第27卷第4期 数 学 研 究 与 评 论 Vol.27, No.4
2007年11月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Nov., 2007

文章编号: 1000-341X(2007)04-0719-09 **文献标识码**: **A**

任意信源的一类 Shannon-Mcmillan 逼近定理

王康康 [1], 杨卫国 [2]
(1. 江苏科技大学数理学院, 江苏 镇江 212003; 2. 江苏大学理学院, 江苏 镇江 212013)
(E-mail: wkk.cn@126.com)

摘 要: 该文通过概率空间上的任意分布与另一任意分布相比较, 研究任意随机序列相对熵密度的小偏差定理. 并由此得出若干任意信源, m 阶马氏信源的 Shannon-Mcmillan 定理. 将已有的关于离散信源的结果加以推广.

关键词: Shannon-Mcmillan 定理; 小偏差定理; 任意信源; 相对熵密度; m 阶马氏信源.
MSC(2000): 60F15
中图分类: O211.6

1 引 言

设 (Ω, F, P) 为一概率空间, 设 $\{X_n, n \geq 0\}$ 是定义在该概率空间上并于字母集 $S = \{s_1, s_2, \ldots\}$ 上取值的任意信源, 其联合分布为

$$P(X_0 = x_0, \ldots, X_n = x_n) = p(x_0, \ldots, x_n) > 0, \quad x_i \in S, 0 \leq i \leq n. \tag{1}$$

其中令

$$f_n(\omega) = -\frac{1}{n+1} \log p(X_0, \ldots, X_n), \tag{2}$$

$f_n(\omega)$ 称为 $\{X_i, 0 \leq i \leq n\}$ 的相对熵密度.

关于 $f_n(\omega)$ 的极限性质是信息论中的重要问题, 在信息论中称为 Shannon-Mcmillan 定理或信源的渐近均匀分割性 (简称 S-M 定理), 它是信息论中编码的基础. Shannon 在其著名的论文 [5] 中首先证明了齐次遍历马氏信源的 S-M 定理. Mcmillan[6] 和 Breiman[7] 则证明了平稳遍历信源的 S-M 定理. 钟开莱 [8] 考虑了字母集为可列集的情况. 以后又有许多作者将上述结果推广到一般的随机过程. 在文献 [1] 中, 刘文通过任意信源与无记忆信源比较, 用特有的分析方法给出了一类任意离散信源相对熵密度用不等式给出的强极限定理, 也称为小偏差定理. 信源相对熵密度小偏差定理是 Shannon-Mcmillan 定理的一种推广, 为 Shannon-Mcmillan 定理的研究开辟了一个全新的研究方向.

本文的目的是将文献 [1] 的结果推广到任意信源的情况. 即通过概率测度下的任意随机变量序列之间相比较, 得到了由 (2) 式定义的相对熵密度的的一类强偏差定理. 将文献 [1] 中所用的方法加以改进, 并作为推论得到了文献 [2] 和 [3] 的主要结果.

如果 Q 为 F 的另一概率测度, 其联合分布为

$$Q(X_0 = x_0, \ldots, X_n = x_n) = q(x_0, \ldots, x_n) > 0, \quad x_i \in S, \ 0 \leq i \leq n. \tag{3}$$

收稿日期: 2005-04-11; **接受日期**: 2006-08-15
基金项目: 国家自然科学基金 (10571076); 江苏省教育厅自然科学基金 (02KJD110003).

定义 设 p 和 q 如 (1) 和 (3) 式所示，令

$$h(P|Q)=\limsup_{n\to\infty}(1/n)\log[p(X_0,\dots,X_n)/q(X_0,\dots,X_n)]. \tag{4}$$

$h(P|Q)$ 称为 P 相对于 Q 的样本散度.

实际上 $h(P|Q)$ 是 $\{X_i,0\le i\le n\}$ 在概率测度 Q 与 P 下的熵密度偏差的上极限.

