第十章　随机过程基础

§10.1　随机过程的基本概念

在初等概率论中所研究的随机现象，基本上可由一个或有限多个随机变量来描述。虽然在极限定理中涉及到了无穷多个随机变量的情况，但我们总是假定它们之间是相互独立的。然而在自然界和科学技术的许多领域中还存在大量随机现象需要用彝族无穷多个相互关联的随机变量来进行描述。通常我们把这一族随机变量称为随机过程。

一、随机过程的定义

定义 1　设 (Ω, F, P) 是一个概率空间，T 是实数集的某个子集，若对每个 $t \in T$, $X(t,\omega)$ 是概率空间 (Ω, F, P) 上的随机变量，则称这族随机变量 $\{X(t,\omega), t \in T\}$ 为 (Ω, F, P) 上的一个随机过程.

易见，我们在概率论中所学过的多维随机向量和随机变量序列，都可以看成随机过程。

如果参数 t 表示时间，$X(t,\omega)$ 表示某系统在时刻 t 的状态，则随机过程 $\{X(t,\omega), t \in T\}$ 表示该系统的状态 $X(t,\omega)$ 随时间 t 的变化而变化的情况，且在时刻 t 的状态不能完全由 t 来确定，它还受随机因素 ω 的影响，于是，称集合 $\{X(t,\omega), t \in T, \omega \in \Omega\}$ 为过程的状态空间,通常用 E 来表示。

对每个固定的 $\omega \in \Omega$，$X(t,\omega)$ 是 T 上的实值函数，称此函数为过程相应于 ω 的样本函数（轨道或现实）。

易见，一个样本函数可看成是对随机过程的一次观察值。

一元随机过程 $\{X(t,\omega), t \in T\}$ 可简写为 $X_T = \{X(t), t \in T\}$。

二、随机过程的有限维分布函数族

在概率论中，我们用联合分布函数（或联合密度函数）来描述一个随机向量 $(X_1, X_2, \cdots, X_n)$ 的概率性质，在过程中我们用有限维分布函数族作为工具来研究随机过程。

定义 2　设 $X_T = \{X(t), t \in T\}$ 是概率空间上的一元随机过程，对任意固定的 $t \in T$, $F(t,x)$ 表示 $X(t)$ 的分布函数，则称 $\{F(t,x), t \in T\}$ 为过程的一维分布函数族；对任意固定的 $t_1, t_2 \in T$, $F(t_1, t_2; x_1, x_2)$ 表示 $\{X(t_1), X(t_2)\}$ 的联合分布函数，则称 $\{F(t_1, t_2; x_1, x_2), t_1, t_2 \in T\}$ 为过程的二维分布函数族。

一般地，对任意固定的自然数 n，任意固定的 $t_1, t_2, \cdots, t_n \in T$，

$F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n)$ 表示 $\{X(t_1),X(t_2),\cdots X(t_n)\}$ 的联合分布函数，则称 $\{F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n),\ t_1,t_2,\cdots,t_n\in T,n\geq 1\}$ 为过程的有限维分布函数族。

固定 $t_1,t_2,\cdots,t_n\in T$，可以得到过程的一个 n 维分布函数，当 $t_1,t_2,\cdots,t_n$ 在 T 中变动，n 在自然数集中变动时，就得到过程的有限维分布函数族。

过程的有限维分布函数族具有下述性质：

1）对 $1,2,\cdots,n$ 的任意排列 $j_1,j_2,\cdots,j_n$，均有

$$F(t_{j_1},t_{j_2}\cdots,t_{j_n};x_{j_1},x_{j_2},\cdots,x_{j_n})=F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n)$$

2）对任意固定的自然数 $m<n$，均有

$$F(t_1,t_2\cdots,t_m;x_1,x_2,\cdots,x_m)=F(t_1,t_2\cdots t_m,\cdots t_n;x_1,x_2,\cdots,x_m,\infty,\cdots,\infty)。$$

我们把这两条性质称为相容性条件，显然，任一有限维分布函数族都满足性质 1），2），反过来，我们有定理

定理 1（柯尔莫果洛夫存在性定理）设 T 是实数集的子集，对自然数 $t_1,t_2,\cdots,t_n\in T$，$F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n)$ 作为 $x_1,x_2,\cdots,x_n$ 的函数是抽象的 n 元分布函数，抽象分布函数族 $\{F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n),\ t_1,t_2,\cdots,t_n\in T,n\geq 1\}$ 满足相容性条件 1），2），则存在一个概率空间上的一个随机过程 $X_T=\{X(t),t\in T\}$，其有限维分布函数族恰为 $\{F(t_1,t_2\cdots,t_n;x_1,x_2,\cdots,x_n),\ t_1,t_2,\cdots,t_n\in T,n\geq 1\}$。

