The Annals of Probability
1974, Vol. 2, No. 3, 522–529

L_1 BOUNDS FOR ASYMPTOTIC NORMALITY OF m-DEPENDENT SUMS USING STEIN'S TECHNIQUE[1]

BY R. V. ERICKSON

Michigan State University

In a recent paper, C. Stein has given a new, direct technique for bounding the error of the normal approximation to the distribution of a sum of dependent random variables, assuming the variables form a stationary sequence with eighth moments. In the present paper we give two L_1 bounds on this error for an arbitrary m-dependent sequence with second moments.

1. Introduction, notation and results. Let $X_1, X_2, \cdots$ be random variables with $EX_k = 0$, $EX_k^2 < \infty$. Set

$$\begin{aligned} s_n^2 &= E(\textstyle\sum_1^n X_k)^2 , \qquad F_n(x) = P(\textstyle\sum_1^n X_k \leqq s_n x) \\ L_{\alpha n} &= \textstyle\sum_1^n E|X_k|^\alpha / s_n^\alpha \qquad \text{and} \qquad R_n = R_{3n} + R_{2n} , \\ R_{3n} &= \textstyle\sum_1^n E(|X_k|^3/s_n^3;\ |X_k| \leqq s_n) , \\ R_{2n} &= \textstyle\sum_1^n E(X_k^2/s_n^2;\ |X_k| > s_n) . \end{aligned}$$

Let $\mathscr{N}$ denote the standard normal distribution with density

$$n(x) = (2\pi)^{-\frac{1}{2}} \exp(-x^2/2) .$$

If $X_1, X_2, \cdots$ are independent, Feller (1968b) shows that $||F_n - \mathscr{N}||_\infty \leqq 6R_n$; and the author (1973) shows $||F_n - \mathscr{N}||_1 < 72R_n$.

In this paper we extend the L_1 result to include m-dependent random variables. To be precise, we prove

THEOREM A. *Suppose $X_1, X_2, \cdots$ are m-dependent. Set $M_3 = (2m + 1)(4m + 1)(6m + 1)$. Then*

$$||F_n - \mathscr{N}||_1 \leqq 13M_3(L_{3n} + L_{4n}) .$$

(See Theorem 3.9 for a somewhat sharper result.)

Since this is of little value when fourth moments fail to exist, we truncate and rescale to obtain

THEOREM B. *Suppose $X_1, X_2, \cdots$ are m-dependent. Set $M_3 = (2m + 1)(4m + 1)(6m + 1)$. Then*

$$||F_n - \mathscr{N}||_1 \leqq 336 M_3 R_n .$$

(See Theorem 4.6 and the paragraph following.)

It should be noted that if $X_1, X_2, \cdots$ are independent, identically distributed

---

Received June 20, 1972; revised October 11, 1973.

[1] Partially supported by NSF GP-28658.

AMS 1970 *subject classifications*. Primary 60F05; Secondary 60F99.

Key words and phrases. L_1 Berry–Esséen, m-dependent, asymptotic normality and error bounds.

then $m = 0$ and we have the "classical rate"

$$||F_n - \mathscr{N}||_1 \leqq \min\{13(A_3/n^{\frac{1}{2}} + A_4/n), 336A_{2+\alpha}/n^{\alpha/2}\}$$

where $A_\beta = E|X_1|^\beta/(EX_1^2)^{\frac{\beta}{2}}$.

The procedure used to prove Theorem 1 is the extremely interesting new technique of Stein (1972), which is direct and makes no use of characteristic functions. It is hoped that our use of this technique to prove an L_1 result will stimulate study of this method. It is noteworthy that the L_1 result proceeds much more simply than the L_∞ result of Stein, and that, in the non independent, identically distributed case, we as yet have found no way to prove the "correct" L_∞ result. All that we have is the trivial

THEOREM C. *Let $X_1, X_2, \cdots$ be m-dependent. Set $M_3 = (2m+1)(4m+1)(6m+1)$. Then*

$$||F_n - \mathscr{N}||_\infty \leqq 17[M_3R_n]^{\frac{1}{2}}\,.$$

In the original version of this paper we included results for the case when $X_1, X_2, \cdots$ satisfy certain mixing conditions. We omit these, however, since in their present rough form they obscure the beautiful technique of Stein, and since they are too crude to yield Theorem A when specialized to the m-dependent case.

