.$$

证明 对任意$z\ge 0$

$$\begin{aligned}\mathbb{P}(B^*(t)\ge z)&=\mathbb{P}(B^*(t)\ge z, B(t)\ge z)+\mathbb{P}(B^*(t)\ge z, B(t)<z)\\&=\mathbb{P}(B(t)\ge z)+\mathbb{P}(B(t)>z)=2\mathbb{P}(B(t)\ge z)\\&=2\int_z^\infty\frac{1}{2\pi t}\mathrm{e}^{-u^2/2t}\mathrm{d}u=\mathbb{P}(|B(t)|\ge z).\end{aligned}$$

推论得证. □

例20.2 假定一只股票价格按标准布朗运动变化，若你以价格$b+c$买入, 而现在的价格恰好为b, 你计划在股票价格回到$b+c$元或最多再观望T 时间后将股票抛出, 问你不能重新获得买入价格的概率($c>0$).

解 从现在开始看股票的价格$S(t)$是一个从b出发的布朗运动, 即$S(t)=b+B(t)$. 所谓不能重新获得买入价格就是指在时刻T之前，$S(t)$最大值小于$b+c$, 即

$$B_T^*=\max_{0\le s\le T}B(s)<c.$$

由推论16.4可知所求概率为

$$p=1-\mathbb{P}(B^*(T)\ge c)=1-\mathbb{P}(|B(T)|>c)=1-2\Phi(-c/\sqrt{T}).$$ □

推论20.5 对任意$z>0$, τ_z的分布密度函数

$$p(t)=\frac{z}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t}.$$

因此$\mathbb{P}(\tau_z<\infty)=1$, $\mathbb{E}(\tau_z)=\infty$.

证明 对任意$z\ge 0$, 由(6.2.2)可知

$$\mathbb{P}(\tau_z\le t)=\mathbb{P}(B^*(t)\ge z)=2\int_z^\infty\frac{1}{\sqrt{2\pi t}}\mathrm{e}^{-u^2/2t}\mathrm{d}u=\int_{z/\sqrt{t}}^\infty\frac{2}{\sqrt{2\pi}}\mathrm{e}^{-u^2/2}\mathrm{d}u$$

两边关于t求导得

$$p(t)=\frac{z}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t}.$$

即

$$\mathbb{P}(\tau_z\le t)=\int_0^t\frac{z}{\sqrt{2\pi s^3}}\mathrm{e}^{-z^2/2s}\mathrm{d}s.$$

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

$$\mathbb{E}(\tau_z)=\int_0^\infty\frac{z}{\sqrt{2\pi s}}\mathrm{e}^{-z^2/2s}\mathrm{d}s=\infty.$$

推论得证. □

注意对任意$z\ge 0$, τ_z与τ_{-z}分布相同(见习题20.5), 因此对任意$z\in R$, τ_z的密度函数

$$p(t)=\frac{|z|}{\sqrt{2\pi t^3}}\mathrm{e}^{-z^2/2t}.$$

因此推论20.5表明对空间R中任意一点z, 布朗运动必然在有限时间内可到达这一点,而到达这一点的平均时长为无穷.

例20.3 求例19.2中在比赛后半程甲都不落后于乙的概率.
解 依题设$Y(t)$是方差为σ的布朗运动, 因而$B(t)=Y(t)/\sigma$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

$$p=\mathbb{P}(B_*\ge 0|B(1)=1).$$

注意到$W=-B$仍为标准布朗运动, 因此

$$p=\mathbb{P}(W^*\le 0|W(1)=-1),$$

其中$W^*=\max\{W(u);\ 1/2\le u\le 1\}$. 记$W(1)=-1$下的条件概率为$\mathbb{P}_1$, 数学期望为$\mathbb{E}_1$ 那么

$$p=\mathbb{P}_1(W^*\le 0)=\mathbb{E}_1\Big(\mathbb{P}_1\big(W^*\le 0|W(1/2)\big)\Big).$$

由推论19.4可知,$W(1)=-1$条件下$W(1/2)\sim N(-1/2,1/4)$, 因此

$$p=\int_{-\infty}^{0}\mathbb{P}_1\big(W^*\le 0|W(1/2)=x\big)f(x+1/2,1/4)\mathrm{d}x.$$

注意到

$$\begin{aligned}\mathbb{P}_1(W^*\le 0|W(1/2)=x)&=\mathbb{P}(W^*\le 0|W(1/2)=x,W(1)=-1)\\&=\mathbb{P}(\max\{U(t),0\le t\le 1/2\}\le -x|U(0)=0,U(1/2)=-1-x),\end{aligned}$$

其中$U(t)=W(t+1/2)-W(1/2)$为标准布朗运动. 将其重新记作B得

$$\mathbb{P}_1(W^*\le 0|W(1/2)=x)=\mathbb{P}\big(B^*(1/2)\le -x|B(1/2)=-1-x\big).$$

将定理20.3中对应的x换成$-1-x$, t换成$1/2$可得$B(1/2)=-1-x$时$B^*(1/2)$的边际分布密度为

$$g(z|-1-x)=4(2z+1+x)\mathrm{e}^{-(2z+1+x)^2+(1+x)^2},\qquad z\ge 0.$$

因此

$$\mathbb{P}_1(W^*\le 0|W(1/2)=x)=\int_0^{-x}4(2z+1+x)\mathrm{e}^{-(2z+1+x)^2+(1+x)^2}\mathrm{d}z=1-\mathrm{e}^{4x}.$$

进而

$$p=\int_{-\infty}^{0}(1-\mathrm{e}^{4x})\frac{1}{\sqrt{\pi/2}}\mathrm{e}^{-2(x+1/2)^2}\mathrm{d}x=2\Phi(1)-1.\qquad\square$$

(C) 零点概率

定理20.6 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

$$p_t(x) = \frac{|x|}{\sqrt{2\pi}} \int_0^t s^{-3/2} \mathrm{e}^{-\frac{x^2}{2s}} \mathrm{d}s.$$

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$x > 0$, 此时

$$\begin{aligned} p_t(x) &= \mathbb{P}(\min_{0 \le u \le t} B(u) + x \le 0) = \mathbb{P}(\min_{0 \le u \le t} B(u) \le -x) \\ &= \mathbb{P}(\max_{0 \le u \le t} (-B(u)) \ge x) = \mathbb{P}(\tau_x \le t) \qquad (\text{因为}-B\text{也是标准布朗运动}) \\ &= \frac{x}{\sqrt{2\pi}} \int_0^t s^{-3/2} \mathrm{e}^{-\frac{x^2}{2s}} \mathrm{d}s. \end{aligned}$$

综合两种情况可知定理成立. □

推论20.7 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20.6

$$\begin{aligned} q &= \mathbb{E}(p_{t-s}(B(s))) = 2\int_0^\infty f(x,s) p_{t-s}(x) \mathrm{d}s \\ &= 2\int_0^{t-s} \mathrm{d}u \int_0^\infty f(x,s) f(x,u) \frac{x}{u} \mathrm{d}x \\ &= 2\int_0^{t-s} \frac{f(0,s+u)}{u} \mathrm{d}u \int_0^\infty x f(x, \frac{su}{s+u}) \mathrm{d}x = \frac{1}{\pi} \int_0^{t-s} \frac{1}{s+u} \sqrt{\frac{s}{u}} \mathrm{d}u. \end{aligned}$$

积分变量代换$v = \sqrt{u/s}$后得

$$\begin{aligned} q &= \frac{2}{\pi} \int_0^{\sqrt{(t-s)/s}} \frac{1}{1+v^2} \mathrm{d}v \qquad (\text{积分变换 }) \\ &= \frac{2}{\pi} \arctan(\sqrt{(t-s)/s}) = \frac{2}{\pi} \arctan \sqrt{\frac{t-s}{s}} = \frac{2}{\pi} \arccos \sqrt{\frac{s}{t}}. \end{aligned}$$

推论得证. □

例20.4 设$c > 0$为常数, 求$p = \mathbb{P}(\text{存在}t_0 \in (1,2)\text{使得}B(t_0) = ct_0 | B(3) = 3c)$.

解 令$W(t)=tB(1/t)$, 那么$W(t)$仍为标准布朗运动, 此时

$$\{存在t_0\in(1,2)使得B(t_0)=ct_0\} = \{存在t_0\in(1/2,1)使得W(t_0)=c\}.$$

进而由推论20.7可知

$$\begin{aligned} p &= \mathbb{P}(存在t_0\in(1/2,1)使得W(t_0)=c|W(1/3)=c) \\ &= \mathbb{P}(存在t_0\in(1/6,2/3)使得W(t_0)=0)=\frac{2}{\pi}\arccos\sqrt{\frac{1}{4}}=2/3. \end{aligned}$$ □

定理20.8* 令$v_t=\sup\{s;0\le s\le t,B(s)=B^*(t)\}$, 那么$v_t$的密度函数为

$$\nu(s)=\frac{1}{\pi\sqrt{s(t-s)}},\qquad 0<s<t.$$

证明 注意到对任意$0<s<t$,

$$\{v_t\le s\}=\{\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)\}.$$

因此

$$\begin{aligned} \mathbb{P}(v_t\le s) &= \mathbb{P}(\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)) \\ &= \mathbb{E}\big(\mathbb{P}(\max_{0\le u\le s}B(u)>\max_{s<u\le t}B(u)|B^*(s),B(s))\big) \\ &= \mathbb{E}\big(\mathbb{P}(B^*(s)-B(s)>\max_{s<u\le t}(B(u)-B(s))|B^*(s),B(s))\big). \end{aligned}$$

由平稳独立增量性可得

$$\begin{aligned} \mathbb{P}(v_t\le s) &= \mathbb{E}\Big(\mathbb{P}\big(B^*(s)-B(s)>\max_{0\le u\le t-s}W(u)\big)\Big), \\ &= \mathbb{E}\Big(\int_0^{B^*(s)-B(s)}2f(u,t-s)\mathrm{d}u\Big) \end{aligned}$$

其中W表示标准布朗运动且与B独立. 由定理19.3,

$$\begin{aligned} \mathbb{P}(v_t\le s) &= \int_0^\infty \mathrm{d}z\int_0^\infty 4f(z+y,s)\frac{z+y}{s}\mathrm{d}y\int_0^y f(u,t-s)\mathrm{d}u \\ &= \int_0^\infty \mathrm{d}z\int_0^\infty 4f(z+u,s)f(u,t-s)\mathrm{d}u \\ &= \int_0^\infty 4f(z,t)\mathrm{d}z\int_0^\infty f(u+\frac{t-s}{t}z,\frac{(t-s)s}{t})\mathrm{d}u \\ &= \int_0^\infty 4f(z,t)\mathrm{d}z\int_{z\sqrt{\frac{(t-s)}{ts}}}^\infty f(u,1)\mathrm{d}u. \end{aligned}$$

两边求导得

$$\begin{aligned}\nu(s) &= \int_0^\infty 4f(z,t)f(z\sqrt{\frac{t-s}{ts}},1)\frac{\sqrt{t}z}{2\sqrt{s^3(t-s)}}\mathrm{d}z\\ &= \int_0^\infty \frac{z}{\pi\sqrt{s^3(t-s)}}\mathrm{e}^{-\frac{z^2}{2s}}\mathrm{d}z = \frac{1}{\pi\sqrt{s(t-s)}}\end{aligned}$$

定理得证. □

练习题

20.1 求$\mathbb{E}(B^*(t)|B(t)=x)$.
20.2 求$\mathbb{P}(B(1)\le x|B(u)\ge 0, 0\le u\le 1)$.
20.3 股票A,B的价格比服从随机过程$Y(t)=\mathrm{e}^{B(t)}$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 股票价格曾超过A 价格的概率
20.4 证明对任意$z\ge 0$, τ_z与τ_{-z}同分布.
20.5 对任意$\lambda\ge 0$, 求$\mathbb{E}(\mathrm{e}^{-\lambda\tau_z})$.
20.6 当$0<z\to\infty$时, 证明$\tau_z\to+\infty$, a.s. ([提示] τ_z单调不降从而极限必然存在, 对任意$t>0$, $\mathbb{P}(\lim_{z\to\infty}\tau_z>t)=\lim_{z\to\infty}\mathbb{P}(\tau_z>t)=1$.)
20.7* 令$\gamma_t=\sup\{s\le t; B(s)=0\}$, 证明

$$\mathbb{P}(\gamma_t\le x)=\frac{2}{\pi}\arcsin\sqrt{\frac{x}{t}},\qquad x\le t.$$

§6.3 几何布朗运动与Black-Scholes公式(待改写)

通常称非负随机过程$\{e^{B(t)};\ t \ge 0\}$为几何布朗运动. 几何布朗运动常常用来刻画金融资产价格的变化，在金融数学中有广泛的应用.

