

二叉树指标随机场关于分枝马氏链的一类强偏差定理

闵 帆, 杨卫国

(江苏大学理学院, 江苏 镇江 212013)

摘要: 本文研究了二叉树指标随机场关于分枝马氏链的一类强偏差定理. 通过引入渐近对数似然比作为二叉树指标任意随机场与分枝马氏链之间偏差的一种度量, 进而构造鞅的方法, 获得了二叉树指标随机场关于分枝马氏链的一类强偏差定理, 推广得到了二叉树指标分枝马氏链的强大数定理和渐近均分性.

关键词: 二叉树; 分枝马氏链; 强偏差定理; 渐近均分性; 强大数定律

MR(2010) 主题分类号: 60J10; 60F10 中图分类号: O211.6

文献标识码: A 文章编号: 0255-7797(2019)05-0780-11

1 引言

设 T 是一个无限树, σ, τ ($\sigma \neq \tau$) 是 T 中任意两个顶点, 则存在唯一的从 σ 到 τ 的路径, $\sigma = z_1, z_2, \dots, z_m = \tau$, 其中 $z_1, z_2, \dots, z_m$ 互不相同且 z_i, z_{i+1} 为相邻两顶点, $m-1$ 称为 σ 到 τ 的距离. 为给 T 中的顶点编号, 我们选定一个顶点作为根顶点 (简称根), 并记为 o . 对于 T 中的任意两个顶点 σ, τ , 如果 τ 是处在从根顶点 o 到 σ 的唯一路径上, 则记为 $\tau \leq \sigma$. 用 $\sigma \wedge \tau$ 表示同时满足 $\sigma \wedge \tau \leq \tau$ 与 $\sigma \wedge \tau \leq \sigma$ 的离 o 最远的顶点.

本文主要研究二叉树 $T_{C,2}$ (二叉树 $T_{C,2}$ 的根点 o 与 2 个支点相连, 其他的支点与 3 个支点相连 (见图 1)). 为了方便, 将 $T_{C,2}$ 简记为 T_2 . 对于 T_2 上的任一顶点 t , 记 $|t|$ 表示根点 o 和顶点 t 之间的距离. 若 $|t| = n$, 则称 t 位于树的第 n 层. L_n 表示 T_2 的第 n 层上所有顶点的集合, $T^{(n)}$ 表示二叉树 T_2 从 0 层 (根) 到 n 层的所有顶点的子图. $|T^{(n)}|$ 记为子图 $T^{(n)}$ 所含顶点数. 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 $(\Omega, \mathcal{F})$ 为一可测空间, $\{X_t, t \in T_2\}$ 是定义在 $(\Omega, \mathcal{F})$ 上且取值于 $G = \{0, 1, \dots, b-1\}$ (b 是正整数) 的随机变量集合, 设 B 为 T_2 的子图, 记 $X^B = \{X_t, t \in B\}$, x^B 表示 X^B 的实现. 设 $\mathbf{P}$ 是可测空间 $(\Omega, \mathcal{F})$ 上的概率测度, 我们称 $\mathbf{P}$ 为树 T_2 上的随机场. 令 $\{X_t, t \in T_2\}$ 在 $\mathbf{P}$ 下的分布为 $\mathbf{P}(X^{T^{(n)}} = x^{T^{(n)}}) = \mathbf{P}(x^{T^{(n)}})$.

定义 1.1 [二叉树指标分枝马氏链]^[1] 设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbf{Q})$ 上在 $G = \{0, 1, \dots, b-1\}$ (b 是正整数) 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, $P = (P(y_1, y_2|x))$ 是定义在 $G \times G^2$ 上的一随机矩阵 (满足 $P(y_1, y_2|x) \geq 0, \forall x, y_1, y_2 \in G$ 及 $\sum_{(y_1, y_2) \in G^2} P(y_1, y_2|x) = 1, \forall x \in G$). 如果对 $n \geq 1$, 有

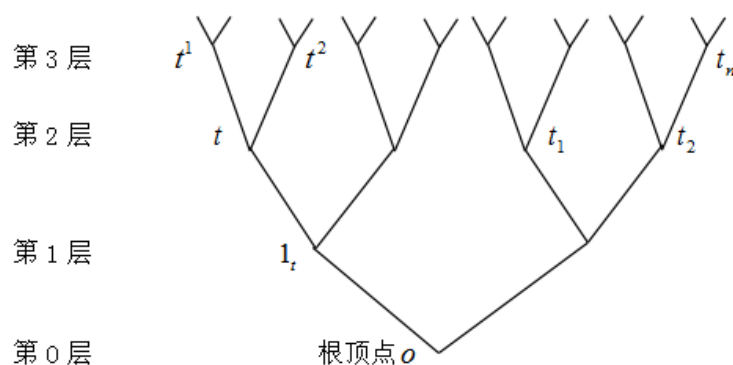
$$\mathbf{Q}(X^{L_n} = x^{L_n} | X^{T^{(n-1)}} = x^{T^{(n-1)}}) = \prod_{t \in L_{n-1}} P(x_{t^1}, x_{t^2} | x_t) \quad (1.1)$$

*收稿日期: 2018-09-29

接收日期: 2018-12-13

基金项目: 国家自然科学基金资助 (11571142).