引理[4] 设 p 和 q 如 (1) 和 (3) 式所示，则有

$$\limsup_{n\to\infty}(1/n)\log[q(X_0,\dots,X_n)/p(X_0,\dots,X_n)]\le 0 \quad P\text{-a.s.} \tag{5}$$

由 (5) 式可推出

$$h(P|Q)\ge\liminf_{n\to\infty}(1/n)\log[p(X_0,\dots,X_n)/q(X_0,\dots,X_n)]\ge 0 \quad P\text{-a.s.} \tag{6}$$

又记

$$q(x_n|x_1,\dots,x_{n-1})=Q(X_n=x_n|X_0=x_0,\dots,X_{n-1}=x_{n-1}) \qquad n\ge 1, \tag{7}$$

则由 (3) 式有

$$q(X_0,\dots,X_n)=q(X_0)\prod_{k=1}^{n}q_k(X_k|X_0,\dots,X_{k-1}). \tag{8}$$

我们记 $X^n=\{X_0,\dots,X_n\}$, $X_m^n=\{X_m,\dots,X_n\}$. x^n, x_m^n 分别为为 X^n 和 X_m^n 的实现.

2 主要结果

定理 1 设 $\{X_n,n\ge 0\}$ 是具有分布 (1) 式的任意信源， $f_n(\omega)$ 和 $h(P|Q)$ 分别由 (2) 和 (4) 式定义，设 $\alpha\ge 0$. $H_k^Q(\omega)$ 为 X_k 在测度 Q 下关于 $X_0,\dots,X_{k-1}$ 的随机条件熵. 即

$$H_k^Q(\omega)=-\sum_{x_k\in S}q_k(x_k|X_0,\dots,X_{k-1})\log q_k(x_k|X_0,\dots,X_{k-1}). \tag{9}$$

令

$$D(c)=\{\omega:h(P|Q)\le c\}, \tag{10}$$

$$b_\alpha=\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}E_Q[(\log(q_k(X_k|X^{k-1}))^2q_k(X_k|X^{k-1})^{-\alpha}|X^{k-1}]<\infty. \tag{11}$$

则，当 $0\le c\le\alpha^2b_\alpha/2$ 时，有

$$\limsup_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\le\sqrt{2cb_\alpha} \quad P\text{-a.s.} \quad D(c), \tag{12}$$

$$\liminf_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\ge-\sqrt{2cb_\alpha}-c \quad P\text{-a.s.} \quad D(c). \tag{13}$$

特别有

$$\lim_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]=0 \quad P\text{-a.s.} \quad D(0). \tag{14}$$

其中 E_Q 表示相对于 Q 的数学期望.

证明 我们取 (Ω, F, P) 为所考虑的概率空间. 设 λ 为任意常数. 令

$$Q_k(\lambda) = E_Q[q_k(X_k|X^{k-1})^{-\lambda}|X^{k-1}] = \sum_{x_k \in S} q_k(x_k|x^{k-1})^{1-\lambda}, \tag{15}$$

$$q_k(\lambda, x_k) = \frac{q_k(x_k|x^{k-1})^{1-\lambda}}{Q_k(\lambda)}, \quad x_k \in S, \tag{16}$$

$$g(\lambda, x_1, \ldots, x_n) = \prod_{k=1}^{n} q_k(\lambda, x_k). \tag{17}$$

则 $g(\lambda, x_1, \ldots, x_n)$, $n = 1, 2, \ldots$ 是 S^n 上的一族相容分布. 令

$$T_n(\lambda, \omega) = \frac{g(\lambda, X_1, \ldots, X_n)}{p(X_0, \ldots, X_n)}. \tag{18}$$

由于 $\{T_n(\lambda, \omega), n \geq 1\}$ 是 a.s. 收敛的非负上鞅 [4,p354]. 故由 Doob 鞅收敛定理 [9] 有

$$\lim_{n\to\infty} T_n(\lambda, \omega) = T_\infty(\lambda, \omega) < \infty \quad P\text{-a.s.} \tag{19}$$

因而有

$$\limsup_{n\to\infty} (1/n) \log T_n(\lambda, \omega) \leq 0 \qquad P\text{-a.s.} \tag{20}$$

