A BOUND FOR THE ERROR IN THE NORMAL APPROXIMATION TO THE DISTRIBUTION OF A SUM OF DEPENDENT RANDOM VARIABLES

CHARLES STEIN
STANFORD UNIVERSITY

1. Introduction

This paper has two aims, one fairly concrete and the other more abstract. In Section 3, bounds are obtained under certain conditions for the departure of the distribution of the sum of n terms of a stationary random sequence from a normal distribution. These bounds are derived from a more abstract normal approximation theorem proved in Section 2. I regret that, in order to complete this paper in time for publication, I have been forced to submit it with many defects remaining. In particular the proof of the concrete results of Section 3 is somewhat incomplete.

A well known theorem of A. Berry [1] and C-G. Esséen [2] asserts that if $X_1, X_2, \cdots$ is a sequence of independent identically distributed random variables with $EX_i = 0, EX_i^2 = 1$, and $\beta = E|X_i|^3 < \infty$, then the cumulative distribution function of $(1/\sqrt{n}) \Sigma_{i=1}^n X_i$ differs from the unit normal distribution by at most $K\beta/\sqrt{n}$ where K is a constant, which can be taken to be 2. It seems likely, but has never been proved and will not be proved here, that a similar result holds for stationary sequences in which the dependence falls off sufficiently rapidly and the variance of $(1/\sqrt{n}) \Sigma_{i=1}^n X_i$ approaches a positive constant. I. Ibragimov and Yu. Linnik ([3], pp. 423–432) prove that, under these conditions, the limiting distribution of $(1/\sqrt{n}) \Sigma X_i$ is normal with mean 0 and a certain variance σ^2. Perhaps the best published results on bounds for the error are those of Phillip [5], who shows that if in addition the X_i are bounded, with exponentially decreasing dependence, then the discrepancy is roughly of the order of $n^{-1/4}$. In Corollary 3.2 of the present paper it is proved that under these conditions the discrepancy is of the order of $n^{-1/2}(\log n)^2$. Actually the assumption of boundedness is weakened to the finiteness of eighth moments. In Corollary 3.1 it is proved that if the assumption of exponential decrease of dependence is strengthened to m dependence, the error in the normal approximation is of the order of $n^{-1/2}$.

The abstract normal approximation theorem of Section 2 is elementary in the sense that it uses only the basic properties of conditional expectation and the elements of analysis, including the solution of a first order linear differential equation. It is also direct, in the sense that the expectation of a fairly arbitrary function of the random variable W in question is approximated directly without going through the characteristic function. Because of the clumsiness of the technique used at this point, it is likely that slightly better results could be obtained by using the basic identity, Lemma 2.1, to approximate the characteristic function of W and, by standard procedures, to go from this to an approximation for the distribution of W. However, I believe that, in the long run, better results will be obtained by direct methods.

I hope it will be helpful to read the following remarks along with Section 2 and the first part of Section 3. The whole paper is based on the moderately simple Lemma 2.1, which is only a recasting of the trivial identity (2.10) in order to make it fairly apparent that the normal approximation applies. The basic idea of this trivial identity is to apply the defining properties of conditional expectation, to obtain an identity containing an arbitrary function, in a situation in which one forgets a random part of the data. Thus we consider three σ-algebras of events, the σ-algebra $\mathscr{F}$ determined by the originally given random variables (in Section 3, a sequence $X_1, \cdots, X_n$), a larger σ-algebra $\mathscr{B}$ allowing one or more other random variables (in Section 3, a random index I uniformly distributed over $\{1, \cdots, n\}$, independent of $X_1, \cdots, X_n$), and a sub σ-algebra $\mathscr{C}$ in which a random part of the data is forgotten. In Section 3, roughly speaking, $\mathscr{C}$ is the σ-algebra determined by I and the $\{X_j\}$ for $|j - I| > m$, where m is an appropriately chosen positive number. Lemma 2.1 is obtained by expressing the arbitrary function f occurring in formula (2.10) in terms of another arbitrary function h by formula (2.9) in such a way as to make the normal approximation apparent. I hope to publish soon identities making other approximations apparent. Theorem 2.2 is included only in order to indicate, in detail but without excessive complications, how Lemma 2.1 leads to a result on the normal approximation. The idea behind the more complicated derivation of Theorem 2.1 from Lemma 2.1 is described immediately after the proof of Theorem 2.2.