作者简介: 闵帆 (1994-), 女, 河南项城, 硕士, 主要研究方向: 概率论极限理论.

图 1: 二叉树 $T_{C,2}$

及

$$\mathbf{Q}(X_0 = x) = p(x), \forall x \in G,$$

则称 $\{X_t, t \in T_2\}$ 为具有初始分布 p 与随机矩阵 P 并在 G 中取值的二叉树指标分枝马氏链.

注 1.2 ^[1] 若 $\{X_t, t \in T_2\}$ 在概率测度 $\mathbf{Q}$ 下为二叉树指标齐次分枝马氏链, 则

$$\mathbf{Q}(X^{T^{(n)}} = x^{T^{(n)}}) = \mathbf{Q}(x^{T^{(n)}}) = p(x_0) \prod_{t \in T^{(n-1)}} P(x_{t^1}, x_{t^2} | x_t). \quad (1.2)$$

为避免技术问题, 总假定 $p(x)$ 和 $P(y_1, y_2 | x)$ 是正的.

定义 1.3 ^[2] 设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 是定义在可测空间 $(\Omega, \mathcal{F})$ 上取值于 G 的随机变量族, $\forall x, y_1, y_2 \in G$, 设 $p(x) > 0$ 且 $P = (P(y_1, y_2 | x))$ 为正的随机矩阵, 设 $\mathbf{Q}, \mathbf{P}$ 为 $(\Omega, \mathcal{F})$ 上的两个概率测度, 其中 $\{X_t, t \in T_2\}$ 在概率测度 $\mathbf{Q}$ 下为具有初始分布 p 与随机矩阵 P 的二叉树指标分枝马氏链. 设 $\mathbf{P}(x^{T^{(n)}}) > 0$, 令

$$h(\mathbf{P} | \mathbf{Q}) = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \ln \frac{\mathbf{P}(X^{T^{(n)}})}{\mathbf{Q}(X^{T^{(n)}})} = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \ln \frac{\mathbf{P}(X^{T^{(n)}})}{p(X_0) \prod_{t \in T^{(n-1)}} P(X_{t^1}, X_{t^2} | X_t)}, \quad (1.3)$$

则称 $h(\mathbf{P} | \mathbf{Q})$ 为 $\mathbf{P}$ 关于 $\mathbf{Q}$ 的样本相对熵率或渐近对数似然比.

类似于参考文献 [2] 中引理 1 的证明可知

$$h(\mathbf{P} | \mathbf{Q}) \geq 0, \quad \mathbf{P} - \text{a.e.},$$

故 $h(\mathbf{P} | \mathbf{Q})$ 可作为 T_2 上的任意随机场与分枝马氏链之间偏差的一种度量.

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. 文献 [1] 研究了二叉树上分枝马氏链的等价性质; 文献 [3] 研究了二叉树有限状态分枝马氏链的强大数定律和 Shannon-McMillan 定理; 文献 [4] 研究了二叉树上有限状态分枝马氏链的强大数定理; 文献

[5] 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; 文献 [6] 研究了关于齐次树上随机场的一类强偏差定理; 文献 [2] 研究了关于 Cayley 树上任意随机场和马氏链场的强偏差定理; 文献 [7] 研究了齐次树指标齐次马氏链随机场的一类强偏差定理; 文献 [8] 研究了 Bethe 树指标随机场关于非齐次马氏链的一类强偏差定理.

本文通过引入渐近对数似然比作为二叉树指标任意随机场与分枝马氏链之间偏差的一种度量, 进而构造鞅的方法, 得出了二叉树指标随机场关于分枝马氏链的一类强偏差定理, 作为推论得到了二叉树指标分枝马氏链的强大数定理和渐近均分性.

2 强偏差定理

强偏差定理是由不等式表示的一类强极限定理, 是由等式表示的一类强极限定理的推广. 本节将建立二叉树指标任意随机场关于分枝马氏链的一类强偏差定理.