§6.4 高斯过程与积分布朗运动

(A) 高斯过程与布朗桥

定义22.1 设$X = \{X(t);\ t \in T\}$是概率空间$(\Omega, \mathfrak{F}, \mathbb{P})$上的随机过程. 称$X$为高斯过程(系), 若对任意的$t_1, \cdots, t_n \in T$,

$$(X(t_1), X(t_2), \cdots, X(t_n))$$

服从正态分布(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$\lambda = (\lambda_1, \cdots, \lambda_n)^T \in R^n$,

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}\sum_{k=1}^n \lambda_k X(t_k)}) = \exp\Big\{\mathrm{i}\lambda^T\mu - \frac{1}{2}\lambda^T\mathbf{D}\lambda\Big\}.$$

由此可知$(X(t_1), X(t_2), \cdots, X(t_n))$的均值向量为$\mu$, 协方差矩阵为$\mathbf{D}$.

显然高斯过程的有限维分布由其均值(函数)和协方差(函数)唯一确定. 由定理19.2, 布朗运动B 是一个高斯过程; 再由推论19.4,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22.1 $X = \{X(t);\ t \in T\}$是高斯系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

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}u(\sum_{k=1}^n \lambda_k X(t_k))}) = \mathbb{E}(\mathrm{e}^{\mathrm{i}(\sum_{k=1}^n u\lambda_k X(t_k))}) = \exp\Big\{\mathrm{i}u\lambda^T\mu - \frac{u^2}{2}\lambda^T\mathbf{D}\lambda\Big\}$$

由此可知$\sum_{k=1}^n \lambda_k X(t_k) \sim N(\lambda^T\mu, \lambda^T\mathbf{D}\lambda)$.

"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,

$$\begin{aligned}
\mathbb{E}(\mathrm{e}^{\mathrm{i}\sum_{k=1}^n \lambda_k X(t_k)}) &= \exp\Big\{\mathrm{i}\mathbb{E}(\sum_{k=1}^n \lambda_k X(t_k)) - \frac{1}{2}\mathrm{Var}(\sum_{k=1}^n \lambda_k X(t_k))\Big\} \\
&= \exp\Big\{\mathrm{i}\sum_{k=1}^n \lambda_k \mathbb{E}(X(t_k)) - \frac{1}{2}\sum_{k,j=1}^n \lambda_k\lambda_j \mathrm{Cov}(X(t_k), X(t_j))\Big\} \\
&= \exp\Big\{\mathrm{i}\lambda^T\mu - \frac{1}{2}\lambda^T\mathbf{D}\lambda\Big\}
\end{aligned}$$

其中$\mu = \{\mathbb{E}(X(t_1), \cdots, \mathbb{E}(X(t_n)))\}$, $\mathbf{D} = \big(\mathrm{Cov}(X(t_i), X(t_j))\big)_{i,j}$正定或半正定. □

例22.1 设W为标准布朗运动, 对任意$x, y \in \mathbb{R}$, 给定$0 \le s < t$, 对任意$u \in [s, t]$, 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$s \le u_1 < \cdots < u_n \le t,\ \lambda_1, \cdots, \lambda_n \in \mathbb{R}$,

$$\begin{aligned}\sum_{k=1}^n \lambda_k W_x^y(t_k)) &= \sum_{k=1}^n \lambda_k \Big[W(u_k) + \Big(\frac{t-u}{t-s}(x - W(s)) + \frac{u-s}{t-s}(y - W(t))\Big)\Big] \\ &= \sum_{k=1}^n \lambda_k W(u_k) - \Big(\sum_{k=1}^n \frac{\lambda_k(t-u_k)}{t-s}\Big)W(s) - \Big(\sum_{k=1}^n \frac{\lambda_k(u_k - s)}{t-s}\Big)W(t) + C.\end{aligned}$$

其中常数$C = \sum_{k=1}^n \frac{\lambda_k(u_k-s)}{t-s}(y-x)$. 由于标准布朗运动是高斯过程, 定理22.1表明布朗桥也是高斯过程. 注意到$W_x^y(s) \equiv x, W_x^y(t) \equiv y$,

$$\mathbb{E}(W_x^y(u)) = x + \frac{u-s}{t-s}(y-x),$$

而且直接计算可知对任意$u \le v \in [s,t]$,

$$\begin{aligned}&\mathrm{Cov}(W_x^y(u), W_x^y(v)) \\ &= \mathrm{Cov}\Big(W(u) - \frac{t-u}{t-s}W(s) - \frac{u-s}{t-s}W(t), W(v) - \frac{t-v}{t-s}W(s) - \frac{v-s}{t-s}W(t)\Big) \\ &= \mathrm{Cov}\Big(W(u), W(v) - \frac{t-v}{t-s}W(s) - \frac{v-s}{t-s}W(t)\Big) = \frac{(t-v)(u-s)}{t-s}.\end{aligned}$$

由于高斯过程有限维分布由其均值函数和协方差函数唯一确定, 对照推论15.4,可知

$$\{W_x^y(u);\ u \in [s,t]\}$$

与给定条件$B(s) = x, B(t) = y$的布朗运动$\{B(u), s \le u \le t\}$有相同的有限维分布,从而可看作是它的一个版本. 注意在$\{W_x^y(u)\}$的构造中,我们并没有对标准布朗运动W设定额外的条件. □

推论22.2 若$\{X(t);\ t \in T\}$是高斯过程, 那么对任意$s_k < t_k, k = 1, 2, \cdots, n$,

$$\{X(t_k) - X(s_k), k = 1, 2, \cdots, n\}$$

是n维正态随机变量.

证明 对任意$\lambda_k \in \mathbb{R}, k = 1, 2, \cdots, n$, 由定理22.1

$$\sum_{k=1}^n \lambda_k (X(t_k) - X(s_k)) = \sum_{k=1}^n \lambda_k X(t_k) + \sum_{k=1}^n (-\lambda_k) X(s_k)$$

为一维正态随机变量, 进而再用一次定理17.1可知推论成立. □

命题22.3 轨道连续的高斯过程$\{X(t), t \ge 0\}$若对任意$0 \le s \le t$都有

$$\mathbb{E}(X(s)) = 0 \text{ 且 } \mathbb{E}(X(s)X(t)) = s,$$

那么$X = \{X(t);\ t \ge 0\}$是标准布朗运动.

证明 只需验证布朗运动的三个条件(a),(b),(c).
(a) X轨道连续且由条件$\mathbb{E}(X(0)^2)=0$可知$X(0)=0$ a.s.
(b) 对任意$0\le s_1<t_1\le s_2<t_2$,

$$\begin{aligned}&\mathbb{E}[(X(t_2)-X(s_2))(X(t_1)-X(s_1))]\\&\quad=\mathbb{E}[X(t_2)X(t_1)]-\mathbb{E}[X(t_2)X(s_1)]-\mathbb{E}[X(s_2)X(t_1)]+\mathbb{E}[X(s_2)X(s_1)]=0.\end{aligned}$$

由于$(X(t_2)-X(s_2),X(t_1)-X(s_1))$为二维正态随机变量, 从而独立. 因此对任意$0\le t_1<t_2<\cdots<t_n$

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

的协方差矩阵为对角矩阵. 这表明$(X(t_1),X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1}))$独立.
(c) 对任意$0\le s<t$, 由$\mathbb{E}(X(t)-X(s))=0$,

$$\mathbb{E}[(X(t)-X(s))^2]=\mathbb{E}[X^2(t)+X^2(s)-2X(s)X(t)]=t-s$$

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22.4 设ξ_n, $n\ge 1$, 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

$$\mathbb{E}(X)=\lim_{n\to\infty}\mathbb{E}(\xi_n),\quad \mathrm{Var}(X)=\lim_{n\to\infty}\mathrm{Var}(X_n).$$

证明 由于ξ_n, $n\ge 1$, 是一列正态随机变量, 存在μ_n, σ_n^2使得对任意$\lambda\in R$

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda\xi_n})=\exp\{\mathrm{i}\lambda\mu_n-\frac{\lambda^2}{2}\sigma_n^2\}.$$

由$\xi_n\xrightarrow{P}X$可知$\mathbb{P}(|X|<\infty)=1$而且$\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda\xi_n})\to\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda\xi}):=\Phi(\lambda)$. 因此

$$|\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda\xi_n})|=\mathrm{e}^{-\frac{\lambda^2}{2}\sigma_n^2}\to|\Phi(\lambda)|,\qquad \mathrm{i}\lambda\mu_n-\frac{\lambda^2}{2}\sigma_n^2\to\ln\Phi(\lambda).$$

注意到$\Phi(0)=1$且由$\Phi(\lambda)$连续可知存在$\delta>0$使得对任意$\lambda\in(-\delta,\delta)$, $|\Phi(\lambda)|>0$. 这表明$n\to\infty$ 时, σ_n^2 极限存在且有限, 记作σ^2. 注意到

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda(\xi_n-\mu_n)})=\exp\{-\frac{\lambda^2}{2}\sigma_n^2\}\to\exp\{-\frac{\lambda^2}{2}\sigma^2\}.$$

因此$\xi_n-\mu_n\xrightarrow{d}N(0,\sigma^2)$. 记$\liminf\limits_{n\to\infty}\mu_n=a$, $\limsup\limits_{n\to\infty}\mu_n=b$, 那么由$\xi_n\xrightarrow{P}X$可知

$$X-a\overset{d}{=}N(0,\sigma^2)\qquad X-b\overset{d}{=}N(0,\sigma^2)$$

因此$a=b<\infty$. μ_n极限也存在, 记作μ. 从而$X\sim N(\mu,\sigma^2)$为正态随机变量. □

由命题22.4容易得到以下推论, 证明请读者自己完成.

推论22.5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$\{X(t);\ t\in T\}\bigcup\{Y_k;\ k=1,2,\cdots,m\}$仍是高斯系, 而且$\mathrm{Cov}(Y_k,Y_n)=\lim\limits_{l\to\infty}\mathrm{Cov}(X(t_l^k),X(t_l^n))$.

