Probab. Th. Rel. Fields 75, 425–430 (1987)

Probability Theory and Related Fields
© Springer-Verlag 1987

A Remark on Almost Sure Convergence of Weighted Sums

Rolf Thrum

Department of Statistics, Humboldt-University at Berlin, Unter den Linden 6, DDR-1086 Berlin, German Democratic Republic

Summary. As a generalization of a theorem of Chow [1] it is shown by an elementary method that for i.i.d. r.v.'s $X_1, \ldots, X_n$, with expectation zero and finite p-th absolute moment $(p \geqq 2)$ the weighted sums $\sum_{i=1}^{n} a_{n,i} X_i / n^{1/p} \left(\sum_{i=1}^{n} a_{n,i}^2 \right)^{1/2}$ converge to zero a.s.

1. Results

The theorem to be presented here is closely related to two theorems of Chow [1].

Theorem 1 (Chow 1966). *If $X_1, \ldots, X_n, \ldots$ are independent identically distributed random variables* (i.i.d. r.v.'s) *with expectation zero and finite second moment and $\{a_{n,i}/i=1, \ldots, n;\ n=1, \ldots\}$ is a sequence of real numbers satisfying $\sum_{i=1}^{n} a_{n,i}^2 = 1$ for $n \geqq 1$ then $\sum_{i=1}^{n} a_{n,i} X_i / n^{1/2} \xrightarrow[n \to \infty]{} 0$ almost surely* (a.s.).

The proof of this and other similar theorems is based on the concept of pseudo-Gaussian variables, on Kolmogorov's strong law of large numbers as well as on truncation techniques (comprehensive theory of pseudo-Gaussian variables can be found in Stout [5]).

Y.S. Chow also modified an idea of Erdös and proved ([1], Theorem 6) the following result, which is a generalization of theorems of Hsu, Robbins, Franck and Hansen.

Theorem 2 (Chow 1966). *Suppose that in addition to the assumptions of Theorem 1 $E|X_1|^p$ is finite for some $p \geqq 2$ and $n^{1/p} \max\{|a_{n,k}| \mid k=1, \ldots, n\}$ is bounded, then $U_n = \sum_{i=1}^{n} a_{n,i} X_i / n^{1/p}$ converges completely to zero, i.e. $\sum_{n=1}^{\infty} P(|U_n| > \delta) < \infty$ for arbitrary positive δ.*

Our aim is to generalize Theorem 1 for the case of r.v.'s with finite p-th moment.

The method of proof seems to be new and is based on two "elementary" tools. An inequality for moments of quadratic forms is used to apply Beppo Levi's lemma. Secondly, a special summation procedure is defined which allows the application of the law of large numbers given in Neveu [3].

Theorem 3. *Suppose that $\{X_i/i=1,\ldots\}$ is a sequence of i.i.d. r.v.'s with expectation zero and finite p-th ($p\geqq 2$) absolute moment and $\{a_{n,i}/i=1,\ldots,n;\ n=1,\ldots\}$ is a sequence of nonrandom weighting coefficients with $\sum_{i=1}^{n} a_{n,i}^2=1$ for all $n\geqq 1$. Then $\sum_{i=1}^{n} a_{n,i}X_i/n^{1/p}\xrightarrow[n\to\infty]{}0$ a.s.*

2. Proof of Theorem 3

The square of $U_n=\sum_{i=1}^{n} a_{n,i}X_i/n^{1/p}$ is the sum of

$$V_n=\sum_{i=1}^{n} a_{n,i}^2 X_i^2/n^{2/p}$$

and

$$W_n=\sum_{i\neq j}^{n} a_{n,i}a_{n,j}X_iX_j/n^{2/p}:\qquad U_n^2=V_n+W_n.$$

By Lemma 2 the inequalities $E|W_n|^p\leqq C/n^2$ hold for every n where C does not depend on n. Applying Beppo-Levi's Theorem (see Révész [6], p. 13) it even follows that $\sum_{n=1} |W_n|^p<\infty$ a.s. and therefore $W_n\xrightarrow[n\to\infty]{}0$ a.s. Hence it suffices to show that $V_n\xrightarrow[n\to\infty]{}0$ a.s. Let us define new r.v.'s $Y_i=X_i^2-EX_i^2$ and coefficients $b_{n,i}=a_{n,i}^2$ $i=1,\ldots,n$,