由 (15)–(18) 式有

$$\begin{aligned}(1/n)\log T_n(\lambda,\omega) =& (1/n)\sum_{k=1}^{n}[(-\lambda \log q_k(X_k|X^{k-1})) - \log E_Q(q_k(X_k|X^{k-1})^{-\lambda}|X^{k-1}) + \\ & \log(\prod_{k=1}^{n} q_k(X_k|X^{k-1}) \Big/ p(X_0, \ldots, X_n))] \qquad P\text{-a.s.}\end{aligned} \tag{21}$$

由 (8), (10), (20) 及 (21) 式有

$$\begin{aligned}&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\lambda\log q_k(X_k|X^{k-1})) - \log E_Q(q_k(X_k|X^{k-1})^{-\lambda}|X^{k-1})] \leq c \\ &\qquad P\text{-a.s.}\omega \in D(c).\end{aligned} \tag{22}$$

由 (22) 式有

$$\begin{aligned}&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\lambda\log q_k(X_k|X^{k-1})) - E_Q(-\lambda\log q_k(X_k|X^{k-1})|X^{k-1})] \\ &\leq \limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[\log E_Q(q_k(X_k|X^{k-1})^{-\lambda}|X^{k-1}) - E_Q(-\lambda\log q_k(X_k|X^{k-1})|X^{k-1})] + c \\ &\qquad P\text{-a.s.}\omega \in D(c)\end{aligned} \tag{23}$$

易知由不等式

$$x^{-\lambda} - 1 + \lambda \log x \leq (1/2)\lambda^2 (\log x)^2 x^{-|\lambda|}, \quad 0 \leq x \leq 1 \tag{24}$$

及 (24), (23) 与 (11) 式，当 $|\lambda|<\alpha$, 时有

$$\begin{aligned}
&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\lambda\log q_k(X_k|X^{k-1}))-E_Q(-\lambda\log q_k(X_k|X^{k-1})|X^{k-1})]\\
&\quad\le\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[E_Q(q_k(X_k|X^{k-1})^{-\lambda}|X^{k-1})-1-E_Q(-\lambda\log q_k(X_k|X^{k-1})|X^{k-1})]+c\\
&\quad\le\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}E_Q[(1/2)\lambda^2(\log(q_k(X_k|X^{k-1}))^2q_k(X_k|X^{k-1})^{-|\lambda|}|X^{k-1}]+c\\
&\quad\le(1/2)\lambda^2\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}E_Q[(\log(q_k(X_k|X^{k-1}))^2q_k(X_k|X^{k-1})^{-\alpha}|X^{k-1}]+c\\
&\quad=(1/2)\lambda^2b_\alpha+c\quad P\text{-a.s.}\quad\omega\in D(c).
\end{aligned}\tag{25}$$

当 $0<\lambda<\alpha$ 时，由 (25) 式有

$$\begin{aligned}
&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\\
&\quad\le(1/2)\lambda b_\alpha+c/\lambda\qquad P\text{-a.s.}\quad\omega\in D(c).
\end{aligned}\tag{26}$$

当 $0<c\le(\alpha^2b_\alpha)/2$ 时，在 (26) 式中令 $\lambda=\sqrt{(2c)/b_\alpha}$, 则有

$$\begin{aligned}
&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\le\sqrt{2cb_\alpha}\\
&\qquad P\text{-a.s.}\qquad\omega\in D(c).
\end{aligned}\tag{27}$$

当 $c=0$ 时，取 $0<\lambda_i<\alpha$ $(i=1,2,\ldots)$, 使得 $\lambda_i\to0$ $(i\to\infty)$ 则对一切 i, 由 (26) 式有

$$\begin{aligned}
&\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\le0\\
&\qquad P\text{-a.s.}\quad\omega\in D(0).
\end{aligned}\tag{28}$$

由 (28) 式知，当 $c=0$ 时， (27) 式仍成立. 类似的，当 $-\alpha<\lambda<0$ 时，利用 (25) 式可证，当 $0\le c\le(\alpha^2b_\alpha)/2$ 时，有

$$\begin{aligned}
&\liminf_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\ge-\sqrt{2cb_\alpha}\\
&\qquad P\text{-a.s.}\quad\omega\in D(c).
\end{aligned}\tag{29}$$

