Wei Guo Yang; Zhong Xing Ye; Wen Liu

A local convergence theorem for partial sums of stochastic adapted sequences

Czechoslovak Mathematical Journal, Vol. 56 (2006), No. 2, 525–532

Persistent URL: http://dml.cz/dmlcz/128083

Czechoslovak Mathematical Journal, 56 (131) (2006), 525–532

A LOCAL CONVERGENCE THEOREM FOR PARTIAL SUMS OF STOCHASTIC ADAPTED SEQUENCES

Weiguo Yang, Zhenjiang, Zhongxing Ye, Shanghai,
and Wen Liu, Tianjin

(Received September 25, 2003)

Abstract. In this paper we establish a new local convergence theorem for partial sums of arbitrary stochastic adapted sequences. As corollaries, we generalize some recently obtained results and prove a limit theorem for the entropy density of an arbitrary information source, which is an extension of case of nonhomogeneous Markov chains.

Keywords: local convergence theorem, stochastic adapted sequence, martingale

MSC 2000: 60F15

1. Introduction and the main results

Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a stochastic adapted sequence on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$, that is, $\{\mathscr{F}_n, n \geqslant 0\}$ is an increasing sequence of sub σ-algebras of $\mathscr{F}$, and X_n is $\mathscr{F}_n$-measurable. Liu, Yan and Yang proved a limit theorem for partial sums of bounded stochastic adapted sequences (see [2]). Liu obtained a limit theorem for multivariate function sequences of discrete random variables (see [3]). The main purpose of this paper is to establish a new limit theorem for partial sums of stochastic adapted sequences. As corollaries, we generalize the above results and establish a limit theorem for the entropy density of an arbitrary information source, which extends the case of nonhomogeneous Markov chains (see [4]).

Theorem 1. *Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a stochastic adapted sequence, and let (a_n) be a sequence of non-negative r.v.'s defined on $(\Omega, \mathscr{F}, \mathbf{P})$. Let $\alpha > 0$, and set*

$$D(\alpha) = \left\{\lim_n a_n = \infty,\ \limsup_n \frac{1}{a_n} \sum_{k=1}^{n} E[X_k^2 \mathrm{e}^{\alpha|X_k|} | \mathscr{F}_{k-1}] < \infty\right\}. \tag{1}$$

Then

$$\lim_n \frac{1}{a_n} \sum_{k=1}^n \{X_k - E[X_k|\mathscr{F}_{k-1}]\} = 0 \quad \text{a.e.,} \quad \omega \in D(\alpha). \tag{2}$$

P r o o f. Define $M_0(\lambda) = 1$ and

$$M_n(\lambda) = \frac{\mathrm{e}^{\lambda \sum_{k=1}^n X_k}}{\prod_{k=1}^n E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]}, \quad n \geqslant 1. \tag{3}$$

Since

$$E[M_n(\lambda)|\mathscr{F}_{n-1}] = M_{n-1}(\lambda) E\left[\frac{\mathrm{e}^{\lambda X_n}}{E[\mathrm{e}^{\lambda X_n}|\mathscr{F}_{n-1}]}\Big|\mathscr{F}_{n-1}\right] = M_{n-1}(\lambda) \quad \text{a.e.,} \tag{4}$$

and $M_n(\lambda) \geqslant 0$, $\{M_n(\lambda), \mathscr{F}_n, n \geqslant 0\}$ is a non-negative martingale. By Doob's Martingale Convergence Theorem, $\lim_n M_n(\lambda) = M_\infty(\lambda) < \infty$ a.e. Let $A = \{\lim_n a_n = \infty\}$. We have

$$\limsup_n \frac{1}{a_n} \log M_n(\lambda) \leqslant 0 \ \text{ a.e.,} \quad \omega \in A. \tag{5}$$

By (3) and (5) we have

$$\limsup_n \frac{1}{a_n} \left\{\lambda \sum_{k=1}^n X_k - \sum_{k=1}^n \log E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]\right\} \leqslant 0 \ \text{ a.e.,} \quad \omega \in A. \tag{6}$$