$$\tilde{a}_{n,i}=a_{n,i}^2\Big/\left(\sum_{i=1}^{n} a_{n,i}^4\right)^{1/2}\qquad i=1,\ldots,n$$

and denote $\tilde{p}=p/2$, then $\sum_{i=1}^{n} b_{n,i}=1$ and $\sum_{i=1}^{n}\tilde{a}_{n,i}^2=1$. Using

$$\left(\sum_{i=1}^{n} a_{n,i}^4\right)^{1/2}\leqq\sum_{i=1}^{n} a_{n,i}^2=1$$

for $\tilde{V}_n=\sum_{i=1}^{n} b_{n,i}Y_i/n^{1/\tilde{p}}$ and $\tilde{U}_n=\sum_{i=1}^{n}\tilde{a}_{n,i}Y_i/n^{1/\tilde{p}}$ one gets:

$$0\leqq V_n=\tilde{V}_n+EX_1^2/n^{1/\tilde{p}},\tag{1}$$

$$|\tilde{V}_n|\leqq|\tilde{U}_n|.\tag{2}$$

If p belongs to the interval $[2,4)$, then $\tilde{p}\in[1,2)\subseteq(0,2)$ and Lemma 1 can be applied which ensures that $\tilde{V}_n$ and by (1) also V_n tend to zero a.s. In the case $p\in[2^k,2^{k+1})$ with a natural number $k>1$ the same argument can be repeated for $\tilde{U}_n$ and $\tilde{p}$ with $\tilde{p}\in[2^{k-1},2^k)$. The conclusion now follows by induction.

3. Remarks and Relationship to Least-Square Estimators

a) Without any additional assumption on the coefficients $a_{n,i}$ the normalizing factor $n^{-1/p}$ cannot be improved. Indeed, if $a_{n,n}=1$ and $a_{n,i}=0$ for $i<n$ then $n^{-1/p}X_n\xrightarrow[n\to\infty]{}0$ a.s. iff $E|X_1|^p<\infty$.

b) From Borel-Cantelli's lemma it follows that the assertion of Theorem 2 is stronger than that of Theorem 3 but the assumption on the $a_{n,i}$ is also more restrictive.

c) The connection to least-square estimators may be of special interest. Under the assumptions of a linear model

$$Y_i=f_i^T b+X_i, \quad i=1,\dots,n,\ n\geqq 1,\ b\in R^k$$

the least-square estimator

$$\hat{b}_n=\left(\sum_{i=1}^n f_i f_i^T\right)^{-1}\sum_{i=1}^n f_i^T Y_i$$

is used to estimate b. Here it is assumed that $n\geqq k$ where k is the rank of $G_n=\sum_{i=1}^n f_i f_i^T$. The difference of the i-th component of $\hat{b}_n$ and b can be written as

$$U_n^i=(\hat{b}_n-b)_i=\sum_{k=1}^n a_{n,k}^i X_k$$

where $a_{n,k}^i=e_i^T G_n^{-1} f_k$ and $e_i^T=(0,\dots,\overset{i}{1},0,\dots,0)$.

Under considerably weaker assumptions on the sequence $\{X_k\}$ than those of Theorem 1 recently it was shown by Chen et al. in [8] that

$$U_n^i\xrightarrow[n\to\infty]{}0 \quad \text{a.s.} \quad \text{iff } \sum_{k=1}^n (a_{n,k}^i)^2=e_i^T G_n^{-1} e_i\xrightarrow[n\to\infty]{}0.$$

They assume that $X_1,\dots,X_n,\dots$ forms a convergence system, i.e. $\sum_{i=1}^\infty c_i X_i$ converges a.s. if $\sum_{i=1}^\infty (c_i)^2<\infty$. This very deep result depends heavily on the special structure of the $a_{n,k}^i$ and cannot be deduced directly from Theorem 3 even if arbitrarily high moments of X_1 exist. From Theorem 3 it follows only that $\hat{b}_n\xrightarrow[n\to\infty]{}b$ a.s. if $e_i^T G_n^{-1} e_i n^{1/p}$ is bounded.

But this remains also true for linear models where f_i may depend on n while in this case the Theorem of Chen et al. cannot be applied.