由 (15)–(18) 式可知，在 (20) 式中令 $\lambda=0$ 则有

$$\limsup_{n\to\infty}(1/n)\log T_n(\omega)=\limsup_{n\to\infty}(1/n)\log\frac{\prod\limits_{k=1}^{n}q_k(X_k|X^{k-1})}{p(X_0,\ldots,X_n)}\le0\qquad P\text{-a.s.}\tag{30}$$

又注意到

$$H_k^Q(\omega)=E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1}),$$

于是由 (2), (30) 及 (27) 式可得

$$\begin{aligned}
&\limsup_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\\
&\le \limsup_{n\to\infty}(1/n)\log[\prod_{k=1}^{n}q_k(X_k|X^{k-1})\Big/p(X_0,\ldots,X_n)]+\\
&\quad \limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\\
&\le \sqrt{2cb_\alpha}\qquad P\text{-a.s.}\quad \omega\in D(c),
\end{aligned}\tag{31}$$

于是 (12) 式成立. 由 (2),(8),(10) 与 (29) 式有

$$\begin{aligned}
&\liminf_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\\
&\ge \liminf_{n\to\infty}(1/n)\log[\prod_{k=1}^{n}q_k(X_k|X^{k-1})\Big/p(X_0,\ldots,X_n)]+\\
&\quad \liminf_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1}))-E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})]\\
&\ge -h(P|Q)-\sqrt{2cb_\alpha}\ge -\sqrt{2cb_\alpha}-c\qquad P\text{-a.s.}\quad \omega\in D(c),
\end{aligned}\tag{32}$$

故 (13) 也成立. 于 (12) 与 (13) 式中令 $c=0$ 可得 (14) 式成立. 证毕. □

3 状态有限空间下的若干 Shannon-Mcmillan 定理

定理 2 设 $\{X_n, n\ge 0\}$ 是具有分布 (1) 并于字母集 $S=\{1,2,\ldots,N\}$ 上取值的任意信源, $f_n(\omega)$, $h(P|Q)$ 和 $D(c)$ 分别由 (2) 和 (4), (10) 式定义, 又设 $0\le\alpha<1, 0\le c\le 2N\alpha^2\big/[(1-\alpha)e]^2$. $H_k^Q(\omega)$ 表示 X_k 在测度 Q 下关于 $X_0,\ldots,X_{k-1}$ 的随机条件熵. 即

$$H_k^Q(\omega)=-\sum_{x_k\in S}q_k(x_k|X_0,\ldots,X_{k-1})\log q_k(x_k|X_0,\ldots,X_{k-1}).$$

则有

$$\limsup_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\le\frac{2e^{-1}}{(1-\alpha)}\sqrt{2cN}\quad P\text{-a.s.}\quad \omega\in D(c),\tag{33}$$

$$\liminf_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]\ge-\frac{2e^{-1}}{(1-\alpha)}\sqrt{2cN}-c\quad P\text{-a.s.}\quad \omega\in D(c).\tag{34}$$

证明 由 $0<\alpha<1$, 考虑函数

$$\varphi(x)=(\log x)^2x^{1-\alpha},\quad 0<x\le 1,\quad 0<\alpha<1.\quad (\varphi(0)=0)\tag{35}$$

易知 $\varphi(x)$ 在 $[0,1]$ 区间上的最大值为

$$\max\{\varphi(x),0\le x\le 1\}=\varphi(e^{2/(\alpha-1)})=(\frac{2}{\alpha-1})^2e^{-2}.\tag{36}$$

由 (11) 和 (36) 式可知

$$\begin{aligned} b_\alpha &= \limsup_{n\to\infty}(1/n)\sum_{k=1}^{n} E_Q[(\log(q_k(X_k|X^{k-1}))^2 q_k(X_k|X^{k-1})^{-\alpha}|X^{k-1}] \\ &= \limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}\sum_{x_k=1}^{N}(\log(q_k(x_k|X^{k-1}))^2 q_k(x_k|X^{k-1})^{1-\alpha}] \\ &\le \limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}\sum_{x_k=1}^{N}[(\frac{2}{\alpha-1})^2 e^{-2}] \le N(\frac{2}{\alpha-1})^2 e^{-2} < \infty, \end{aligned} \tag{37}$$