引理 2.1 设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 是定义在可测空间 $(\Omega, \mathcal{F})$ 上的随机变量族. $\mathbf{P}$ 与 $\mathbf{Q}$ 是 $(\Omega, \mathcal{F})$ 上的两个概率测度, 其中 $\{X_t, t \in T_2\}$ 在 $\mathbf{Q}$ 下是具有严格正的随机矩阵 $P = (P(y_1, y_2|x))$ 的二叉树指标分枝马氏链. $\mathbf{P}(X^{T(n)}) > 0$, 设 $\{g_t(x, y_1, y_2), t \in T_2\}$ 是定义在 G^3 上的函数族, $L_0 = \{0\}$, $\mathcal{F}_n = \sigma(X^{T(n)}), n \geq 1$, 令

$$t_n(\lambda, \omega) = \frac{e^{\lambda \sum_{t \in T^{(n-1)}} g_t(X_t, X_{t^1}, X_{t^2})}}{\prod_{t \in T^{(n-1)}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \cdot \frac{\mathbf{Q}(X^{T(n)})}{\mathbf{P}(X^{T(n)})}, \quad (2.1)$$

其中 λ 为实数, $E_{\mathbf{Q}}$ 表示在 $\mathbf{Q}$ 下的数学期望, 则 $\{t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ 为在概率测度 $\mathbf{P}$ 下的非负鞅.

证 由式 (1.1), (2.1) 可知

$$t_n(\lambda, \omega) = t_{n-1}(\lambda, \omega) \cdot \frac{e^{\lambda \sum_{t \in L_{n-1}} g_t(X_t, X_{t^1}, X_{t^2})}}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \cdot \frac{\prod_{t \in L_{n-1}} P(X_{t^1}, X_{t^2} | X_t)}{\mathbf{P}(X^{L_n} | X^{T(n-1)})}. \quad (2.2)$$

令

$$I_n = \frac{e^{\lambda \sum_{t \in L_{n-1}} g_t(X_t, X_{t^1}, X_{t^2})}}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \cdot \frac{\prod_{t \in L_{n-1}} P(X_{t^1}, X_{t^2} | X_t)}{\mathbf{P}(X^{L_n} | X^{T(n-1)})}, \quad (2.3)$$

则

$$\begin{aligned} & E_{\mathbf{P}}(I_n | \mathcal{F}_{n-1}) \\ &= \sum_{x^{L_n}} \frac{e^{\lambda \sum_{t \in L_{n-1}} g_t(X_t, x_{t^1}, x_{t^2})}}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \cdot \frac{\prod_{t \in L_{n-1}} P(x_{t^1}, x_{t^2} | X_t)}{\mathbf{P}(X^{L_n} = x^{L_n} | X^{T(n-1)})} \cdot \mathbf{P}(X^{L_n} = x^{L_n} | X^{T(n-1)}) \\ &= \sum_{x^{L_n}} \frac{\prod_{t \in L_{n-1}} e^{\lambda g_t(X_t, x_{t^1}, x_{t^2})} P(x_{t^1}, x_{t^2} | X_t)}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \end{aligned}$$

$$\begin{aligned}
&= \frac{\prod_{t \in L_{n-1}} \sum_{(x_{t^1}, x_{t^2}) \in G^2} e^{\lambda g_t(X_t, x_{t^1}, x_{t^2})} \cdot P(x_{t^1}, x_{t^2} | X_t)}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} \\
&= \frac{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]}{\prod_{t \in L_{n-1}} E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})} | X_t]} = 1, \quad \mathbf{P} - \text{a.e.}
\end{aligned} \quad (2.4)$$

由式 (2.2), (2.4) 可知

$$E_{\mathbf{P}}(t_n(\lambda, \omega) | \mathcal{F}_{n-1}) = t_{n-1}(\lambda, \omega), \quad \mathbf{P} - \text{a.e.}$$

又易知 $E_{\mathbf{P}}[t_1(\lambda, \omega)] = 1$, 因此

$$E_{\mathbf{P}}[t_n(\lambda, \omega)] = E_{\mathbf{P}}[E_{\mathbf{P}}(t_n(\lambda, \omega) | \mathcal{F}_{n-1})] = E_{\mathbf{P}}[t_{n-1}(\lambda, \omega)] = \cdots = E_{\mathbf{P}}[t_1(\lambda, \omega)] = 1,$$

故 $\{t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ 为在概率测度 $\mathbf{P}$ 下的非负鞅.

引理 2.2 设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 为二叉树指标分枝马氏链, 设 $\{g_t(x, y_1, y_2), t \in T_2\}$ 是定义在 G^3 上的函数族, $\{a_n, n \geq 1\}$ 为正随机变量序列. 设 $\alpha > 0$,

$$D(\alpha) = \left\{ \lim_n a_n = \infty, \limsup_n \frac{1}{a_n} \sum_{t \in T^{(n-1)}} E[e^{\alpha |g_t(X_t, X_{t^1}, X_{t^2})|} | X_t] < \infty \right\},$$

则

$$\lim_n \frac{1}{a_n} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} = 0 \quad \text{a.e.,} \quad \omega \in D(\alpha).$$

