# 2.9 马尔可夫信源的信息熵

## 2.9.1 马尔可夫链

### 一、概念

设随机序列$\{X_n, n \in T\}$

为一马尔可夫过程，

$T=\{0,1,2,\cdots\}$为离散的时间参数集合，

$X_n \in$状态空间集$S=\{S_1, S_2, \cdots, S_J\}$，

若对所有正整数$n \in T$,

如果条件概率均满足

$$P\{X_n = S_{i_n} \mid X_{n-1} = S_{i_{n-1}}, X_{n-2} = S_{i_{n-2}}, \cdots, X_1 = S_{i_1}\}$$
$$= P\{X_n = S_{i_n} \mid X_{n-1} = S_{i_{n-1}}\}$$

则称随机过程$\{X_n, n \in T\}$为一个马尔可夫链。

**直观含义：**如果系统在$n$-1时刻处于状态$S_{n-1}$,则在将来时刻$n$的状态$S_n$与过去时刻$n$-2,…,1的状态$S_{n-2}$,$S_{n-1}$,…,$S_1$无关，仅与现在时刻$n$-1的状态$S_{n-1}$有关。

即

已知系统的现在，系统的将来与过去无关。

## 1、马尔可夫链的初始分布：

在马尔可夫链中，

记$\{p_i, i\in S\}, p_i = p\{X_0 = i\}\geq 0, i\in S$

且满足$\sum_{i\in S} p_i = 1,$

为马尔可夫链的初始分布

## 2、马尔可夫链的*k*步转移概率：

$$p_{ij}^{(k)}(m) = P\left\{X_{m+k} = j \mid X_m = i\right\} \quad i, j \in S$$

当*k*=1时称为一步转移概率：

$$p_{ij}^{(1)}(m) = p_{ij}(m)$$

3、齐次马尔可夫链：转移概率与***m***（起始时刻）无关

$$p_{ij}^{(k)}(m) = p_{ij}^{k}$$

4、遍历性：

对于齐次马尔可夫链$\{X_n, n \in T\}$,

若对所有$i, j$,均存在不依赖于$i$的极限

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

且满足 $$p_j = \sum_i p_i p_{ij}, \sum_j p_j = 1$$

# 遍历性的直观含义：

（1）无论随机点从那一个状态$S_i$出发，当转移步数$n$足够大时，转移到状态$S_j$的概率$p^{(n)}{}_{ij}$都近似等于一个常数$p_j$。即如果转移步数$n$充分大，就可以用$p_j$作为$n$步转移概率$p(n)_{ij}$的近似值。

（2）马尔可夫链在初始时刻可以处在任意状态，经过足够长的状态转移后，它所处的状态与初始状态无关；此时，每种状态出现的概率已经到达一种平稳分布。

# 二、马尔可夫链的数学描述

## 1、有限维分布函数族

若对任意的正整数$n$及

$$t_1 < t_2 < \cdots t_k < \cdots < t_n, t_k \in T(k = 1, 2, \cdots, n)$$

马尔可夫链的有限维分布函数族可表示为

$$P(X_{t_1} = i_1, \cdots X_{t_n} = i_n)$$

$$= \sum_{i \in x} p_i P(X_{t_1} = i_1 \mid X_0 = i) P(X_{t_2} = i_2 \mid X_{t_1} = i_1)$$

$$\cdots P(X_{t_n} = i_n \mid X_{t_{n-1}} = i_{n-1})$$

## 2、求法

(1) 马尔可夫链的初始分布:$\{p_i = P(X_0 = i), i \in S\}$

(2) 条件概率: $P(X_{t_k} = i_k \mid X_{t_{k-1}} = i_{k-1})$

## 三、切普曼－可尔莫哥夫方程（简称C－K方程）

若对任意的正整数$m,r,k$,具有$m+r$步转移概率的齐次马尔可夫链

均有 $$p_{ij}^{(m+r)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(r)}$$

## 证明：

$$p_{ij}^{(m+r)} = p\left\{X_{n+m+r} = S_j \mid X_n = S_i\right\} = \frac{p\left\{X_{n+m+r} = S_j, X_n = S_i\right\}}{P\{X_n = S_i\}}$$

乘法定理

$$= \sum_{k\in S} \frac{P\{X_{n+m+r} = S_j, X_{n+m} = S_k, X_n = S_i\}}{P\left\{X_{n+m} = S_k, X_n = S_i\right\}} \cdot \frac{P\left\{X_{n+m} = S_k, X_n = S_i\right\}}{P\{X_n = S_i\}}$$

$$= \sum_{k\in S} P\{X_{n+m+r} = S_j \mid X_{n+m} = S_k, X_n = S_i\} \cdot P\left\{X_{n+m} = S_k \mid X_n = S_i\right\}$$

$$= \sum_{k\in S} P\{X_{n+m+r} = S_j \mid X_{n+m} = S_k\} \cdot P\left\{X_{n+m} = S_k \mid X_n = S_i\right\}$$

由马尔可夫定义及齐次性

$$= \sum_{k\in S} p_{ik}^{(m)} p_{kj}^{(r)}$$

## 直观含义：

**C-K**方程表达了

在$X$的随机点$n$时刻处于状态$i$，然后经$m+r$步转移，于$n+m+r$时刻到达状态$j$的转移概率，等于随机点在$n$时刻由状态$i$出发，先经$m$步转移于$n+m$时刻到达$X$中任一状态$S$的转移概率乘以随机点在$n+m$时刻，由状态$S$出发再经$r$步转移于$n+m+r$时刻到达状态$j$的转移概率的乘积对$X$中所有状态之和。

# 四、齐次马尔可夫链的概率分布

由$C-K$方程，

$$P\{X(t_1)=i_1, X(t_2)=i_2, \cdots, X(t_n)=i_n,\}$$

$$=\sum_{i\in S} p_i p_{ii_1}^{(t_1)} p_{i_1 i_2}^{(t_2-t_1)} \cdots p_{i_{n-1}i_n}^{(t_n-t_{n-1})}$$ **n维情况**

$$P\{X_n=j\}=\sum_{i\in S} p_i p_{ij}^{(n)}$$ **1维情况**