The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1502–1506

ALMOST SURE CONVERGENCE OF QUADRATIC FORMS IN INDEPENDENT RANDOM VARIABLES[1]

DALE E. VARBERG

Hamline University

In [3], we began a study of convergence of quadratic forms in independent random variables. Simultaneously, Fau Dyk Tin and G. E. Silov [2] initiated their study of this problem but restricted to the case of quadratic mean convergence and normal variables. Our aim in this paper is to consider carefully the problem of almost sure convergence (convergence with probability one). Several of our results will generalize well known theorems for series of independent random variables.

We shall assume throughout that $X_1, X_2, \cdots$ is a sequence of independent real random variables with $E(X_k) = 0$ and $E(X_k^2) = 1$, $k = 1, 2, \cdots$. Note that we do not assume that the X_k's are identically distributed or place conditions on the higher moments. Let (a_{jk}), $j, k = 1, 2, \cdots$, be a real (not necessarily symmetric) matrix and let

$$S_n = \sum_{j,k=1}^{n} a_{jk} X_j X_k .$$

At various times, we shall place special restrictions on (a_{jk}). We say that (a_{jk}) is *Hilbert-Schmidt* if $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$, that it is *nuclear* if $a_{jk} = \sum_{i=1}^{\infty} b_{ji} c_{ik}$ where (b_{ji}) and (c_{ik}) are Hilbert-Schmidt, and that it is *positive semi-definite* if it is symmetric and $\sum_{j,k=1}^{n} a_{jk} u_j u_k \geqq 0$ for all choices of n and $u_1, \cdots, u_n$. We observe that the class of Hilbert-Schmidt matrices contains all nuclear matrices (use the Schwarz inequality) and all those positive semi-definite matrices with finite trace (use the inequality $\sum_{j,k=1}^{n} a_{jk}^2 \leqq (\sum_{k=1}^{n} a_{kk})^2$).

THEOREM 1. *If (a_{jk}) is Hilbert-Schmidt and $\sum_{k=1}^{n} |a_{kk}| < \infty$, then S_n converges almost surely.*

REMARK. This theorem is the best possible in the following sense. For any Hilbert-Schmidt matrix (a_{jk}) with $\sum_{k=1}^{\infty} |a_{kk}| = \infty$, there is a sequence $X_1, X_2, \cdots$ of independent random variables with $E(X_k) = 0$ and $E(X_k^2) = 1$ for which S_n diverges almost surely. We omit this fairly simple construction.

PROOF. Let K_n, L_n, and M_n be defined in the obvious manner by

$$\begin{aligned} S_n &= \sum_{j=1}^{n} X_j \sum_{k=1}^{j-1} a_{jk} X_k + \sum_{k=1}^{n} X_k \sum_{j=1}^{k-1} a_{jk} X_j + \sum_{k=1}^{n} a_{kk} X_k^2 \\ &= K_n + L_n + M_n . \end{aligned}$$

Now K_n is a martingale and since

$$[E(|K_n|)]^2 \leqq E(K_n^2) = \sum_{j=1}^{n} \sum_{k=1}^{j-1} a_{jk}^2 \leqq \sum_{j,k=1}^{\infty} a_{jk}^2 ,$$

it follows that K_n converges almost surely (martingale convergence theorem).

Received 2 January 1968.

[1] This research was supported by the National Science Foundation under Grant GP-7843.

A similar argument applies to L_n. On the other hand,

$$M_n = \sum_{k=1}^{n} a_{kk}(X_k^2 - 1) + \sum_{k=1}^{n} a_{kk} .$$

The first term, call it P_n, is a martingale and since

$$E(|P_n|) \leqq \sum_{k=1}^{n} | a_{kk} | E(|X_k^2 - 1|) \leqq 2 \sum_{k=1}^{\infty} |a_{kk}| < \infty,$$

it also converges almost surely. That S_n has this property is now an immediate consequence.

