The Annals of Mathematical Statistics
1971, Vol. 42, No. 3, 1079-1083

A BOUND ON TAIL PROBABILITIES FOR QUADRATIC FORMS IN INDEPENDENT RANDOM VARIABLES[1]

BY D. L. HANSON AND F. T. WRIGHT[2]

University of Missouri and University of Iowa

Let X_i for $i = 0, \pm 1, \cdots$ be independent random variables whose distributions are symmetric about zero and such that

$$(1) \qquad P\{|X_i| \geqq x\} \leqq M \int_x^\infty e^{-\gamma t^2} dt$$

for all i and all $x \geqq 0$ where M and γ are positive constants. Suppose a_{ij} for $i, j = 0, \pm 1, \cdots$ are real numbers such that $a_{ij} = a_{ji}$ for all i and j, and such that

$$(2) \qquad \Lambda^2 = \sum_{i,j} a_{ij}^2 < \infty.$$

Let A denote the matrix $((|a_{ij}|))_{i,j=0,\pm 1,\cdots}$, and let $\|A\|$ be the norm of A considered as an operator on l_2, the index on the sequences in l_2 taking on the values $0, \pm 1, \cdots$. Define

$$(3) \qquad S = \sum_{i,j} a_{ij}(X_i X_j - EX_i X_j).$$

The purpose of this paper is to prove:

THEOREM. *Under the assumptions stated above, S exists as a limit, both in quadratic mean and almost surely, of the sequence*

$$\{S_N = \sum_{i,j=-N}^{N} a_{ij}(X_i X_j - EX_i X_j)\},$$

and there exist constants C_1 and C_2 depending on M and γ (but not on the coefficients a_{ij}) such that for every $\varepsilon > 0$

$$(4) \qquad P\{S \geqq \varepsilon\} \leqq \exp(-\min\{C_1\varepsilon/\|A\|, C_2\varepsilon^2/\Lambda^2\}).$$

If $\{Y_k = \sum_v \alpha_v X_{k-v}\}$ is a moving average, then quadratic sums of the form (3) occur naturally when estimating its spectral density. This was the original motivation behind our work.

We would very much like to remove the restriction that the distributions of the X's be symmetric. Unfortunately, our proof depends heavily on this symmetry.

Received August 17, 1970.

[1] Research partially sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Grant No. AFOSR-68-1394.

[2] Part of this research was done while the author held a summer fellowship from the Old Gold Foundation at the University of Iowa.

PROOF OF THEOREM. Let

(5) $$a_{ij}^{(N)} = a_{ij} \quad \text{if} \quad -N \leqq i,j \leqq N,$$
$$=0 \quad \text{otherwise};$$

and let

(6) $$A_N = ((|a_{ij}^{(N)}|))_{i,j=0,=1,\cdots}$$

Define

(7) $$\Lambda_N{}^2 = \sum_{i,j} (a_{ij}^{(N)})^2.$$

A minor re-indexing of our X's enables us to use Theorems 1 and 2 of Varberg [2] to obtain both types of convergence of S_N to S.

The main part of our proof consists, first, of showing that (4) holds with S, A, and Λ replaced by S_N, A_N, and Λ_N respectively, and second, of a "continuity" argument which removes the subscripts.

Before proceeding with the main part of the proof we establish several lemmas.

LEMMA 1. *If* $EX^2 = \sigma^2 < \infty$, *then for* $n = 0, 1, 2, \cdots$ *and* $k = 1, 3, 5, \cdots$ *we have*

$$EX^{2n}(X^2-\sigma^2)^k \geqq \sigma^{2n}E(X^2-\sigma^2)^k$$

if both sides exist and are finite.

LEMMA 2. *If* Z *is* $N(0, 1)$ (*i.e., normal with* $EZ = 0$ *and* $EZ^2 = 1$), *then*

$$EZ^{2n} = \frac{(2n)!}{n!2^n}$$

for $n = 0, 1, 2, \cdots$.

LEMMA 3. *If* Z *is* $N(0, 1)$ *and* $n \geqq 2$, *then* $E(Z^2-1)^n \geqq 1$.