定理的证明见相关文献。

三、随机过程的常见类型

以随机过程的有限维分布函数的特性为标志，随机过程有以下一些常见类型。

1。二阶矩过程：这种过程的每个随机变量均存在二阶矩。常见的二阶矩过程有正交增量过程、宽平稳过程和正态过程等。

2。独立增量过程：这种过程的状态的增量是相互独立的随机变量，即对任给的 $t_1,t_2,\cdots,t_n\in T$，$X(t_2)-X(t_1),\cdots,X(t_n)-X(t_{n-1})$ 相互独立。

3。马尔可夫过程：在已知现在的状态的情况下，过程的将来的状态与过去无关。

§10.2　二阶矩过程

一、二阶矩过程的定义

定义 3　设 $\{X(t),t\in T\}$ 是复值随机过程，若对每个 $t\in T$，$E(|X(t)|^2)$ 存在，则称此过程为二阶矩过程。

我们知道，有限维分布函数族可以全面刻画二阶矩过程的概率概率性质和变化规律，

但在随机过程中，只有极少数非常特殊的情形才能给出其有限维分布函数的具体形式，而对于绝大多数场合，我们无法得到其有限维分布函数族。受概率论的启发，我们首先讨论随机过程的数字特征。

二、二阶矩过程的数字特征

定义 4 设 $\{X(t), t\in T\}$ 是复值随机过程，若对每个 $t\in T$，$E[|X(t)|]$ 存在，则称 $E[X(t)]$ 为随机过程 $\{X(t), t\in T\}$ 的均值函数。记为

$$m(t)=E[X(t)],\quad t\in T\}$$

定义 5 设 $\{X(t), t\in T\}$ 是复值随机过程，若对任意的 $s,t\in T$，

$E[|(X(s)-m(s))\overline{(X(t)-m(t)}|]$ 存在，则称 $E[(X(s)-m(s))\overline{(X(t)-m(t))}]$ 为随机过程 $\{X(t), t\in T\}$ 的协方差函数。记为

$$\Gamma(s,t)=E[(X(s)-m(s))\overline{(X(t)-m(t))}],\quad s,t\in T,$$

特殊地，当 $s=t$ 时，我们可以定义方差函数 $D(t)=E[X(t)-m(t)]^2,\quad t\in T$

显然，二阶矩过程的均值函数、方差函数均存在。

均值函数描述了随机过程状态平均取值的演化，协方差函数描述了随机过程状态相关性的演化，而方差函数描述的是随机过程状态分散程度的演化。

设 $\{X(t), t\in T\}$ 是一个二阶矩过程，均值函数为 $m(t)$，则 $\{\overline{X}(t)=X(t)-m(t), t\in T\}$ 也是一个二阶矩过程，且显然其均值函数 $E[\overline{X}(t)]=0$，所以我们在以后的讨论中都假设二阶矩过程的均值函数为零而不是一般性。这样一来，对二阶矩过程来说，协方差函数就成了它最重要的数字特征。我们下面讨论协方差函数的性质。

性质 1 二阶矩过程 $\{X(t), t\in T\}$ 的协方差函数 $\Gamma(s,t)$ 具有 Hermite 性，即对任意的 $s,t\in T$，有 $\Gamma(s,t)=\overline{\Gamma(t,s)}$。

利用协方差函数的定义和复数的性质易证，证明留给读者。

显然，若二阶矩过程 $\{X(t), t\in T\}$ 是实值随机过程，则 Hermite 性将变为对称性，即对任意的 $s,t\in T$，有 $\Gamma(s,t)=\Gamma(t,s)$。

性质 2 二阶矩过程 $\{X(t), t\in T\}$ 的协方差函数 $\Gamma(s,t)$ 是非负定的，即对任意固定的自然数 n，任意的 $t_1,t_2,\cdots,t_n\in T$，以及 T 上任意复值函数 $\theta(t)$，有

$$\sum_{k=1}^{n}\sum_{i=1}^{n}\Gamma(t_k,t_i)\theta(t_k)\overline{\theta(t_i)}\geq 0$$

证明： $$\sum_{k=1}^{n}\sum_{i=1}^{n}\Gamma(t_k,t_i)\theta(t_k)\overline{\theta(t_i)}$$

$$=\sum_{k=1}^{n}\sum_{i=1}^{n}\{E[X(t_k)-m(t_k)]\overline{[X(t_i)-m(t_i)]}\theta(t_k)\overline{\theta(t_i)}$$

$$=E\{\sum_{k=1}^{n}\theta(t_k)[X(t_k)-m(t_k)]\}\overline{\{\sum_{i=1}^{n}\theta(t_i)[X(t_i)-m(t_i)]}$$

$$=E\left|\sum_{k=1}^{n}\theta(t_k)[X(t_k)-m(t_k)]\right|^2\geq 0\text{。}$$

性质 1 和性质 2 是紧密联系的，由性质 2 可以推出性质 1。

（1）先取 $n=1,\theta(t)=1$，则由性质 2，对任意的 $t\in T$，有 $\Gamma(t,t)\geq 0$；再取 $n=2,\theta(t_1)=1$，$\theta(t_2)=i$，有

$$\Gamma(t_1,t_1)-i\,\Gamma(t_1,t_2)+i\,\Gamma(t_2,t_1)+\Gamma(t_2,t_2)\geq 0$$

由此式及 $\Gamma(t_1,t_1)\geq 0$，$\Gamma(t_2,t_2)\geq 0$ 得 $-i\,\Gamma(t_1,t_2)+i\,\Gamma(t_2,t_1)$ 是实数，即 $\Gamma(t_1,t_2)$ 与 $\Gamma(t_2,t_1)$ 的实部相等。