We wish to thank Professor Stein for stimulating discussion during the summer of 1972, especially for helping with the bound of E_{11} below which improves our original version and leads to the "correct rate" in Theorem A.

In the final section of this paper we give an example which shows that the so-called "correct rate" for m-dependent variables may be far from correct.

2. Stein's equation. Stein (1969) gives an interesting characterization of the standard normal random variable: W has distribution $\mathscr{N}$ iff $E[f'(W) - Wf(W)] = 0$ for sufficiently many functions f. Thus, to see if W is approximately normal, check the magnitude of $E[f'(W) - Wf(W)]$. It is exactly this that Stein does.

To proceed let us first verify the above characterization. This is done most simply by defining the *normal transform* of any function h in $\mathbf{B} = \{h: R \to R \mid h$ is bounded, measurable$\}$:

$$\begin{aligned} h^{\mathscr{N}}(x) &= \textstyle\int_{-\infty}^{x} [h(y) - \nu h]\varphi(y)\,dy/\varphi(x) \\ &= -\textstyle\int_{x}^{\infty} [h(y) - \nu h]\varphi(y)\,dy/\varphi(x)\,, \end{aligned}$$

where $\nu h = \int_{-\infty}^{\infty} h(y)\varphi(y)\,dy = Eh(\mathscr{N})$. (We shall consistently use $\mathscr{N}$ for both the standard normal distribution and its corresponding random variable on any appropriate probability space.) Further, define

$$h^{\mathscr{N}d}(x) = xh^{\mathscr{N}}(x) + h(x) - \nu h\,.$$

(2.1) PROPOSITION. *For any h in $\mathbf{B}$ and any finite a, b, we have*

$$||h^{\mathscr{N}}||_\infty \leqq 4||h||_\infty\,, \qquad ||h^{\mathscr{N}d}||_\infty \leqq 4||h||_\infty$$

and $\int_a^b h^{\mathscr{N}d}(x)\,dx = h^{\mathscr{N}}(b) - h^{\mathscr{N}}(a)$.

PROOF. When $x<0$, $|h^{\mathscr{N}}(x)| \leqq 2\|h\|_\infty \mathscr{N}(x)/\varphi(x)$, and $\mathscr{N}(x)/\varphi(x) \leqq \min(1/|x|, 1/2\varphi(x)) \leqq 2$ (for the last use $x = -\frac{1}{2}$ and for the first notice that $\int_{-\infty}^x \varphi(y)\,dy \leqq \int_{-\infty}^x -y\varphi(y)\,dy/(-x) = \varphi(x)/(-x)$). This same reasoning shows that $|x|\mathscr{N}(x)/\varphi(x) \leqq 1$, and hence $\|h^{\mathscr{N}d}\|_\infty \leqq 4\|h\|_\infty$. The last assertion follows by using Fubini's theorem. □

The characterization of normal is now obvious: for each h in $\mathbf{B}$ set $f = h^{\mathscr{N}}$; if $W = \mathscr{N}$ then $E[f'(\mathscr{N}) - \mathscr{N}f(\mathscr{N})] = \nu h - \nu h = 0$, and conversely if for all h in $\mathbf{B}$, $0 = E(f'(W) - Wf(W)) = Eh(W) - Eh(\mathscr{N})$ then $W = \mathscr{N}$.

The next step in Stein's technique is to bound $E[f'(W) - Wf(W)] = Eh(W) - Eh(\mathscr{N})$, h in $\mathbf{B}$, $f = h^{\mathscr{N}}$, $f' = h^{\mathscr{N}d}$. This is done in great generality in Stein (1972), but here we shall deal only with the m-dependent case.