所以 (11) 式自然成立. 由 (37) 与 (25) 式有

$$\begin{aligned} &\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\lambda\log q_k(X_k|X^{k-1})) - E_Q(-\lambda\log q_k(X_k|X^{k-1})|X^{k-1})] \\ &\le N\lambda^2\frac{2e^{-2}}{(\alpha-1)^2} + c \qquad P\text{-a.s.} \quad \omega\in D(c). \end{aligned} \tag{38}$$

当 $0<\lambda<\alpha$ 时, 由 (38) 式有

$$\begin{aligned} &\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1})) - E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})] \\ &\le N\lambda\frac{2e^{-2}}{(\alpha-1)^2} + \frac{c}{\lambda} \qquad P\text{-a.s.} \quad \omega\in D(c). \end{aligned} \tag{39}$$

当 $0<c\le 2N\alpha^2/[(1-\alpha)e]^2$ 时, 在 (39) 式中令 $\lambda=(1-\alpha)e\sqrt{c/2N}$, 便有

$$\begin{aligned} &\limsup_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1})) - E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})] \\ &\le \frac{2e^{-1}}{(1-\alpha)}\sqrt{2cN} \qquad P\text{-a.s.} \quad \omega\in D(c). \end{aligned} \tag{40}$$

类似可证, 当 $c=0$, 时 (42) 式仍成立. 仿照 (12) 式的证明, 由 (40), (2) 与 (30) 式可得 (33) 式成立.

类似地, 利用 $-\alpha<\lambda<0$ 的 (38) 式可得, 当 $0\le c\le 2N\alpha^2/[(1-\alpha)e]^2$ 时, 有

$$\begin{aligned} &\liminf_{n\to\infty}(1/n)\sum_{k=1}^{n}[(-\log q_k(X_k|X^{k-1})) - E_Q(-\log q_k(X_k|X^{k-1})|X^{k-1})] \\ &\ge -\frac{2e^{-1}}{(1-\alpha)}\sqrt{2cN} \qquad P\text{-a.s.} \quad \omega\in D(c) \end{aligned} \tag{41}$$

仿照 (13) 式的证明, 由 (2), (4), (10) 和 (41) 式可得 (34) 式成立.

注 定理 2 的上下界比文献 [1] 定理 1 的上下界更精确.

推论 1[2] 设 $\{X_n, n\ge 0\}$ 是具有分布 (1) 的任意信源, 其中 $f_n(\omega)$ 由 (2) 定义. $H_k(\omega)$ 为 X_k 在测度 P 下关于 $X_0,\ldots,X_{k-1}$ 的随机条件熵. 即

$$H_k(\omega) = -\sum_{x_k\in S} p_k(x_k|X_0,\ldots,X_{k-1})\log p_k(x_k|X_0,\ldots,X_{k-1}). \tag{42}$$

则有

$$\lim_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k(\omega)]=0 \qquad P\text{-a.s.} \tag{43}$$

证明 由定理 2 中 (33) 和 (34) 式知当 $c=0$ 有

$$\lim_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H_k^Q(\omega)]=0 \qquad P\text{-a.s.} \quad \omega\in D(0). \tag{44}$$

这时令 $P\equiv Q$, 则有 $H_k^Q(\omega)=H_k(\omega)$, $h(P|Q)\equiv 0$. 这表明当 $c=0$ 时 (10) 式恒成立. 因而有 $D(0)=\Omega$. 从而由 (44) 式可得 (43) 式成立. 该推论即是刘文 [2] 的主要结果.