（2）再取 $n=2,\theta(t_1)=\theta(t_2)\equiv 1$，有

$$\Gamma(t_1,t_1)+\Gamma(t_1,t_2)+\Gamma(t_2,t_1)+\Gamma(t_2,t_2)\geq 0$$

从而 $\Gamma(t_1,t_2)+\Gamma(t_2,t_1)$ 也是实数，即 $\Gamma(t_1,t_2)$ 与 $-\Gamma(t_2,t_1)$ 的实部相等。

综合（1）、2〕得 $\Gamma(t_1,t_2)=\overline{\Gamma(t_2,t_1)}$，任意的 $t_1,t_2\in T$ 均成立。

从上述论证中可得如下一般结论:如果二元实变复值函数是非负定的,则它必是 Hermite 性的。

容易直接验证，二阶矩过程 $\{X(t),t\in T\}$ 的协方差函数 $\Gamma(s,t)$ 与均值函数 $m(t)$ 之间有下述关系：

$$\Gamma(s,t)=E[X(s)\overline{X(t)}]-m(s)\overline{m(t)}\text{。}$$

三、正交增量过程

定义 6 若二阶矩过程 $\{X(t),t\in T\}$ 满足：对任意的 $t_1,t_2,t_3,t_4\in T$，且 $t_1\leq t_2\leq t_3\leq t_4$，均有

$$E[X(t_2)-X(t_1)]\overline{[X(t_4)-X(t_3)]}=0$$

则称此过程为正交增量过程。

定理 2 设$\{X(t), t\in T\}$是正交增量过程，T是有限区间$[a,b)$，或半直线$[a,+\infty)$，$X(a)=0$，并设均值函数$m(t)\equiv 0$，则

（1）$\Gamma(s,t)=D[\min(s,t)]$；

（2）方差函数$D(t)$是单调不减函数。

证明：（1）对$a\le s\le t\in T$，令$t_1=a\le t_2=t_3\le t=t_4$，由定义6有

$$E[X(s)-X(a)][\overline{X(t)-X(s)}]=0$$

于是 $\Gamma(s,t)=E[(X(s)-m(s))\overline{(X(t)-m(t)}]$

$$=E[(X(s))\overline{(X(t)}]$$

$$=E[(X(s))\overline{(X(t)-X(s)+X(s)}]$$

$$=E[(X(s))\overline{(X(s)}]$$

$$=D(s)\text{。}$$

（2）对任意$a\le s<t\in T$，有

$$0\le E[|X(t)-X(s)|^2]$$

$$=E[(X(t)-X(s))\overline{(X(t)-X(s)}]$$

$$=E[X(t)\overline{X(t)}]-E[X(t)\overline{X(s)}]-E[X(s)\overline{X(t)}]+E[X(s)\overline{X(t)}]$$

$$=\Gamma(t,t)-\Gamma(t,s)-\Gamma(s,t)+\Gamma(t,t)$$

$$=D(t)-D(s)-D(s)+D(s)$$

$$=D(t)-D(s)$$

于是$D(s)\le D(t)$。

下面讨论正交增量过程和独立增量过程之间的关系。

定义7 设$\{X(t), t\in T\}$是一个复值随机过程，且对任意的$n\ge 3$，任意$t_1,t_2,\cdots,t_n\in T$，

当$t_1\le t_2\le\cdots\le t_n$时，总有$X(t_2)-X(t_1)$，$X(t_3)-X(t_2)$，$\cdots, X(t_n)-X(t_{n-1})$相互独立，则称此过程为独立增量过程。

定理 3 设二阶矩过程$\{X(t),t\in T\}$是独立增量过程，且均值函数$m(t)$恒为常数，则此过程为正交增量过程。

证明留给读者。

§10.3 马尔可夫链

独立随机试验模型最直接的推广就是Markov链模型。因早在1906年俄国数学家Markov就对它进行研究而得名。以后 Kolmogorov,Feller,Doob 等数学家发展了这一理论。