(B)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, 以$B(t)$为被积函数, 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16.4可知$\int_a^b B(u)\mathrm{d}f(u)$ 服从正态分布. 进一步由推论16.5还可得随机过程

$$\Big\{\int_a^t B(u)\mathrm{d}f(u);\ t\geq a\Big\}$$

是高斯过程. 直接计算可知对任意$s,t\geq a$,

$$\begin{aligned}\mathbb{E}\Big(\int_a^t B(u)\mathrm{d}f(u)\Big)&=\int_0^t\mathbb{E}(B(u))\mathrm{d}f(u)=0,\\ \mathrm{Cov}\Big(\int_a^t B(u)\mathrm{d}f(u),\int_a^s B(v)\mathrm{d}f(v)\Big)&=\int_a^t\int_a^s\mathbb{E}(B(u)B(v))\mathrm{d}f(u)\mathrm{d}f(v)\\ &=\int_a^s\int_a^t(s\wedge t)\mathrm{d}f(s)\mathrm{d}f(u).\end{aligned}$$

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积分. 设$f(t)$是具有连续一阶导数的非随机函数, 对任意$0\leq a\leq b$, 定义区间$[a,b]$上随机积分

$$\int_a^b f(s)\mathrm{d}B(s)=\lim_{\delta\to 0}\sum_{\substack{i=1\\ \delta=\max\{t_i-t_{i-1}\}}}^n f(t_{i-1})[B(t_i)-B(t_{i-1})]. \tag{6.4.1}$$

其中$a = t_0 < t_1 < \cdots < t_n = b$为$[a, b]$的任一分割. 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$f(t)$为非随机函数时, 极限中的部分和仍为正态随机变量, 因此$\int_a^b f(s)\mathrm{d}B(s)$也是正态随机变量. 进一步由推论22.5还可知随机过程

$$\Big\{ \int_a^t f(u)\mathrm{d}B(u);\ t \geq a \Big\}$$

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$0 = t_0 < t_1 < \cdots < t_m = s < \cdots < t_n = t$

$$\begin{aligned}
\mathbb{E}\Big(\int_a^t f(u)\mathrm{d}B(u) \Big) &= \lim_{\delta \to 0} \sum_{i=1}^{n} f(t_{i-1})(\mathbb{E}[B(t_i)] - \mathbb{E}[B(t_{i-1})]) \\
&= \int_a^t f(u)\mathrm{d}\mathbb{E}(B(u)) = 0,
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Cov}\Big(\int_a^t f(u)\mathrm{d}B(u), \int_a^s f(u)\mathrm{d}B(u) \Big) \\
&\qquad = \lim_{\delta \to 0} \mathbb{E}\Big(\sum_{i=1}^{n} f(t_{i-1})[B(t_i) - B(t_{i-1})] \sum_{i=1}^{m} f(t_{i-1})[B(t_i) - B(t_{i-1})] \Big) \\
&\qquad = \lim_{\delta \to 0} \mathbb{E}\Big(\sum_{i=1}^{m} f^2(t_{i-1})[B(t_i) - B(t_{i-1})]^2 \Big), \quad \text{由布朗运动独立增量性} \\
&\qquad = \lim_{\delta \to 0} \sum_{i=1}^{m} f^2(t_{i-1})(t_i - t_{i-1}) = \int_a^s f^2(u)\mathrm{d}u.
\end{aligned}$$

例22.2 设B是标准布朗运动，试求出$\int_0^1 u\mathrm{d}B(u)$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19.4以及习题19.2, 仿上面的计算过程可知

$$\begin{aligned}
&\mathbb{E}(\int_0^1 u\mathrm{d}B(u)|B(1) = 1) = \int_0^1 u\mathrm{d}\mathbb{E}(B(u)|B(1) = 1) = \int_0^1 u\mathrm{d}u = \frac{1}{2}, \\
&\mathbb{E}\Big[\Big(\int_0^1 u\mathrm{d}B(u) \Big)^2 \Big| B(1) = 1 \Big] = \lim_{\delta \to 0} \sum_{i=1}^{m} t_{i-1}^2 \mathbb{E}\Big([B(t_i) - B(t_{i-1})]^2 \Big| B(1) = 1 \Big) \\
&\qquad = \lim_{\delta \to 0} \sum_{i=1}^{m} t_{i-1}^2 (t_i - t_{i-1}) = \int_0^1 t^2 \mathrm{d}t = 1/3.
\end{aligned}$$

即$(\int_0^1 u\mathrm{d}B(u)|B(1)=1) \sim N(1/2,1/3)$. □

要注意的是在随机积分表达式$\int f(t)\mathrm{d}B(t)$中, 虽然我们借用了微分符号$\mathrm{d}B(t)$, 但它并不是$B(t)$的真实微分, 因为由定理19.7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).

练习题

22.1 设$B=\{B(t),t\geq 0\}$为标准布朗运动, 试确定$Y=\int_0^1 B(s)ds$的分布.若$B(1)=x$, 试在此条件下再确定Y的分布.

22.2 设B是标准布朗运动，试求出$\int_0^1 u\mathrm{d}B(u)$在$B(1)=x$条件下的分布.

22.3* 证明推论22.5.

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Levy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. 本章介绍最简单的一类鞅模型–离散时间鞅.

§7.1　鞅与停时

(A) 鞅的定义及举例

定义23.1 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) $\mathbb{E}(|X_n|)<\infty$,

(2) X_n是$X_0,Y_0,\cdots,Y_n$的函数.

(3) $\mathbb{E}(X_{n+1}|X_0,Y_0,\cdots,Y_n)=X_n(\geq X_n,\leq X_n)$,

注23.1 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$X_0,Y_0,\cdots,Y_n$条件下, X_{n+1}的条件期望与X_n相同(比X_n大, 比X_n小). 更形象些, 若将时刻n理解为现在, $X_0,Y_0,\cdots,Y_n$ 理解为的现在的信息, 那么(下, 上)鞅就是在掌握现在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23.1 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$n$局后改赌徒拥有的资本.

(1) 若对任意$n\geq 0$, $\mathbb{E}(X_{n+1}|S_0,Y_0,\cdots,Y_n)=0$, 那么

$$\begin{aligned}\mathbb{E}(S_{n+1}|S_0,Y_0,\cdots,Y_n) &= \mathbb{E}(S_n+X_{n+1}|S_0,Y_0,\cdots,Y_n)\\ &= \mathbb{E}(S_n|S_0,Y_0,\cdots,Y_n)+\mathbb{E}(X_{n+1}|S_0,Y_0,\cdots,Y_n)=S_n,\end{aligned}$$

即$S=\{S_n,n\geq 0\}$关于$Y=\{Y_n,n\geq 1\}$是鞅.

(2) 若对任意$n\geq 0$, $\mathbb{E}(X_{n+1}|S_0,Y_0,\cdots,Y_n)>0$, 那么

$$\begin{aligned}\mathbb{E}(S_{n+1}|S_0,Y_0,\cdots,Y_n) &= \mathbb{E}(S_n+X_{n+1}|S_0,Y_0,\cdots,Y_n)\\ &= \mathbb{E}(S_n|S_0,Y_0,\cdots,Y_n)+\mathbb{E}(X_{n+1}|S_0,Y_0,\cdots,Y_n)\geq S_n\end{aligned}$$

即S关于Y是下鞅.

(3) 若对任意$n \geq 0$, $\mathbb{E}(X_{n+1}|S_0, Y_0, \cdots, Y_n) < 0$, 那么

$$\begin{aligned}\mathbb{E}(S_{n+1}|S_0, Y_0, \cdots, Y_n) &= \mathbb{E}(S_n + X_{n+1}|S_0, Y_0, \cdots, Y_n) \\ &= \mathbb{E}(S_n|S_0, Y_0, \cdots, Y_n) + \mathbb{E}(X_{n+1}|S_0, Y_0, \cdots, Y_n) \leq S_n,\end{aligned}$$

即S关于Y是上鞅. □

直观而言, 在赌博模型中S为鞅表示赌局(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"庄家", 下鞅则有利于赌徒. 由赌徒模型还可以看出, 由于赌本S_n 的变化不足以完全反映赌场信息的变化, 在鞅的定义中引进"信息"过程Y 在有些时候是很有必要的.

下面例子表明, 鞅(现象)在随机过程中可以自然地出现.

例23.2 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

$$X_0 = g(Y_0), \quad X_n = g(Y_n) - \sum_{k=0}^{n-1}[\phi(Y_k) - g(Y_k)], \quad n \geq 1.$$

那么由条件期望性质与马氏性, 对任意$n \geq 1$

$$\begin{aligned}&\mathbb{E}(X_{n+1}|Y_0, \cdots, Y_n) = \mathbb{E}(g(Y_{n+1})|Y_n) - \sum_{k=0}^{n}[\phi(Y_k) - g(Y_k)] \\ &= \sum_{j\in S} p_{i,j} g(j)|_{i=Y_n} - \sum_{k=0}^{n}[\phi(Y_k) - g(Y_k)] = \phi(Y_n) - \sum_{k=0}^{n}[\phi(Y_k) - g(Y_k)] \\ &= g(Y_n) - \sum_{k=0}^{n-1}[\phi(Y_k) - g(Y_k)] = X_n.\end{aligned}$$

这表明$X = \{X_n; n \geq 0\}$关于Y为鞅. 特别, 若对任意$i \in S$, $\phi(i) = (\geq, \leq) g(i)$, 那么$X = \{g(Y_n); n \geq 0\}$为鞅(下鞅, 上鞅). □

例23.3 设$X = \{X_n; n \geq 1\}$为独立同分布随机变量.

(1) 若$\mathbb{E}(X_n) = \mu$, 令$S_0 = 0$, $S_n = \sum\limits_{k=1}^{n} X_i - n\mu$, 则$S = \{S_n, n \geq 0\}$关于$X$为鞅.

(2) 若存在$\lambda \in R$使得$\phi(\lambda) = \mathbb{E}(\mathrm{e}^{\lambda X_n}) < \infty$, 令

$$Y_n = [\phi(\lambda)]^{-n} \mathrm{e}^{\lambda \sum\limits_{k=1}^{n} X_i}, \quad n \geq 0, \ (\text{规定} Y_0 = 1)$$

则$\{Y_n; n \geq 0\}$ 为鞅. 特别, 若$\phi(\lambda_0) = 1$, 则$\big\{\exp\{\lambda_0 \sum\limits_{i=1}^{n} X_i\}; n \geq 0\big\}$为鞅.

(3) 设f_0, f_1为两概率密度函数, 令

$$L_n = \prod_{i=1}^{n} \frac{f_1(X_n)}{f_0(X_n)}, \qquad n \geq 0, \text{ (规定}L_0 = 1)$$

在假设检验中以X_n的分布密度函数为f_0为原假设, 此时$\{L_n, n \geq 0\}$为鞅.

解 (1) 对任意$n \geq 0$

$$\begin{aligned}\mathbb{E}(S_{n+1}|S_0, X_1, \cdots, X_n) &= \mathbb{E}(S_n + X_{n+1} - \mu|S_0, X_1, \cdots, X_n)\\ &= S_n + \mathbb{E}(X_{n+1} - \mu|S_0, X_1, \cdots, X_n) = S_n.\end{aligned}$$

所以$\{S_n, n \geq 0\}$为鞅.

(2)对任意$n \geq 0$

$$\begin{aligned}\mathbb{E}(Y_{n+1}|Y_0, X_1, \cdots, X_n) &= \mathbb{E}(Y_n \phi^{-1}(\lambda)\mathrm{e}^{\lambda X_{n+1}}|Y_0, X_1, \cdots, X_n)\\ &= Y_n \phi^{-1}(\lambda)\mathbb{E}(\mathrm{e}^{\lambda X_{n+1}}) = Y_n.\end{aligned}$$

所以$\{Y_n, n \geq 0\}$为鞅.