Note added in proof. It has come to my attention that some results giving sharp bounds for the order of magnitude of the error in the normal approximation in similar problems have been obtained by Statulevičius. His method is entirely different, and in some ways his results go beyond those of this paper, including results for large deviations. Several of his papers are published in *Lietuvos Matematikos Rinkinys*. See also his paper in the *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes*, published in 1964 by the Publishing House of the Czechoslovak Academy of Sciences.

2. The general problem of normal approximation

The purpose of this section is to derive a bound for the error in the normal approximation to the distribution of a random variable in a fairly abstract setting.

THEOREM 2.1. *Let $(\Omega, \mathscr{B}, P)$ be a probability space, and $\mathscr{F}$ and $\mathscr{C}$ be sub σ-algebras of $\mathscr{B}$. Let G be a $\mathscr{B}$-measurable random variable such that $EG^8 < \infty$,* *and let*

$$W = E^{\mathscr{F}} G. \tag{2.1}$$

Let W^* be a $\mathscr{C}$-measurable random variable such that

$$E(W - W^*)^8 < \infty \tag{2.2}$$

and

$$EG(W - W^*) = 1. \tag{2.3}$$

Then, for all real a,

$$|P\{W \leqq a\} - \Phi(a)| \leqq R, \tag{2.4}$$

where

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx \tag{2.5}$$

and

$$\begin{aligned} R = {} & 6\{\operatorname{Var} E^{\mathscr{F}}[G(W - W^*)]\}^{1/2} + 3E\,|G(W - W^*)^3| \\ & + 3\{E(|W| + \tfrac{1}{2})^2 E[E^{\mathscr{F}}|G(W - W^*)^2|]^2\}^{1/2} \\ & + 15(E[G(W - W^*)]^4)^{1/4} \left\{\frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^2|}{(E[G(W - W^*)^2])^2} \right. \\ & \left. + \frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^3|}{(E|G(W - W^*)^3|)^2}\right\}^{3/4} \\ & + \max\{54E|G(W - W^*)^2|, 23E|G(W - W^*)^3|\} + 3E|E^{\mathscr{C}}G|. \end{aligned} \tag{2.6}$$

The proof of this theorem is based on a fairly simple identity.

LEMMA 2.1. *Under the hypotheses of Theorem 2.1, if h is a bounded measurable function, then*

$$\begin{aligned} & EE^{\mathscr{F}}\left[G \int_{W^*}^{W} h(z)\, dz\right] \\ & \quad = Nh + E\left\{Wf(W) - E^{\mathscr{F}}\left[G \int_{W^*}^{W} zf(z)\, dz\right] - (E^{\mathscr{C}}G)f(W^*)\right\}, \end{aligned} \tag{2.7}$$

where

$$Nh = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(x)\, e^{-x^2/2}\, dx, \tag{2.8}$$

and

$$f(w) = e^{w^2/2} \int_{-\infty}^{w} [h(x) - Nh]\, e^{-x^2/2}\, dx. \tag{2.9}$$

Of course in (2.7), if $W^* > W$, the integral $\int_{W^*}^{W} h(z)\,dz$ is to be interpreted as $-\int_{W}^{W^*} h(z)\,dz$, and similarly with $h(z)$ replaced by $zf(z)$. The conditional expectation signs $E^{\mathscr{F}}$ and $E^{\mathscr{C}}$ in (2.7) could be dropped, but they help suggest the way the lemma will be applied.

PROOF OF LEMMA 2.1. Let f (otherwise arbitrary) be a bounded function, the integral of a bounded measurable function f'. Then

$$\begin{aligned} E[Wf(W)] &= E[(E^{\mathscr{F}}G)f(W)] \\ &= E[Gf(W)] \\ &= E\{G[f(W) - f(W^*)] + (E^{\mathscr{C}}G)f(W^*)\}. \end{aligned} \tag{2.10}$$

We can rewrite this in the form

$$\begin{aligned} &E[f'(W) - Wf(W)] \\ &\quad = E\{f'(W) - G[f(W) - f(W^*)]\} + E\{G[f(W) - f(W^*)] - Wf(W)\} \\ &\quad = E\{f'(W) - G[f(W) - f(W^*)] - (E^{\mathscr{C}}G)f(W^*)\}. \end{aligned} \tag{2.11}$$