Fix $n > m$, and from now on *Assume*

$$\{X_1\}_1^n \text{ is } m\text{-dependent, i.e. for all } i < n - m$$
$$\sigma\{X_1, \cdots, X_i\} \text{ is independent of } \sigma\{X_{i+m+1}, \cdots, X_n\}$$

and use the Notation

$$A_i = X_i/s_n\,, \qquad B_i = \textstyle\sum_{|j-i|\leqq m} A_j\,, \qquad C_i = \sum_{|j-i|\leqq 2m} A_j$$
$$D_i = \textstyle\sum_{|j-i|\leqq 3m} A_j\,, \qquad Q_i = A_iB_i - EA_iB_i\,, \qquad S = \sum A_j$$
$$b_i = S - B_i\,, \qquad c_i = S - C_i\,, \qquad d_i = S - D_i\,,$$
$$f = h^{\mathscr{N}}\,, \qquad f' = h^{\mathscr{N}} \qquad \text{for } h \text{ in } \mathbf{B}\,.$$

Summation is always taken over the indicated range and between 1 and n.

Under the assumptions it is clear that the entries in the pair's (A_i, b_i), (B_i, c_i), (C_i, d_i) and (Q_i, c_i) are independent.

(2.2) PROPOSITION (Stein's equation). *Suppose $L_{4n} < \infty$. Under the above assumptions for each h in $\mathbf{B}$, $Eh(S) - Eh(\mathscr{N}) = \sum_1^4 E_4$, where $E_k = E_k(h)$ is given by*

$$E_1 = \textstyle\sum E[f'(S) - f'(c_i)]Q_i$$
$$E_2 = \tfrac{1}{2} \textstyle\sum EA_iB_i^2 f(S)$$
$$E_3 = \textstyle\sum EA_i \int_S^{b_i} z \int_S^z f'(u)\,du\,dz$$
$$E_4 = \textstyle\sum EA_i \int_S^{b_i} [h(z) - h(S)]\,dz\,.$$

PROOF. Observe that

$$Eh(S) - Eh(\mathscr{N}) = E[f'(S) - Sf(S)] = E\{f'(S) + \textstyle\sum A_i[f(b_i) - f(S)]\}\,.$$

Now write $f(b) - f(s) = \int_s^b [f'(z) - f'(s)]\,dz + (b - s)f'(s)$ and $f'(z) - f'(s) = (z - s)f(s) + z\int_s^z f'(u)\,du + h(z) - h(s)$.

Formal substitution produces terms E_2, E_3 and E_4 and would force $E_1 = Ef'(S)[1 - \sum A_iB_i]$. The expression for E_1 now follows because m-dependence implies both that $1 = ES^2 = \sum EA_iB_i$ and that c_i and Q_i are independent. Under the assumption that L_{4n} is finite it is easy to show that each E_k is finite (more details are given in the succeeding propositions), and the formal substitution is justified. □

3. Proof of Theorem A. We would have a nice L_∞ Berry–Esséen result for m-dependent random variables if we could show $|E_k(h)| \leqq$ const. $||h||_\infty(L_{3n} + L_{4n})$. We cannot do this; nor do we see how this can be done replacing $(L_{3n} + L_{4n})$ by something of the form $\sum_{k=3}^{8} L_{kn}$, as in Stein (1972).

But we can do something for an L_1 bound, $||F_n - \mathscr{N}||_1$. In order to obtain such a bound take $h = h_x = I_{(-\infty,x]}$ so that $Eh(S) - Eh(\mathscr{N}) = F_n(x) - \mathscr{N}(x) = \sum_1^4 E_k(h_x)$. We show how to bound

$$E_{k1} = \int_{-\infty}^{\infty} |E_k(h_x)|\, dx\,.$$

First we must prove something like (2.1).