推论 2 设 $\{X_n,n\geq 0\}$ 为 m 阶非齐次马氏信源, 其中 $f_n(\omega)$ 由 (2) 定义. 设 $H(p_k(X_{k-m}^{k-1},1),\ldots,p_k(X_{k-m}^{k-1},N))$ 为在测度 P 下 m 阶非齐次马氏链的随机条件熵. 即

$$H(p_k(X_{k-m}^{k-1},1),\ldots,p_k(X_{k-m}^{k-1},N))=-\sum_{x_k\in S}p_k(x_k|X_{k-m}^{k-1})\log p_k(x_k|X_{k-m}^{k-1}), \tag{45}$$

则有

$$\lim_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=1}^{n}H(p_k(X_{k-m}^{k-1},1),\ldots,p_k(X_{k-m}^{k-1},N))]=0 \qquad P\text{-a.s.} \tag{46}$$

证明 根据推论 1 知这时 $H_k(\omega)=H(p_k(X_{k-m}^{k-1},1),\ldots,p_k(X_{k-m}^{k-1},N))$, 由 (43) 式即得 (46) 式. 当 $m=1$ 时，该推论便是刘文，杨卫国 [3] 的定理 2, 所以刘和杨的结果是该推论的特例.

设 $\{X_n,n\geq 0\}$ 是以 $S=\{1,2,\ldots,N\}$ 为状态空间的 m 阶非齐次马氏信源. 其在测度 P 下的 m 维初始分布与 m 阶转移矩阵分别为

$$p_o(i_0,\ldots,i_{m-1})=P(X_0=i_0,\ldots X_{m-1}=i_{m-1}), \tag{47}$$

$$P_n=(p_n(i_m|i_0,\ldots,i_{m-1})), \tag{48}$$

其中 $p_n(i_m|i_0,\ldots,i_{m-1})=P(X_n=i_m|X_{n-m}=i_0,\ldots X_{n-1}=i_{m-1})$.

这时 $\{X_n,n\geq 0\}$ 在测度 P 下的联合分布和相对熵密度分别为

$$p(x_0,\ldots,x_n)=p_o(x_0,\ldots,x_{m-1})\prod_{k=m}^{n}p_k(x_k|x_{k-m},\ldots,x_{k-1}), \tag{49}$$

$$f_n(\omega)=-\frac{1}{n}[\log p_o(X_0,\ldots,X_{m-1})+\sum_{k=m}^{n}\log p_k(X_k|X_{k-m},\ldots,X_{k-1})]. \tag{50}$$

又设 $\{X_n,n\geq 0\}$ 在测度 Q 下的 m 维初始分布与 m 阶转移矩阵分别为

$$q_o(j_0,\ldots,j_{m-1})=Q(X_0=j_0,\ldots,X_{m-1}=j_{m-1}), \tag{51}$$

$$Q_n=(q_n(j_m|j_0,\ldots,j_{m-1})), \tag{52}$$

其中 $q_n(j_m|j_0,\ldots,j_{m-1})=Q(X_n=j_m|X_{n-m}=j_0,\ldots,X_{n-1}=j_{m-1})$.

推论 3 设 $\{X_n, n\geq 0\}$ 是 m 阶非齐次马氏信源. 其在测度 P 和 Q 下的 m 维初始分布和 m 阶转移矩阵分别为 (47),(48) 式和 (51),(52) 式. 如果有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=m}^{n}\Big[\frac{p_k(i_m|i_0,\ldots,i_{m-1})}{q_k(i_m|i_0,\ldots,i_{m-1})}-1\Big]^{+}=0,\quad \forall i_0,\ldots,i_m\in S, \tag{53}$$

则有

$$\lim_{n\to\infty}[f_n(\omega)-(1/n)\sum_{k=m}^{n}H_Q(q_k(X_{k-m}^{k-1},1),\ldots,q_k(X_{k-m}^{k-1},N))]=0\qquad P\text{-a.s.} \tag{54}$$

证明 由 (2), (4) 与 (50) 式有

$$h(P|Q)=\limsup_{n\to\infty}\frac{1}{n+1}[\log p_o(X^{m-1})-\log q_o(X^{m-1})+\sum_{k=m}^{n}\log\frac{p_k(X_k|X_{k-m}^{k-1})}{q_k(X_k|X_{k-m}^{k-1})}]. \tag{55}$$