一、马尔可夫链的概念

1、马尔可夫链的定义

定义 8 设有随机过程$\{X(t),t\in T\}$，$T=\{0,1,2,\cdots\}$，其状态空间为$E=\{0,1,2,\cdots\}$，若对任意的正整数k，任意的$t_i\in T$，$t_i<t_{i+1},i=0,1,2,\cdots,k$，及任意非负整数$i_0,i_1,\cdots,i_{k+1}$，有

$$P\{X(t_{k+1})=i_{k+1}\mid X(t_0)=i_0,X(t_1)=i_1,\cdots,X(t_k)=i_k\}$$

$$=P\{X(t_{k+1})=i_{k+1}\mid X(t_k)=i_k\} \qquad (1)$$

（1）式表示的性质称为 Markov 性或无后效性。其直观含义是：如果把时刻t_k看作现在，那么t_{k+1}是将来的时刻，而$t_0,t_1,\cdots,t_{k-1}$则是过去的时刻，在确切知道系统现在状态的条件下，系统将来的状况与过去的状况无关。

2、转移概率

条件概率$P\{X(n+k)=j\mid X(n)=i\}$表示系统在时刻n处于状态i的条件下，在时刻$n+k$处于状态j的概率。如果把状态的变化看作转移，则称条件概率$P\{X(n+k)=j\mid X(n)=i\}$为时刻n系统从状态i经过k步转移到状态j的k步转移概率，记为

$$p_{ij}^{(k)}(n)=P\{X(n+k)=j\mid X(n)=i\},i,j\in E$$

一般地，转移概率$p_{ij}^{(k)}(n)$不仅与状态i,j有关，而且和时刻n有关。当$p_{ij}^{(k)}(n)$与时刻n无关时，表明 Markov 链具有平稳的转移概率，此时称 Markov 链为（离散时间的）齐次的 Markov 链，并把$p_{ij}^{(k)}(n)$简记为$p_{ij}^{(k)}$。

当$k=1$时，把$p_{ij}^{(1)}$记为p_{ij}，称为 Markov 链的一步转移概率。

用$P^{(k)}=(p_{ij}^{(k)})$表示 Markov 链的k步转移概率所组成的矩阵，称为 Markov 链得k步

转移概率矩阵，$P = (p_{ij})$ 称为一步转移概率矩阵。

显然，转移概率矩阵具有下列性质

（1）$p_{ij}^{(k)} \geq 0, i, j \in E$；

（2）$\sum_{j \in E} p_{ij}^{(k)} = 1, i \in E, k = 1,2,3,\cdots$。

通常称满足（1），2⁽ 得矩阵为随机矩阵。

Markov 链的 k 步转移概率满足重要的 Chapman-Kolmogorov 方程（简称 C-K 方程）。

定理 4 （C-K 方程）对任意正整数 k, l 及 $i, j \in E$，有

$$p_{ij}^{(k+l)} = \sum_{r \in E} p_{ir}^{(k)} p_{rj}^{(l)} \tag{2}$$

证明： $p_{ij}^{(k+l)} = P\{X(m+k+l) = j \mid X(m) = i\}$

$$= \frac{P\{X(m) = i, X(m+k+l) = j\}}{P\{X(m) = i\}}$$

$$= \frac{\sum_{r \in E} P\{X(m) = i, X(m+k) = r, X(m+k+l) = j\}}{P\{X(m) = i\}}$$

$$= \sum_{r \in E} \frac{P\{X(m) = i, X(m+k) = r\}}{P\{X(m) = i\}}$$

$$\cdot \frac{P\{X(m) = i, X(m+k) = r, X(m+k+l) = j\}}{P\{X(m) = i, X(m+k) = r\}}$$

$$= \sum_{r \in E} P\{X(m+k) = r \mid X(m) = i\}$$

$$P\{X(m+k+l) = j \mid X(m) = i, X(m+k) = r\}$$

$$= \sum_{r \in E} P\{X(m+k) = r \mid X(m) = i\}$$

$$\cdot P\{X(m+k+l) = j \mid X(m+k) = r\}$$

$$= \sum_{r \in E} p_{ir}^{(k)} p_{rj}^{(l)}$$

C-K 方程的直观含义是：要想由状态 i 经 $k+l$ 步到达状态 j，需先经过 k 步到达任意状态 r，然后再经过 l 步由状态 r 转移到状态 j 上去。

C-K 方程也可以用矩阵形式表示为

$$P^{(k+l)} = P^{(k)} \cdot P^{(l)}$$

由 C-K 方程可得

$$p_{ij}^{(k+1)} = \sum_{r \in E} p_{ir}^{(k)} p_{rj}$$

递推可得到

$$p_{ij}^{(k+1)} = \sum_{j_1, j_2, \cdots, j_k \in E} p_{ij_1} p_{j_1 j_2} \cdots p_{j_k j}$$

