应用概率统计 第 33 卷
第 4 期 2017 年 8 月

Chinese Journal of Applied Probability and Statistics
Aug., 2017, Vol. 33, No. 4, pp. 408-416
doi: 10.3969/j.issn.1001-4268.2017.04.006

# 二叉树分枝马氏链的强大数定律和Shannon-McMillan定理*

张 艳 杨卫国

(江苏大学理学院, 镇江, 212013)

**摘 要:** 首先研究二叉树有限状态分枝马氏链的随机序偶出现频率的强大数定律, 之后研究二叉树有限状态分枝马氏链函数的强大数定律, 作为推论得到二叉树有限状态分枝马氏链的Shannon-McMillan定理.

**关键词:** 二叉树; 非齐次马氏链; 强大数定律; Shannon-McMillan定理

**中图分类号:** O211.4; O211.6

**英文引用格式:** Zhang Y, Yang W G. The strong law of large numbers and the Shannon-McMillan theorem for nonhomogeneous bifurcating Markov chains indexed by a binary tree [J]. Chinese J. Appl. Probab. Statist., 2017, 33(4): 408–416. (in Chinese)

## §1. 引 言

树图$T$是一个没有回路的连通图, 对于任意两个顶点$\alpha \neq \beta \in T$. 设$\overline{\alpha\beta}$是连接$\alpha$和$\beta$的唯一路径, 路径$\overline{\alpha\beta}$中含有的边数记为$d(\alpha,\beta)$, 称为$\alpha$到$\beta$的距离.

设$T$是一个以$o$为根顶点的局部有限的无限树图, 对于$T$中的任意两个顶点$\sigma$和$t$, 如果$\sigma$是处在从根顶点$o$到$t$的唯一路径上, 则记为$\sigma \wedge t$. 我们用$\sigma \wedge t$表示同时满足$\sigma \wedge t \leqslant t$与$\sigma \wedge t \leqslant \sigma$的离$o$最远的顶点. 对于$T$上的任一顶点$t$, $|t|$表示$o$和$t$之间的距离. 一个顶点如果其与根顶点的距离为$n$, 则称该顶点位于第$n$层. $L_n$表示$T$的第$n$层上所有顶点的子图, $L_n^m$表示$T$的含有从$n$层到$m$层的所有顶点的子图, 特别$T^{(n)} = L_0^n$表示$T$的含有从0层(根)到$n$层的所有顶点子图. 如果树图$T$的根顶点有$N$个相邻顶点, 而其他顶点有$N+1$个相邻顶点, 我们称此树为Cayley树, 记为$T_{C,N}$. 对于Cayley树$T_{C,N}$上的每一个顶点$t$, 在它的下一层都有$N$个相邻顶点, 我们称这$N$个顶点为$t$的子代, $t$为这$N$个顶点的父代. 对$\forall\, t \in T_{C,N}$, 记$1_t$为$t$的父代. 本文主要研究二叉树$T_{C,2}$ (见图1). 为了方便, 我们将$T_{C,2}$简记为$T_2$. 对于二叉树上任一顶点$t$, 记$t^1$和$t^2$为$t$的两个子代.

设$(\Omega, \mathscr{F}, \mathsf{P})$为一概率空间, $\{X_t, t \in T_2\}$为定义在$(\Omega, \mathscr{F}, \mathsf{P})$上的树指标随机过程, 设$A$为$T_2$的子图, 记$X^A = \{X_t, t \in A\}$, 用$|A|$表示$A$中顶点的个数, $x^A$表示$X^A$的实现.

*国家自然科学基金项目(批准号: 11571142)资助.
本文2016年1月20日收到, 2016年4月20日收到修改稿.