(3.1) PROPOSITION. *Set* $h = h_x = I_{(-\infty,x]}$, $f_x = h^{\mathscr{N}}$, $f_x' = h^{\mathscr{N}d}$. *For each real* y *we have*

$$\int_{-\infty}^{\infty} |f_x(y)|\, dx = 1\,, \qquad \int_{-\infty}^{\infty} |f_x'(y)|\, dx \leqq 3\,.$$

PROOF. From the definition of the normal transform it is clear that

$$\begin{aligned} f_x(y) &= \mathscr{N}(y)[1 - \mathscr{N}(x)]/n(x) \qquad &\text{if} \quad y \leqq x \\ &= \mathscr{N}(x)[1 - \mathscr{N}(y)]/n(x) \qquad &\text{if} \quad y \geqq x\,. \end{aligned}$$

The first equality now follows by direct calculation, using Fubini's theorem and expressing $\mathscr{N}(y) = \int_{-\infty}^{y} n(u)\, du$. Let $L_1(y)$ denote the left side of the second inequality. Similarly it is shown that

$$L_1(y) = 2n(y)[1 + y\mathscr{N}(y)/n(y)][1 - y(1 - \mathscr{N}(y))/n(y)]\,.$$

If use is made of the fact (Feller (1968a), page 193) $p_{2k+1}(x) \leqq [1 - \mathscr{N}(x)]/n(x) \leqq p_{2k}(x)$, $k = 0, 1, \cdots, x > 0$, where $p_k(x) = x^{-1} + \sum_1^k (-1)^j[1\cdot 3\cdots(2j-1)]x^{-(2j+1)}$, it is easy to show that $L_1(y) \leqq 3$, $|yL_1(y)| \leqq 3$ and $\lim_{|y|\to\infty} |yL_1(y)| = 2$. $\square$

Before proceeding to give bounds for E_{k1} we state an extension of an inequality of Loéve (1963), page 155:

(3.2) PROPOSITION (c_r-inequality). *Given random variables* $U_1, \cdots, U_k$, *set* $c_{rk} = \max(1, k^{r-1})$. *Then* $E|U_1 + \cdots + U_k|^r \leqq c_{rk} \sum_1^k E|U_j|^r$.

To get the best bounds we randomize, as in Stein (1972). Let J be a random variable independent of all $X_1, \cdots, X_n$ and let J be uniform on the set of indices $\{1, \cdots, n\}$. Now consider the randomized random variables A_J, B_J, Q_J.

(3.3) PROPOSITION. *For* $\alpha \geqq 1$ *we have* $E|A_J|^\alpha \leqq L_{\alpha n}/n$, $E|B_J|^\alpha \leqq (2m + 1)^\alpha L_{\alpha n}/n$, $E|Q_J|^\alpha \leqq 2^\alpha E|A_J B_J|^\alpha$.

PROOF. Note first that $E|A_J|^\alpha = n^{-1} \sum E|A_j|^\alpha = n^{-1}L_{\alpha n}$. Next, by (3.2), $E|B_J|^\alpha = \sum E|B_j|^\alpha/n \leqq (2m+1)^{\alpha-1} \sum_j \sum_{|k-j|\leqq 2m} E|A_k|^\alpha/n \leqq (2m+1)^\alpha L_{\alpha n}/n$. Finally, again by (3.2) and also the moment inequality, $E|Q_J|^\alpha \leqq 2^{\alpha-1}[E|A_J B_J|^\alpha + E|q_J|^\alpha]$, where $|q_j|^\alpha = |E(A_j B_j)|^\alpha \leqq E|A_j B_j|^\alpha$. $\square$

Similar results are true for $E|C_J|^\alpha$, $E|C_J - D_J|^\alpha$, $\cdots$ and these will be used without further comment. Another inequality typical of what will be used below is

(3.4) PROPOSITION. *Set* $K_J = C_J - D_J$. *Then* $nE|C_J K_J Q_J| \leqq 2(2m+1)(4m+1)(2m)L_{4n}$.

PROOF. Use a three term Hölder inequality to get

$$\|C_J K_J Q_J\|_1 \leqq \|C_J\|_4 \|K_J\|_4 \|Q_J\|_2 \leqq 2\|C_J\|_4\|K_J\|_4\|A_J\|_4\|B_J\|_4$$
$$\leqq 2[(4m+1)^4(L_{4n}/n)(2m)^4(L_{4n}/n)(L_{4n}/n)(2m+1)^4 L_{4n}/n]^{\frac{1}{4}} . \quad \square$$

We are now in a position to find appropriate bounds for E_{k1}.

