Internat. J. Math. & Math. Sci.
Vol. 8 No. 3 (1985) 563-569

AN EDGEWORTH EXPANSION FOR A SUM OF M-DEPENDENT RANDOM VARIABLES

WAN SOO RHEE

Faculty of Management Sciences
The Ohio State University
Columbus, Ohio 43210

(Received February 20, 1985)

ABSTRACT. Given a sequence $X_1, X_2, \ldots, X_n$ of m-dependent random variables with moments of order $3+\alpha$ $(0<\alpha\leq 1)$, we give an Edgeworth expansion of the distribution of $S\sigma^{-1}$ ($S = X_1 + X_2 + \ldots + X_n$, $\sigma^2 = ES^2$) under the assumption that $E[\exp(itS\sigma^1)]$ is small away from the origin. The result is of the best possible order.

KEY WORDS AND PHRASES. Edgeworth Expansion, m-dependent, Berry-Esseen bound, Central Limit Theorem.

1980 MATHEMATICS SUBJECT CLASSIFICATION CODE. P60F05, S60G50

1. INTRODUCTION.

A sequence of random variables (r.v.) $X_1, X_2, \ldots, X_n$ is said to be m-dependent if for each $i\leq j\leq n-m-1$ the two sequences $(X_i)_{i\leq j}$ and $(X_i)_{i>j+m}$ are independent. Let $S = X_1 + \ldots + X_n$, $\sigma^2 = ES^2$. A Berry-Esseen bound of the exact order for the distribution of $S\sigma^{-1}$ has been obtained by V.V. Shergin [1] under the assumption of existence of moments of order $2+\alpha$ $(0<\alpha\leq 1)$.

The purpose of this work is to establish an Edgeworth expansion for the distribution of $S\sigma^{-1}$, under the assumption of the existence of moments of order $3+\alpha$ $(0<\alpha\leq 1)$. Provided that the characteristic function of $E\exp(itS\sigma^{-1})$ is small away from the origin, this bound is of the best possible order ($O(n^{-(1+\alpha)/2})$ in the stationary case). The result is stated and discussed in section 2, section 3 outlines the proof and section 4 contains the various estimates needed.

2. MAIN RESULT AND DISCUSSION.

Let $\Phi(t)$ and $\Psi(t)$ be the distribution and density function respectively of a standard normal random variable. Let $\mu = E(S^3)\sigma^{-3}$,

$$L = (m+1)^2\sigma^{-3}\sum_{j\le n} E|X_j|^3, \quad M = (m+1)^{2+\alpha}\sigma^{-3-\alpha}\sum_{j\le n} E|X_j|^{3+\alpha},$$

$$N = \sum_{|j-k|\le 12m} \sigma^{-5}E|X_j|^3E(X_k^2)$$

THEOREM 1. There exists an universal constant K such that if we set

$$A = M + L^2 + N\log^{3/2}L^{-1} + \exp-(MK)^{-1} + L\exp-(LK)^{-2}$$

and

$$\delta = \operatorname{Sup}\{|E\exp(itS\sigma^{-1})|,\ (KL)^{-1} < |t| < A^{-1}\}$$

then

$$\operatorname*{Sup}_{t\in R}|P(S\sigma^{-1}\le t) - \Phi(t) - \mu^3/6\,(1-t^2)\Psi(t)| \le K(A+\delta A^{-1}).$$

To see what is the order of A, let us specialize to the case when we have a stationary sequence (X_j) such that $\sigma_n^2 = ES_n^2 \to \infty$, where $S_n = \sum_{i=1}^n X_i$. In this case $\sigma_n >$ $\beta\sqrt{n}$, where β is some constant, so we get M of the order $n^{-(1+\alpha)/2}$, L of the order $n^{-1/2}$, and N of the order $n^{-5/2}$. So, in A, the main term is M (or the main terms are L^2 and M if $\alpha = 1$). So, provided δ is small enough the bound given by theorem A is of order $n^{-(1+\alpha)/2}$, the best possible order.