In order to make conditions for the validity of the normal approximation to the distribution of W more apparent, we express f in terms of an arbitrary bounded measurable function h by (2.9). This function f is the unique bounded solution of the differential equation

$$f'(w) - wf(w) = h(w) - Nh. \tag{2.12}$$

Then equation (2.11) yields

$$\begin{aligned} Eh(W) &= Nh + E[f'(W) - Wf(W)] \\ &= Nh + E\{f'(W) - G[f(W) - f(W^*)] - (E^{\mathscr{C}}G)f(W^*)\} \\ &= Nh + E\left\{f'(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} f'(z)\,dz\right] - (E^{\mathscr{C}}G)f(W^*)\right\} \\ &= Nh + E\left\{\left(h(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right]\right)\right. \\ &\qquad \left. + \left(Wf(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} zf(z)\,dz\right]\right) - (E^{\mathscr{C}}G)f(W^*)\right\}. \end{aligned} \tag{2.13}$$

If we subtract $Eh(W) - E^{\mathscr{F}}[G\int_{W^*}^{W} h(z)\,dz]$ from both sides of this equation, we obtain (2.7). This completes the proof of Lemma 2.1. We observe that we have not used all of the assumptions of Theorem 2.1 in proving this lemma. All of the manipulations of expected values are justified if we assume only that

$$EG^2 < \infty, \qquad E(W - W^*)^2 < \infty \tag{2.14}$$

instead of $EG^8 < \infty$ and $E(W - W^*)^8 < \infty$, and we have not used (2.3), although the proposition is of little interest unless (2.3) is approximately true.

The derivation of Theorem 2.1 from Lemma 2.1 will be somewhat tedious.

In order to indicate the relevance of Lemma 2.1 to questions of normal approximation without considerable complications, we first prove Theorem 2.2 below. The inadequacy of the bound obtained in this way helps to provide some motivation for the complications involved in the proof of Theorem 2.1.

THEOREM 2.2. *Let* $(\Omega, \mathscr{B}, P)$ *be a probability space, and* $\mathscr{F}$ *and* $\mathscr{C}$ *be sub* σ*-algebras of* $\mathscr{B}$. *Let* G *be a* $\mathscr{B}$*-measurable random variable and* W^* *a* $\mathscr{C}$*-measurable random variable such that, with*

$$W = E^{\mathscr{F}} G, \tag{2.15}$$

we have

$$EG(W - W^*) = 1 \tag{2.16}$$

and

$$EG^2(W - W^*)^4 < \infty. \tag{2.17}$$

Let h *be the indefinite integral of a measurable function* h', *and suppose that, for all* w,

$$|h'(w)| \leqq K, \qquad 0 \leqq h(w) \leqq 1, \tag{2.18}$$

where K *is a positive constant. Then*

$$\begin{aligned} |Eh(W) - Nh| \leqq E|E^{\mathscr{C}}G| &+ 2E|1 - E^{\mathscr{F}}[G(W - W^*)]| \\ &+ (E(|W| + K + 1)^2 E[G^2(W - W^*)^4])^{1/2} \\ &+ \tfrac{1}{3}E[|G|\cdot|W - W^*|^3] = R', \end{aligned} \tag{2.19}$$

say.

Also, provided

$$EW^2 \leqq 4, \tag{2.20}$$

we have, for all real a,

$$\begin{aligned} |P\{W \leqq a\} - \Phi(a)| \leqq E|E^{\mathscr{C}}G| &+ 2E|1 - E^{\mathscr{F}}[G(W - W^*)]| \\ &+ 3\{E[G^2(W - W^*)^4]\}^{1/2} + \tfrac{1}{3}E[|G|\cdot|W - W^*|^3] \\ &+ \frac{2}{(2\pi)^{1/4}}(E[G^2(W - W^*)^4])^{1/4}. \end{aligned} \tag{2.21}$$

Ordinarily this result will be of interest only when EW^2 is close to 1, so that condition (2.20) is very weak.