(3.5) PROPOSITION.

$$E_{11} \leqq 10(2m+1)(4m+1)L_{3n} + 6(2m+1)(4m+1)(8m+1)L_{4n} .$$

PROOF. We may certainly assume $L_{4n} < \infty$, else there is nothing to prove. From the definition of f_x' we have $E_1(h_x) = nE[f_x'(S) - f_x'(c_J)]Q_J = nE\{C_J f_x(S) Q_J + c_J Q_J \int_{cJ}^{S} f_x'(u)\,du + [h_x(S) - h_x(c_J)]Q_J\}$. Now $\int_{-\infty}^{\infty} |h_x(S) - h_x(c_J)|\,dx = |C_J|$, and (3.1) implies

$$E_{11} \leqq nE|C_J Q_J| + 3nE|c_J C_J Q_J| + nE|C_J Q_J| .$$

As in (3.4), $E|C_J Q_J| \leqq \|C_J\|_3 \|Q_J\|_{\frac{3}{2}} \leqq 2\|C_J\|_3\|A_J\|_3\|B_J\|_3 \leqq 2(2m+1)(4m+1)L_{3n}/n$. Next $c_J = d_J + (D_J - C_J)$ and d_j is independent of $C_j Q_j$. Thus

$$E|c_J C_J Q_J| \leqq E|d_J C_J Q_J| + E|K_J C_J Q_J| , \qquad K_J = D_J - C_J .$$

Finally, since $E|d_j| \leqq E|S| + E|D_j| \leqq 1 + \delta_j$, $\delta_j = E|D_j|$, we have $E|d_J C_J Q_J| = n^{-1} \sum E|C_j Q_j| E|d_j| \leqq E|C_J Q_J| + E|\delta_J C_J Q_J|$. But $E|\delta_J|^4 \leqq n^{-1} \sum E|D_j|^4 \leqq (6m+1)^4 L_{4n}/n$, and we argue as in (3.4) to obtain $E|\delta_J C_J Q_J| \leqq \|\delta_J\|_4 \|C_J\|_4 \|Q_J\|_2 \leqq 2(6m+1)(4m+1)(2m+1)L_{4n}/n$. Adding all of the estimates gives the result. $\square$

(3.6) PROPOSITION. $E_{21} \leqq \frac{1}{2}(2m+1)^2 L_{3n}$.

PROOF. $2E_{21} \leqq nE|A_J B_J^2| \leqq n\|A_J\|_3 \|B_J^2\|_{\frac{3}{2}} = n\|A_J\|_3\|B_J\|_3^2 \leqq n(L_{3n}/n)^{\frac{1}{3}}((2m+1)^3 L_{3n}/n)^{\frac{2}{3}}$. $\square$

(3.7) PROPOSITION. $E_{31} \leqq \frac{3}{2}(2m+1)^2 L_{3n} + 3[\frac{5}{6}(2m+1)^3 + \frac{1}{2}(2m+1)^2(6m+1)]L_{4n}$.

PROOF. Since $|z| \leqq |z - S| + |S|$ we have

$$\tfrac{1}{3}E_{31} \leqq nE|A_J \textstyle\int_S^{bJ} |z|\,|z-S|\,dz|$$
$$\leqq \frac{n}{3} E|A_J B_J^3| + \frac{n}{2} E|A_J B_J^2 S| .$$

Write $S = B_J + c_J + K_J$, $K_J = C_J - B_J$ to get $\frac{1}{3}E_{31} \leqq \frac{5}{6}H_1 + \frac{1}{2}H_2 + \frac{1}{2}H_3$, where $H_1 = nE|A_J B_J^3| \leqq n\|A_J\|_4\|B_J\|_4^3 \leqq (2m+1)^3 L_{4n}$, $H_2 = nE|A_J B_J^2 K_J| \leqq (2m+1)^2(2m)L_{4n}$ and (as in the proof of (3.5)) $H_3 = nE|A_J B_J^2 c_J| \leqq nE|A_J B_J^2| + nE|\gamma_J A_J B_J^2| \leqq (2m+1)^2 L_{3n} + (2m+1)^2(4m+1)L_{4n}$, where $\gamma_j = E|C_j|$. $\square$

(3.8) PROPOSITION. $E_{41} \leqq \frac{1}{2}(2m+1)^2 L_{3n}$.