定理 2.3 设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 为定义在可测空间 $(\Omega, \mathcal{F})$ 上且取值于 $G = \{0, 1, \dots, b-1\}$ (b 是正整数) 的随机变量族, $\mathbf{P}$ 与 $\mathbf{Q}$ 是定义在 $\mathcal{F}$ 上的两个概率测度, 并且 $\{X_t, t \in T_2\}$ 是在概率测度 $\mathbf{Q}$ 下的二叉树指标分枝马氏链, 也即式 (1.2) 成立. 令 $h(\mathbf{P} | \mathbf{Q})$ 如 (1.3) 式定义, 且 $\{g_t(x, y_1, y_2), t \in T_2\}$ 是定义在 G^3 上的函数族. 设 $c \geq 0$ 是一个常数, 令

$$D(c) = \{\omega : h(\mathbf{P} | \mathbf{Q}) \leq c\}. \quad (2.5)$$

假设存在 $\alpha > 0, \forall i \in G$, 有

$$b_\alpha(i) = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[e^{\alpha |g_t(X_t, X_{t^1}, X_{t^2})|} | X_t = i] \leq \tau.$$

令

$$H_\beta^\alpha = \frac{2\tau}{e^2(\beta - \alpha)^2}, \quad (2.6)$$

其中 $0 < \beta < \alpha$. 则当 $0 \leq c \leq \beta^2 H_\beta^\alpha$ 时, 有

$$\begin{aligned}
&\limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \left| \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} \right| \\
&\leq 2\sqrt{cH_\beta^\alpha}, \quad \mathbf{P} - \text{a.e.,} \quad \omega \in D(c).
\end{aligned} \quad (2.7)$$

特别地, 有

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \quad (2.8)$$

证 令 $t_n(\lambda, \omega)$ 如 (2.1) 式定义. 由引理 2.1 知, $\{t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ 是非负 $\mathbf{P}$ -鞅. 则由 Doob 鞅收敛定理可知

$$\lim_{n \rightarrow \infty} t_n(\lambda, \omega) = t(\lambda, \omega) < \infty, \quad \mathbf{P} - \text{a.e.} \quad (2.9)$$

因此有

$$\limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \ln t_n(\lambda, \omega) \leq 0, \quad \mathbf{P} - \text{a.e.}$$

由式 (2.1), (2.9) 可知

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \ln t_n(\lambda, \omega) \\ &= \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \left\{ \sum_{t \in T^{(n-1)}} [\lambda g_t(X_t, X_{t^1}, X_{t^2}) - \ln E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})}|X_t] + \ln \frac{\mathbf{Q}(x^{T^{(n)}})}{\mathbf{P}(x^{T^{(n)}}) \right\} \\ &\leq 0, \quad \mathbf{P} - \text{a.e.} \end{aligned} \quad (2.10)$$

由式 (1.3) 及 (2.10) 可知

$$\limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{\lambda g_t(X_t, X_{t^1}, X_{t^2}) - \ln E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})}|X_t]\} \leq h(\mathbf{P}|\mathbf{Q}), \quad \mathbf{P} - \text{a.e.} \quad (2.11)$$

由式 (2.11) 可知

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} \\ &= \limsup_{n \rightarrow \infty} \frac{\lambda}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - \frac{1}{\lambda} \ln E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})}|X_t]\} \\ & \quad + \limsup_{n \rightarrow \infty} \frac{\lambda}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \left\{ \frac{1}{\lambda} \ln E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})}|X_t] - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t] \right\} \quad (2.12) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{\ln E_{\mathbf{Q}}[e^{\lambda g_t(X_t, X_{t^1}, X_{t^2})}|X_t] - E_{\mathbf{Q}}[\lambda g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} \\ & \quad + h(\mathbf{P}|\mathbf{Q}), \quad \mathbf{P} - \text{a.e.} \end{aligned}$$

令 $|\lambda| \leq \beta$. 由不等式 $\ln x \leq x - 1$ ($x > 0$) 和 $e^x - 1 - x \leq \frac{x^2}{2}e^{|x|}$, 并注意到

$$\max\{x^2 e^{-hx}, x \geq 0\} = \frac{4}{e^2 h^2} \quad (h > 0),$$

则对于所有的 $\omega \in \Omega$, 有

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{\ln E_{\mathbf{Q}}[e^{\lambda g_t} | X_t] - E_{\mathbf{Q}}[\lambda g_t | X_t]\} \\
 & \leq \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{E_{\mathbf{Q}}[e^{\lambda g_t} | X_t] - 1 - E_{\mathbf{Q}}[\lambda g_t | X_t]\} \\
 & = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[e^{\lambda g_t} - 1 - \lambda g_t | X_t] \\
 & \leq \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[\frac{1}{2} \lambda^2 g_t^2 e^{|\lambda g_t|} | X_t] \\
 & = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[\frac{1}{2} \lambda^2 g_t^2 e^{\alpha |g_t|} e^{(|\lambda| - \alpha) |g_t|} | X_t] \\
 & \leq \frac{\lambda^2}{2} \limsup_{n \rightarrow \infty} \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[\frac{4}{e^2 (|\lambda| - \alpha)^2} e^{\alpha |g_t|} | X_t] \\
 & \leq \frac{2\lambda^2 \tau}{e^2 (\beta - \alpha)^2}.
 \end{aligned} \tag{2.13}$$