Theorem A gives a first order expansion, but it is clear that the same type of methods will apply for higher order expansion, though the computation becomes rather complex.

3. METHODS.

We first suppose m=1. We will use the following estimate, proved by the author in [2], using ideas of V.V. Shergin:

LEMMA 1. There exists a universal constant K_2, such that if S is a sum of m-dependent random variables then for $K_2|t|L < 1$, we have

$$|E\exp(itS\sigma^{-1})| \le (1+K_2|t|)\operatorname{Max}\{\exp(-t^2/80),\ (K_2|t|L)^{-1/4}\log L\}. \quad (3.1)$$

We shall use Esseen smoothing inequality,

$$\operatorname*{Sup}_{t\in R}|P(S\sigma^{-1}\le t) - \Phi(t) - \mu/6\,(1-t^2)\Psi(t)| \le 24A + \int_{-A^{-1}}^{A^{-1}} J(t)dt, \quad (3.2)$$

where

$$J(t) = |E\exp(itS\sigma^{-1}) - \exp(-t^2/2)(1-i\mu^3t^3/6\sigma^3)|. \quad (3.3)$$

The integral for $(K_2L\,e^{20})^{-1} \le |t| < A^{-1}$ is bounded by

$$2\delta A^{-1} + K_3|\mu|\sigma^{-3}\exp-(K_3L)^{-2}$$

for some constant K_3. Let $T_0 = \operatorname{Inf}(20\log^{1/2}L^{-1},\ M^{-1/2})$.

For $T_0 \leq |t| \leq (K_2 L e^{20})^{-1}$, we have $(K_2|t|L)^{-1/4} \log L \leq L^5$, so it follows easily that the integral is bounded by $K_3(L^2+\exp\{-(MK_3)^{-1}\})$.

To study J(t) for $|t| \leq T_0$, let $f(t) = E \exp(itS\sigma^{-1})$.

In order to simplify notations, we set $U = S\sigma^{-1}$ and $Y_i = X_i\sigma^{-1}$. We have $f'(t) = i\, E\, U\exp(itu)$. For $1 \leq j \leq n$ and $1 \leq \ell \leq 6$, let $U_{j,\ell} = \sum_{k-j > \ell} Y_k$ and let $U_{j,0} = U$. So we have

$$\begin{aligned} f'(t) &= i\, E(\sum_{j\leq n} Y_j \exp(itU)) \qquad (3.4)\\ &= i\, E\{ \sum_{j\leq n} Y_j \exp(itU_{j,1}) \\ &+ \sum_{j\leq n} Y_j(\exp(it(U-U_{j,1}))-1)\exp(itU_{j,2}) \\ &+ \sum_{j\leq n} Y_j(\exp(it(U-U_{j,1})-1)(\exp(it(U_{j,1}-U_{j,2})-1)\exp(itU_{j,3}) \\ &+ \sum_{r=2,3,4} \sum_{j\leq n} Y_j \prod_{0\leq\ell\leq r} (\exp(it(U_{j,\ell}-U_{j,\ell+1}))-1)\exp(itU_{j,r+1}) \\ &+ \sum_{j\leq n} Y_j \prod_{0\leq\ell\leq 5} (\exp(it(U_{j,\ell}-U_{j,\ell+1})-1)\exp(itU_{j,6}) \} \end{aligned}$$

Except for the last term, the last exponential in each term is independent of the first part of the term. The first term has expectation zero. For the second and third, we expand the expectation of the first part, then replace $E(\exp(itU_{j,k}))$, (k=2,3), by $E(\exp(itU))$, modulo a perturbation. We use several time the estimate (3.1) for these computations. The last two terms are bounded more directly. The result is a relation of the type

$$f'(t) = (-t/2 - i\mu t^3/6 + R(t))f(t) + H(t)$$

where R(t) and H(t) are small. Integration of this relation yield the needed estimate for f(t).