PROOF. $\int_{-\infty}^{\infty} |h_x(z) - h_x(S)|\,dx = |z - S|$, and thus $E_{41} \leqq (n/2)E|A_J B_J^2|$. $\square$

Addition gives the next result, of which Theorem A is a simple corollary.

(3.9) THEOREM. *Suppose $X_1, X_2, \cdots$ are m-dependent. Then for all n*

$$\|F_n - \mathscr{N}\|_1 \leqq (12.5)(2m+1)(3.6m+1)L_{3n} + 10(2m+1)(22m^2+9.4m+1)L_{4n}.$$

Let us point out the problem in getting L_∞ bounds by the above methods. For the L_1 bound we encountered $\int |h_x(z) - h_x(S)|\,dx = |z-S|$, $h_x = I_{(-\infty,x]}$ both in E_{11} and E_{41}; in the L_∞ case this is instead $\sup_x |h_x(z) - h_x(S)| = 1$. This is exactly the problem. Some bounds can be given, but they lead to the square root of the "correct bound" and a proof of Theorem C which is not as nice as our geometric proof using Theorem B. (See the last paragraphs of Section 4.)

4. Proof of Theorem B and C. We prove Theorem B by giving an L_1 bound for the error committed by truncation. We also give results which apply to the L_∞ case. To do this we first prove

(4.1) PROPOSITION. *Let U and V be random variables with distributions G and H. Then*

$$\|G-H\|_1 \leqq E|U-V|, \qquad \|G-H\|_\infty \leqq P(U \neq V).$$

PROOF. The L_∞ inequality follows from the identity $G(x) - H(x) = P(U \leqq x, U \neq V) - P(V \leqq x, U \neq V)$. For the L_1 inequality define $I_a = I_{(-\infty,a]}$, so that $\|G-H\|_1 = \int_{-\infty}^{\infty} |EI_a(U) - EI_a(V)|\,da \leqq E\int_{-\infty}^{\infty} |I_a(U) - I_a(V)|\,da = E|U-V|$. (This trick now has been used three times.) □

(4.2) PROPOSITION. *Let $a > 0$ and define $\mathscr{N}_a(x) = \mathscr{N}(ax)$. Then $\|\mathscr{N}_a - \mathscr{N}\|_1 \leqq \frac{4}{5}|1-a^{-1}|$ and $\|\mathscr{N}_a - \mathscr{N}\|_\infty \leqq \frac{4}{5}|1-a^{-1}|$.*

PROOF. Let U have distribution $\mathscr{N}$ and $V = U/a$ have distribution $\mathscr{N}_a$. Then $E|U-V| = |1-a^{-1}|E|U| \leqq \frac{4}{5}|1-a^{-1}|$ and the L_1 part follows from (4.1). For the L_∞ part, first take $a \geqq 1$ and notice that $|\mathscr{N}(ax) - \mathscr{N}(x)| = |\int_x^{ax} n(y)\,dy| \leqq |a-1|\,|xn(x)|$. Since $|xn(x)|$ is maximal at $|x| = 1$ and $n(1) < \frac{4}{5}$ we have the result, for $\|\mathscr{N}_a - \mathscr{N}\|_\infty = \|\mathscr{N} - \mathscr{N}_{a^{-1}}\|_\infty$. □

Let us use these in the following truncation setup. Set $\tau^2 = s_n^2 = E(\sum X_k)^2$; define $X_k^\tau = X_k I(|X_k| \leqq \tau)$, and

$$Y_k = X_k^\tau - EX_k^\tau, \qquad Z_k = X_k - Y_k.$$

Set $T = \sum Y_k$, $t_n^2 = ET^2$, $r = s_n/t_n$. Finally, write $F_n - \mathscr{N} = H_1 + H_2 + H_3$, $H_1(x) = F_n(x) - P(T \leqq s_n x)$, $H_2(x) = P(T \leqq t_n rx) - \mathscr{N}(rx)$ and $H_3(x) = \mathscr{N}_r(x) - \mathscr{N}(x)$.