PROOF. For $n \geqq 2$ we have (using Lemma 2)

$$\begin{aligned} E(Z^2-1)^n &= \sum_\alpha [\tbinom{n}{2\alpha}EZ^{2(n-2\alpha)} - \tbinom{n}{2\alpha+1}EZ^{2(n-2\alpha-1)}] \\ &\geqq \tbinom{n}{0}EZ^{2n} - \tbinom{n}{1}EZ^{2n-2} \\ &= [(2n-1)-n]EZ^{2n-2} \geqq 1. \end{aligned}$$

LEMMA 4. *If* Z *is* $N(0, 1)$ *and* X *satisfies* (1) *then there exists* λ *depending only on* M and γ *such that*

$$|EX^{2n}(X^2-EX^2)^v| \leqq \lambda^{2n+2v}EZ^{2n}(Z^2-1)^v$$

for $n = 0, 1, \cdots$ *and* $v = 0, 1, \cdots$.

PROOF. *Case* $v = 0$. Setting $x = y/(2\gamma)^{\frac{1}{2}}$ we get

$$\begin{aligned} EX^{2n} &\leqq \textstyle\int_0^\infty x^{2n}Me^{-\gamma x^2}dx = (2\gamma)^{-n-\frac{1}{2}}M\int_0^\infty y^{2n}e^{-y^2/2}dy \\ &= (2\gamma)^{-n}(M\pi^{\frac{1}{2}}/2\gamma^{\frac{1}{2}})\textstyle\int_{-\infty}^\infty (2\pi)^{-\frac{1}{2}}y^{2n}e^{-y^2/2}\,dy. \end{aligned}$$

The desired inequality holds if

$$\lambda \geqq \lambda_1 = (2\gamma)^{-\frac{1}{2}}(1+M\pi^{\frac{1}{2}}/2\gamma^{\frac{1}{2}}).$$

Case $v = 1, n = 0$. The conclusion of this lemma is $0 = 0$.

Case $v = 1, n \geqq 1$. Using Lemma 2 we get

$$EZ^{2n}(Z^2-1)^1 \geqq 1 > P\{|Z| \geqq 2\}.$$

[Proof continued below.]

Case $v \geqq 2$, *v odd*. Using Lemma 1 and then Lemma 3 we get

$$EZ^{2n}(Z^2-1)^v \geqq E(Z^2-1)^v \geqq 1 > P\{|Z| \geqq 2\}.$$

[Proof continued below.]

Case $v \geqq 2$, v even. $EZ^{2n}(Z^2-1)^v \geqq P\{|Z| \geqq 2\}$. We now continue the proof for these last three cases. The symbols a, b, and C will denote various constants whose exact values do not matter, and which may depend on M and γ but not on n and v. The inequality $1+C \leqq C$ demonstrates our usage. Setting $\tau^2 = EX^2$ and $x = y/(2\gamma)^{\frac{1}{2}}$ we get

$$\begin{aligned}
|EX^{2n}(X^2-\tau^2)^v| &\leqq C^{n+v} + \textstyle\int_\tau^\infty x^{2n}|x^2-\tau^2|^v M e^{-\gamma x^2}\,dx \\
&\leqq C^{n+v}[1+a\textstyle\int_{\tau(2\gamma)^{1/2}}^\infty y^{2n}|y^2-2\gamma\tau^2|^v e^{-y^2/2}\,dy] \\
&\leqq C^{n+v}\left[b+a\int_{2+2\tau\gamma^{1/2}}^\infty y^{2n}(y^2-1)^v\left(\frac{y^2-2\gamma\tau^2}{y^2-1}\right)^v e^{-y^2/2}\,dy\right] \\
&\leqq aC^{n+v}[1+\textstyle\int_{2+2\tau\gamma^{1/2}}^\infty y^{2n}(y^2-1)^v e^{-y^2/2}\,dy] \\
&\leqq bC^{n+v}[1+EZ^{2n}(Z^2-1)^v+\textstyle\int_0^{2+2\tau\gamma^{1/2}} y^{2n}|y^2-1|^v e^{-y^2/2}\,dy]. \\
&\leqq aC^{n+v}[1+EZ^{2n}(Z^2-1)^v].
\end{aligned}$$