Letting $\lambda > 0$ and $\lambda < 0$, dividing both sides of (6) by λ, we have respectively

$$\limsup_n \frac{1}{a_n} \sum_{k=1}^n \left\{X_k - \frac{\log E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]}{\lambda}\right\} \leqslant 0 \ \text{ a.e.,} \quad \omega \in A, \ \lambda > 0, \tag{7}$$

$$\liminf_n \frac{1}{a_n} \sum_{k=1}^n \left\{X_k - \frac{\log E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]}{\lambda}\right\} \geqslant 0 \ \text{ a.e.,} \quad \omega \in A, \ \lambda < 0. \tag{8}$$

Using the inequalities $\log x \leqslant x - 1$ $(x > 0)$, $0 \leqslant \mathrm{e}^x - 1 - x \leqslant \frac{1}{2} x^2 \mathrm{e}^{|x|}$ and the properties of the superior and inferior limits,

$$\limsup_n (a_n + b_n) \leqslant \limsup_n a_n + \limsup_n b_n,$$
$$\liminf_n (a_n + b_n) \geqslant \liminf_n a_n + \liminf_n b_n,$$

it follows by (7) that when $0 < \lambda < \alpha$,

$$\begin{aligned}(9)\qquad \limsup_n \frac{1}{a_n}\sum_{k=1}^n \{X_k - E[X_k|\mathscr{F}_{k-1}]\} &\leqslant \limsup_n \frac{1}{a_n}\sum_{k=1}^n \left\{\frac{\log E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]}{\lambda} - E[X_k|\mathscr{F}_{k-1}]\right\}\\ &\leqslant \limsup_n \frac{1}{a_n}\sum_{k=1}^n \left\{\frac{E[\mathrm{e}^{\lambda X_k}|\mathscr{F}_{k-1}]-1}{\lambda} - E[X_k|\mathscr{F}_{k-1}]\right\}\\ &= \limsup_n \frac{1}{a_n}\sum_{k=1}^n \left\{\frac{E[(\mathrm{e}^{\lambda X_k}-1-\lambda X_k)|\mathscr{F}_{k-1}]}{\lambda}\right\}\\ &\leqslant \frac{\lambda}{2}\limsup_n \frac{1}{a_n}\sum_{k=1}^n E[X_k^2\mathrm{e}^{\lambda|X_k|}|\mathscr{F}_{k-1}]\\ &\leqslant \frac{\lambda}{2}\limsup_n \frac{1}{a_n}\sum_{k=1}^n E[X_k^2\mathrm{e}^{\alpha|X_k|}|\mathscr{F}_{k-1}] \text{ a.e.,}\quad \omega\in D(\alpha);\end{aligned}$$

and when $-\alpha < \lambda < 0$, it similarly follows from (8) that

$$\begin{aligned}(10)\qquad \liminf_n \frac{1}{a_n}\sum_{k=1}^n \{X_k - E[X_k|\mathscr{F}_{k-1}]\} &\geqslant \frac{\lambda}{2}\limsup_n \frac{1}{a_n}\sum_{k=1}^n E[X_k^2\mathrm{e}^{\alpha|X_k|}|\mathscr{F}_{k-1}] \text{ a.e.,}\quad \omega\in D(\alpha).\end{aligned}$$

Letting $\lambda \downarrow 0$ and $\lambda \uparrow 0$ in (9) and (10) respectively, we have

$$(11)\qquad \limsup_n \frac{1}{a_n}\sum_{k=1}^n \{X_k - E[X_k|\mathscr{F}_{k-1}]\} \leqslant 0 \text{ a.e.,}\quad \omega\in D(\alpha),$$

$$(12)\qquad \liminf_n \frac{1}{a_n}\sum_{k=1}^n \{X_k - E[X_k|\mathscr{F}_{k-1}]\} \geqslant 0 \text{ a.e.,}\quad \omega\in D(\alpha),$$

which implies (2). □

Corollary 1 (see [2]). *Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a bounded stochastic adapted sequence, that is, there exists $K > 0$, such that $|X_n| \leqslant K$ for all $n \geqslant 0$. Let $\{a_n, n \geqslant 0\}$ be a sequence of non-negative r.v.'s. Put*

$$(13)\qquad \Omega_0 = \left\{\lim_n a_n = \infty,\ \limsup_n \frac{1}{a_n}\sum_{k=1}^n E[|X_k||\mathscr{F}_{k-1}] < \infty\right\}.$$