We will bound H_k using $L_{\alpha n}(Z) = \sum E|Z_k|^\alpha/s_n^\alpha$ and $L_{\alpha n}(Y) = \sum E|Y_k|^\alpha/s_n^\alpha$. From (4.1) follows

(4.3) PROPOSITION. $\|H_1\|_1 \leqq L_{1n}(Z)$.

Define $M_k = \prod_{j=1}^k (2jm+1)$.

From (3.9) and the change of variable $y = rx$ follows

(4.4) PROPOSITION. *We have*

$$||H_2||_1 \leqq (12.5)r^2 M_2 L_{3n}(Y) + 10r^3 M_3 L_{4n}(Y) .$$

Clearly $||H_3||_1 \leqq \frac{4}{5}|1 - r^{-2}|$, because of (4.2). Let us bound this.

(4.5) PROPOSITION. $|1 - r^{-2}| \leqq M_1[4L_{1n}(Z) + L_{2n}(Z)]$.

PROOF. Notice that $1 - r^{-2} = (s_n^2 - t_n^2)/s_n^2$ and that $s_n^2 - t_n^2 = \mathrm{Var}(\sum X_k) - \mathrm{Var}(T) = \mathrm{Var}(\sum Z_k) + 2\,\mathrm{Cov}(T, \sum Z_k)$. Because of m-dependence $\mathrm{Cov}(Z_k, Z_j) = 0 = \mathrm{Cov}(Y_k, Z_j)$ if $|j - k| > m$. In general, $|2\,\mathrm{Cov}(Z_k, Z_j)| \leqq EZ_k^2 + EZ_j^2$, and $|\mathrm{Cov}(Y_k, Z_j)| \leqq E|Y_k Z_j| \leqq 2s_n E|Z_j|$. The result follows by writing $s_n^2 - t_n^2 = \sum [EZ_k^2 + 2\,\mathrm{Cov}(Y_k, Z_k)] + \sum_{k=1}^m \sum_{j=1}^{n-k} [2\,\mathrm{Cov}(Z_j, Z_{k+j}) + 2\,\mathrm{Cov}(Z_j, Y_{k+j}) + 2\,\mathrm{Cov}(Z_{k+j}, Y_j)]$. □

Combining these we prove

(4.6) THEOREM. $||F_n - \mathscr{N}||_1 \leqq 14(K_1 + K_2)$, *where* $K_1 = M_1[4L_{1n}(Z) + L_{2n}(Z)]$ *and* $K_2 = M_2 L_{3n}(Y) + M_3 L_{4n}(Y)$.

PROOF. By (4.1) we know that $||F_n - \mathscr{N}||_1 \leqq E|S - \mathscr{N}| \leqq ||S - \mathscr{N}||_2 = [\mathrm{Var}(S) + \mathrm{Var}\,\mathscr{N}]^{\frac{1}{2}} \leqq 2^{\frac{1}{2}}$. Thus we may assume $K_1 \leqq 2^{\frac{1}{2}}/14$, and from $1 - r^{-2} \leqq K_1$ it follows that $r^2 \leqq 1.12$, $r^3 \leqq 1.2$. Substitution of these values in the bound for $||H_2||_1$ and addition of the bounds for $||H_k||_1$ give the result. □

To prove Theorem B use the c_r-inequality (3.2) and the moment inequality to see that $L_{3n}(Y) \leqq 8R_{3n}$, $L_{4n}(Y) \leqq 16R_{3n}$, while $L_{1n}(Z) \leqq 2R_{2n}$, $L_{2n}(Z) \leqq R_{2n}$.