(3)在X_n的概率分布密度函数为f_0的条件假设下, 对任意$n \geq 0$

$$\begin{aligned}\mathbb{E}(L_{n+1}|L_0, X_1, \cdots, X_n) &= \mathbb{E}\Big(L_n \frac{f_1(X_{n+1})}{f_0(X_{n+1})}\Big|L_0, X_1, \cdots, X_n\Big)\\ &= L_n \mathbb{E}\Big(\frac{f_1(X_{n+1})}{f_0(X_n)}\Big) = L_n \int_{-\infty}^{\infty} f_1(x)dx = L_n.\end{aligned}$$

所以$\{L_n, n \geq 0\}$为鞅. □

例23.4 (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

$$\begin{aligned}&\mathbb{E}(M_{n+1}|M_0, \cdots, M_n) = M_n + \mathbb{E}(M_{n+1} - M_n|M_0, \cdots, M_n)\\ &\quad = M_n + \mathbb{E}(N(t_{n+1}) - N(t_n) - \lambda(t_{n+1} - t_n)|N(t_0), \cdots, N(t_n))\\ &\quad = M_n + \mathbb{E}(N(t_{n+1}) - N(t_n) - \lambda(t_{n+1} - t_n)) = M_n.\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}(M_{n+1}|M_0,\cdots,M_n)\\
&\quad=M_n\mathbb{E}\Big(\exp\big\{\lambda(B(t_{n+1})-B(t_n))-\frac{\lambda\sigma^2(t_{n+1}-t_n)}{2}\big\}\Big|B(t_0),\cdots,B(t_n)\Big)\\
&\quad=M_n\mathbb{E}(\exp\{\lambda(B(t_{n+1})-B(t_n))-\frac{\lambda\sigma^2(t_{n+1}-t_n)}{2}\})=M_n.
\end{aligned}$$

因此$\big\{\exp\big\{\lambda B(t_n)-\frac{\lambda\sigma^2 t_n}{2}\big\};n\geq 0\big\}$ 也是鞅. □

例23.4表明对连续时间随机过程我们也能构造鞅. 事实上我们此时能得到更严格的所谓连续时间鞅. 对它的详细讨论由于需要更多的专门知识, 本书从略.

(B) 鞅的简单性质

由定义23.1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X既是上鞅又是下鞅.

性质23.1 若随机过程$X=\{X_n;n\geq 0\}$, $Y=\{Y_n;n\geq 0\}$关于$\{Z_n;n\geq 0\}$都是上鞅(下鞅), 则

(1) $\mathbb{E}(X_n),\mathbb{E}(Y_n)$ 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(3), 其他显然成立. 对任意$n\geq 0$

$$\mathbb{E}(X_{n+1}\wedge Y_{n+1}|Z_0,\cdots,Z_n)\leq\mathbb{E}(Y_{n+1}|Z_0,\cdots,Z_n)\leq Y_n,$$

而且

$$\mathbb{E}(X_{n+1}\wedge Y_{n+1}|Z_0,\cdots,Z_n)\leq\mathbb{E}(X_{n+1}|Z_0,\cdots,Z_n)\leq X_n.$$

因此

$$\mathbb{E}(X_{n+1}\wedge Y_{n+1}|Z_0,\cdots,Z_n)\leq\min\{X_n,Y_n\}=X_n\wedge Y_n.$$

所以$\{X_n\wedge Y_n,n\geq 0\}$ 仍是上鞅. □

引理23.2(Jessen不等式) 设$\psi(x)$ 是$\mathbb{R}$上凸函数, X为随机变量,

$$\mathbb{E}(|X|)<\infty \text{ 且 } \mathbb{E}(|\psi(X)|)<\infty,$$

那么$\psi(\mathbb{E}(X))\leq\mathbb{E}(\psi(X))$. 进一步, 任取随机变量序列$\{Y_n,n\geq 0\}$, 对任意$n\geq 0$

$$\psi(\mathbb{E}(X|Y_0,\cdots,Y_n))\leq\mathbb{E}(\psi(X)|Y_0,\cdots,Y_n).$$

证明 只证后一个不等式. 由于ψ是凸函数, 由数学分析知识可知ψ的左右导数都存在且对任意$x,y\in\mathbb{R}$

$$\psi(y)-\psi(x)\geq\psi'_+(x)(y-x),$$

其中$\psi'_+(y)$表示y的右导数. 因此

$$\psi(x)-\psi(\mathbb{E}(X|Y_0,\cdots,Y_n))\geq\psi'_+(\mathbb{E}(X|Y_0,\cdots,Y_n))(X-\mathbb{E}(X|Y_0,\cdots,Y_n)).$$

两边关于$Y_0,\cdots,Y_n$求条件期望得

$$\begin{aligned}&\mathbb{E}(\psi(x)-\psi(\mathbb{E}(X|Y_0,\cdots,Y_n))|Y_0,\cdots,Y_n)\\&\quad\geq\psi'_+(\mathbb{E}(X|Y_0,\cdots,Y_n))\mathbb{E}(X-\mathbb{E}(X|Y_0,\cdots,Y_n)|Y_0,\cdots,Y_n)\\&\quad\geq\psi'_+(\mathbb{E}(X|Y_0,\cdots,Y_n))[\mathbb{E}(X|Y_0,\cdots,Y_n)-\mathbb{E}(X|Y_0,\cdots,Y_n)]=0.\end{aligned}$$

因此$\mathbb{E}(\psi(x)|Y_0,\cdots,Y_n)\geq\psi(\mathbb{E}(X|Y_0,\cdots,Y_n))$. □

注意到对任意$r>1$,$x^r,x>0$为凸函数, 由Jessen不等式可得hölder不等式：

$$\mathbb{E}(|X|^r)\geq[\mathbb{E}(|X|)]^r\ \text{即}\ \mathbb{E}(|X|)\leq[\mathbb{E}(|X|^r)]^{1/r}.$$

性质23.3 设$\{X_n\}$关于$\{Y_n\}$为(下)鞅, $\psi(x)$为R上(非降)凸函数. 若$\mathbb{E}(|\psi(X_n)|)<\infty$对所有$n\geq0$都成立, 那么$\{\psi(X_n),n\geq0\}$是下鞅.

证明 由Jessen不等式, 对任意$n\geq0$

$$\mathbb{E}(\psi(X_{n+1})|Y_0,\cdots,Y_n)\geq\psi(\mathbb{E}(X_{n+1}|Y_0,\cdots,Y_n)=\psi(X_n)(\geq\psi(X_n)).$$

因此$\{\psi(X_n),n\geq0\}$是下鞅. □

由性质23.3我们容易构造出下鞅. 若$\{X_n;n\geq0\}$是鞅, 只要下面的随机变量序列有意义且可积, 那么

$$\{|X_n|;n\geq0\},\quad\{\mathrm{e}^{X_n};n\geq0\},\quad\{-\ln(X_n),n\geq0\}$$

都是下鞅, 而当$r\geq1$时$\{|X_n|^r;n\geq0\}$也是下鞅.

性质23.4 若$Y=\{Y_n;n\geq0\}$关于$X=\{X_n;n\geq0\}$是下鞅, 那么存在一个关于X的鞅$\{M_n;n\geq0\}$使得

$$Y_n=M_n+N_n,$$

其中$\{N_n;n\geq0\}$是一个单调不降的随机序列, $N_0=0$且N_n由$Y_0,X_0,\cdots,X_{n-1}$完全给定.

证明 对任意$n\geq1$, 令

$$N_n=\sum_{k=0}^{n-1}[\mathbb{E}(Y_{k+1}|Y_0,X_0,\cdots,X_k)-Y_k].$$

由条件数学期望与下鞅定义可知N_n是$Y_0,X_0,\cdots,X_{n-1}$的函数, 而且$\{N_n;n\geq1\}$是一个单调不降的随机序列. 再令$M_n=Y_n-N_n$, 那么

$$\begin{aligned}&\mathbb{E}(M_{n+1}|Y_0,X_0,\cdots,X_n)\\&\quad=\mathbb{E}(Y_{n+1}|Y_0,X_0,\cdots,X_n)-\sum_{k=0}^{n}[\mathbb{E}(Y_{k+1}|Y_0,X_0,\cdots,X_k)-Y_k]\\&\quad=Y_n-\sum_{k=0}^{n-1}[\mathbb{E}(Y_{k+1}|Y_0,X_0,\cdots,X_k)-Y_k]=M_n.\end{aligned}$$

由此可知M_n为鞅进而结论成立. □

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$(Y_0, X_0, \cdots, X_{n-1})$唯一确定的随机过程为可料的(Predictable)或可容许的(Admissible).

例23.5 一个赌徒参加一种掷色子的赌博, 以X_i表示第i局该赌徒的胜负情况, 若赌徒胜则$X_i = 1$, 否则$X_i = -1$. 假设每一局胜负结果相互独立, 而且$\mathbb{P}(X_i = 1) \le 1/2$. 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$k \le n-1$,

$$\{H_n = 2^k\} = \{X_{n-k-1} = 1, X_{n-k} = -1, \cdots, X_{n-1} = -1\}$$

因此H_n为$X_0, X_1, \cdots, X_{n-1}$的函数, 即存在函数$\phi_n \le 2^n$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

$$\begin{aligned}\mathbb{E}(M_n|M_0, X_0, \cdots, X_{n-1}) &= \sum_{k=1}^{n} \mathbb{E}(H_k X_k|M_0, X_0, \cdots, X_{n-1}) \\ &= \sum_{k=1}^{n-1} H_k X_k + H_n \mathbb{E}(X_n|M_0, X_0, \cdots, X_{n-1}) \le \sum_{k=1}^{n-1} H_k X_k = M_{n-1}\end{aligned}$$

即M_n关于X_n为上鞅, 从而$\mathbb{E}(M_n)$为非增序列. 这表明倍增策略不能实现稳赢. □

(C) 停时及其简单性质

定义23.2 设T为Ω上非负整数值的随机变量(可以取$+\infty$). 称T 为关于$Y = \{Y_n, n \ge 0\}$的停时(Stopping time)或马氏时间(Markov time), 若对任意非负整数n, 事件$\{T = n\}$可以通过$Y_0, \cdots, Y_n$表示, 等价地事件$\{T = n\}$的示性函数是随机量$Y_0, \cdots, Y_n$的函数, 也即存在函数ϕ_n使得$\mathbf{1}_{\{T=n\}} = \phi_n(Y_0, \cdots, Y_n)$. 当无需强调"信息"过程$Y$时, 我们也简称$T$为停时.

直观地看, 停时定义的本质要求就是随机事件$\{T = n\}$是否发生由$Y_0, \cdots, Y_n$的信息就可判断,不受n之后Y值的影响.

例23.6 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23.7 设$X = \{X_n; n \ge 0\}$是状态空间为S的时齐马氏链, 对任意$i \in S$, $\tau_i = \inf\{n \ge 1; X_n = i\}$为停时. 事实上,若$n = 0$, $\{\tau_i = 0\} = \emptyset$; 若$n > 0$,

$$\{\tau_i = n\} = \{X_k \ne i, k = 1, \cdots, n-1, X_n = i\}.$$

例23.8 设$N(t)$是以$W_i, i \geq 1$为等待时间序列的更新过程, T_n为第n 次更新发生的时间, 那么对任意$t \geq 0$, $N(t)$关于$T = \{T_n; n \geq 0\}$ 不是停时,但$N(t)+1$关于T 是停时. 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 及之前的信息.