上式表明，Markov 链的多步转移概率可由其一步转移概率推得，即是说，系统状态的转移概率完全由一步转移概率确定。

3、初始分布与绝对分布

定义 9 Markov 链 $X_T = \{X(n), n = 0,1,2,\cdots\}$ 在时刻 $n(n \ge 0)$ 取各状态的概率

$$P\{X(n) = i\} = p_i^{(n)}, i \in E$$

称为它在时刻 n 的绝对概率分布，简称为绝对分布。

在上述定义中，当 $n = 0$ 时的绝对分布称为 Markov 链的初始分布

$$P\{X(0) = i\} = p_i, i \in E$$

在这里绝对概率是与转移概率（条件概率）相对而言的。

利用全概率公式，容易得到

$$p_j^{(n)} = \sum_{i \in E} p_i p_{ij}^{(n)}, j \in E$$

该式表明，任一时刻的绝对概率分布完全由初始分布和转移概率所确定 。更一般地，Markov 链的有限维分布完全由初始分布和一步转移概率所确定。

例 1 直线上的随机游动

考虑在直线上作随机运动的质点。如果在某个时刻质点位于整数 i 上，则下一步质点以概率 p 向右移动一格到 $i+1$，以概率 $q=1-p$ 向左移动一格到 $i-1$。以 $X(n)$ 表示时刻 n 质点的位置，则 $\{X(n), n = 0,1,2,\cdots\}$ 是一个随机过程。它满足 Markov 性，因此，这是一个齐次的 Markov 链。

（1）设状态空间为 $E = \{0, \pm 1, \pm 2, \cdots\}$，则其一步转移概率为

$$p_{i,i+1} = p, p_{i,i-1} = q, i \in E$$

$$p_{ij} = 0, |i - j| > 1$$

我们把这样的随机游动称为自由随机游动。

将随机游动在边界上的状态加以不同限制，可得到不同特性的随机游动。

（2）带有一个吸收壁的随机游动。设状态空间 $E = \{0,1,2,\cdots\}$，其一步转移概率为

$$p_{i,i+1} = p, p_{i,i-1} = q, i \ge 1 \quad p_{00} = 1,$$

$$p_{ij}=0,|i-j|>1$$

则当质点一旦到达状态 0 点后，将永远停留在 0 这个状态上，把 0 这样的状态称为吸收状态（吸收壁），把这种随机游动称为带有一个吸收壁的随机游动。

（3）带有两个吸收壁的随机游动。设状态空间 $E=\{0,1,2,\cdots,n\}$，其一步转移概率为

$$p_{i,i+1}=p,p_{i,i-1}=q,1\le i<n \quad p_{00}=1,p_{nn}=1,$$

$$p_{ij}=0,|i-j|>1$$

则当质点一旦到达状态 0 点（或状态 n 点）后，将永远停留在 0 这个状态（状态 n）上，把这种随机游动称为带有两个吸收壁（0 与 n）的随机游动。

（4）带有一个反射壁的随机游动。设状态空间 $E=\{0,1,2,\cdots,\}$，其一步转移概率为

$$p_{i,i+1}=p,p_{i,i-1}=q,1\le i \quad p_{00}=0,p_{01}=1,$$

$$p_{ij}=0,|i-j|>1$$

注意到质点一旦到达状态 0 点，则马上反射到状态 1 点，即质点在 0 点不可能停留，把这种随机游动称为带有一个反射壁（状态 0）的随机游动。

(5) 带有一个弹性壁的随机游动。设状态空间 $E=\{0,1,2,\cdots\}$，其一步转移概率为

$$p_{i,i+1}=p,p_{i,i-1}=q,i\ge 1; \quad p_{00}=\alpha,p_{01}=\beta,\alpha+\beta=1,$$

$$p_{ij}=0,|i-j|>1$$

把这种随机游动称为带有一个吸收壁（状态 0）的随机游动。

随机游动除了本身具有一定的物理意义以外，现实生活中的很多现象可用随机游动来解释。例如，甲，乙两赌徒赌博，设甲有无限多的赌资，乙的赌资为 a 元。每赌一局，输者给赢者 1 元，连续进行。如果乙赌徒输光泽赌博停止，则显然这是带有一个吸收壁 0 的随机游动；若规定，要么乙赌徒输光，要么乙赌徒赢得 b 元后赌博停止，则这又是带有两个吸收比（0 与 $a+b$）的随机游动；更进一步，如果允许乙赌徒进行欠账式的赌博，则这一过程又变成了自由随机游动，即修改赌博规则，可以得到不同类型的随机游动。

二、状态的分类

为了揭示 Markov 链的基本结构，需要对其状态按照某些概率特性进行分类。为此，我们先定义几个概念。

定义 10 若存在正整数 n，使 $p_{ij}^{(n)}>0$，则称状态 i 可达状态 j，记为 $i\to j$；反之，若对一切正整数 n，都有 $p_{ij}^{(n)}=0$，则从状态 i 不能到达状态 j。