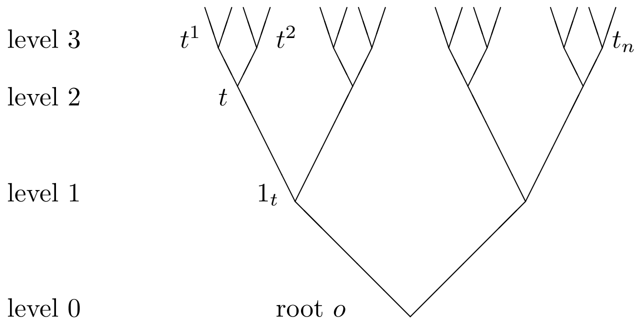


**图1　二叉树$T_{C,2}$**

**定义 1** ([1; 定义1])　设$T_2$为二叉树, $\{X_t, t\in T_2\}$是定义在概率空间$(\Omega,\mathscr{F},\mathsf{P})$上可列状态空间$G$中取值的随机变量集合, 设$p=\{p(x), x\in G\}$是$G$上一概率分布, $P_t=(P_t(y_1,y_2\,|\,x))$, $x,y_1,y_2\in G$, $t\in T_2$是定义在$G\times G^2$上的一随机矩阵, 满足$P_t(y_1,y_2\,|\,x)\geqslant 0$, $\forall\, y_1,y_2,x\in G$, 及 $\sum\limits_{(y_1,y_2)\in G^2} P_t(y_1,y_2\,|\,x)=1$, $\forall\, x\in G$, 如果$\forall\, n\geqslant 1$,

$$\mathsf{P}\big(X^{L_n}=x^{L_n}\,|\,X^{T^{(n-1)}}=x^{T^{(n-1)}}\big)=\prod_{t\in L_{n-1}} P_t(x_{t^1},x_{t^2}\,|\,x_t),$$

和

$$\mathsf{P}(X_o=x)=p(x),\qquad \forall\, x\in G,$$

则称$\{X_t, t\in T_2\}$为具有初始分布$p$与随机矩阵$P$在$G$中取值的二叉树上非齐次分枝马氏链.

如果$\forall\, t\in T_2$, $P_t=P$, 这里$P=(P(y_1,y_2\,|\,x))$, $x,y_1,y_2\in G$是定义在$G\times G^2$上的随机矩阵, $\{X_t, t\in T_2\}$为在$G$中取值的二叉树上齐次分枝马氏链. 且由文献[1]已证明$\{X_t, t\in T_2\}$为由定义1定义的二叉树非齐次分枝马氏链的充要条件是$\forall\, n\geqslant 1$和$\forall\, x^{T^{(n)}}\in S^{T^{(n)}}$,

$$\mathsf{P}\big(X^{T^{(n)}}=x^{T^{(n)}}\big)=p(x_o)\prod_{t\in T^{(n-1)}} P_t(x_{t^1},x_{t^2}\,|\,x_t). \tag{1}$$

树指标随机过程是近年来发展起来的概率论的一个新的研究方向. Berger和Ye [2]研究了齐次树上某些平稳随机场的熵率. Benjamini和Peres [3]给出了树指标马氏链的定义并研究其常返性及角常返性. Yang [4]研究了齐次树指标有限状态马氏链的强大数定理和渐近均分性(AEP). Yang和Ye [5]研究了齐次树指标层非齐次马氏链的强大数定理和渐近均分性. Huang和Yang [6]研究了一致有界树指标马氏链的强大数定理和渐近均分性. Dong等 [7]研究了Cayley树指标有限状态非齐次马氏链的强大数定理和渐近均分性. 陈晓雪等 [8]研究了树指标马氏链的等价定义. Guyon [9]定义了在任意状态空间取值的二叉树上分枝马氏链, 并研究了其极限定理. Dang等 [1]研究了二叉树非齐次分枝马氏链的等价定义、强大数定理和熵遍历定理. 杨卫国 [10]总结了近年来树指标马氏链强极限定理的若干研究.

本篇文章首先研究二叉树有限状态分枝马氏链的随机序偶集出现频率的强大数定律, 之后研究二叉树有限状态分枝马氏链函数的强大数定理, 作为推论得到二叉树有限状态分枝马氏链的Shannon-McMillan定理.

## §2. 强极限定理

**引理 2** ([11; 引理3.1]) 设$T_2$为二叉树, $\{X_t, t\in T_2\}$是如定义1所定义的在可列状态空间$G$中取值的二叉树非齐次分枝马氏链, $\{g_t(x,y_1,y_2), t\in T_2\}$是定义在$G^3$上的函数. $\forall t\in T_2$, $\exists \alpha>0$, s.t. $\mathsf{E}[\mathrm{e}^{\alpha|g_t(X_t,X_{t^1},X_{t^2})|}]<\infty$, 设$L_0=0$, $\mathscr{F}_n=\sigma(X^{T^{(n)}})$, $n\geqslant 1$. 当$|\lambda|\leqslant\alpha$, 令

$$t_n(\lambda,\omega)=\frac{\mathrm{e}^{\lambda\sum\limits_{t\in T^{(n-1)}}g_t(X_t,X_{t^1},X_{t^2})}}{\prod\limits_{t\in T^{(n-1)}}\mathsf{E}[\mathrm{e}^{\lambda g_t(X_t,X_{t^1},X_{t^2})}\,|\,X_t]},\qquad n\geqslant 1,$$