In the three cases under consideration we have $1 \leqq EZ^{2n}(Z^2-1)^v/P\{|Z| \geqq 2\}$ so in these cases

$$|EX^{2n}(X^2-\tau^2)^v| \leqq bC^{n+v}EZ^{2n}(Z^2-1)^v.$$

If $\lambda \geqq \max\{\lambda_1, (1+b)C^{\frac{1}{2}}\}$ where b and C come from the inequality above, and λ_1 comes from the case $v = 0$, then the conclusion of the lemma holds.

LEMMA 5. *If A is a real symmetric $n\times n$ matrix, then there exists a real $n\times n$ orthogonal matrix D such that*

$$D^TAD \equiv B = \operatorname{diag}(b_1, \cdots, b_n)$$

where $b_1, \cdots, b_n$ are the eigenvalues of A. It follows that $\operatorname{tr} A = \operatorname{tr} B$, $\operatorname{tr} A^2 = \operatorname{tr} B^2$, *and* $\|A\| = \max |b_i|$.

LEMMA 6. *If X is a random variable with $EX = 0$ and $Ee^{\theta X}$ defined and finite for θ in some open interval about zero, then there exist $\tau > 0$ and $C > 0$ such that for $|\theta| \leqq \tau$ we have $Ee^{\theta X} \leqq e^{C\theta^2}$.*

PROOF. See Lemma 3 in [1] and note that the tails of a distribution go down at least exponentially fast if and only if its moment generating function is finite in some open interval about zero.

We now proceed with the main part of our proof. We will use A_N both as defined in (6) and as the matrix $((|a_{ij}|))_{i,j=-N,\cdots,N}$.

For all $\theta \geqq 0$, $P\{S_N \geqq \varepsilon\} \leqq E\, e^{\theta(S_N-\varepsilon)}$ which equals

$$e^{-\varepsilon\theta} \sum_{k=0}^{\infty} \frac{\theta^k}{k!} E\left[\sum_{i,j=-N}^{N} a_{ij}(X_iX_j - EX_iX_j)\right]^k \tag{8}$$

for $|\theta|$ less than the radius of convergence of the series. Terms in the expansion of $E[\sum_{i,j=-N}^{N} a_{ij}(X_iX_j-EX_iX_j)]^k$ are of the form $\prod_{v=1}^{k} a_{i_v,j_v} \prod_{i=-N}^{N} EX_i^{\alpha_i}(X_i^2-EX_i^2)^{\beta_i}$ since the X_i's are independent and $EX_i \equiv 0$. Using λ from Lemma 4 and letting $\{Z_i\}$ be an i.i.d. sequence of $N(0, 1)$ random variables we have

(i) $EX_i^{\alpha_i}(X_i^2-EX_i^2)^{\beta_i} = \lambda^{\alpha_i+2\beta_i}EZ_i^{\alpha_i}(Z_i^2-1)^{\beta_i} = 0$ from symmetry if α_i is odd,

and

(ii) $|EX_i^{\alpha_i}(X_i^2-EX_i^2)^{\beta_i}| \leqq \lambda^{\alpha_i+2\beta_i}EZ_i^{\alpha_i}(Z_i^2-1)^{\beta_i}$ from Lemma 4 if α_i is even

so that (8) is bounded by

$$e^{-\varepsilon\theta} \sum_{k=0}^{\infty} \frac{\theta^k\lambda^{2k}}{k!} E\left[\sum_{i,j=-N}^{N} |a_{ij}|(Z_iZ_j - EZ_iZ_j)\right]^k \tag{9}$$

$$= e^{-\varepsilon\theta} E \exp[\theta\lambda^2 \textstyle\sum_{i,j=-N}^{N} |a_{ij}|(Z_iZ_j - EZ_iZ_j)] \tag{10}$$