对于停时我们有如下的性质

性质23.6 若S, T是关于$X = \{X_n; n \geq 0\}$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

$$\mathbf{1}_{\{S=k\}} = \phi_k(X_0, \cdots, X_k), \quad \mathbf{1}_{\{T=k\}} = \psi_k(X_0, X_1, \cdots X_k).$$

因此对任意$n \geq 0$

$$\begin{aligned} \mathbf{1}_{\{S+T=n\}} &= \sum_{k=0}^{n} \mathbf{1}_{\{S=k\}} \mathbf{1}_{\{T=n-k\}} \\ &= \sum_{k=0}^{n} \phi_k(X_0, \cdots, X_k) \psi_{n-k}(X_0, X_1, \cdots X_{n-k}). \end{aligned}$$

由此可知$S+T$为停时. 同理可证$S \wedge T$, $S \vee T$都是停时, 细节请读者自己补充. □

定理23.7(Wald等式) 设$X = \{X_n; n \geq 1\}$为独立同分布随机变量序列, $\mathbb{E}(|X_1|) < \infty$, 设$T$关于$X$为停时且$\mathbb{E}(T) < \infty$. 那么

$$\mathbb{E}(\sum_{i=1}^{T} X_i) = \mathbb{E}(X_1)\mathbb{E}(T). \tag{7.1.1}$$

证明 对任意$k \geq 1$, 由于T为停时, 可记$\mathbf{1}_{\{T=k\}} = \phi_k(X_1, \cdots, X_k)$.

$$\begin{aligned} \mathbb{E}\Big(\sum_{k=1}^{T} X_k\Big) &= \sum_{k=1}^{\infty} \mathbb{E}\Big(\sum_{i=1}^{k} X_i \mathbf{1}_{\{T=k\}}\Big) = \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} \mathbb{E}\Big(X_i \mathbf{1}_{\{T=k\}}\Big) \\ &= \sum_{i=1}^{\infty} \mathbb{E}\Big(X_i - X_i \sum_{k=1}^{i-1} \mathbf{1}_{\{T=k\}}\Big) = \sum_{i=1}^{\infty} \mathbb{E}\Big(X_i - X_i \sum_{k=1}^{i-1} \phi_k(X_1, \cdots, X_k)\Big) \end{aligned}$$

利用条件数学期望性质并注意到$\{X_n, n \geq 1\}$是独立同分布的,

$$\begin{aligned} \mathbb{E}\Big(\sum_{k=1}^{T} X_k\Big) &= \sum_{i=1}^{\infty} \mathbb{E}\Big(\mathbb{E}\big(X_i - X_i \sum_{k=1}^{i-1} \phi_k(X_1, \cdots, X_k) \big| X_1, \cdots, X_{i-1}\big)\Big) \\ &= \sum_{i=1}^{\infty} \mathbb{E}(X_i | X_1, \cdots, X_{i-1}) \Big(1 - \mathbb{E}(\sum_{k=1}^{i-1} \phi_k(X_1, \cdots, X_k))\Big) \\ &= \mathbb{E}(X_1) \sum_{i=1}^{\infty} (1 - \mathbb{E}(\sum_{k=1}^{i-1} \mathbf{1}_{\{T=k\}})) = \mathbb{E}(X_1) \sum_{i=1}^{\infty} \mathbb{P}(T \geq i). \end{aligned}$$

注意到$\sum\limits_{i=1}^{\infty}\mathbb{P}(T\geq i)=\mathbb{E}(T)$, (7.1.1)成立. □

练习题

23.1 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.

23.2 设$X_1,X_2,\cdots$为独立同分布随机变量满足

$$\mathbb{P}(X_1=1)=p,\mathbb{P}(X=-1)=q=1-p.$$

对任意$n\geq 0$ 令$Y_0=1$以及

$$Y_n=\left(\frac{q}{p}\right)^{S_n}, \text{ 其中 } S_n=\sum_{k=1}^{n}X_k,\ \ n\geq 1.$$

证明$\{Y_n,n\geq 0\}$关于X为鞅.

23.3 设$X_1,X_2,\cdots$为独立同分布随机变量, $\mathbb{E}(X_n)=0$, $\mathbb{E}(X_n^2)=a^2$. 对任意$n\geq 0$, 令$Y_0=0$,

$$Y_n=(\sum_{k=1}^{n}X_k)^2-na^2\ \ n\geq 1.$$

证明$\{Y_n,n\geq 0\}$关于X为鞅.

23.4 设$X=\{X_n,n\geq 0\}$为任一随机过程, $\tau=\min\{n\geq 0;X_n=a\text{或}b\}$, $\eta=\inf\{n\geq 0;X_n\leq a\}$. 证明$\tau,\eta$都是停时.

23.5 利用Jessen不等式证明Lyapunov不等式: 对任意随机变量X, 若$0<s<t$, 则

$$(\mathbb{E}(|X|^s))^{1/s}\leq(\mathbb{E}(|X|^t))^{1/t}.$$

23.6* 请补充证明若$\psi(x)$是区间(a,b)内凸函数,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$\psi'_+(y)$表示y的右导数.

§7.2 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(上, 下)鞅, 所有停时也以该信息过程为参考. 对任意一个停时T, 定义随机变量X_T 如下:

$$\text{对任意}\omega\in\Omega,\ \text{若}T(\omega)=n,\ \text{则}X_T(\omega)=X_n(\omega).$$

以下理论结果只涉及鞅或上鞅的情形, 下鞅的结果可通过上鞅得到, 故不重复.

(A) 有界停时定理

引理24.1 设$\{X_n;n\geq 0\}$是鞅(上鞅), T是一个停时, 那么对任意$0\leq k\leq n$

$$\mathbb{E}(X_n\mathbf{1}_{\{T=k\}})=(\leq)\mathbb{E}(X_k\mathbf{1}_{\{T=k\}}).$$

证明 由条件数学期望性质

$$\begin{aligned}\mathbb{E}(X_n\mathbf{1}_{\{T=k\}})&=\mathbb{E}(\mathbb{E}(X_n\mathbf{1}_{\{T=k\}}|X_0,Y_0,\cdots,Y_k))\\&=\mathbb{E}(\mathbf{1}_{\{T=k\}}\mathbb{E}(X_n|X_0,Y_0,\cdots,Y_k))=(\leq)\mathbb{E}(X_k\mathbf{1}_{\{T=k\}}).\end{aligned}$$

引理得证. □

引理24.2 设$\{X_n;n\geq 0\}$是鞅(上鞅), T是一个有界停时, $0\leq n\leq T\leq m$, 那么

$$\mathbb{E}(X_T|X_0,Y_0,\cdots,Y_n)=(\leq)X_n.$$

证明 注意到X_n是$X_0,Y_1,\cdots,Y_n$的函数, 由条件数学期望的平滑性,

$$\begin{aligned}\mathbb{E}(X_T|X_0,Y_0,\cdots,Y_n)&=\mathbb{E}(\mathbb{E}(X_T|X_0,Y_0,\cdots,Y_{m-1})|X_0,Y_0,\cdots,Y_n)\\&=\mathbb{E}(\sum_{k=n}^{m}\mathbb{E}(X_k\mathbf{1}_{\{T=k\}}|X_0,Y_0,\cdots,Y_{m-1})|X_0,Y_0,\cdots,Y_n)\\&=\mathbb{E}(\sum_{k=n}^{m-1}X_k\mathbf{1}_{\{T=k\}}|X_0,Y_0,\cdots,Y_n)\\&\quad+\mathbb{E}(\mathbb{E}(X_m\mathbf{1}_{\{T=m\}}|X_0,Y_0,\cdots,Y_{m-1})|X_0,Y_0,\cdots,Y_n)\end{aligned}$$

注意到

$$\mathbf{1}_{\{T=m\}}=1-\mathbf{1}_{\{T\leq m-1\}}$$

是$Y_0,\cdots,Y_{m-1}$的函数, 由条件分布性质以及(上)鞅性质可知

$$\begin{aligned}&\mathbb{E}(\mathbb{E}(X_m\mathbf{1}_{\{T=m\}}|X_0,Y_0,\cdots,Y_{m-1})|X_0,Y_0,\cdots,Y_n)\\&\quad=\mathbb{E}(\mathbf{1}_{\{T=m\}}\mathbb{E}(X_m|X_0,Y_0,\cdots,Y_{m-1})|X_0,Y_0,\cdots,Y_n)\\&\quad=(\leq)\mathbb{E}(X_{m-1}\mathbf{1}_{\{T=m\}}|X_0,Y_0,\cdots,Y_n).\end{aligned}$$

因此

$$\begin{aligned}\mathbb{E}(X_T|X_0,Y_0,\cdots,Y_n) &= (\le)\mathbb{E}(\sum_{k=n}^{m-1} X_k\mathbf{1}_{\{T=k\}}|X_0,Y_0,\cdots,Y_n)\\ &\qquad\qquad +\mathbb{E}(X_{m-1}\mathbf{1}_{\{T=m\}}|X_0,Y_0,\cdots,Y_n)\\ &= \mathbb{E}(X_{T\wedge(m-1)}|X_0,Y_0,\cdots,Y_n).\end{aligned}$$

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

$$\mathbb{E}(X_T|X_0,Y_0,\cdots,Y_n) = (\le)\mathbb{E}(X_{T\wedge n}|X_0,Y_0,\cdots,Y_n) = X_n.$$

引理24.2得证. □

定理24.3(有界停时定理) 设$\{X_n; n\ge 0\}$是(上)鞅，若S,T为两有界停时且$S\le T$ a.s. 那么

$$\mathbb{E}(X_T) = (\le)\mathbb{E}(X_S).$$

证明 不妨设$S,T\le m$, 由引理24.2,

$$\begin{aligned}\mathbb{E}(X_T) &= \sum_{n=0}^{m}\mathbb{E}(X_T\mathbf{1}_{\{S=n\}}) = \sum_{n=0}^{m}\mathbb{E}(X_{T\vee n}\mathbf{1}_{\{S=n\}})\\ &= \sum_{n=0}^{m}\mathbb{E}(\mathbb{E}(X_{T\vee n}\mathbf{1}_{\{S=n\}}|Y_0,\cdots,Y_n))\\ &= \sum_{n=0}^{m}\mathbb{E}(\mathbf{1}_{\{S=n\}}\mathbb{E}(X_{T\vee n}|Y_0,\cdots,Y_n))\\ &= (\le)\sum_{n=0}^{m}\mathbb{E}(\mathbf{1}_{\{S=n\}}X_n) = \mathbb{E}(X_S).\end{aligned}$$

定理得证. □

(B) 一般停时定理

引理24.4 设随机变量W满足$\mathbb{E}(|W|)<\infty$, 若T为有限停时, 即$\mathbb{P}(T<\infty)=1$. 那么

$$\lim_{n\to\infty}\mathbb{E}(W\mathbf{1}_{\{T>n\}}) = 0 \ \text{ 且 } \lim_{n\to\infty}\mathbb{E}(W\mathbf{1}_{\{T\le n\}}) = \mathbb{E}(W).$$

证明 由正项级数的收敛性,当$n\to\infty$时

$$\mathbb{E}(|W|) \ge \mathbb{E}(|W|\mathbf{1}_{\{T\le n\}}) = \sum_{k=0}^{n}\mathbb{E}(|W|\mathbf{1}_{\{T=k\}}) \to \sum_{k=0}^{\infty}\mathbb{E}(|W|\mathbf{1}_{\{T=k\}}).$$

由于$\mathbb{P}(T<\infty)=1$, 由引理3.13可知

$$\mathbb{E}(|W|\mathbf{1}_{\{T\ge n+1\}}) = \sum_{k=n+1}^{\infty}\mathbb{E}(|W|\mathbf{1}_{\{T=k\}}) \to 0.$$

从而

$$0 \le |\mathbb{E}(W) - \mathbb{E}(W\mathbf{1}_{\{T\le n\}})| \le |\mathbb{E}(W\mathbf{1}_{\{T\ge n+1\}})| \le \mathbb{E}(|W|\mathbf{1}_{\{T\ge n+1\}}) \to 0.$$

由此可得引理结论. □

定理24.5 设$\{X_n; n\ge 0\}$为鞅，T是一个有限停时,若

$$\mathbb{E}(\sup_{n\ge 0}|X_{T\wedge n}|) < \infty,$$

那么$\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

证明 令$W = \sup\limits_{n\ge 0}|X_{T\wedge n}|$. 由于$\mathbb{P}(T<\infty)=1$,

$$X_T = \sum_{k=0}^{\infty} X_k \mathbf{1}_{\{T=k\}} = \sum_{k=0}^{\infty} X_{T\wedge k}\mathbf{1}_{\{T=k\}}.$$

从而$|X_T| \le W$. 因此$\mathbb{E}(|X_T|) \le \mathbb{E}(W) < \infty$. 此时由引理18.4

$$|\mathbb{E}(X_{T\wedge n}) - \mathbb{E}(X_T)| \le \mathbb{E}(|X_{T\wedge n} - X_T|\mathbf{1}_{\{T\ge n\}}) \le 2\mathbb{E}(W\mathbf{1}_{\{T\ge n\}}) \to 0.$$