如果 $i\to j$，且 $j\to i$，则称状态 i 与 j 相通，记为 $i\leftrightarrow j$。

定义 11 设 $X_T=\{X(n),n=0,1,2,\cdots\}$ 是 Markov 链，状态空间为 E，对任意的 $i,j\in E$，

令 $T_{ij}=\min\{n:X(0)=i,X(n)=j,n\ge 1\}$

则称T_{ij}为从状态i出发首次到达状态j的时刻，或称为从i到j的首达时。

按照定义 11，T_{ij}显然是一个随机变量，它的取值是指系统从状态i出发使$X(n)=j$的最小正整数n。如果这样的正整数n不存在，则记$T_{ij}=\infty$。

定义 12 设$X_T=\{X(n),n=0,1,2,\cdots\}$是 Markov 链，状态空间为$E$，对任意的$i,j\in E$，

令 $f_{ij}^{(n)}=P\{T_{ij}=n\mid X(0)=i\},n=1,2,\cdots$

则称$f_{ij}^{(n)}$为从状态i出发经过n步转移首次到达状态j的概率，简称为从i到j的首达概率。

又令 $f_{ij}=\sum_{n=1}^{\infty}f_{ij}^{(n)}=P\{T_{ij}<\infty\mid X(0)=i\}$

则f_{ij}表示系统自从状态i出发（在有限时间内）迟早到达状态j的概率。

显然有$0\le f_{ij}^{(n)}\le f_{ij}\le 1,n=1,2,\cdots$ 。

定义 13 设$X_T=\{X(n),n=0,1,2,\cdots\}$是 Markov 链，状态空间为$E$，对任意的$i\in E$，

令 $\mu_i=E(T_{ii})=\sum_{n=1}^{\infty}nf_{ii}^{(n)}$

则称μ_i为状态i的平均返回时间。

转移概率与首达概率有如下的基本关系定理。

定理 5 对任何状态$i,j\in E,n\ge 1$，有

$$p_{ij}^{(n)}=\sum_{l=1}^{n}f_{ij}^{(l)}p_{jj}^{(n-l)}\text{。}$$

证明： $p_{ij}^{(n)}=P\{X(n)=j\mid X(0)=i\}$

$$=P\{T_{ij}\le n,X(n)=j\mid X(0)=i\}$$

$$=\sum_{l=1}^{n}P\{T_{ij}=l,X(n)=j\mid X(0)=i\}$$

$$=\sum_{l=1}^{n}P\{T_{ij}=l\mid X(0)=i\}\cdot P\{X(n)=j\mid T_{ij}=l,X(0)=i\}$$

$$=\sum_{l=1}^{n}f_{ij}^{(l)}\cdot P\{X(n)=j\mid X(0)=i,X(1)\ne j,\cdots,X(l-1)\ne j,X(l)=j\}$$

$$=\sum_{l=1}^{n} f_{ij}^{(l)} \cdot P\{X(n)=j \mid X(l)=j\}$$

$$=\sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

定义 14 对状态i，若正整数集合$\{n: n \geq 1, p_{ii}^{(n)}>0\}$非空，则称该集合的最大公约数$t$为状态$i$的周期。若$t>1$，则称状态$i$是周期的；若$t=1$，则称状态$i$是非周期的。

有了以上的准备，我们现在对状态进行分类。

定义 15 设$X_T=\{X(n), n=0,1,2,\cdots\}$是 Markov 链，状态空间为$E$，

（1）若f_{ii}=1，则称状态i是常返的；若$f_{ii}<1$，则称状态i是非常返的。

（2）若f_{ii}=1，且$\mu_i<\infty$，则称状态i是正常返的；若f_{ii}=1，且$\mu_i=\infty$，则称状态i是零常返的。

（3）若状态i是非周期正常返的，则称i是遍历的。

定理 6 状态i常返的充分必要条件是

$$\sum_{n=1}^{\infty} p_{ii}^{(n)}=\infty 。$$

（证明略）

由定理 6，易得，若状态i非常返，则必有$\lim_{n \to \infty} p_{ii}^{(n)}=0$。

进一步，我们可以得到状态的分类判别法：

（1）i非常返$\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)}<\infty$；

（2）i零常返$\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)}=\infty$，且$\lim_{n \to \infty} p_{ii}^{(n)}=0$；

（3）i正常返$\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)}=\infty$，且$\lim_{n \to \infty} p_{ii}^{(n)}>0$；

（4）i是遍历的$\Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)}=\infty$，且$\lim_{n \to \infty} p_{ii}^{(n)}=\dfrac{1}{\mu_i}>0$。