其中为了方便, 将用 g_t 代替 $g_t(X_t, X_{t^1}, X_{t^2})$, 由式 (2.6), (2.12) 和 (2.13) 知

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} \\
 & \leq \lambda^2 H_{\beta}^{\alpha} + h(\mathbf{P} | \mathbf{Q}), \quad \mathbf{P} - \text{a.e.}
 \end{aligned} \tag{2.14}$$

由式 (2.5) 与 (2.14) 可知

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} \\
 & \leq \lambda^2 H_{\beta}^{\alpha} + c, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(c).
 \end{aligned} \tag{2.15}$$

当 $0 < \lambda \leq \beta < \alpha$ 时, 将式 (2.15) 左右两边同时除以 λ 可知

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} \\
 & \leq \lambda H_{\beta}^{\alpha} + \frac{c}{\lambda}, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(c).
 \end{aligned} \tag{2.16}$$

容易看出, 当 $0 < c < \beta^2 H_{\beta}^{\alpha}$ 时, 函数 $f(\lambda) = \lambda H_{\beta}^{\alpha} + \frac{c}{\lambda}$ 在 $\lambda = \sqrt{\frac{c}{H_{\beta}^{\alpha}}}$ 时获得最小值 $f(\sqrt{\frac{c}{H_{\beta}^{\alpha}}}) = 2\sqrt{c H_{\beta}^{\alpha}}$. 令式 (2.16) 中 $\lambda = \sqrt{\frac{c}{H_{\beta}^{\alpha}}}$, 可知

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2}) | X_t]\} \\
 & \leq 2\sqrt{c H_{\beta}^{\alpha}}, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(c).
 \end{aligned} \tag{2.17}$$

当 $c = 0$, 由式 (2.16) 可知

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} \\ \leq \lambda H_{\beta}^{\alpha}, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \end{aligned} \quad (2.18)$$

令式 (2.18) 中 $\lambda \rightarrow 0^+$, 可知

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} \\ \leq 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \end{aligned} \quad (2.19)$$

因此当 $c = 0$ 时, 公式 (2.17) 成立. 当 $-\alpha < -\beta \leq \lambda < 0$ 时, 由式 (2.15) 类似可知

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \{g_t(X_t, X_{t^1}, X_{t^2}) - E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t]\} \\ \geq -2\sqrt{cH_{\beta}^{\alpha}}, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(c). \end{aligned} \quad (2.20)$$

由式 (2.17), (2.20) 可知式 (2.7), 继而可知式 (2.8). 故此定理 2.3 得证.

3 强大数定理及渐近均分性 (AEP)

本节研究二叉树指标随机场关于分枝马氏链的强大数定理与渐近均分性 (AEP), 作为推论得到了二叉树指标分枝马氏链的强大数定理和渐近均分性.

设 T_2 为二叉树, $\{X_t, t \in T_2\}$ 是定义在可测空间 $(\Omega, \mathcal{F})$ 上在 G 中取值的随机变量族, 可得结论如下.

定义 3.1 [3] 设 $\mathbf{P}$ 是 $(\Omega, \mathcal{F})$ 上的一个概率测度, $x^{T^{(n)}}$ 为 $X^{T^{(n)}}$ 的实现. 记 $\{X_t, t \in T_2\}$ 在概率测度 $\mathbf{P}$ 下的分布为 $\mathbf{P}(X^{T^{(n)}} = x^{T^{(n)}}) = \mathbf{P}(x^{T^{(n)}}) > 0$. 设

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \ln \mathbf{P}(X^{T^{(n)}}), \quad (3.1)$$

则称 $f_n(\omega)$ 为 $X^{T^{(n)}}$ 在概率测度 $\mathbf{P}$ 下的熵密度. 如果 $\{X_t, t \in T_2\}$ 在概率测度 $\mathbf{Q}$ 下是具有初始分布 $p(x)$ 和随机矩阵 P 的二叉树指标分枝马氏链, 则

$$g_n(\omega) = -\frac{1}{|T^{(n)}|} \ln \mathbf{Q}(X^{T^{(n)}}) = -\frac{1}{|T^{(n)}|} [\ln p(X_0) + \sum_{t \in T^{(n-1)}} \ln P(X_{t^1}, X_{t^2}|X_t)]. \quad (3.2)$$

熵密度 $f_n(\omega)$ 收敛 (L_1 收敛, 依概率收敛, a.e. 收敛) 到一个常数的性质在信息论中被称为渐近均分性 (AEP) 或者是 Shannon-McMillan 定理.