所以$\mathbb{E}(X_T) = \lim\limits_{n\to\infty}\mathbb{E}(X_{T\wedge n}) = \mathbb{E}(X_0)$, 其中最后一个等号由有界停时定理可得. □

推论24.6 设$\{X_n; n\ge 0\}$为鞅，T是一个有限停时, 若$\mathbb{E}(T)<\infty$ 且存在$K<\infty$使得

$$\mathbb{E}(|X_{n+1} - X_n||X_0, Y_0, \cdots, Y_n) \le K,$$

那么$\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

证明 令$Z_0 = |X_0|$, $Z_n = |X_n - X_{n-1}|$, $W = Z_0 + \cdots + Z_T$. 则$W \ge |X_T|$且

$$\mathbb{E}(W) = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\mathbb{E}(Z_k\mathbf{1}_{\{T=n\}}) = \sum_{k=0}^{\infty}\mathbb{E}(Z_k\mathbf{1}_{\{T\ge k\}}).$$

注意到$\mathbf{1}_{\{T\ge k\}} = 1 - \mathbf{1}_{\{T\le k-1\}}$为$Y_0,\cdots,Y_{k-1}$的函数, 由条件假设

$$\begin{aligned}\mathbb{E}(Z_k\mathbf{1}_{\{T\ge k\}}) &= \mathbb{E}(\mathbf{1}_{\{T\ge k\}}\mathbb{E}(Z_k|X_0,Y_0,\cdots,Y_{k-1})) \\ &= \mathbb{E}\big(\mathbf{1}_{\{T\ge k\}}\mathbb{E}\big(|X_k - X_{k-1}|\big|X_0,Y_0,\cdots,Y_{k-1}\big)\big) \le K\mathbb{P}(T\ge k)\end{aligned}$$

从而

$$\mathbb{E}(W) \le \sum_{k=0}^{\infty} K\mathbb{P}(T\ge k) \le K(1+\mathbb{E}(T)) < \infty.$$

因为$|X_{T\wedge n}| \le W$对一切n成立, 由定理23.5 可知结论成立. □

定理24.7 设$\{X_n; n\ge 0\}$为鞅，T是一个有限停时, 若$\mathbb{E}(|X_T|)<\infty$且

$$\lim_{n\to\infty}\mathbb{E}(X_n\mathbf{1}_{\{T\ge n\}}) = 0,$$

那么$\mathbb{E}(X_T)=\mathbb{E}(X_0)$.

证明 注意到

$$\begin{aligned}\mathbb{E}(X_T) &= \mathbb{E}(X_T\mathbf{1}_{\{T<n\}})+\mathbb{E}(X_T\mathbf{1}_{\{T\ge n\}})=\mathbb{E}(X_{T\wedge n}\mathbf{1}_{\{T<n\}})+\mathbb{E}(X_T\mathbf{1}_{\{T\ge n\}})\\ &= \mathbb{E}(X_{T\wedge n})-\mathbb{E}(X_{T\wedge n}\mathbf{1}_{\{T\ge n\}})+\mathbb{E}(X_T\mathbf{1}_{\{T\ge n\}})\end{aligned}$$

由条件$\lim\limits_{n\to\infty}\mathbb{E}(X_{T\wedge n}\mathbf{1}_{\{T\ge n\}})=0$及引理18.4可知$\lim\limits_{n\to\infty}\mathbb{E}(X_T\mathbf{1}_{\{T\ge n\}})=0$. 因此

$$\mathbb{E}(X_T)=\lim_{n\to\infty}\mathbb{E}(X_{T\wedge n})=\mathbb{E}(X_0),$$

其中最后等号由有界停时定理可得. □

定理24.8 设$\{X_n;n\ge 0\}$为上鞅, T为有限停时, 若存在随机变量W使得$\mathbb{E}(|W|)<\infty$且对任意$n\ge 0$, $X_{T\wedge n}\ge -W$ 成立, 那么$\mathbb{E}(X_0)\ge\mathbb{E}(X_T)$.

证明 任意取定一个正整数N,令$X_n^N=X_n\wedge N$, 那么$\{X_n^N,n\ge 0\}$仍是上鞅, 从而由有界停时定理

$$\mathbb{E}(X_0^N)\ge\mathbb{E}(X_{T\wedge n}^N).$$

由条件假设可知$|X_{T\wedge n}^N|\le N\vee|W|<|W|+N$. 因为$T$为有限停时,

$$X_T^N=\lim_{n\to\infty}X_{T\wedge n}^N,\quad a.s.$$

从而$|X_T^N|\le N+|W|$. 当$n\to\infty$时,由引理18.4,

$$\mathbb{E}|X_{T\wedge n}^N-X_T^N|\le\mathbb{E}(|X_{T\wedge n}^N-X_T^N|\mathbf{1}_{\{T>n\}})\le 2\mathbb{E}((N+W)\mathbf{1}_{\{T>n\}})\to 0.$$

因此$\lim\limits_{n\to\infty}\mathbb{E}(X_{T\wedge n}^N)=\mathbb{E}(X_T^N)$.由此可得

$$\mathbb{E}(X_0)\ge\mathbb{E}(X_0^N)\ge\mathbb{E}(X_T^N).$$

因为$\mathbb{E}(X_T^N)$随着N的增加而单调增加,

$$\mathbb{E}(X_0)\ge\lim_{N\to\infty}\mathbb{E}(X_T^N).$$

另一方面, 由$X_T=\lim\limits_{n\to\infty}X_{T\wedge n}$及条件假设可知$X_T\ge -W$. 所以

$$\mathbb{E}(X_T\mathbf{1}_{\{X_T\le 0\}})\ge\mathbb{E}(-W)>-\infty.$$

从而

$$\begin{aligned}\mathbb{E}(X_T) &= \mathbb{E}(X_T\mathbf{1}_{\{X_T\le 0\}})+\mathbb{E}(X_T\mathbf{1}_{\{X_T>0\}})\\ &= \mathbb{E}(X_T\mathbf{1}_{\{X_T\le 0\}})+\sum_{k=1}^{\infty}\mathbb{E}(X_T\mathbf{1}_{\{k-1<X_T\le k\}})\\ &\le \mathbb{E}(X_T\mathbf{1}_{\{X_T\le 0\}})+\lim_{N\to\infty}\Big(\sum_{k=1}^{N}\mathbb{E}(X_T\mathbf{1}_{\{k-1<X_T\le k\}})+N\mathbb{E}(\mathbf{1}_{\{X_T>N\}})\Big)\\ &= \lim_{N\to\infty}\mathbb{E}(X_T^N).\end{aligned}$$

综上所述可得$\mathbb{E}(X_0) \geq \mathbb{E}(X_T)$. □

推论24.9 若$\{X_n\}$为非负上鞅, T为有限停时, 那么$\mathbb{E}(X_0) \geq \mathbb{E}(X_T)$.

(C) 停时定理应用举例

下面我们看两个应用停时定理的例子.

例24.1 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, $p < q$. 定义

$$T = \inf\{n \geq 0, S_n = 0 \text{ 或 } S_n = b\}.$$

求$\mathbb{E}(T)$以及$\mathbb{E}(S_T)$.

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

$$\mathbb{P}(X_i = 1) = p, \quad \mathbb{P}(X_i = -1) = q.$$

由定义可知T关于$\{S_n\}$为停时. 对任意$k > b$(不妨设$k = mb + r$其中$m \geq 1$, $r = 0, 1, \cdots, b-1$), 当事件$\{T = k\}$发生时, 在k 之前任意相邻的b 步中游动方向X_i 不能完全相同(否则就提前到访了0 或b). 将k步按每b步组合成

$$\{1, 2, \cdots, b\}, \{b+1, \cdots, 2b\}, \cdots \{(m-1)b+1, \cdots, mb\}, \{mb+1, \cdots, mb+r\}.$$

由X_i的独立同分布性可知$\mathbb{P}(T = k) \leq (1 - p^b - q^b)^m$, 从而

$$\begin{aligned}\mathbb{E}(T) &= \sum_{m=0}^{\infty} \sum_{r=0}^{b-1} (mb + r)\mathbb{P}(T = mb + r) \\ &\leq \sum_{m=0}^{\infty} (1 - p^b)^m (m+1) b^2 = \frac{b^2}{p^{2b}} < \infty.\end{aligned}$$

令$Y_n = S_n - n(p - q)$, 显然$\{Y_n; n \geq 0\}$关于$\{S_n\}$为鞅. 由

$$\mathbb{E}(|Y_{n+1} - Y_n||S_0, \cdots, S_n) = \mathbb{E}(|X_{n+1} - (p - q)||S_0, \cdots, S_n) \leq 1 + (p - q)$$

以及推论24.6可得

$$\mathbb{E}(Y_T) = \mathbb{E}(Y_0),$$

即

$$\mathbb{E}(S_T) = a + (p - q)\mathbb{E}(T). \tag{7.2.1}$$

另一方面, 令$M_n = \left(\frac{q}{p}\right)^{S_n}$. 则$\{M_n, n \geq 0\}$也是鞅,而且对任意$n \geq 0$, $M_{T \wedge n} \leq 1$, 由定理24.5,

$$\mathbb{E}(M_T) = \mathbb{E}(M_0) = \left(\frac{q}{p}\right)^a.$$

即

$$\mathbb{P}(S_T = 0) + \left(\frac{q}{p}\right)^b \mathbb{P}(S_T = b) = \left(\frac{q}{p}\right)^a. \tag{7.2.2}$$

将$\mathbb{P}(S_T=0)=1-\mathbb{P}(S_T=b)$代入(7.2.2)解得

$$\mathbb{P}(S_T=b)=\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}.$$

从而

$$\mathbb{E}(S_T)=b\mathbb{P}(S_T=b)=b\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}.$$

将其代入(7.2.1)得

$$\mathbb{E}(T)=\frac{b}{p-q}\frac{1-\left(\frac{q}{p}\right)^a}{1-\left(\frac{q}{p}\right)^b}-\frac{a}{p-q}.$$

例24.2 设X_i, $i\geq 1$是独立同分布的随机变量服从正态分布$N(\mu,\sigma^2)$, $S_0=x>0$, $S_n=x+\sum\limits_{k=1}^{n}X_k$, 令$A=\{$存在$n>0$使得$S_n\leq 0\}$, 估计$\mathbb{P}(A)$的上界.

解 令$\tau=\inf\{n>0;S_n\leq 0\}$, 那么$\tau$为停时且$A=\{\tau<\infty\}$. 由$X_i\sim N(\mu,\sigma^2)$可知

$$\mathbb{E}(\mathrm{e}^{-\lambda X_i})=\mathrm{e}^{-\lambda\mu+\lambda^2\sigma^2/2}.$$

取$\lambda=2\mu/\sigma^2\triangleq\theta$, 由例23.3(2)可知$W_n=\exp\{-\theta S_n\}$为鞅. 由有界停时定理, 对任意$n\geq 1$

$$\mathrm{e}^{-\theta x}=\mathbb{E}(W_0)=\mathbb{E}(W_{\tau\wedge n})=\mathbb{E}(W_\tau\mathbf{1}_{\{\tau\leq n\}})+\mathbb{E}(W_\tau\mathbf{1}_{\{\tau>n\}})\geq\mathbb{P}(\tau\leq n).$$

令$n\to\infty$得$\mathbb{P}(A)\leq\mathrm{e}^{-\theta x}=\mathrm{e}^{-2\mu x/\sigma^2}$. □

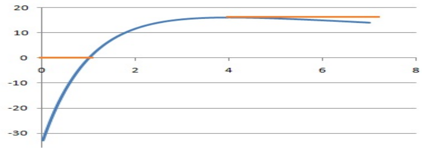
例24.3 以Z_n表示第n年开始时水库蓄水量, Y_{n+1}表示第n年的净入库流量(Y_{n+1}为负表示净出库流量), b表示水库的最大库容, 显然$Z_{n+1}=(Z_n+Y_{n+1})^+\wedge b$. 以$a$表示水库最低警戒库容, 令$T=\min\{n,Z_n\leq a\}$, 试求$\mathbb{E}(T)$的下界估计, 其中我们假定$Z_0=z\in(a,b)$以及

$$\mathbb{E}(Y_{n+1}|Z_0,Y_1,\cdots,Y_n)\leq m,\ \ \mathbb{E}(\exp\{-\lambda Y_{n+1}\}|Z_0,Y_1,\cdots,Y_n)\leq 1.$$

解 对任意$z\geq 0$, 令

$$g(z)=\frac{1}{m}\left(\mathrm{e}^{\lambda(b-a)}\frac{1-\mathrm{e}^{-\lambda(z-a)}}{\lambda}-(z-a)\right)$$

以及$f(z)=\begin{cases}g(z)\vee 0, & 0\leq z\leq b,\\ g(b); & z>b.\end{cases}$ 易知当$0\leq z<\infty$时$f(z)\geq g(z)$. (参见下图.)