The method just described has been used in the stationary case by A.N. Tikhomirov [3]. It does not seem possible to extend his method directly to the general case. However, the estimate (3.1) made this possible. It should be noted that the method used to obtain (3.1) does not seem to extend to establish theorem A.

4. ESTIMATES.

We shall use the fact that for $x \in R$, $|\exp(ix)-1| \leq |x|^{\alpha}$, $|\exp(ix) - 1 - ix| \leq |x|^{1+\alpha}$, $|\exp(ix) - 1 - ix + x^2/2| \leq |x|^{2+\alpha}$.

Let $a_j = i\, E(Y_j \exp(it(U-U_{j,1}))-1)$. Then the above formula give

$$\begin{aligned} a_j &= -tE(Y_j(U-U_{j,1})) - (it^2/2)\, E\, (Y_j(U-U_{j,1})^2) \qquad (4.1)\\ &+ t^{2+\alpha} R_j^1(t) \end{aligned}$$

where $|R_j^1(t)| \leq E|Y_j(U-U_{j,1})|^{2+\alpha}$. Let

$b_j = iE(Y_j(\exp(it(U-U_{j,1}))-1)(\exp(it(U_{j,1}-U_{j,2}))-1))$.

Then

$$b_j = -it^2 E(Y_j(U-U_{j,1})(U_{j,1}-U_{j,2})) + |t|^{2+\alpha} R_j^2(t) \quad (4.2)$$

where

$$|R_j^2(t)| \le E(|Y_j||U-U_{j,1}||U-U_{j,2}|(|U-U_{j,1}|^{\alpha} + |U-U_{j,2}|^{\alpha})).$$

To prove theorem A, we can as well assume $L \le 10^{-3}$ for otherwise by taking $K \ge 10^6$, the inequality will be automatically satisfied. Then, for each j, we have

$$E|Y_j|^2 \le (E|Y_j|^3)^{2/3} \le L^{2/3} \le 10^{-2}.$$

From 1-dependence, for $j \le n$, $\ell \le 4$, we have

$$E\, U_{j,\ell}^2 = E\, U^2 - \sum_{|k-j|\le \ell} E\, Y_k^2 - \sum_{k=j-\ell-1}^{j+\ell} E\, Y_k Y_{k+1}.$$

(this assumes $\ell+1 \le j \le n-\ell-1$; the proof of the estimate below is similar in the other cases). So finally $E\, U_{j,\ell}^2 \ge 1/2$. By using (3.1), where t is changed in $tEU_{j,\ell}^2$, we get that for $8K_2|t|L < 1$, we have

$$\text{for each } j \text{ and } \ell \le 4, \quad |E\exp(itU_{j,\ell})| \le a(t) \quad (4.3)$$

where

$$a(t) = (1+K_2|t|)\operatorname{Max}\{\exp(-t^2/320),\ (8K_2tL)^{-1/4}\log L\}. \quad (4.4)$$

We have

$$\begin{aligned} \exp(itU)-\exp(itU_{j,2}) &= (\exp(it(U-U_{j,2}))-1)\exp(itU_{j,3}) \\ &+ (\exp(it(U-U_{j,2}))-1)(\exp(it(U_{j,2}-U_{j,3}))-1)(\exp(itU_{j,4})) \\ &+ (\exp(it(U-U_{j,2}))-1)(\exp(it(U_{j,2}-U_{j,3}))-1) \\ &\quad (\exp(it(U_{j,3}-U_{j,4}))-1)(\exp(itU_{j,5})) \end{aligned}$$

so we get

$$\begin{aligned} &|E(\exp(itU)-\exp(itU_{j,2})| \quad (4.5) \\ &\le t^2 a(t)\{E(U-U_{j,2})^2 + E|U-U_{j,2}||U_{j,2}-U_{j,3}|\} \\ &+ t^3 E|U-U_{j,2}||U_{j,2}-U_{j,3}||U_{j,3}-U_{j,4}|. \end{aligned}$$