则$\{t_n(\lambda,\omega),\mathscr{F}_n,n\geqslant 1\}$是非负鞅, 且$\mathsf{E}[t_n(\lambda,\omega)]=1$.

**引理 3** ([1; 定理3]) 设$T_2$, $\{X_t, t\in T_2\}$, $\{g_t(x,y_1,y_2), t\in T_2\}$如引理2所定义, 令$\{a_n, n\geqslant 1\}$是正随机变量序列, 设

$$B=\Big\{\omega:\lim_{n\to\infty}a_n(\omega)=\infty\Big\},$$

$$D(\alpha)=\Big\{\omega:\limsup_{n\to\infty}\frac{1}{a_n}\sum_{t\in T^{(n-1)}}\mathsf{E}\big[g_t^2(X_t,X_{t^1},X_{t^2})\mathrm{e}^{\alpha|g_t(X_t,X_{t^1},X_{t^2})|}\,|\,X_t\big]=M(\omega)<\infty\Big\}\cap B,$$

$$H_n(\omega)=\sum_{t\in T^{(n-1)}}g_t(X_t,X_{t^1},X_{t^2}),$$

$$G_n(\omega)=\sum_{t\in T^{(n-1)}}\mathsf{E}[g_t(X_t,X_{t^1},X_{t^2})\,|\,X_t],$$

那么有

$$\lim_{n\to\infty}\frac{H_n(\omega)-G_n(\omega)}{a_n}=0,\qquad \text{a.e. }\ \omega\in D(\alpha).$$

**引理 4** ([1; 推论1]) 设$T_2$为二叉树, $\{X_t, t\in T_2\}$是如定义1所定义的在可列状态空间$G$中取值的二叉树非齐次分枝马氏链, $\{g_t(x,y_1,y_2), t\in T_2\}$在$G^3$上一致有界, 那么

$$\lim_{n\to\infty}\frac{H_n(\omega)-G_n(\omega)}{|T^{(n)}|}=0,\qquad \text{a.e.}$$

**引理 5** ([7; 引理3]) 设$T_2$是二叉树, $\phi(x)$是定义在区间$\triangle$上的有界函数, $\phi$在$x=a$处连续($a\in\triangle$). 设$\{a_t, t\in T_2\}$是实数列. 如果

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|a_t-a|=0,$$

那么

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|\phi(a_t)-\phi(a)|=0. \tag{2}$$

## §3. 强大数定律和Shannon-McMillan定理

**定义 6**　设$T_2$为二叉树, $\{X_t,t\in T_2\}$是如定义1所定义的在可列状态空间$G$中取值的二叉树非齐次分枝马氏链, $S_k(T^{(n)})\,(k\in G)$是随机变量集$\{X_t,t\in T^{(n)}\}$中$k$出现的次数, $S_{l,k_1,k_2}(T^{(n)})$是随机序偶集$\{(X_t,X_{t^1},X_{t^2}),t\in T^{(n-1)}\}$中序偶$(l,k_1,k_2)\,(l,k_1,k_2\in G)$出现的次数, 即

$$S_k(T^{(n)})=|\{t\in T^{(n)}:X_t=k\}|,$$
$$S_{l,k_1,k_2}(T^{(n)})=|\{t\in T^{(n-1)}:(X_t,X_{t^1},X_{t^2})=(l,k_1,k_2)\}|,$$

易知

$$S_k(T^{(n)})=\sum_{t\in T^{(n)}}I_k(X_t), \tag{3}$$
$$S_{l,k_1,k_2}(T^{(n)})=\sum_{t\in T^{(n-1)}}I_l(X_t)I_{k_1}(X_{t^1})I_{k_2}(X_{t^2}), \tag{4}$$