for $0 \leqq \theta <$ the radius of convergence of the series in (9). Now let Z be the vector $(Z_{-N}, \cdots, Z_N)$, let D be a matrix such as is guaranteed by Lemma 5 for A_N, let $W = (W_{-N}, \cdots, W_N) = ZD$, and let $B = D^TA_ND$. Then $W_{-N}, \cdots, W_N$ are independent $N(0, 1)$ random variables and

$$\begin{aligned}\textstyle\sum_{i,j=-N}^{N} |a_{ij}|(Z_iZ_j - EZ_iZ_j) &= ZA_NZ^T - \operatorname{tr} A_N \\ &= (ZD)(D^TA_ND)(D^TZ^T) - \operatorname{tr}(D^TA_ND) = WBW^T - \operatorname{tr} B \\ &= \textstyle\sum_{i=-N}^{N} b_i(W_i^2-1).\end{aligned}$$

Expression (10) becomes

$$e^{-\varepsilon\theta} \textstyle\prod_{k=-N}^{N} E \exp[\theta\lambda^2 b_k(W_k^2-1)] \tag{11}$$

since the random variables W_k^2 are independent, each having the Chi-squared distribution with one degree of freedom. Since $E(W_k^2-1) = 0$ we can apply Lemma 6 to obtain $\tau > 0$ and $C > 0$ such that (11) is bounded by

$$e^{-\varepsilon\theta} \exp\{C\theta^2\lambda^4 \textstyle\sum_{k=-N}^{N} b_k^2\} \tag{12}$$

for $0 \leqq \theta\lambda^2 \max_k |b_k| \leqq \tau$. (Incidentally, a "backwards" inspection of our proof shows that everything we have done so far is valid for θ in this range.) Since $\sum_{i,j=-N}^{N} a_{ij}^2 = \operatorname{tr} A_N{}^2$, from Lemma 5 we see that (12) can be rewritten

$$\exp[-\varepsilon\theta + C\theta^2\lambda^4\Lambda_N{}^2] \tag{13}$$

which is minimized if $\theta = \varepsilon/2C\lambda^4\Lambda_N{}^2$ if that is a permissible value of θ. Set

$$\theta_0 = \min\{\varepsilon/2C\lambda^4\Lambda_N{}^2, \tau/\lambda^2||A_N||\}$$

and recall from Lemma 5 that $||A_N|| = \max_{|i| \leqq N} |b_i|$. Then (13), and therefore (8), is bounded by

$$\exp\{-\theta_0(\varepsilon - C\theta_0\lambda^4\Lambda_N{}^2)\} \leqq e^{-\theta_0\varepsilon/2}.$$

Set $C_1 = \tau/2\lambda^2$ and $C_2 = 1/4C\lambda^4$ and the theorem is proved in the special case $S = S_N$.

Since $S_N \to S$ in quadratic mean, $S_N \to S$ in probability so

$$P\{S \geqq \varepsilon\} \leqq \liminf_N \exp[-\min\{C_1(\varepsilon-\delta)/||A_N||, C_2(\varepsilon-\delta)^2/\Lambda_N{}^2\}]$$

for every $0 < \delta < \varepsilon$. We know $\Lambda_N{}^2 \to \Lambda^2$ and if we can show that $\|A_N\| \to \|A\|$ we can let $N \to \infty$ and then $\delta \downarrow 0$ and be done. Now $\|A_N\|$ has the same value whether A_N is considered to be an operator on $2N+1$ dimensional Hilbert space or an operator on l_2. If $T = ((t_{ij}))$ then $||T|| \leqq [\sum_{i,j} t_{ij}^2]^{\frac{1}{2}}$. Thus

$$||A - A_N||^2 \leqq \sum_{|i|>N \text{ and/or } |j|>N} a_{ij}^2 \to 0.$$

It follows that $\|A_N\| \to \|A\|$.

REFERENCES

[1] HANSON, D. L. (1967). Some results relating moment generating functions and convergence rates in the law of large numbers. *Ann. Math. Statist.* **38** 742–750.

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