To prove Theorem C, if $||F_n - \mathscr{N}||_\infty = h$, draw a triangle with height h and base $5h/2$ completely between the graphs of F_n and $\mathscr{N}$. This can be done since $\mathscr{N}$ has maximum slope $n(0) < \frac{2}{5}$. This triangle has area equal $5h^2/4 \leqq ||F_n - \mathscr{N}||_1$ and we see that $||F_n - \mathscr{N}||_\infty \leqq 2[||F_n - \mathscr{N}||_1/5]^{\frac{1}{2}} \leqq 16.5(M_3 R_n)^{\frac{1}{2}}$.

5. An example. Suppose $X_1, X_2, \cdots$ are identically distributed and 1-dependent and that $EX_1^4 < \infty$. Then (3.9) implies that

$$||F_n - \mathscr{N}||_1 \leqq 1000Kn(s_n^{-3} + s_n^{-4}) = B_n$$

where $E|X_1|^3, EX_1^4 \leqq K < \infty$. We will construct a sequence so that

$$B_n > Cn^{\frac{1}{4}}\,, \qquad ||F_n - \mathscr{N}||_1 \leqq Cn^{-\frac{1}{4}}$$

for some constant C, $0 < C < \infty$. This will show that while our bounds in terms of $L_{\alpha n}$ look effective, in fact they are not the proper generalization of the classical Berry–Esséen theorem.

In order to construct such a sequence we merely have to choose the X_i in such a way that

$$s_n^2 = E(\textstyle\sum_1^n X_k)^2 \sim n^{\frac{1}{2}}$$

while $S_n = \sum_1^n X_k$ consists only of essentially $n^{\frac{1}{2}}$ independent, identically distributed (i.i.d.) terms.

To do this let $Y_1, Y_2, \cdots$ be i.i.d. with $EY_1 = 0$, $EY_1^2 = 1$, $EY_1^4 = K < \infty$. Notice that if $X_{2k-1} = Y_k$, $X_{2k} = -Y_k$ then $s_n^2 = 0$ if n even, $= 1$ if n odd. If we do not force quite so much cancellation we have the result we want.

More precisely, define X_k inductively as follows. Set $X_1 = Y_1$, $X_2 = Y_2$. After k terms are defined, when $X_k = Y_j$, define $X_{k+1} = -Y_j$ if $s_k^2 > k^{\frac{1}{2}}$ and $X_{k+1} = +Y_{j+1}$ if $s_k^2 \leqq k^{\frac{1}{2}}$; on the other hand, if $X_k = -Y_j$ just define $X_{k+1} = Y_{j+1}$. A graph makes it clear that $|s_n^2 - n^{\frac{1}{2}}| \leqq 2$. Therefore $\sum_1^n X_k$ contains approximately $n^{\frac{1}{2}}$ nonzero i.i.d. terms. If we apply (3.9) with $m = 0$ to this i.i.d. sequence we get $\|F_n - \mathscr{N}\|_1 \leqq C' n^{\frac{1}{2}}(n^{-\frac{3}{4}} + n^{-\frac{1}{4}}) < Cn^{-\frac{1}{4}}$. The result now follows from the definition of B_n.

REFERENCES

ERICKSON, R. V. (1973). On an L_p version of the Berry-Esséen theorem for independent and m-dependent variables. *Ann. Probability* **1** 497-503.

FELLER, W. (1968a). *An Introduction to Probability Theory and Its Applications* **1** 3rd. ed. Wiley, New York.

FELLER, W. (1968b). On the Berry-Esséen theorem. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **10** 261-268.

LOÈVE, M. (1963). *Probability Theory.* 3rd ed. Van Nostrand, Princeton.

STEIN, C. (1969). *Multivariate Analysis I*, Technical report No. 42, Department of Statistics, Stanford Univ.

STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 583-602.

DEPARTMENT OF STATISTICS AND PROBABILITY
WELLS HALL
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48823