这里

$$I_k(i)=\begin{cases}1, & i=k;\\ 0, & i\neq k.\end{cases}$$

**定义 7**　设$\{X_t,t\in T\}$是在$G$上取值的树指标随机过程, 记

$$P(x^{T^{(n)}})=\mathsf{P}(X^{T^{(n)}}=x^{T^{(n)}}).$$

令

$$f_n(\omega)=-\frac{1}{|T^{(n)}|}\ln P(X^{T^{(n)}}),$$

$f_n(\omega)$称为$X^{T^{(n)}}$的熵密度, 若$\{X_t,t\in T_2\}$是如定义1所定义的在可列状态空间$G$中取值的二叉树非齐次分枝马氏链, 由(1)有

$$f_n(\omega)=-\frac{1}{|T^{(n)}|}\Big[\ln P(X_o)+\sum_{t\in T^{(n-1)}}\ln P_t(X_{t^1},X_{t^2}\,|\,X_t)\Big], \tag{5}$$

$f_n(\omega)$在一定意义上收敛于常数($L_1$收敛, 概率收敛, a.e.收敛), 这在信息论中称为熵密度定理或Shannon-McMillan定理或渐近等分性(AEP).

**引理 8** ([1; 定理4]) 设$G=\{0,1,2,\ldots,b-1\}$是有限状态空间, $\{X_t,t\in T_2\}$是在$G$中取值随机矩阵$\{P_t,t\in T_2\}$如定义1所定义的二叉树非齐次分枝马氏链, $S_l(T^{(n)})$如(3)所定义. 设$P=(P(y_1,y_2\,|\,x))$, $x,y_1,y_2\in G$是另一个随机矩阵, $P_1(y_1\,|\,x)=\sum\limits_{y_2\in G}P(y_1,y_2\,|\,x)$, $P_2(y_2\,|\,x)=\sum\limits_{y_1\in G}P(y_1,y_2\,|\,x)$, $P_1=(P_1(y\,|\,x))$, $P_2=(P_2(y\,|\,x))$, 令$Q=(P_1+P_2)/2$, 假设转移矩阵$Q$遍历, 如果$\forall\, x,y_1,y_2\in G$,

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}}|P_t(y_1,y_2\,|\,x)-P(y_1,y_2\,|\,x)|=0, \tag{6}$$

那么

$$\lim_{n\to\infty}\frac{S_l(T^{(n)})}{|T^{(n)}|}=\pi(l)\qquad \text{a.e.},\ \ l\in G, \tag{7}$$

其中$\pi=(\pi(0),\pi(1),\ldots,\pi(b-1))$为由$Q$确定的唯一平稳分布.

**定理 9** 在引理8的条件下, $S_{l,k_1,k_2}(T^{(n)})$如(4)所定义, 那么

$$\lim_{n\to\infty}\frac{S_{l,k_1,k_2}(T^{(n)})}{|T^{(n)}|}=\frac{1}{2}\pi(l)P(k_1,k_2\,|\,l)\qquad \text{a.e.},\ \ l,k_1,k_2\in G, \tag{8}$$

其中$\pi=(\pi(0),\pi(1),\ldots,\pi(b-1))$为由$Q$确定的唯一平稳分布.

**证明:** 在引理4中令$g_t(x,y_1,y_2)=I_l(x)I_{k_1}(y_1)I_{k_2}(y_2)$, 则

$$\begin{aligned}
H_n(\omega)&=\sum_{t\in T^{(n-1)}}g_t(X_t,X_{t^1},X_{t^2})\\
&=\sum_{t\in T^{(n-1)}}I_l(X_t)I_{k_1}(X_{t^1})I_{k_2}(X_{t^2})=S_{l,k_1,k_2}(T^{(n)}),\\
G_n(\omega)&=\sum_{t\in T^{(n-1)}}\mathsf{E}[g_t(X_t,X_{t^1},X_{t^2})\,|\,X_t]\\
&=\sum_{t\in T^{(n-1)}}\sum_{(x_{t^1},x_{t^2})\in G^2}g_t(X_t,x_{t^1},x_{t^2})P_t(x_{t^1},x_{t^2}\,|\,X_t)\\
&=\sum_{t\in T^{(n-1)}}\sum_{(x_{t^1},x_{t^2})\in G^2}I_l(X_t)I_{k_1}(x_{t^1})I_{k_2}(x_{t^2})P_t(x_{t^1},x_{t^2}\,|\,X_t)\\
&=\sum_{t\in T^{(n-1)}}I_l(X_t)\sum_{(x_{t^1},x_{t^2})\in G^2}I_{k_1}(x_{t^1})I_{k_2}(x_{t^2})P_t(x_{t^1},x_{t^2}\,|\,X_t)\\
&=\sum_{t\in T^{(n-1)}}I_l(X_t)P_t(k_1,k_2\,|\,l),
\end{aligned}$$