Then

$$\lim_n \frac{1}{a_n}\sum_{k=1}^{n}\{X_k - E[X_k|\mathscr{F}_{k-1}]\} = 0 \text{ a.e.}, \quad \omega \in \Omega_0. \tag{14}$$

P r o o f. Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a bounded stochastic adapted sequence. Since for $\alpha > 0$

$$E[X_k^2 \mathrm{e}^{\alpha|X_k|}|\mathscr{F}_{k-1}] \leqslant K\mathrm{e}^{\alpha K}E[|X_k|\big|\mathscr{F}_{k-1}],$$

we have $\Omega_0 \subset D(\alpha)$. The corollary follows directly from Theorem 1. $\square$

Theorem 2. *Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a non-negative stochastic adapted sequence for which there exist $\alpha > 0$ and $K > 0$ such that*

$$E[X_n^2 \mathrm{e}^{\alpha X_n}|\mathscr{F}_{n-1}] \leqslant KE[X_n|\mathscr{F}_{n-1}] \text{ a.e.} \tag{15}$$

Set

$$A = \left\{\sum_{n=1}^{\infty} X_n = \infty\right\}, \quad B = \left\{\sum_{n=1}^{\infty} E[X_n|\mathscr{F}_{n-1}] = \infty\right\}. \tag{16}$$

Then $A = B$ a.e., and

$$\lim_n \frac{\sum\limits_{k=1}^{n} X_k}{\sum\limits_{k=1}^{n} E[X_k|\mathscr{F}_{k-1}]} = 1 \text{ a.e.}, \quad \omega \in B. \tag{17}$$

Remark. If $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ is a bounded non-negative stochastic adapted sequence, then (15) holds; if $\{X_n, n \geqslant 0\}$ is a sequence of non-negative r.v.'s such that

$$\sup_n E[X_n^2 \mathrm{e}^{\alpha X_n}|\mathscr{F}_{n-1}] < \infty, \quad \inf_n E[X_n|\mathscr{F}_{n-1}] > 0,$$

then (15) also holds.

P r o o f. If we set $a_n = \sum\limits_{k=1}^{n} E[X_k|\mathscr{F}_{k-1}]$, then by (15) and Theorem 1 we obtain (17), which implies $B \subset A$ a.e. If we let $a_n = \sum\limits_{k=1}^{n} X_k$, by the definition of the sets A and B and Theorem 1 we have

$$\limsup_n \frac{\sum_{k=1}^{n} E[X_k|\mathscr{F}_{k-1}]}{\sum_{k=1}^{n} X_k} = 0, \quad \omega \in AB^c,$$

and

$$\lim_n \frac{\sum_{k=1}^{n} E[X_k|\mathscr{F}_{k-1}]}{\sum_{k=1}^{n} X_k} = 1 \ \text{a.e.}, \quad \omega \in AB^c.$$

Thus we have $AB^c = \emptyset$ a.e. Hence we must have $A = B$ a.e. $\square$

This theorem implies immediately

Corollary 2. *Let $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ be a bounded non-negative stochastic adapted sequence such that $0 \leqslant X_n \leqslant K$, for all $n \geqslant 0$. Put*

$$A = \left\{\sum_{n=1}^{\infty} X_n = \infty\right\}, \quad B = \left\{\sum_{n=1}^{\infty} E[X_n|\mathscr{F}_{n-1}] = \infty\right\}.$$

Then $A = B$ a.e., and (17) holds.

This corollary is an extension of the Extended Borel-Cantelli Lemma (see [2]).