定义 3.2 [3] 设 $S_k(T^{(n)})(k \in G)$ 是随机变量集 $\{X_t, t \in T^{(n)}\}$ 中 k 出现的次数, $S_n(k, l_1, l_2) = S_{k, l_1, l_2}(T^{(n)})$ 是随机序偶集 $\{(X_t, X_{t^1}, X_{t^2}), t \in T^{(n-1)}\}$ 中序偶 (k, l_1, l_2) ($k, l_1, l_2 \in G$) 出现的次数, 即

$$S_k^1(T^{(n)}) = |\{t \in T^{(n-1)}, X_{t^1} = k\}| = \sum_{t \in T^{(n-1)}} I_k(X_{t^1}),$$

$$\begin{aligned}
S_k^2(T^{(n)}) &= |\{t \in T^{(n-1)}, X_{t^2} = k\}| = \sum_{t \in T^{(n-1)}} I_k(X_{t^2}), \\
S_k(T^{(n)}) &= |\{t \in T^{(n)}, X_t = k\}| = \sum_{t \in T^{(n)}} I_k(X_t), \\
S_n(k, l_1, l_2) &= |\{t \in T^{(n-1)}, (X_t, X_{t^1}, X_{t^2}) = (k, l_1, l_2)\}|.
\end{aligned}$$

显然

$$\begin{aligned}
S_k(T^{(n)}) &= S_k^1(T^{(n)}) + S_k^2(T^{(n)}) + I_k(X_0), \\
S_n(k, l_1, l_2) &= \sum_{t \in T^{(n-1)}} I_k(X_t) I_{l_1}(X_{t^1}) I_{l_2}(X_{t^2}),
\end{aligned}$$

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

定理 3.3 设 $\mathbf{P}$ 与 $\mathbf{Q}$ 是定义在 $\mathcal{F}$ 上的两个概率测度, 并且 $\{X_t, t \in T_2\}$ 是由定义 1.1 定义的二叉树指标分枝马氏链 $S_k^1(T^{(n)})$, $S_k^2(T^{(n)})$, $S_k(T^{(n)})$ 如定义 3.2 所定义, 令

$$P_1(y_1|x) = \sum_{y_2} P(y_1, y_2|x), \quad P_2(y_2|x) = \sum_{y_1} P(y_1, y_2|x).$$

设转移矩阵 $P_1 = (P_1(y_1|x)), P_2 = (P_2(y_2|x))$, 令转移矩阵 $R = \frac{1}{2}(P_1 + P_2)$, 则

$$\lim_{n \rightarrow \infty} \frac{S_k(T^{(n)})}{|T^{(n)}|} = \pi(k), \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0), \quad (3.3)$$

其中 $(\pi(0), \dots, \pi(b-1))$ 为由矩阵 R 决定的平稳分布.

证 首先转移矩阵 $R = \frac{1}{2}(P_1 + P_2)$ 是遍历的, 这是因为 $P(y_1, y_2|x) > 0$, 可知 $P_1(y_1|x) > 0$, $P_2(y_2|x) > 0$, 所以 $\frac{1}{2}(P_1(y_1|x) + P_2(y_2|x)) > 0$. 即 R 是遍历的.

令 $g_t(x, y_1, y_2) = I_k(y_1)$, 则

$$\begin{aligned}
\sum_{t \in T^{(n-1)}} g_t(X_t, X_{t^1}, X_{t^2}) &= \sum_{t \in T^{(n-1)}} I_k(X_{t^1}) = S_k^1(T^{(n)}), \\
\sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t] &= \sum_{t \in T^{(n-1)}} \sum_{(x_{t^1}, x_{t^2}) \in G^2} g_t(X_t, x_{t^1}, x_{t^2}) P(x_{t^1}, x_{t^2}|X_t) \\
&= \sum_{t \in T^{(n-1)}} \sum_{(x_{t^1}, x_{t^2}) \in G^2} I_k(x_{t^1}) P(x_{t^1}, x_{t^2}|X_t) = \sum_{t \in T^{(n-1)}} P_1(k|X_t) = \sum_{t \in T^{(n-1)}} \sum_{l=0}^{b-1} I_l(X_t) P_1(k|l) \\
&= \sum_{l=0}^{b-1} S_l(T^{(n-1)}) P_1(k|l),
\end{aligned}$$

故由式 (2.8) 可知

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} [S_k^1(T^{(n)}) - \sum_{l=0}^{b-1} S_l(T^{(n-1)}) P_1(k|l)] = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \quad (3.4)$$

同理令 $g_t(x, y_1, y_2) = I_k(y_2)$. 类似可知

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} [S_k^2(T^{(n)}) - \sum_{l=0}^{b-1} S_l(T^{(n-1)})P_2(k|l)] = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \quad (3.5)$$

由式 (3.5) 与式 (3.6) 相加, 结合 $\lim_{n \rightarrow \infty} \frac{|T^{(n)}|}{|T^{(n-1)}|} = 2$, $R = \frac{1}{2}(P_1 + P_2)$, 可知

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_k(T^{(n)})}{|T^{(n)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} R(k|l) \right\} = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \quad (3.6)$$