Corollary 1. *If* $\sum_{j,k=1}^{\infty} |a_{jk}| < \infty$, *then* S_n *converges almost surely.*

Corollary 2. *If* (a_{jk}) *is nuclear, then* S_n *converges almost surely.*

Corollary 3. *If* (a_{jk}) *is positive semi-definite and* $\sum_{k=1}^{\infty} a_{kk} < \infty$, *then* S_n *converges almost surely.*

A recent result of Gundy [1] allows us to relate the convergence of S_n to that of other random variables. Let $V_k = \sum_{j=1}^{k-1} (a_{kj} + a_{jk})X_j$ and note that

$$S_n = \sum_{k=1}^{n} V_k X_k + \sum_{k=1}^{n} a_{kk} X_k^2.$$

Theorem 2. *Suppose there are positive constants* δ *and* ϵ *such that* $P(|X_k| > \delta) \geqq \epsilon$, $k = 1, 2, \cdots$. *If* $\sum_{k=1}^{\infty} |a_{kk}| < \infty$, *then* S_n, $\sum_{k=1}^{n} V_k^2$, *and* $\sum_{k=1}^{n} V_k^2 X_k^2$ *converge on equivalent sets*, i.e., *on sets which differ at most by null sets.*

Proof. Since $\sum_{k=1}^{\infty} |a_{kk}| < \infty$, $\sum_{k=1}^{n} a_{kk} X_k^2$ converges almost surely just as in the proof of Theorem 1. Thus S_n and $\sum_{k=1}^{n} V_k X_k$ converge on equivalent sets. The result is now a direct application of Gundy's main theorem [1], p. 731. Theorem 2 will play an important role later in the proof of Theorem 7.

For the rest of our results, we shall want to impose a condition on the fourth moment of X_k. Let us suppose that $E(X_k^4) \leqq C < \infty$, $k = 1, 2, \cdots$. Now

$$\sum_{k=1}^{n} a_{kk} X_k^2 = \sum_{k=1}^{n} a_{kk}(X_k^2 - 1) + \sum_{k=1}^{n} a_{kk} ,$$

and the first term on the right of this equality is a sum of independent random variables with means zero and variances

$$a_{kk}^2 E([X_k^2 - 1]^2) \leqq a_{kk}^2 (C + 1)$$

which therefore converges almost surely by the Kolmogorov-Khintchine theorem when $\sum_{k=1}^{\infty} a_{kk}^2$ converges. This means, for example, that in Theorem 2 we may replace the requirement of convergence of $\sum |a_{kk}|$ by that of the two series $\sum a_{kk}$ and $\sum a_{kk}^2$. We may also improve Theorem 1 obtaining

Theorem 3. *If* $E(X_k^4) \leqq C < \infty$, $k = 1, 2, \cdots$, *and if* (a_{jk}) *is Hilbert-Schmidt, then* S_n *converges with probability one or zero according as* $\sum_{k=1}^{\infty} a_{kk}$ *converges or diverges.*

Though it is somewhat out of the context of this paper, we mention that we can do still better if we consider convergence in quadratic mean rather than almost sure convergence.

Theorem 4. *Let* (a_{jk}) *be symmetric and suppose that* $1 < c \leqq E(X_k^4) \leqq C < \infty$, $k = 1, 2, \cdots$. *Then* S_n *converges in quadratic mean if and only if* (a_{jk}) *is Hilbert-Schmidt and* $\sum_{k=1}^{\infty} a_{kk}$ *converges.*

Remark. This result should be compared with Theorem 1 of [2]. We mention also that the special case which occurs as Theorem 1 of [3] is not quite correct. We forgot to note that for the only if part of that theorem, we need $E(X_k^4)$ bounded away from one. Of course, $E(X_k^4) = E([X_k^2 - 1]^2) + 1 \geqq 1$ so this is not a strong condition.