定理 7 如果$i \leftrightarrow j$，则

（1）i与j或同为常返，或同为非常返；

（2）在常返的情况下，i与j或同为正常返，或同为零常返；

（3）i与j或同为非周期的，或同为周期的妾周期相同。

证明：（1）由$i \leftrightarrow j$，按照可达的定义，必存在正整数l与n，使得

$$p_{ij}^{(l)} = \alpha > 0, \quad p_{ji}^{(n)} = \beta > 0$$

由 C-K 方程，有

$$p_{ii}^{(l+m+n)} \ge p_{ij}^{(l)} p_{jj}^{(m)} p_{ji}^{(n)} = \alpha\beta p_{jj}^{(m)} \qquad \text{（a）}$$

$$p_{jj}^{(n+m+l)} \ge p_{ji}^{(n)} p_{ii}^{(m)} p_{ij}^{(l)} = \alpha\beta p_{ii}^{(m)} \qquad \text{(b)}$$

对（a），b） 两式两端关于m从 1 到∞求和得

$$\sum_{m=1}^{\infty} p_{ii}^{(m+n+l)} \ge \alpha\beta \sum_{m=1}^{\infty} p_{jj}^{(m)}, \quad \sum_{m=1}^{\infty} p_{jj}^{(n+m+l)} \ge \alpha\beta \sum_{m=1}^{\infty} p_{ii}^{(m)}$$

所以级数$\sum_{m=1}^{\infty} p_{ii}^{(m)}$与$\sum_{m=1}^{\infty} p_{jj}^{(m)}$同时收敛或同时发散，故$i$ 与 j同为常返或同为非常返。

（2）对（a），b） 两式关于m取极限，有

$$\lim_{m\to\infty} p_{ii}^{(l+m+n)} \ge \alpha\beta \lim_{m\to\infty} p_{jj}^{(m)}, \quad \lim_{m\to\infty} p_{jj}^{(n+m+l)} \ge \alpha\beta \lim_{m\to\infty} p_{ii}^{(m)}$$

由此可见$\lim_{m\to\infty} p_{ii}^{(m)}$与$\lim_{m\to\infty} p_{jj}^{(m)}$同时为零或同时为正，由状态的分类判别法知状态$i$与$j$在常返情形同为零常返或同为正常返。

（3）由于$p_{ii}^{(l+n)} \ge p_{ij}^{(l)} p_{ji}^{(n)} = \alpha\beta > 0$，因此状态$i$的周期$t_i$整除$l+n$。

设s是使$p_{jj}^{(s)} > 0$的任意正整数，则由 C-K 方程得

$$p_{ii}^{(l+s+n)} \ge p_{ij}^{(l)} p_{jj}^{(s)} p_{ji}^{(n)} > 0,$$

所以，t_i整除$l+s+n$，从而t_i整除s。

而s是使$p_{jj}^{(s)} > 0$的任意正整数，所以t_i是正整数集$\{n : n \ge 1, p_{jj}^{(n)} > 0\}$的公约数。

若记状态j的周期为t_j，则有 $t_i \le t_j$。

同理由对称性可证得$t_j \le t_i$，故得$t_i = t_j$。

所以，当$t_i = t_j > 1$时，i 与 j都是周期的，且周期相同；当$t_i = t_j = 1$时，i 与 j都是非周期的。

以上都是对单个状态的类型所作的讨论，下面我们将从整体上研究 Markov 链的状态空间。

定义 16 设C是状态空间E的一个子集，若对任意的$i \in C, j \notin C$，及任意的$n \ge 1$，都有 $p_{ij}^{(n)} = 0$，则称C是一个闭集。

定义 16 说明，闭集内的任一状态，不论转移多少步，都不能转移到闭集之外的状态上

去，即是说，随着时间的推移，闭集内任一状态只能在闭集内部的状态间转移。

显然，任一 Markov 链的状态空间E是一个闭集，而且是最大的闭集，若状态i是一个吸收态，则单点集$\{i\}$是一个闭集，且是最小的闭集。

定理 8 Markov 链的所有常返状态构成的集合是一闭集。

定理 9 （状态空间的分解定理）Markov 链的状态空间E必可分解为

$$E = N + C_1 + C_2 + \cdots + C_k + \cdots$$

其中N为全体非常返状态组成的集合，$C_1, C_2, \cdots, C_k, \cdots$是互不相交的由常返状态组成的闭集，且满足

（1）对每一确定的k, C_k内任意两个状态都相通；

（2）C_k与$C_j (j \neq k)$中的状态之间不相通。

证：先将状态空间E中的状态按常返和非常返分成两类，非常返状态构成N，常返状态构成一闭集C。

在C中再按相通关系分类：在C中任取状态i_1，凡与i_1相通的状态组成集合C_1；再从剩余的状态中任取i_2，凡与i_2相通的状态组成集合C_2，如果还有剩余的状态，再继续这种过程，…。这样最终将C分解为$C_1, C_2, \cdots, C_k, \cdots$闭集之和，由$C_k$的构造可知，它们必然满足（1），2$\emptyset$ 。

称$C_1, C_2, \cdots, C_k, \cdots$为基本常返闭集。

定义 17 设闭集A是状态空间E的子集，如果A不包含任何非空的真闭子集，则称闭集A是不可约的；一个 Markov 链，如果其状态空间是不可约的，则称该 Markov 链是不可约的。