而且对任意$n \geq 0$, $f(Z_{n+1}) = f(Z_n + Y_{n+1})$. 进一步由假设条件, 对任意$n \geq 0$, 当$a \leq Z_n \leq b$时,

$$\begin{aligned}
&\mathbb{E}(f(Z_{n+1})|Z_0, Y_1, \cdots, Y_n) = \mathbb{E}(f(Z_n + Y_{n+1})|Z_0, Y_1, \cdots, Y_n) \\
&\quad \geq \mathbb{E}(g(Z_n + Y_{n+1})|Z_0, Y_1, \cdots, Y_n) \\
&\quad = \frac{1}{m}\mathbb{E}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1-\mathrm{e}^{-\lambda(Z_n+Y_{n+1}-a)}}{\lambda} - (Z_n - a + Y_{n+1})\Big| Z_0, Y_1, \cdots, Y_n\Big) \\
&\quad \geq \frac{1}{m}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1-\mathrm{e}^{-\lambda(Z_n-a)}}{\lambda} - (Z_n - a)\Big) - 1 \\
&\quad \geq f(Z_n) - 1;
\end{aligned}$$

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

$$X_n = f(Z_n) + n$$

为下鞅. 由下鞅有界停时定理,

$$\mathbb{E}(f(Z_0)) = \mathbb{E}(X_0) \leq \mathbb{E}(X_{T\wedge n}) = \mathbb{E}(f(Z_{T\wedge n})) + T \wedge n.$$

即$f(z) \leq \mathbb{E}(f(Z_{T\wedge n})) + T \wedge n$. 由于$f$为有界连续函数, 存在常数$M > 0$使得

$$|\mathbb{E}(f(Z_{T\wedge n})) - \mathbb{E}(f(Z_T))| \leq \mathbb{E}(|f(Z_{T\wedge n}) - f(Z_T)|) \leq M\mathbb{P}(T > n) \to 0.$$

因此

$$\lim_{n\to\infty} \mathbb{E}(f(Z_{T\wedge n})) = \mathbb{E}(f(Z_T)) = 0.$$

又由

$$\mathbb{E}(T \wedge n) = \lim_{n\to\infty} \sum_{k=1}^{\infty} \mathbb{P}(T \wedge n \geq k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(T \geq k) = \mathbb{E}(T).$$

可得$\mathbb{E}(T) \geq f(z)$, 即

$$\mathbb{E}(T) \geq \frac{1}{m}\Big(\mathrm{e}^{\lambda(b-a)}\frac{1-\mathrm{e}^{-\lambda(z-a)}}{\lambda} - (z-a)\Big), \quad z \in (a, b).$$

练习题

24.1 试构造一个鞅$\{X_n, n \geq 0\}$和停时T使得$\mathbb{E}(X_T) \neq \mathbb{E}(X_0)$.
24.2 证明对任意鞅或非负下鞅$\{X_n, n \geq 0\}$及停时τ, $\mathbb{E}(|X_\tau|) \leq \varliminf_{n\to\infty} \mathbb{E}(|X_n|)$.
24.3 若例24.1中$p = 1/2$, 求相应的$\mathbb{E}(T), \mathbb{E}(S_T)$.

§7.3 鞅不等式与极限

作为本章以及本书的结束, 本节我们简单介绍几个鞅的不等式以及一个关于鞅收敛的基本定理.

(A) 极大不等式与Doob不等式

定理25.1 设$\{X_n; n \ge 0\}$为上鞅, 则对任意$\lambda > 0$有

$$\lambda \mathbb{P}(\max_{0\le k\le n} X_k \ge \lambda) \le \mathbb{E}(X_0) - \mathbb{E}(X_n \mathbf{1}_{\{\max_{0\le k\le n} X_k < \lambda\}}); \tag{7.3.1}$$

$$\lambda \mathbb{P}(\min_{0\le k\le n} X_k \le -\lambda) \le -\mathbb{E}(X_n \mathbf{1}_{\{\min_{0\le k\le n} X_k < -\lambda\}}); \tag{7.3.2}$$

$$\lambda \mathbb{P}(\max_{0\le k\le n} |X_k| \ge \lambda) \le \mathbb{E}(X_0) + 2\mathbb{E}(X_n^-), \qquad \text{其中}\ \ X_n^- = (-X_n) \vee 0. \tag{7.3.3}$$

证明 (1) 令$T = \min\{k; X_k \ge \lambda\} \wedge n$, 则$T$为有界停时. 由有界停时定理

$$\begin{aligned}\mathbb{E}(X_0) \ge \mathbb{E}(X_T) &= \mathbb{E}(X_T \mathbf{1}_{\{\max_{0\le k\le n} X_k < \lambda\}}) + \mathbb{E}(X_T \mathbf{1}_{\{\max_{0\le k\le n} X_k \ge \lambda\}}) \\ &\ge \lambda \mathbb{P}(\max_{0\le k\le n} X_k \ge \lambda) + \mathbb{E}(X_n \mathbf{1}_{\{\max_{0\le k\le n} X_k < \lambda\}}),\end{aligned}$$

移项得(7.3.1).

(2)令$S = \min\{k; X_k \le -\lambda\} \wedge n$, 则$S$为有界停时,且$S \le n$. 由有界停时定理

$$\begin{aligned}\mathbb{E}(X_n) \le \mathbb{E}(X_S) &= \mathbb{E}(X_S \mathbf{1}_{\{\min_{0\le k\le n} X_k \le -\lambda\}}) + \mathbb{E}(X_S \mathbf{1}_{\{\min_{0\le k\le n} X_k > -\lambda\}}) \\ &\le -\lambda \mathbb{P}(\min_{0\le k\le n} X_k \le -\lambda) + \mathbb{E}(X_n \mathbf{1}_{\{\max_{0\le k\le n} X_k > -\lambda\}}),\end{aligned}$$

移项得(7.3.2).

(3) 由(7.3.1),(7.3.2)两项直接相加可得(7.3.3). □

注25.1 由$\{X_n; n \ge 0\}$为下鞅可推出$\{|X_n|; n \ge 0\}$仍是下鞅. 由(2)可得, 此时

$$\mathbb{P}(\max_{0\le k\le n} |X_k| \ge \lambda) \le \frac{1}{\lambda} \mathbb{E}(|X_n| \mathbf{1}_{\{\max_{0\le k\le n} |X_k| \ge \lambda\}}).$$

引理25.2 若随机变量ζ, η满足$\mathbb{E}(|\zeta|^p) < \infty, \mathbb{E}(|\eta|^q) < \infty$, 其中$1/p + 1/q = 1$. 那么$E|\zeta\eta| < \infty$且

$$\mathbb{E}(|\zeta\eta|) \le [\mathbb{E}(|\zeta|^p)]^{1/p} [\mathbb{E}(|\eta^q|)]^{1/q}. \tag{7.3.4}$$

证明 由$\ln x$为凹函数可知, 对任意$x, y \in (0, \infty)$, $\lambda \in [0, 1]$,

$$\ln(\lambda x + (1-\lambda) y) \ge \lambda \ln x + (1-\lambda) \ln y.$$

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. 令$x = |\zeta|^p/\mathbb{E}(|\zeta|^p)$, $y = |\eta|^q/\mathbb{E}(|\eta|^q)$, 那么

$$\frac{|\zeta\eta|}{\mathbb{E}(|\zeta|^p)^{1/p}(\mathbb{E}(|\eta|^q))^{1/q}} \le \frac{1}{p}\frac{|\zeta|^p}{\mathbb{E}(|\zeta|^p)} + \frac{1}{q}\frac{|\eta|^q}{\mathbb{E}(|\eta|^q)}.$$

两边求期望得

$$\mathbb{E}\Big(\frac{|\zeta\eta|}{(\mathbb{E}(|\zeta|^p))^{1/p}(\mathbb{E}(|\eta|^q))^{1/q}}\Big) \le \frac{1}{p}\mathbb{E}\Big(\frac{|\zeta|^p}{\mathbb{E}(|\zeta|^p)}\Big) + \frac{1}{q}\mathbb{E}\Big(\frac{|\eta|^q}{\mathbb{E}(|\eta|^q}\Big) = 1.$$

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25.3 (鞅极大不等式)设$\{X_n; n \ge 0\}$为鞅或非负下鞅, $X_n^* = \max\limits_{0\le k\le n}|X_k|$. 那么

(1) 对任意$\lambda > 0$及$r \ge 1$, $\mathbb{P}(X_n^* \ge \lambda) \le \lambda^{-r}\mathbb{E}(|X_n|^r)$. (极大不等式)

(2) 对任意$r > 1$, $\mathbb{E}(|X_n^*|^r) \le (\frac{r}{r-1})^r\mathbb{E}(|X_n|^r)$. (Doob 不等式)

证明 若$\mathbb{E}(|X_n|^r) = +\infty$, 则结论显然成立. 下设$\mathbb{E}(|X_n|^r) < \infty$.

由$\{|X_n|^r, n \ge 0\}$为非负下鞅可得, 对任意$k \le n$,

$$\mathbb{E}(|X_n|^r) \ge \mathbb{E}(|X_k|^r).$$

从而令

$$Y_k = \begin{cases} |X_k|^r, & k \le n; \\ |X_n|^r, & k > n. \end{cases}$$

那么$\{Y_k; k \ge 0\}$仍为非负下鞅. 令

$$T = \min\{k, |Y_k| \ge \lambda\} \wedge n.$$

显然T为有界停时且$T \le n$,

$$\mathbb{E}(|X_n|^r) = \mathbb{E}(|Y_n|^r) \ge \mathbb{E}(|Y_T|^r) = \mathbb{E}(|X_T|^r) \ge \lambda^r\mathbb{P}(X_T \ge \lambda) = \lambda^r\mathbb{P}(X_n^* \ge \lambda).$$

由此立即可得

$$\mathbb{P}(X_n^* \ge \lambda) \le \lambda^{-r}\mathbb{E}(|X_n|^r).$$

注意到

$$\begin{aligned}\mathbb{E}(|X_n^*|^r) &= \int_0^\infty r\lambda^{r-1}\mathbb{P}(|X_n^*| \ge \lambda)d\lambda \le \int_0^\infty r\lambda^{r-2}\mathbb{E}(|X_n|\mathbf{1}_{\{|X_n^*|\ge\lambda\}})d\lambda \\ &= \mathbb{E}(|X_n|\int_0^{X_n^*} r\lambda^{r-2}d\lambda) = \frac{\lambda}{\lambda-1}\mathbb{E}(|X_n||X_n^*|^{r-1}).\end{aligned}$$

由Hölder不等式得

$$\mathbb{E}(|X_n^*|^r) \le (\frac{r}{r-1})^r(\mathbb{E}(|X_n|^r))^{1/r}(\mathbb{E}(|X_n^*|^r))^{1-1/r}.$$

由此可得Doob不等式成立. □

例25.1 设$X_n, n \ge 1$为独立同分布的随机变量且

$$\mathbb{E}(X_1) = 0, \quad \mathrm{Var}(X_1) = \sigma^2.$$

则$S_n = \sum_{k=1}^n X_k$为鞅, 由极大不等式可得

$$\mathbb{P}(\max_{0\le k\le n} S_n \ge \lambda) \le \frac{1}{\lambda^2}\mathbb{E}(S_n^2) = \frac{n\sigma^2}{\lambda^2}. \qquad \text{(Kolomogrov 不等式)}$$

以及由Doob不等式得

$$\mathbb{E}(|\max_{0\le k\le n} S_k|^2) \le 4\mathbb{E}(S_n^2) = 4n\sigma^2.$$

(B) 上穿不等式

对任意随机序列$X = \{X_n; n \ge 0\}$以及闭区间$[a, b]$, 令

$$T_0 = \inf\{n \ge 0, X_n \le a\}, \qquad T_1 = \inf\{n > T_0, X_n > b\};$$

$$T_{2j} = \inf\{n > T_{2j-1}, X_n \le a\}, \qquad T_{2j+1} = \inf\{n > T_{2j}, X_n > b\},\ j = 1, 2, \cdots.$$

显然$\{T_k, k \ge 1\}$为一列单调不降的关于X_n的停时列.