Corollary 3 (see [3])**.** *Let $\{X_n, n \geqslant 0\}$ be a sequence of arbitrary discrete r.v.'s taking values in $S = \{t_0, t_1, \ldots\}$, and let $g_n(x_0, \ldots, x_n)$ be real functions defined on S^{n+1}. Let $\{a_n, n \geqslant 1\}$ be a sequence of non-negative r.v.'s. Let $\alpha > 0$ and put*

$$D(\alpha) = \Big\{ \lim_n a_n = \infty,$$
$$\limsup_n \frac{1}{a_n} \sum_{k=1}^{n} E[g_k^2(X_0, \ldots, X_k) \mathrm{e}^{\alpha |g_k(X_0, \ldots, X_k)|} | X_0, \ldots, X_{k-1}] < \infty \Big\}.$$

Then

$$\lim_n \frac{1}{a_n} \sum_{k=1}^{n} \{g_k(X_0, \ldots, X_k) - E[g_k(X_0, \ldots, X_k)|X_0, \ldots, X_{k-1}]\} = 0 \ \text{a.e.}, \quad \omega \in D(\alpha).$$

Proof. Let $Y_n = g_n(X_0, \ldots, X_n)$ and $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$. Then $\{Y_n, \mathscr{F}_n, n \geqslant 0\}$ is a stochastic adapted sequence. This corollary immediately follows from Theorem 1. $\square$

2. A limit theorem for the entropy density of an arbitrary information source

Let $\{X_n, n \geqslant 0\}$ be arbitrary information source taking values in alphabet $S = \{1, 2, \ldots, N\}$ with finite distributions

$$p(x_0, \ldots, x_n) = P(X_0 = x_0, \ldots, X_n = x_n), \quad x_i \in S, \ 0 \leqslant i \leqslant n, \ n \geqslant 0. \tag{18}$$

Let

$$f_n(\omega) = -\frac{1}{n} \log p(X_0, \ldots, X_n), \quad n \geqslant 0. \tag{19}$$

$f_n(\omega)$ is called the entropy density of $\{X_n, n \geqslant 0\}$. Denote

$$p_n(x_n|X_0, \ldots, X_{n-1}) = P(X_n = x_n|X_0, \ldots, X_{n-1}). \tag{20}$$

Then

$$p(x_0, \ldots, x_n) = p(x_0) \prod_{k=1}^{n} p_k(x_k|x_0, \ldots, x_{k-1}). \tag{21}$$

In this case

$$f_n(\omega) = -\frac{1}{n} \left[\log p(X_0) + \sum_{k=1}^{n} \log p_k(X_k|X_0, \ldots, X_{k-1}) \right]. \tag{22}$$

If $\{X_n, n \geqslant 0\}$ is a nonhomogeneous Markov chain taking values in the state space S with the initial distribution

$$p = (p(1), \ldots, p(N)), \tag{23}$$

and transition matrices

$$P_n = (p_n(j|i)), \quad i, j \in S, \ n \geqslant 0, \tag{24}$$

where $p_n(j|i) = P(X_n = j|X_{n-1} = i)$, then

$$p(x_0, \ldots, x_n) = p(x_0) \prod_{k=1}^{n} p_k(x_k|x_{k-1}) \tag{25}$$

and

$$f_n(\omega) = -\frac{1}{n} \left[\log p(X_0) + \sum_{k=1}^{n} \log p_k(X_k|X_{k-1}) \right]. \tag{26}$$

The limit property of $f_n(\omega)$ plays an important role in information theory (see [1]). From Theorem 1 we can obtain easily a limit theorem for $f_n(\omega)$ which holds for an arbitrary information source.