4. Auxiliary Results

Lemma 1. *Suppose that $X_1, \ldots, X_n, \ldots$ are i.i.d. r.v.'s with $EX_1 = 0$ and $E|X_1|^p < \infty$ for some $p \in (0, 2)$. Assume that the nonrandom coefficients $b_{n,i}$ fulfill*

$$\sup\left\{\sum_{i=1}^{n} |b_{n,i}| \,\middle|\, n \geqq 1\right\} < \infty. \quad \textit{Then} \quad \sum_{i=1}^{n} b_{n,i} X_i / n^{1/p} \xrightarrow[n\to\infty]{} 0 \ \textit{a.s.}$$

Proof. Define $S_i = \sum_{j=1}^{i} X_j / i^{1/p}$

$$c_{n,i} = (i/n)^{1/p} (b_{n,i} - b_{n,i+1}) \quad \text{if } 1 \leqq i \leqq n-1$$

and

$$c_{n,n} = b_{n,n}.$$

Then

$$\sum_{i=1}^{n} b_{n,i} X_i / n^{1/p} = \sum_{i=1}^{n} c_{n,i} S_i, \tag{3}$$

$$\sum_{i=1}^{n} |c_{n,i}| \leqq 2 \sup\left\{\sum_{i=1}^{n} |b_{n,i}| \,\middle|\, n \geqq 1\right\}, \tag{4}$$

$$\lim_{n\to\infty} |c_{n,i}| = 0 \quad \text{for every fixed } i. \tag{5}$$

From (4) and (5) one obtains easily that for every sequence of real numbers r_n with $r_n \xrightarrow[n\to\infty]{} 0$ it holds

$$\sum_{i=1}^{n} c_{n,i} r_i \xrightarrow[n\to\infty]{} 0. \tag{6}$$

By the law of large numbers (see Neveu [3], p. 216, Theorem IV.7.1) we have $S_n \xrightarrow[n\to\infty]{} 0$ a.s.

Combining this with (3) and (6) the assertion follows.

Lemma 2. *Suppose that $\{X_n | n \geqq 1\}$ is a sequence of independent r.v.'s with $EX_i = 0$, $E|X_i|^{2s} < \infty$ for $i \geqq 1$ and some $s \geqq 1$. Assume that $\{a_{i,j} | i, j = 1, 2, \ldots\}$ is a double array of real numbers with $a_{i,j} = a_{j,i}$ for $i \neq j$ and $a_{i,i} = 0$ for all i. Then the r.v.'s $Z_n = \sum_{i,j=1}^{n} a_{i,j} X_i X_j$ have the following properties:*

$$E|Z_n|^{2s} < \infty \quad \textit{for all } n \geqq 1. \tag{7}$$

The sequence $Z_1, \ldots, Z_n, \ldots$ is a martingale with respect to the sigma-fields $\mathscr{F}_n$ generated by $X_1, X_2, \ldots, X_n$.

$$\begin{aligned} E|Z_n|^{2s} &\leqq B(2s) \left(\sum_{i<j}^{n} a_{i,j}^2 (E|X_i|^{2s})^{1/s} (E|X_j|^{2s})^{1/s}\right)^s \\ &\leqq (EZ_n^2)^s (B(2s)/2^{2s}) (\max\{E|X_i|^{2s}/(EX_i^2)^2 \mid i = 1, \ldots, n\})^2. \end{aligned} \tag{8}$$

$B(2s)$ is a constant which depends neither on n nor on the distribution of $(X_1, \ldots, X_n)$.

Proof. Using the inequality

$$|Z_n|^{2s} \leqq 2^{2s} \binom{n}{2}^{2s-1} \sum_{1 \leqq i \leqq j \leqq n} |X_i|^{2s} |X_j|^{2s} |a_{i,j}|^{2s}$$

and the independence of X_i and X_j $(i \neq j)$ one obtains that $E|Z_n|^{2s} < \infty$.