定义25.1 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$U_a^b[X, n]$.

直观地理解, 时刻n之前的上穿次数就是序列X在n之前按时间顺序, X的取值从区间$[a, b]$的下方升到上方这种现象先后发生的次数. 按定义,

$$\{U_a^b[X, n] = j\} = \{T_{2j-1} \le n < T_{2j+1}\}.$$

定理25.4 设$\{X_n; n \ge 0\}$为上鞅,则对任意$n \ge 1$,

$$\mathbb{E}(U_a^b[X, n]) \le \frac{1}{b-a}\mathbb{E}[(X_n - a)^-].$$

证明 由有界停时定理, 对任意$k \ge 0$,

$$\begin{aligned}
0 &\ge \mathbb{E}(X_{T_{2k+1}\wedge n} - X_{T_{2k}\wedge n}) \\
&= \mathbb{E}\Big((X_{T_{2k+1}\wedge n} - X_{T_{2k}\wedge n})(\mathbf{1}_{\{T_{2k}\le n<T_{2k+1}\}} + \mathbf{1}_{\{n\ge T_{2k+1}\}})\Big) \\
&\ge \mathbb{E}((X_n - a)\mathbf{1}_{\{T_{2k}\le n<T_{2k+1}\}}) + (b-a)\mathbb{P}(T_{2k+1} \le n) \\
&\ge \mathbb{E}((X_n - a)\mathbf{1}_{\{T_{2k}\le n<T_{2k+1}\}}) + (b-a)\mathbb{P}(U_a^b(x, n) \ge k+1).
\end{aligned}$$

从而

$$
\begin{aligned}
\mathbb{P}(U_a^b(X,n) \geq k+1) &\leq \frac{1}{b-a}\mathbb{E}(-(X_n-a)\mathbf{1}_{\{T_{2k}\leq n<T_{2k+1}\}}) \\
&\leq \frac{1}{b-a}\mathbb{E}((X_n-a)^-\mathbf{1}_{\{T_{2k}\leq n<T_{2k+1}\}}).
\end{aligned}
$$

两边关于k求和得

$$
\mathbb{E}(U_a^b(X,n)) \leq \frac{1}{b-a}\mathbb{E}((X_n-a)^-\mathbf{1}_{\{T_0\leq n\}}) \leq \frac{1}{b-a}\mathbb{E}((X_n-a)^-).
$$

定理证毕. □

对下鞅我们有如下的上穿不等式, 证明参见练习25.4.

定理25.5 设$\{X_n; n\geq 0\}$为下鞅, 则对任意$n\geq 1$, $\mathbb{E}(U_a^b(X,n)) \leq \frac{1}{b-a}\mathbb{E}((X_n-a)^+)$.

(C) Doob收敛定理

定理25.6 (Doob收敛定理) 设$\{X_n; n\geq 0\}$为(上，下)鞅，若$M=\sup\limits_n \mathbb{E}(|X_n|)<\infty$. 则存在随机变量$X_\infty$使得当$n\to\infty$时$X_n\to X_\infty$ a.s. 且$\mathbb{E}|X_\infty|<\infty$.

证明 令$\mathbb{Q}$表示有理数全体, 设$a,b\in\mathbb{Q}$, $a<b$. 令$U_a^b(X)$为$\{X_n\}$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

$$
\begin{aligned}
\mathbb{E}(U_a^b(X)) &\leq \frac{1}{b-a}\sup_n \mathbb{E}((X_n-a)^-)(\text{或 } \frac{1}{b-a}\sup_n \mathbb{E}((X_n-a)^+)) \\
&\leq \frac{1}{b-a}(a+\sup_n \mathbb{E}|X_n|)<\infty.
\end{aligned}
$$

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$\mathbb{P}(W_{a,b})=0$, 从而$\mathbb{P}(W)=0$. 显然对任意$\omega\notin W$,

$$
\liminf_{n\to\infty} X_n(\omega) = \limsup_{n\to\infty} X_n(\omega),
$$

即$\lim\limits_{n\to\infty} X_n(\omega)$存在, 记作$X_\infty(\omega)$. 若$\omega\in W$则补充定义$X_\infty(\omega)=0$. 因此

$$
X_n \to X_\infty, a.s.
$$

令$p=\mathbb{P}(|X_\infty(\omega)|=\infty)$, 那么由

$$
\lim_{n\to\infty} \mathbb{P}(|X_n(\omega)|\geq c) \to \mathbb{P}(|X_\infty(\omega)|\geq c) \geq p
$$

可知, 对$c=3M/p$时，存在N, 当$n>N$时, $\mathbb{P}(|X_n|\geq c)\geq p/2$. 因此若$p>0$,则

$$\mathbb{E}(|X_n|)\geq \frac{3M}{2}.$$

这与条件$M=\sup_n \mathbb{E}(|X_n|)$矛盾, 因此$p=0$, 即$\mathbb{P}(|X_\infty|<\infty)=1$.

另一方面,对任意$c>0, n>0, \epsilon>0$可得

$$\begin{aligned}\mathbb{E}(|X_\infty|\mathbf{1}_{\{|X_\infty|\leq c\}}) &= \mathbb{E}(|X_\infty|\mathbf{1}_{\{|X_\infty|\leq c,|X_n|\leq c\}})+\mathbb{E}(|X_\infty|\mathbf{1}_{\{|X_\infty|\leq c,|X_n|>c\}})\\ &\leq \mathbb{E}(|X_\infty|\mathbf{1}_{\{|X_\infty|\leq c,|X_n|\leq c\}})+\mathbb{E}(|X_n|\mathbf{1}_{\{|X_\infty|\leq c,|X_n|>c\}})\\ &\leq \mathbb{E}(|X_\infty-X_n|\mathbf{1}_{\{|X_\infty|\leq c,|X_n|\leq c\}})+\mathbb{E}(|X_n|\mathbf{1}_{\{|X_\infty|\leq c\}})\\ &\leq \epsilon\mathbb{P}(|X_\infty|\leq c,|X_n|\leq c)+2c\mathbb{P}(|X_n-X_\infty|>\epsilon)+\mathbb{E}(|X_n|)\\ &\leq \epsilon+2c\mathbb{P}(|X_n-X_\infty|>\epsilon)+a\end{aligned}$$

先令$n\to\infty$再令$\epsilon\to 0$得$\mathbb{E}(|X_\infty|\mathbf{1}_{\{|X_\infty|\leq c\}})\leq a$. 再令$c\to\infty$得$\mathbb{E}(|X_\infty|)\leq a$. □

例25.2 设非负随机过程$X=\{X_n,\ n\geq 0\}$满足

$$\mathbb{E}(X_{n+1}|X_0,\cdots,X_n)=2X_n+1,\qquad n\geq 0.$$

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

$$\mathbb{E}(\frac{X_{n+1}}{2^{n+1}}|X_0,\cdots,X_n)=\frac{X_n}{2^n}+\frac{1}{2^n}\geq\frac{X_n}{2^n}$$

由此可知$\{X_n/2^n\}$是下鞅, 而且

$$\begin{aligned}\mathbb{E}(\frac{X_{n+1}}{2^{n+1}}) &= \mathbb{E}(\mathbb{E}(\frac{X_{n+1}}{2^{n+1}}|X_0,\cdots,X_n))=\mathbb{E}(\frac{X_n}{2^n}+\frac{1}{2^n})\\ &= \mathbb{E}(X_1/2)+\frac{1}{2}+\cdots+\frac{1}{2^n}\leq \mathbb{E}(X_1/2)+1<+\infty.\end{aligned}$$

由定理19.6可知$n\to\infty$时$X_n/2^n$几乎必然收敛. □

例25.3 设X是时齐的不可约常返马氏链, 转移概率矩阵$P=(p_{ij})_{i,j\in S}$. 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

$$\begin{aligned}\mathbb{E}(Y_{n+1}|X_0,\cdots,X_n) &= \mathbb{E}(y(X_{n+1})|X_0,\cdots,X_n)=\mathbb{E}(y(X_{n+1})|X_n)\\ &= \sum_{j\in S}p_{X_n,j}y(j)=y(X_n)=Y_n.\end{aligned}$$

因此$\{Y_n, n \geq 0\}$为鞅且有界. 由鞅收敛定理可知存在随机变量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$i, j \in S$,

$$\mathbb{P}(X_n = i, \text{i.o.}) = \mathbb{P}(X_n = j, \text{i.o.}) = 1,$$

因此存在$\omega \in \Omega$使得$Y_n(\omega) \to U(\omega)$ 且存在子列n'_k, n''_k

$$X_{n'_k}(\omega) = i, \qquad X_{n''_k}(\omega) = j.$$

这表明$y(i) = U(\omega) = y(j)$. 因此$y(i), i \in S$, 为恒常数. □

练习题

25.1 设$\{X_n, n \geq 0\}$为鞅, $\mathbb{E}(X_n) = 0$且$\mathbb{E}(X_n^2) < \infty$, $n \geq 0$, 证明

$$\mathbb{P}(\max_{0 \leq k \leq n} X_k \geq \lambda) \leq \frac{\mathbb{E}(X_n^2)}{\mathbb{E}(X_n^2) + \lambda^2}.$$

25.2 给定$\lambda > 0$, 若$X_1, X_2, \cdots$,满足对任意$n \geq 1$, $\mathbb{E}(\exp(\lambda X_{n+1}) | X_1, \cdots, X_n) \leq 1$, 令$S_n = X_1 + \cdots + X_n$, $S_0 = 0$. 证明

$$\mathbb{P}(\max_{n \geq 0}(x + S_n) > l) \leq \mathrm{e}^{-\lambda(l-x)}, \quad , x \leq l.$$

25.3 设$X_1, X_2, \cdots$,为独立随机变量且存在$t > 0$使得对任意$k \geq 1$,

$$\psi_k(t) = \mathbb{E}(\exp(tX_k)) < \infty.$$

若$\Psi_n(t) = \prod_{k=1}^n \psi_k(t) \to \Psi(t) \in (0, \infty)$,证明$S_n = X_1 + \cdots + X_n$几乎必然收敛.
25.4* 证明定理25.5.

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] 钱敏平, 龚光鲁. 随机过程论(第二版). 北京: 北京大学出版社, 1997.

[8] 方兆本, 廖柏其. 随机过程(第二版). 北京: 科学出版社，2004.

[9] 何声武. 随机过程引论. 北京: 高等教育出版社, 1999.

[10] 严士健, 刘秀芳. 测度与概率. 北京: 北京师范大学出版社, 2003.