PROOF. Let $Y_k = 2X_k \sum_{j=1}^{k-1} a_{jk} X_j + a_{kk}(X_k^2 - 1)$ so that $S_n - E(S_n) = \sum_{k=1}^{n} Y_k$. Now $Y_1, Y_2 \cdots$ is easily seen to be a sequence of orthogonal random variables with

$$E(Y_k^2) = 4\sum_{j=1}^{k-1} a_{jk}^2 + a_{kk}^2[E(X_k^4) - 1].$$

By a well-known theorem for series of orthogonal random variables, $S_n - E(S_n)$ converges in quadratic mean if and only if $\sum_{k=1}^{\infty} E(Y_k^2) < \infty$, which in turn is seen to be true, if and only if, $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$. Turning to S_n, we note that if $\sum_{j,k=1}^{\infty} a_{jk}^2$ and $\sum_{k=1}^{\infty} a_{kk}$ both converge, then $S_n \equiv S_n - E(S_n) + \sum_{k=1}^{n} a_{kk}$ will converge in quadratic mean. On the other hand, if S_n converges in quadratic mean, then $E(S_n) \equiv \sum_{k=1}^{n} a_{kk}$ must converge and hence $S_n - E(S_n)$ converges in quadratic mean implying in turn that $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$.

We conclude with some results for normal variables.

THEOREM 5. *If X_k is normal and if (a_{jk}) is positive semi-definite, then S_n converges with probability one or zero according as $\sum_{k=1}^{\infty} a_{kk}$ converges or diverges.*

PROOF. If $\sum_{k=1}^{\infty} a_{kk}$ converges, then S_n converges almost surely by Corollary 3. Contrariwise, suppose that $\sum_{k=1}^{\infty} a_{kk}$ diverges (necessarily to $+\infty$ since $a_{kk} \geqq 0$). If S_n did converge on a set M of positive probability say to a random variable S, then

$$E(I_M \exp(-\tfrac{1}{2}S)) > 0,$$

I_M being the indicator of the set M. On the other hand,

$$E(I_M \exp(-\tfrac{1}{2}S)) = \lim_{n\to\infty} E(I_M \exp(-\tfrac{1}{2}S_n)) \leqq \lim_{n\to\infty} E(\exp(-\tfrac{1}{2}S_n)) = 0,$$

a clear contradiction. That the last limit is actually zero is a consequence of the following calculation. Let $\delta_{1n}, \delta_{2n}, \cdots, \delta_{nn}$ be the eigen values of the matrix $(a_{jk})_{j,k=1}^{n}$. Then

$$\begin{aligned} E(\exp(-\tfrac{1}{2}S_n)) &= E(\exp(-\tfrac{1}{2}\textstyle\sum_{k=1}^{n} \delta_{kn} X_k^2)) \\ &= [\textstyle\prod_{k=1}^{n} (1 + \delta_{kn})]^{-\frac{1}{2}} \\ &= [1 + \textstyle\sum_{k=1}^{n} \delta_{kn} + \text{nonnegative terms}]^{-\frac{1}{2}} \\ &= [1 + \textstyle\sum_{k=1}^{n} a_{kk} + \text{nonnegative terms}]^{-\frac{1}{2}} \end{aligned}$$

and the latter goes to zero as n goes to infinity.

COROLLARY 4. *Let X_k be normal. Then $\sum_{k=1}^{n} |a_k| X_k^2$ converges with probability one or zero according as $\sum_{k=1}^{\infty} |a_k|$ converges or diverges.*

COROLLARY 5. *If X_k is normal and $a_{jk} = \sum_{i=1}^{m} b_{ji} b_{ki}$ $(1 \leqq m \leqq \infty)$, then S_n converges with probability one or zero according as $\sum_{j=1}^{\infty} \sum_{i=1}^{m} b_{ji}^2$ converges or diverges. In particular, $\sum_{j,k=1}^{n} b_j b_k X_j X_k$ converges with probability one or zero according as $\sum_{k=1}^{\infty} b_k^2$ converges or diverges.*

A problem suggested by the latter result is the case $a_{jk} = b_j c_k$ which however does not fall under Theorem 5 since (a_{jk}) is not positive semi-definite. We still have the following result.