Theorem 3. *Let $\{X_n,\, n \geqslant 0\}$ be arbitrary information source taking values in alphabet $S = \{1, 2, \ldots, N\}$ defined as before, and let $f_n(\omega)$ be its entropy density. Then*

$$(27)\quad \lim_n \left\{ f_n(\omega) - \frac{1}{n} \sum_{k=1}^{n} H[p_k(1|X_0, \ldots, X_{k-1}), \ldots, p_k(N|X_0, \ldots, X_{k-1})] \right\} = 0 \ \text{a.e.},$$

where $H(p(1), \ldots, p(N))$ is the entropy of the distribution $(p(1), \ldots, p(N))$, that is

$$H(p(1), \ldots, p(N)) = -\sum_{k=1}^{N} p(k) \log p(k).$$

P r o o f. Setting $Y_n = -\log p_n(X_n|X_0, \ldots, X_{n-1})$, $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$ and $a_n = n$, then $\{Y_n, \mathscr{F}_n,\, n \geqslant 0\}$ is a stochastic adapted sequence. Using the inequality

$$x^{\frac{1}{2}} (\log x)^2 \leqslant 16\mathrm{e}^{-2}, \quad 0 \leqslant x \leqslant 1,$$

we have

$$\begin{aligned}
&E[Y_n^2 \mathrm{e}^{\frac{1}{2}|Y_n|} | X_0, \ldots, X_{n-1}] \\
&\quad = E[(\log p_n(X_n|X_0, \ldots, X_{n-1}))^2 \mathrm{e}^{-\frac{1}{2} \log p_n(X_n|X_0, \ldots, X_{n-1})} | X_0, \ldots, X_{n-1})] \\
&\quad = \sum_{x_n} (p(x_n|X_0, \ldots, X_{n-1}))^{\frac{1}{2}} (\log p_n(x_n|X_0, \ldots, X_{n-1}))^2 \\
&\quad \leqslant 16N\mathrm{e}^{-2}, \quad \omega \in \Omega.
\end{aligned}$$

Thus $D(\frac{1}{2}) = \Omega$. It follows from Theorem 1 that

$$(28)\qquad \lim_n \left\{ \frac{1}{n} \sum_{k=1}^{n} Y_k - \frac{1}{n} \sum_{k=1}^{n} E[Y_k|X_0, \ldots, X_{k-1}] \right\} = 0 \ \text{a.e.}$$

Since

$$\begin{aligned}
(29)\quad E[Y_n|X_0, \ldots, X_{n-1}] &= E[-\log p_n(X_n|X_0, \ldots, X_{n-1})|X_0, \ldots, X_{n-1}] \\
&= -\sum_{x_n} p_n(x_n|X_0, \ldots, X_{n-1}) \log p(x_n|X_0, \ldots, X_{n-1}) \\
&= H[p_n(1|X_0, \ldots, X_{n-1}), \ldots, p_n(N|X_0, \ldots, X_{n-1})],
\end{aligned}$$

(27) follows from (22), (28) and (29). □

Corollary 4 (see [4]). *Let* $\{X_n, n \geqslant 0\}$ *be a nonhomogeneous Markov chain taking values in the state space* S *with initial distribution* (23) *and transition matrices* (24), *let* $f_n(\omega)$ *be its entropy density. Then*

$$\lim_n \left\{ f_n(\omega) - \frac{1}{n} \sum_{k=1}^{n} H[p_k(1|X_{k-1}), \ldots, p_k(N|X_{k-1})] \right\} = 0 \ \text{a.e.}$$

References

[1] *P. Algoat, T. Cover*: A sandwich proof of the Shannon-McMillan-Breiman theorem. Annals of Probability *16* (1988), 899–909. Zbl 0653.28013

[2] *W. Liu, J. A. Yan, and W. G. Yang*: A limit theorem for partial sums of random variables and its applications. Statistics and Probability Letters *62* (2003), 79–86. Zbl pre02041714

[3] *W. Liu*: Some limit properties of the multivariate function sequences of discrete random variables. Statistics and Probability Letters *61* (2003), 41–50. Zbl 1016.60034

[4] *W. Liu, W. G. Yang*: A limit theorem for the entropy density of nonhomogeneous Markov information source. Statistics and Probability Letters *22* (1995), 295–301. Zbl 0833.60034

Authors' addresses: Weiguo Yang, Faculty of Science, Jiangsu University, Zhenjiang 212013, P.R. China, e-mail: wgyang@ujs.edu.cn; Zhongxing Ye, Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, P.R. China; Wen Liu, Department of Mathematics, Hebei University of Technology, Tianjin 300130, P.R. China.