由引理4知

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\Big[S_{l,k_1,k_2}(T^{(n)})-\sum_{t\in T^{(n-1)}}I_l(X_t)P_t(k_1,k_2\,|\,l)\Big]=0,\qquad \text{a.e.},$$

由条件(6)易证

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|I_l(X_t)P_t(y_1,y_2\,|\,x)-I_l(X_t)P(y_1,y_2\,|\,x)|=0,$$

所以

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}|S_{l,k_1,k_2}(T^{(n)})-\sum_{t\in T^{(n-1)}}I_l(X_t)P(k_1,k_2\,|\,l)|=0,\qquad \text{a.e.},\tag{9}$$

因为

$$S_l(T^{(n)})=\sum_{t\in T^{(n)}}I_l(X_t),\qquad \lim_{n\to\infty}\frac{|T^{(n-1)}|}{|T^{(n)}|}=\frac{1}{2},$$

由式(7)知

$$\lim_{n\to\infty}\frac{\sum\limits_{t\in T^{(n-1)}}I_l(X_t)}{|T^{(n)}|}=\lim_{n\to\infty}\frac{S_l(T^{(n-1)})}{|T^{(n-1)}|}\frac{|T^{(n-1)}|}{|T^{(n)}|}=\frac{1}{2}\pi(l),\qquad l\in G,\ \ \text{a.e.},\tag{10}$$

由式(9)及(10)可得(8).　　□

**定理 10**　在引理8的条件下, $f_t(x,y_1,y_2)$是定义在$G^3$上的函数, 满足$\exists\,\alpha>0$, s.t.

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}\mathsf{E}\big[f_t^2(X_t,X_{t^1},X_{t^2})\mathrm{e}^{\alpha|f_t(X_t,X_{t^1},X_{t^2})|}|X_t\big]<\infty,$$

设$g_t(x,y_1,y_2)=f_t(x,y_1,y_2)p_t(y_1,y_2\,|\,x)$, $g$是定义在$G^3$上的另一个函数, 如果$\forall\,x,y_1,y_2\in G$,

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|g_t(x,y_1,y_2)-g(x,y_1,y_2)|=0,\tag{11}$$

且$g$有限, 则

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}f_t(X_t,X_{t^1},X_{t^2})=\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{2}\pi(l)g(l,k_1,k_2),\qquad \text{a.e.},\tag{12}$$

其中$\pi=(\pi(0),\pi(1),\ldots,\pi(b-1))$为由$Q$确定的唯一平稳分布.

**证明:**　由于

$$\begin{aligned}
H_n(\omega)&=\sum_{t\in T^{(n-1)}}f_t(X_t,X_{t^1},X_{t^2}),\\
G_n(\omega)&=\sum_{t\in T^{(n-1)}}\mathsf{E}[f_t(X_t,X_{t^1},X_{t^2})\,|\,X_t]\\
&=\sum_{t\in T^{(n-1)}}\sum_{k_1,k_2\in G^2}f_t(X_t,x_{t^1},x_{t^2})P_t(x_{t^1},x_{t^2}\,|\,X_t)\\
&=\sum_{t\in T^{(n-1)}}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}f_t(X_t,k_1,k_2)p_t(k_1,k_2\,|\,X_t)\\
&=\sum_{t\in T^{(n-1)}}\sum_{l=0}^{b-1}I_l(X_t)\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}f_t(l,k_1,k_2)p_t(k_1,k_2\,|\,l)\\
&=\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}}I_l(X_t)g_t(l,k_1,k_2)
\end{aligned}$$

$$= \sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g_t(l,k_1,k_2),$$

所以由引理4有

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\Big[\sum_{t\in T^{(n-1)}} f_t(X_t,X_{t^1},X_{t^2}) - \sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g_t(l,k_1,k_2)\Big] = 0,\quad \text{a.e.,} \tag{13}$$