由不等式 $\log x\leq x-1(x>0)$ 及 $a\leq[a]^+$ 有

$$\begin{aligned}\frac{1}{n}\sum_{k=m}^{n}\log\frac{p_k(X_k|X_{k-m}^{k-1})}{q_k(X_k|X_{k-m}^{k-1})}&=\frac{1}{n}\sum_{k=m}^{n}\sum_{i_0,\ldots,i_m\in S}\delta_{i_0}(X_{k-m})\ldots\delta_{i_m}(X_k)\log\frac{p_k(i_m|i_0,\ldots,i_{m-1})}{q_k(i_m|i_0,\ldots,i_{m-1})}\\&\leq\frac{1}{n}\sum_{k=m}^{n}\sum_{i_0,\ldots,i_m\in S}\delta_{i_0}(X_{k-m})\ldots\delta_{i_m}(X_k)[\frac{p_k(i_m|i_0,\ldots,i_{m-1})}{q_k(i_m|i_0,\ldots,i_{m-1})}-1]^{+}\\&\leq\sum_{i_0,\ldots,i_m\in S}\frac{1}{n}\sum_{k=m}^{n}[\frac{p_k(i_m|i_0,\ldots,i_{m-1})}{q_k(i_m|i_0,\ldots,i_{m-1})}-1]^{+}.\end{aligned} \tag{56}$$

由 (53) 与 (56) 式有

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{k=m}^{n}\log\frac{p_k(X_k|X_{k-m}^{k-1})}{q_k(X_k|X_{k-m}^{k-1})}\leq 0\qquad P\text{-a.s.} \tag{57}$$

由 (55) 与 (57) 式有

$$h(P|Q)\leq 0\qquad P\text{-a.s.} \tag{58}$$

由 (10) 与 (58) 式有 $D(0)=\Omega$. 于是由推论 1 中的 (44) 式知 (54) 式成立.

参考文献:

[1] 刘文, 杨卫国. 关于 Shannon-Mcmillan 定理的若干研究 [J]. 数学物理学报, 1994, **14**(3): 337–345.
LIU Wen, YANG Wei-guo. *Some investigations related to the Shannon-McMillan theorem* [J]. Acta Math. Sci. (Chinese), 1994, **14**(3): 337–345. (in Chinese)

[2] 刘文. 一个对离散信源普遍成立的强极限定理 [J]. 数学物理学报, 1996, **16**(4): 440–443.
LIU Wen. *A strong limit theorem for discrete information sources that holds universally* [J]. Acta Math. Sci. (Chinese), 1996, **16**(4): 440–443. (in Chinese)

[3] LIU Wen, YANG Wei-guo. *An extension of Shannon-McMillan theorem and some limit properties for nonhomogeneous Markov chains* [J]. Stochastic Process. Appl., 1996, **61**(1): 129–145.

[4] CHUNG K L. *A Course in Probability Theory* [M]. Academic Press, New York, 1974.

[5] SHANNON C. *A mathematical Theory of Communication* [M]. Bell System Tech J., 1948, **27**: 379–423.

[6] MCMILLAN B. *The Basic Theorem of information theory* [J]. Ann. Math. Statist., 1953, **24**: 196–219.

[7] BREIMAN L. *The individual ergodic theorem of information theory* [J]. Ann. Math. Statist., 1957, **28**: 809–811.
[8] CHUNG K L. *The ergodic theorem of information theory* [J]. Ann. Math. Statist, 1961, **32**: 612–614.
[9] DOOB J L. *Stochastic Process* [M]. New York, Wiley, 1953.
[10] GRAY R M. *Entropy and Information Theory* [M]. New York, Springer, 1990.

A Class of Shannon-Mcmillan Approximation Theorems for Arbitrary Discrete Information Source

WANG Kang-kang[1], *YANG Wei-guo*[2]
(1. School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China;
2. Faculty of Science, Jiangsu University, Zhenjiang 212013, China)

Abstract: In this paper, a class of small-deviation theorems for the relative entropy densities of the arbitrary stochastic sequence are discussed by comparing two joint distributions. As corollaries, some Shannon-Mcmillan theorems for arbitrary information source, m-order Markov information source are obtained and some results for the discrete information source are extended.

Key words: Shannon-Mcmillan theorem; arbitrary information source; small-deviation theorems; relative entropy density; m-order Markov information source.