It should be noted that if one uses at this point the cruder estimate $|E\exp(itU_{j,3})| \le 1$, the order of the bound obtained in the stationary case drops from $O(1/n)$ to $O(\log^{1/2} n/n)$. In a similar but simpler way, we get

$$|E(\exp(itU)-\exp(itU_{j,3})| \le t^2\{E(U-U_{j,3})^2 + E|U-U_{j,3}||U_{j,3}-U_{j,4}|\}.$$

The expectation of the term in (3.4) obtained for r=2 is bounded by

$$H_1(t) = t^{3+\alpha} a(t)\ E|Y_j(U-U_{j,1})(U_{j,1}-U_{j,2})||U_{j,2}-U_{j,3}|^{\alpha}.$$

The expectation of the terms obtained for r=3 and 4 is similarly bounded by $2H_1(t)$ and $4H_1(t)$ respectively using the fact that $|\exp(itZ)-1| \le 2$ for all Z. Finally the expectation of the last term is bounded by

$$2|t|^5\ E|Y_j(U-U_{j,1})(U_{j,1}-U_{j,2})|E|U_{j,3}-U_{j,4}||U_{j,4}-U_{j,5}|.$$

It remains to combine these estimates. By 1-depndence, we have

$$\sum_j E\, Y_j(U-U_{j,1}) = \sum_j E\, Y_j U = E\, U^2 = 1$$

and similarly,

$$\sum_j (EY_j(U-U_{j,1})^2 + 2EY_j(U-U_{j,1})(U_{j,1}-U_{j,2})$$

$$= \sum_j EY_j U^2 = EU^3 = \mu.$$

So, from (3.4) we get,

$$f'(t) = (-t-i\mu t^2/2 + t^3R(t))f(t) \qquad (4.6)$$
$$+ (1+t^4)a(t)H_2(t) + (1+t^5)H_3(t),$$

where $|R(t)| \le \sum_j R_j^1(t) + R_j^2(t)$,

and $$|H_2(t)| \le \sum_j E|Y_j(U-U_{j,2})|E((U-U_{j,2})^2 + |U-U_{j,2}||U_{j,2}-U_{j,3}|)$$

$$+ 7\sum_j E|Y_j(U-U_{j,1})(U_{j,1}-U_{j,2})||U_{j,2}-U_{j,3}|^\alpha$$

$$H_3(t) \le \sum_j E|Y_j(U-U_{j,1})|E|U-U_{j,2}||U_{j,2}-U_{j,3}||U_{j,3}-U_{j,4}|$$

$$+ \sum_j E|Y_j(U-U_{j,1})(U-U_{j,2})|E((U-U_{j,3})^2 + |U-U_{j,3}||U_{j,3}-U_{j,4}|)$$

$$+ 2\sum_j E|Y_j(U-U_{j,1})(U_{j,1}-U_{j,2})|E|U_{j,3}-U_{j,4}||U_{j,4}-U_{j,5}| .$$

Now using (4.3) and the c_r-inequality of [4],

$$\sum_j R_j^1(t) \le \sum_j E|Y_j||U-U_{j,1}|^{2+\alpha}$$

$$\le \sum_j (E\,|Y_j|^{3+\alpha})^{1/(3+\alpha)}(E|U-U_{j,1}|^{3+\alpha})^{1/(3+\alpha)}$$

$$\le (\sum_j E|Y_j|^{3+\alpha})^{1/(3+\alpha)}(\sum_j E|U-U_{j,1}|^{3+\alpha})^{1/(3+\alpha)}$$

$$\le K_4M.$$

Similar computations give

$$|R(t)| \le K_5M, \qquad |H_2(t)| \le K_5M, \qquad |H_3(t)| \le K_5N.$$

Let $H(t) = t^3a(t)H_2(t) + (1+t^5)H_4(t)$. We now assume $t\ge 0$, the case $t\le 0$ is similar.