又由(11)有

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\Big|\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g_t(l,k_1,k_2)\\
&-\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g(l,k_1,k_2)\Big|\\
\leqslant{}&\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}}|g_t(l,k_1,k_2)-g(l,k_1,k_2)| = 0,
\end{aligned}\tag{14}$$

因为

$$\begin{aligned}
&\Big|\frac{1}{|T^{(n)}|}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g_t(l,k_1,k_2) - \sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{2}\pi(l)g(l,k_1,k_2)\Big|\\
\leqslant{}&\Big|\frac{1}{|T^{(n)}|}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g_t(l,k_1,k_2)\\
&-\frac{1}{|T^{(n)}|}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g(l,k_1,k_2)\Big|\\
&+\Big|\frac{1}{|T^{(n)}|}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\sum_{t\in T^{(n-1)}} I_l(X_t)g(l,k_1,k_2) - \sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{2}\pi(l)g(l,k_1,k_2)\Big|\\
\leqslant{}&\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|g_t(l,k_1,k_2)-g(l,k_1,k_2)|\\
&+\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}g(l,k_1,k_2)\Big|\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}I_l(X_t)-\frac{1}{2}\pi(l)\Big|\\
={}&\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|g_t(l,k_1,k_2)-g(l,k_1,k_2)|\\
&+\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}g(l,k_1,k_2)\left|\frac{\sum\limits_{t\in T^{(n-1)}}I_l(X_t)}{|T^{(n-1)}|}\frac{|T^{(n-1)}|}{|T^{(n)}|}-\frac{1}{2}\pi(l)\right|\\
={}&\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|g_t(l,k_1,k_2)-g(l,k_1,k_2)|\\
&+\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}g(l,k_1,k_2)\Big|\frac{1}{2}\frac{S_l(T^{(n-1)})}{|T^{(n-1)}|}-\frac{1}{2}\pi(l)\Big|,
\end{aligned}\tag{15}$$

由引理8、(13)、(14)和(15)可得(12), 定理证毕. □

**推论 11** ([1; 定理5])　在引理8条件下, $f_n(\omega)$是如(5)所定义的熵密度, 那么

$$\lim_{n\to\infty} f_n(\omega) = -\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\frac{1}{2}\pi(l)P(k_1,k_2\,|\,l)\ln P(k_1,k_2\,|\,l), \qquad \text{a.e.}, \tag{16}$$

其中$\pi=(\pi(0),\pi(1),\ldots,\pi(b-1))$是由$Q$确定的唯一平稳分布.

**证明:**　设$f_t(x,y_1,y_2)=-\ln P_t(y_1,y_2\,|\,x)$, 令$\alpha=1/2$, $\forall\, t\in T_2$,

$$\begin{aligned}
&\mathsf{E}\big[f_t^2(X_t,X_{t^1},X_{t^2})\mathrm{e}^{\alpha|f_t(X_t,X_{t^1},X_{t^2})|}\,|\,X_t\big]\\
=&\sum_{(x_{t^1},x_{t^2})\in G^2}\ln^2 P_t(x_{t^1},x_{t^2}\,|\,X_t)\mathrm{e}^{-[\ln P_t(x_{t^1},x_{t^2}|X_t)]/2}P_t(x_{t^1},x_{t^2}\,|\,X_t)\\
=&\sum_{(x_{t^1},x_{t^2})\in G^2}\ln^2 P_t(x_{t^1},x_{t^2}\,|\,X_t)[P_t(x_{t^1},x_{t^2}\,|\,X_t)]^{1/2}\leqslant 16b^2\mathrm{e}^{-2},
\end{aligned}$$

所以$\forall\, t\in T_2$,

$$\mathsf{E}\big[\mathrm{e}^{|f_t(X_t,X_{t^1},X_{t^2})|/2}\,|\,X_t\big]<\infty,$$

及

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}\mathsf{E}\big[f_t^2(X_t,X_{t^1},X_{t^2})\mathrm{e}^{|f_t(X_t,X_{t^1},X_{t^2})|/2}\,|\,X_t\big]\leqslant 16b^2\mathrm{e}^{-2}.$$