THEOREM 6. *Let X_k be normal and $a_{jk} = b_j c_k$ where neither of the sequences b_j and c_k consists of all zeros. Then S_n converges almost surely if and only if $\sum_{j=1}^{\infty} b_j^2$ and $\sum_{k=1}^{\infty} c_k^2$ both converge.*

PROOF. The if part follows from Corollary 2. To handle the converse, let us suppose that S_n converges almost surely. Then the characteristic function of S_n is given by (see [3], Example 2)

$$\begin{aligned}\phi_n(t) &= E(\exp(itS_n)) \\ &= [1 - 2it\textstyle\sum_{k=1}^{n} b_k c_k - t^2((\sum_{k=1}^{n} b_k c_k)^2 - \sum_{k=1}^{n} b_k^2 \sum_{k=1}^{n} c_k^2)]^{-\frac{1}{2}}\end{aligned}$$

which must converge to a characteristic function, say $\phi(t)$, as n goes to infinity. But being a characteristic function, there is a $t_0 > 0$ such that $\phi(t_0) \neq 0$. Hence $[\phi(t_0)]^{-2} \equiv \lim_{n\to\infty} [\phi_n(t_0)]^{-2}$ must exist. However, a careful look at the expression for $[\phi_n(t_0)]^{-2}$ convinces one that this limit will not exist if either $\sum b_k^2$ or $\sum c_k^2$ diverge. We remark that we do not know whether a zero-one law holds in this case.

Returning to a more general situation, we prove a zero-one law similar to Theorem 5 but without the positive definiteness condition.

THEOREM 7. *Let X_k be normal and suppose that (a_{jk}) is symmetric with $\sum_{k=1}^{\infty} |a_{kk}| < \infty$. Then S_n converges with probability one or zero according as $\sum_{j,k=1}^{\infty} a_{jk}^2$ converges or diverges.*

PROOF. If $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$, then S_n converges with probability one by Theorem 1. Suppose that $\sum_{j,k=1}^{\infty} a_{jk}^2$ diverges but that S_n converges on a set M of positive probability. Now $\sum_{k=1}^{\infty} a_{kk} X_k^2$ converges with probability one (see proof of Theorem 1) and since

$$S_n = 2\sum_{k=1}^{n} X_k \sum_{j=1}^{k-1} a_{jk} X_j + \sum_{k=1}^{n} a_{kk} X_k^2 = \sum_{k=1}^{n} V_k X_k + \sum_{k=1}^{n} a_{kk} X_k^2,$$

$\sum_{k=1}^{n} V_k X_k$ must converge almost everywhere on M. But by Theorem 2, $\sum_{k=1}^{n} V_k^2$ must then also converge almost everywhere on M, say to a random variable V. We argue now just as in the proof of Theorem 5, noting that $\sum_{k=1}^{n} V_k^2$ is a positive semi-definite form. Thus if I_M is the indicator of M, $E(I_M \exp(-\frac{1}{8}V)) > 0$. But

$$E(I_M \exp(-\tfrac{1}{8}V)) \leq \lim_{n\to\infty} E(\exp(-\tfrac{1}{8}\textstyle\sum_{k=1}^{n} V_k^2)) = 0,$$

a contradiction. To see that the latter limit is zero, consider the following calculation.

$$\begin{aligned}E(\exp(-\tfrac{1}{8}\textstyle\sum_{k=1}^{n} V_k^2)) &= E(\exp(-\tfrac{1}{2}\textstyle\sum_{k=1}^{n} \sum_{j=1}^{k-1} \sum_{i=1}^{k-1} a_{jk} a_{ik} X_j X_i)) \\ &= E(\exp(-\tfrac{1}{2}\textstyle\sum_{i,j=1}^{n} \sum_{k=\max(i,j)+1}^{n} a_{ik} a_{jk} X_i X_j)) \\ &= [1 + \textstyle\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{jk}^2 + \text{nonnegative terms}]^{-\frac{1}{2}}\end{aligned}$$

which goes to zero as n goes to infinity.

REFERENCES

[1] Gundy, Richard F. (1967). The martingale version of a theorem of Marcinkiewicz and Zygmund. *Ann. Math. Statist.* **38** 725–734.

[2] Tin, Fau Dyk and Silov, G. E. (1966). Quadratic functionals in a space with Gaussian measure (Russian). *Uspehi Mat. Nauk.* **21** 226–229.

[3] Varberg, D. E. (1966). Convergence of quadratic forms in independent random variables. *Ann. Math. Statist.* **37** 567–576.