From (4.6), we get $f(t) = G(t)f_1(t)$ where

$$f_1(t) = \exp(-t^2/2 - i\mu t^3/6 + \int_0^t u^3R(u)du)$$

and $G(t) = 1 + \int_0^t H(u)\exp(u^2/2 + i\mu u^3/6 - \int_0^u s^3R(s)ds)du.$

So we have $f(t) = f_1(t) + f_2(t)$, where

$$f_2(t) = \int_0^t H(u)\exp(-t^2/2 + u^2/2 + \int_0^t s^3R(s)ds)du. \qquad (4.7)$$

We have $2M^{1/2}T_0 \le 1$. So for $t \le T_0$, we have

$$\left|\int_u^t s^3 R(s)ds\right| \le (t^2-u^2)(t^2+u^2)K_5 M/4 \le t^2/4 - u^2/4.$$

Hence

$$|f_2(t)| \le \int_0^t H(u)\exp(-t^2/4 + u^2/4)\, du. \tag{4.8}$$

We can also assume that L is small enough so that $80K_2 L \log^{1/2} L^{-1} \le e^{-8}$ for otherwise by taking K lage enough, theorem A will be automatically saisfied. But for $|t| \le T_0 = 20 \log^{1/2} L^{-1}$ we have $\exp(-t^2/320) \ge L^{400/320}$ and $(8K_2|t|L)^{-1/4} \log L \le L^{-2}$, so we have $a(t) \le (1+K_2|t|)\exp(-t^2/320)$. It follows then easily from (4.8) that

$$|f_2(t)| \le K_6\{M(1+t^2)e^{-t^2/320} + (1+t^4)N\} \tag{4.9}$$

On the other hand,

$$|f_1(t) - \exp(-t^2/2)(1+i\mu t^3/6)|$$
$$\le \exp(-t^2/2)\ |\{\exp(i\mu t^3/6 -1 -it^3/6\}|$$
$$+ |\exp(-t^2/2 + i\mu t^3/6)(\exp(K_5 t^4 M) - 1)|.$$

Since $t \le T_0$ and, as already used, we can then suppose $K_5 t^2 M \le t^2/4$.

We get, using the fact that $|e^{-a}-e^{-b}| \le |b-a|\exp(-\inf(a,b))$,

$$|f_1(t) - \exp(-t^2/2)(1+it^3/6)| \tag{4-10}$$
$$\le \exp(-t^2/4)(eK_5 t^4 M + \mu^2 t^6).$$

We have $\mu = EU^3 = \sum_{i,j,k\le n} EY_iY_jY_k$. Considering that $EY_j = 0$ for each j, and that the variables are 1-dependent, $EY_iY_jY_k$ is zero unless there is an ℓ with $i,j,k \in \{\ell, \ell+1, \ell+2\}$. It follows easily that $EU^3 \le K_7 \sum E|Y_i|^3 \le K_7 L$. Estimation of $\int_{-T_0}^{T_0} J(t)dt$ using (4.9) and (4.10) gives the result in the case of 1-dependence.

We reduce the case of m-dependence to 1-dependence by using the standard blocking argument. If $X_1, X_2, \ldots, X_n$ is a m-dependent sequence, for $j \le [n/m]$, we set

$$Z_j = \sum_{i=(j-1)m+1}^{jm} X_i$$

and we set for $jm < n$,

$$Z_{n,m+1} = \sum_{i=jm+1}^{m} X_i.$$

The (Z_j) are 1-dependent; we apply the bound of theorem A to the Z_j, then compute the moments of the Z_j in function of the moments of the X_i using the c_r-inequality.

REFERENCES

[1] SHERGIN, V.V. On the convergence rate in the central limit theorem for m-dependent random variables, Theor. Prob. appl. XXIV (1979), 782-796.

[2] RHEE, W. On the characteristic function of a sum of m-dependent random variables, to appear.

[3] Tikhomirov, A.N. On the convergence rate in the central limit theorem for weakly dependent random variables, Theor. Prob. appl. XXV (1980), 790-809.

[4] LOEVE, M. Probability theory, 3rd edition, VanNostrand, Princeton (1963).

[5] FELLER, W. An introduction to probability theory and its applications, Vol. 2, Wiley, New York (1971).