用 $R(m|k)$ 乘以式 (3.7), 将对 $k \in \{0, 1, \dots, b-1\}$ 求和, 并再次利用式 (3.7) 可知

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left\{ \sum_{k=0}^{b-1} \frac{S_k(T^{(n)})R(m|k)}{|T^{(n)}|} - \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} + \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} R^{(2)}(m|l) \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} R^{(2)}(m|l) \right\} = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \end{aligned} \quad (3.7)$$

由归纳法可知

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_m(T^{(n+N)})}{|T^{(n+N)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} R^{(N+1)}(m|l) \right\} = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0). \quad (3.8)$$

注意到

$$\sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} = \frac{1}{|T^{(n-1)}|} \sum_{l=0}^{b-1} S_l(T^{(n-1)}) = 1,$$

且 $\lim_{N \rightarrow \infty} R^{(N)}(k|l) = \pi(k)$, $k \in G$, 则式 (3.4) 得证.

定理 3.4 在定理 3.3 的条件下, 令 $S_n(k, l_1, l_2) = S_{k, l_1, l_2}(T^{(n)})$ 如定义 3.2 所述, 则

$$\lim_{n \rightarrow \infty} \frac{S_n(k, l_1, l_2)}{|T^{(n)}|} = \frac{1}{2} \pi(k) P(l_1, l_2|k), \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0), \quad (3.9)$$

其中 $(\pi(0), \dots, \pi(b-1))$ 为由矩阵 R 决定的平稳分布.

证 令 $g_t(x, y_1, y_2) = I_k(x)I_{l_1}(y_1)I_{l_2}(y_2)$, 则

$$\begin{aligned} \sum_{t \in T^{(n-1)}} g_t(X_t, X_{t^1}, X_{t^2}) &= \sum_{t \in T^{(n-1)}} I_k(X_t)I_{l_1}(X_{t^1})I_{l_2}(X_{t^2}) = S_n(k, l_1, l_2), \\ \sum_{t \in T^{(n-1)}} E_{\mathbf{Q}}[g_t(X_t, X_{t^1}, X_{t^2})|X_t] &= \sum_{t \in T^{(n-1)}} \sum_{(x_{t^1}, x_{t^2}) \in G^2} g_t(X_t, x_{t^1}, x_{t^2})P(x_{t^1}, x_{t^2}|X_t) \\ &= \sum_{t \in T^{(n-1)}} \sum_{(x_{t^1}, x_{t^2}) \in G^2} I_k(X_t)I_{l_1}(x_{t^1})I_{l_2}(x_{t^2})P(x_{t^1}, x_{t^2}|X_t) \\ &= \sum_{t \in T^{(n-1)}} I_k(X_t) \sum_{(x_{t^1}, x_{t^2}) \in G^2} I_{l_1}(x_{t^1})I_{l_2}(x_{t^2})P(x_{t^1}, x_{t^2}|X_t) \\ &= \sum_{t \in T^{(n-1)}} I_k(X_t)P(l_1, l_2|k) = S_{n-1}(k)P(l_1, l_2|k). \end{aligned} \quad (3.10)$$

由式 (2.8) 可知

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} [S_n(k, l_1, l_2) - S_{n-1}(k)P(l_1, l_2|k)] = 0, \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0).$$

因为 $\lim_{n \rightarrow \infty} \frac{|T^{(n-1)}|}{|T^{(n)}|} = \frac{1}{2}$, 所以

$$\lim_{n \rightarrow \infty} \frac{S_{n-1}(k)}{|T^{(n)}|} = \lim_{n \rightarrow \infty} \frac{S_{n-1}(k)}{|T^{(n-1)}|} \frac{|T^{(n-1)}|}{|T^{(n)}|} = \frac{1}{2} \pi(k),$$

因此式 (3.10) 得证.

定理 3.5 在定理 3.3 的条件下, $f_n(\omega)$ 是如 (3.1) 所定义的熵密度, 则

$$\lim_{n \rightarrow \infty} f_n(\omega) = - \sum_{k \in G} \sum_{l_1 \in G} \sum_{l_2 \in G} \frac{1}{2} \pi(k) P(l_1, l_2|k) \ln P(l_1, l_2|k), \quad \mathbf{P} - \text{a.e.}, \quad \omega \in D(0), \quad (3.11)$$

其中 $(\pi(0), \dots, \pi(b-1))$ 为由矩阵 R 决定的平稳分布.