令$\phi(x)=x\ln x$, $(\phi(0)=0)$, 易知$\phi(x)$是定义在$[0,1]$上的连续函数. 令

$$\begin{aligned}
g_t(l,k_1,k_2)&=-P_t(k_1,k_2\,|\,l)\ln P_t(k_1,k_2\,|\,l),\\
g(l,k_1,k_2)&=-P(k_1,k_2\,|\,l)\ln P(k_1,k_2\,|\,l).
\end{aligned}$$

由引理5和(6)知$\forall\, k_1,k_2,l\in G$,

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}|P_t(k_1,k_2\,|\,l)\ln P_t(k_1,k_2\,|\,l)-P(k_1,k_2\,|\,l)\ln P(k_1,k_2\,|\,l)|=0,$$

故$g_t(l,k_1,k_2)$满足(11). 所以由定理9, 有

$$\begin{aligned}
&\lim_{n\to\infty}-\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n-1)}}\ln P_t(X_t,X_{t^1},X_{t^2})\\
=&-\frac{1}{2}\sum_{l=0}^{b-1}\sum_{k_1=0}^{b-1}\sum_{k_2=0}^{b-1}\pi(l)P(k_1,k_2\,|\,l)\ln P(k_1,k_2\,|\,l), \qquad \text{a.e.},
\end{aligned}\tag{17}$$

由式(5)和(17)可得(16).　　□

## 参 考 文 献

[1] Dang H, Yang W G, Shi Z Y. The strong law of large numbers and the entropy ergodic theorem for nonhomogeneous bifurcating Markov chains indexed by a binary tree [J]. *IEEE Trans. Inform. Theory*, 2015, **61(4)**: 1640–1648.

[2] Berger T, Ye Z X. Entropic aspects of random fields on trees [J]. *IEEE Trans. Inform. Theory*, 1990, **36(5)**: 1006–1018.

[3] Benjamini I, Peres Y. Markov chains indexed by trees [J]. *Ann. Probab.*, 1994, **22(1)**: 219–243.

[4] Yang W G. Some limit properties for Markov chains indexed by a homogeneous tree [J]. *Statist. Probab. Lett.*, 2003, **65(3)**: 241–250.

[5] Yang W G, Ye Z X. The asymptotic equipartition property for nonhomogeneous Markov chains indexed by a homogeneous tree [J]. *IEEE Trans. Inform. Theory*, 2007, **53(9)**: 3275–3280.

[6] Huang H L, Yang W G. Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree [J]. *Sci. China Ser. A*, 2008, **51(2)**: 195–202.

[7] Dong Y, Yang W G, Bai J F. The strong law of large numbers and the Shannon-McMillan theorem for nonhomogeneous Markov chains indexed by a Cayley tree [J]. *Statist. Probab. Lett.*, 2011, **81(12)**: 1883–1890.

[8] 陈晓雪, 杨卫国, 王豹. 树指标马氏链的等价定义 [J]. *数学研究*, 2012, **45(4)**: 411–414.

[9] Guyon J. Limit theorems for bifurcating Markov chains, application to the detection of cellular aging [J]. *Ann. Appl. Probab.*, 2007, **17(5/6)**: 1538–1569.

[10] 杨卫国. 关于树指标马氏链强极限定理的若干研究 [J]. *数学进展*, 2014, **43(2)**: 206–218.

[11] 党慧, 杨卫国, 高荣, 等. 二叉树上分枝马氏链的强大数定理 [J]. *纯粹数学与应用数学*, 2013, **29(5)**: 529–535.

# The Strong Law of Large Numbers and the Shannon-McMillan Theorem for Nonhomogeneous Bifurcating Markov Chains Indexed by a Binary Tree

ZHANG Yan YANG WeiGuo

(*Faculty of Science, Jiangsu University, Zhenjiang, 212013, China*)

**Abstract:** In this paper, we first study the strong law of large numbers for the frequencies of occurrence of random ordered couples of states for nonhomogeneous bifurcating Markov chains indexed by a binary tree. Then the strong law of large numbers are studied for functions of the nonhomogeneous bifurcating Markov chains indexed by a binary tree. As a corollary, we obtain the shannon-McMillan theorem for these Markov chains with finite state space.

**Keywords:** two rooted tree; nonhomogeneous bifurcating Markov chains; strong law of large numbers; Shannon-McMillan theorem

**2010 Mathematics Subject Classification:** 60J10; 60F15