The martingale property follows from the recursive formula

$$Z_n = Z_{n-1} + 2X_n \sum_{j=1}^{n-1} a_{n,j} X_j \quad \text{and} \quad E(X_n | \mathscr{F}_{n-1}) = 0.$$

Define $Y_k = \sum_{j=1}^{k-1} a_{k,j} X_j$ if $k = 2, \ldots, n$, then $(\triangle Z)_k = Z_k - Z_{k-1} = 2X_k Y_k$ if $k = 2, \ldots, n$. Burkholder's martingale inequality (see Shirjajev [4], p. 489–490) implies

$$E|Z_n|^{2s} \leqq (B_{2s})^{2s} E\left(\sum_{k=2}^{n} (\triangle Z)_k^2\right)^s$$

where B_{2s} is a universal constant smaller than $18(2s)^{3/2}/(2s-1)^{1/2}$.

By Minkowski's inequality and because X_k and Y_k are independent r.v.'s it follows

$$E\left(\sum_{k=2}^{n} (\triangle Z)_k^2\right)^s = 2^{2s} E\left(\sum_{k=2}^{n} Y_k^2 X_k^2\right)^s$$

$$\leqq 2^{2s} \left(\sum_{k=2}^{n} (E|X_k|^{2s})^{1/s} (E|Y_k|^{2s})^{1/s}\right)^s.$$

Now an inequality for linear forms due to Whittle [7] can be applied to Y_k which yields

$$E|Y_k|^{2s} \leqq W(2s) \left(\sum_{j=1}^{k-1} (E|X_j|^{2s})^{1/s} a_{k,j}^2\right)^s$$

where $W(2s) \leqq 2^{3s} \Gamma((2s+1)/2)/\pi^{1/2}$ and Γ denotes the Gamma-function. Finally, let $B(2s) = (2B_{2s})^{2s} W(2s)$ then

$$E|Z_n|^{2s} \leqq B(2s) \left(\sum_{k=2}^{n} \sum_{j=1}^{k-1} (E|X_j|^{2s})^{1/s} (E|X_k|^{2s})^{1/s} a_{j,k}^2\right)^s$$

$$= B(2s) \Big(\sum_{1 \leqq j < k \leqq n} (E|X_j|^{2s})^{1/s} (E|X_k|^{2s})^{1/s} a_{j,k}^2\Big)^s$$

$$\leqq (B(2s)/2^{2s}) (\max\{E|X_j|^{2s}/(EX_j^2)^s \; j = 1, \ldots, n\})^2 (EZ_n^2)^s$$

because of $EZ_n^2 = 4 \sum_{1 \leqq i < j \leqq n} a_{i,j}^2 (EX_i^2)(EX_j^2)$.

Remark. If there exists $4s$-th moment of X_i then an upper bound for $E|Z_n|^{2s}/(EZ_n^2)^s$ can be derived directly from an inequality for quadratic forms given in Whittle [7]. In the case $s = 2$ a refined calculation is possible and yields sharper bounds (see [2]).

Acknowledgements. I am grateful to Prof. O. Bunke, Prof. P. Révész, Dr. M. Nußbaum and Dr. W. Schmidt for useful comments. I am also indebted to Prof. H. Läuter for stimulating uniform versions of Lemma 1 and Theorem 3 on the basis of the concept given here (see [9]).

References

1. Chow, Y.S.: Some convergence theorems for independent random variables. Ann. Math. Statist. **37**, 1482–1492 (1966)
2. Kleffe, J., Thrum, R.: Inequalities for moments of quadratic forms with applications to a.s. convergence. MOS, Series Statist. **14**, 211–216 (1983)
3. Neveu, J.: Bases mathématiques du calcul de probalités. Paris: Masson (1964). (In russ. Moscow: Nauka 1969)
4. Shirjajev, A.N.: Verojatnost. Moscow: Nauka 1980
5. Stout, W.F.: Almost sure convergence. New York: Academic Press 1974
6. Révész, P.: The laws of large numbers. Budapest: Akadémia kiadó 1967
7. Whittle, P.: Bounds for the moments of linear and quadratic forms in independent variables. Teorija Yerojatnost. i primenen, **5**, 331–334 (1960)
8. Chen Gui-Jing, Lai, T.L., Wei, C.Z.: Convergence system and strong consistency of least square estimates in regression models. J. Multivariate Anal. **11**, 319–333 (1981)
9. Thrum, R.: Uniform almost sure convergence of Weighted sums and applications in regression models. Paper at the 5-th Pannonian Symposium on Mathematical Statistics. May 20–24, 1985, Visegrád, Hungary (1985)

Received February 20, 1984