证 因为

$$- \frac{1}{|T^{(n)}|} \sum_{t \in T^{(n-1)}} \ln P(X_{t^1}, X_{t^2}|X_t) = - \lim_{n \rightarrow \infty} \sum_{k \in G} \sum_{l_1 \in G} \sum_{l_2 \in G} \ln P(l_1, l_2|k) \frac{S_n(k, l_1, l_2)}{|T^{(n)}|}, \quad (3.12)$$

由式 (3.2) 与式 (3.10)、(3.13) 可知

$$\lim_{n \rightarrow \infty} g_n(\omega) = - \sum_{k \in G} \sum_{l_1 \in G} \sum_{l_2 \in G} \frac{1}{2} \pi(k) P(l_1, l_2|k) \ln P(l_1, l_2|k), \quad \mathbf{P} - \text{a.e.} \quad \omega \in D(0). \quad (3.13)$$

由式 (1.3), (3.1), (3.2) 可知

$$\lim_{n \rightarrow \infty} g_n(\omega) = \lim_{n \rightarrow \infty} f_n(\omega), \quad \mathbf{P} - \text{a.e.} \quad \omega \in D(0). \quad (3.14)$$

故综合式 (3.14), (3.15) 可知式 (3.12) 得证.

定理 3.6 设 $\{X_t, t \in T_2\}$ 在概率测度 $\mathbf{Q}$ 下是具有严格正的随机矩阵 $P = (P(y_1, y_2|x))$ 的二叉树指标分枝马氏链, $\{X_t, t \in T_2\}, S_n(k), S_n(k, l_1, l_2), f_n(\omega)$ 如定义 3.2 所述, 令

$$P_1(y_1|x) = \sum_{y_2} P(y_1, y_2|x), P_2(y_2|x) = \sum_{y_1} P(y_1, y_2|x).$$

设转移矩阵 $P_1 = (P_1(y_1|x))P_2 = (P_2(y_2|x))$, 假设转移矩阵 $R = \frac{1}{2}(P_1 + P_2)$, 则

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{S_n(k)}{|T^{(n)}|} &= \pi(k), \quad \mathbf{Q} - \text{a.e.}, \\ \lim_{n \rightarrow \infty} \frac{S_n(k, l_1, l_2)}{|T^{(n)}|} &= \frac{1}{2} \pi(k) P(l_1, l_2|k), \quad \mathbf{Q} - \text{a.e.}, \\ \lim_{n \rightarrow \infty} f_n(\omega) &= - \sum_{k \in G} \sum_{l_1 \in G} \sum_{l_2 \in G} \frac{1}{2} \pi(k) P(l_1, l_2|k) \ln P(l_1, l_2|k), \quad \mathbf{Q} - \text{a.e.}, \end{aligned} \quad (3.15)$$

其中 $(\pi(0), \dots, \pi(b-1))$ 为由矩阵 R 决定的平稳分布.

证 令定理 3.3 中 $\mathbf{P} \equiv \mathbf{Q}$, 则 $g_n(\omega) = f_n(\omega)$, 进而知 $D(0) = \Omega$. 因此由式 (3.4), (3.10), (3.12) 可知式 (3.16), (3.17), (3.18) 得证.

参 考 文 献

- [1] 党慧, 杨卫国. 二叉树上分支马氏链的等价性质 [J]. 应用概率统计, 2014, 30(5): 491–496.
- [2] Liu W, Wang L Y. The Markov approximation of the random field on Cayley tree and a class of small deviation theorems[J]. Stati. Prob. Lett., 2003, 63: 113–121.
- [3] 张艳. 二叉树分枝马氏链的强大数定律和 Shannon-McMillan 定理 [J]. 应用概率统计, 2017, 33(4): 408–416.
- [4] 党慧, 杨卫国, 高荣, 石志岩. 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.
- [5] 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. Inf. Theory, 2015, 61(4): 1640–1648.
- [6] 彭维才, 杨卫国. 关于齐次树上随机场的一类强偏差定理 [J]. 系统科学与数学, 2009, 29(5): 577–584.
- [7] Peng W C, Yang W G, Wang B. A class of small deviation theorems for functionals of random fields on a homogeneous tree[J]. Math. Anal. Appl., 2010, 361: 293–301.
- [8] Yang W G. A class of strong deviation theorems for the random fields associated with nonhomogeneous Markov chains indexed by a Bethe tree[J]. Stoch. Anal. Appl., 2012, 30(2): 220–237.

A CLASS OF STRONG DEVIATION THEOREMS FOR THE RANDOM FIELDS ASSOCIATED WITH BIFURCATING MARKOV CHAINS INDEXED BY A BINARY TREE

MIN Fan, YANG Wei-Guo

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

Abstract: In this paper, we study a class of strong deviation theorems for the random fields which are associated with bifurcating Markov chains indexed by a binary tree. By introducing the asymptotic logarithmic likelihood ratio as a measure of the deviation between the arbitrary random fields and the bifurcating Markov chains on a binary tree, and constructing the martingale, a class of strong deviation theorems for the random fields which are associated with bifurcating Markov chains indexed by a binary tree is established. As corollaries, we obtain the strong law of large numbers and the AEP for the bifurcating Markov chains indexed by a binary tree.

Keywords: binary tree; bifurcating Markov chains; strong deviation theorem; asymptotic equipartition property; strong law of large numbers

2010 MR Subject Classification: 60J10; 60F10