由定理 8 易知，一个不可约的 Markov 链，或者没有非常返态，或者没有常返状态，而在只有常返状态的不可约 Markov 链中，所有的状态都是相通的。

三、Markov 链的平稳分布

在实际问题中，常常要研究当n很大时$p_{ij}^{(n)}$的性质，在数学上就是要研究$\lim\limits_{n\to\infty} p_{ij}^{(n)}$是否存在？如果存在是否和初始状态$i$有关？在 Markov 链的理论中，有关这一类问题的定理称为遍历性定理。

定义 18 设 Markov 链的状态空间为E，若对一切$i, j \in E$存在不依赖于i的极限

$$\lim_{n\to\infty} p_{ij}^{(n)} = p_j,$$

则称 Markov 链具有遍历性。

该极限的直观含义是：如果把具有遍历性的 Markov 链看作一个系统，则系统不论从哪一个状态i出发，当转移步数n充分大后，转移到状态j的概率都接近于p_j，换句话说，经

历足够长的一段时间后，系统将达到平稳状态。

定理 10 如果状态 j 非常返或零常返，则对一切 i 有

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0$$

定理 11 如果状态 j 为非周期的正常返态（遍历态），则有

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j} f_{ij}$$

由定理 10 及定理 11，若 Markov 链不可约， j 为非周期的常返态，则有

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j} \geq 0 \text{。}$$

我们称 $\{\frac{1}{\mu_j}, j\in E\}$ 为极限分布。

定义 19 设 Markov 链 $\{X(n), n\geq 0\}$ 的转移概率矩阵为 $P=(p_{ij})$ ，若非负数列 $\{\pi_j\}$ 满足

$$\sum_{j=0}^{\infty} \pi_j = 1,$$

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, j = 0,1,2,\cdots,$$

则称 $\{\pi_j, j\in E\}$ 为 Markov 链 $\{X(n), n\geq 0\}$ 的平稳分布。

为了理论证明的需要，我们常常把上述第二个条件改写为

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}^{(n)}, j = 0,1,2,\cdots; n = 1,2,\cdots$$

显然，这是一个更一般的表达式，两者的等价性留给读者自证。

如果 Markov 链的初始分布为

$$P\{X(0)=j\} = p_j, j = 0,1,2,\cdots$$

是平稳分布，则容易证明 Markov 链的绝对分布也是平稳分布，且其平稳分布就是初始分布 $\{p_j, j=0,1,2,\cdots\}$ ，这正是“平稳”二字的由来。

定理 12 非周期不可约常返链是正常返链的充分必要条件是它存在平稳分布，且这个平稳分布就是极限分布 $\{\frac{1}{\mu_j}, j\in E\}$ 。

证明 充分性：若存在平稳分布 $\{\pi_j, j\in E\}$ ，则

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}^{(n)}$$

由于 $\pi_i \geq 0, \sum_{i=0}^{\infty} \pi_i = 1$，所以极限符号和求和符号可以交换位置，两边取 $n \to \infty$ 得

$$\pi_j = \lim_{n\to\infty} \sum_{i=0}^{\infty} \pi_i p_{ij}^{(n)} = \sum_{i=0}^{\infty} \pi_i (\lim_{n\to\infty} p_{ij}^{(n)}) = \sum_{i=0}^{\infty} \pi_i (\frac{1}{\mu_j}) = \frac{1}{\mu_j}。$$

由 $\sum_{i=0}^{\infty} \pi_i = 1$，至少存在一个 $\pi_k > 0$，即 $\frac{1}{\mu_k} > 0$，故状态 k 正常返，从而整个马氏链正常返，所有的 $\pi_j = \frac{1}{\mu_j} > 0$。

必要性：设马氏链是正常返的，则有

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_j} > 0，$$

由 C-K 方程

$$p_{ij}^{(m+n)} = \sum_{k\in E} p_{ik}^{(m)} p_{kj}^{(n)} = \sum_{k=0}^{\infty} p_{ik}^{(m)} p_{kj}^{(n)}$$

令 $m \to \infty$ 得

$$\frac{1}{\mu_j} = \sum_{k=0}^{\infty} \frac{1}{\mu_k} p_{kj}^{(n)} ,$$

再令 $n \to \infty$ 得

$$\frac{1}{\mu_j} = \sum_{k=0}^{\infty} \frac{1}{\mu_k} (\lim_{n\to\infty} p_{kj}^{(n)}) = (\sum_{k=0}^{\infty} \frac{1}{\mu_k}) \frac{1}{\mu_j} ,$$

即得 $$\sum_{k=0}^{\infty} \frac{1}{\mu_k} = 1,$$

由平稳分布的定义可知极限分别 $\{\frac{1}{\mu_j}\}$ 就是平稳分布。

由定理 12 立即可得

齐次有限 Markov 链必存在平稳分布。