PROOF. We start with an intermediate form of (2.13):

$$Eh(W) = Nh + E\left\{\left(f'(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} f'(z)\,dz\right]\right) - (E^{\mathscr{C}}G)f(W^*)\right\}. \tag{2.22}$$

We shall need some inequalities that are fairly clear intuitively, at least qualitatively:

$$|f(w)| \leqq 1, \qquad |wf(w)| \leqq 1, \tag{2.23}$$

$$(2.24)\qquad |f'(w)| \leqq 2, \qquad |f''(w)| \leqq |w| + K + 1.$$

These follow easily from Lemma 2.5, proved later in this section. Then, with ϑ denoting a constant or random variable less than or equal to 1 in absolute value

$$(2.25)\qquad E\left\{f'(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} f'(z)\,dz\right]\right\}$$
$$= E\left\{f'(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} (f'(W) + \vartheta(|z| + K + 1)(z - W))\,dz\right]\right\}$$
$$= E\{f'(W)(1 - E^{\mathscr{F}}[G(W - W^*)])\}$$
$$+ \vartheta E\left[(|W| + K + 1)\frac{|G|(W - W^*)^2}{2} + \frac{|G|\cdot|W - W^*|^3}{3}\right].$$

But, by Schwarz's inequality,

$$(2.26)\qquad |E[(|W| + K + 1)|G|(W - W^*)^2]| \leqq \{E(|W| + K + 1)^2 E[G^2(W - W^*)^4]\}^{1/2}.$$

$$(2.27)\qquad \left|E\left\{f'(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} f'(z)\,dz\right]\right\}\right| \leqq 2E|1 - E^{\mathscr{F}}[G(W - W^*)]|$$
$$+ \{E(|W| + K + 1)^2 E[G^2(W - W^*)^4]\}^{1/2} + \tfrac{1}{3}E[|G|\cdot|W - W^*|^3].$$

Together with (2.22) this yields (2.19).

To prove (2.21), we apply (2.19) with h given by

$$(2.28)\qquad h(w) = \begin{cases} 1 & \text{for } w \leqq a, \\ 1 - \dfrac{w - a}{\lambda} & \text{for } a \leqq w \leqq a + \lambda, \\ 0 & w \geqq a + \lambda, \end{cases}$$

where

$$(2.29)\qquad \lambda = \{(2\pi)E[G^2(W - W^*)^4]\}^{1/4}.$$

Thus we obtain

$$(2.30)\qquad P\{W \leqq a\} \leqq Eh(W) \leqq Nh + R'$$
$$\leqq \Phi(a + \lambda) + R' \leqq \Phi(a) + \frac{\lambda}{\sqrt{2\pi}} + R'.$$

Similarly,

$$(2.31)\qquad P\{W \leqq a\} \geqq \Phi(a) - \frac{\lambda}{\sqrt{2\pi}} - R'.$$

Here the K in R' can be taken to be $1/\lambda$. Then (2.21) follows immediately from (2.30) and (2.31).

There is a conspicuous loss of accuracy in going from (2.19) to (2.21), in that the derivative of h is equal to the large bound $1/\lambda$ only on a very small interval, and 0 elsewhere. Thus it is plausible that the bound for $|P\{W \leqq a\} - \Phi(a)|$ in (2.21) ought to be something like the bound R' in (2.19). In order to use Lemma 2.1 to prove something like this, we proceed as follows.

(i) We choose a positve number λ that is roughly of the order of magnitude of the bound we want to obtain for $|P\{W \leqq a\} - \Phi(a)|$. We then take for h, in (2.7), one minus the c.d.f. of a random variable with mean a and expected absolute deviation from a of the order of λ, say $h = h_{a,\lambda}$. By a tedious calculation we approximate the left side of (2.7) by

$$(2.32)\qquad \begin{aligned} E\underline{h}_{a,\lambda}(W) - R_1 &\leqq EE^{\mathscr{F}}\left[G\int_{W^*}^{W} h_{a,\lambda}(z)\,dz\right] \\ &\leqq E\bar{h}_{a,\lambda}(W) + R_1, \end{aligned}$$

where R_1 will be part of the bound and $\underline{h}_{a,\lambda}$ and $\bar{h}_{a,\lambda}$ are functions having roughly the same behavior as $h_{a,\lambda}$ shifted by a small multiple of λ.

(ii) We obtain a bound for

$$(2.33)\qquad E\left\{Wf(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} zf(z)\,dz\right]\right\},$$

which occurs on the right side of (2.7). Because f is much less abrupt than h this causes less difficulty than (i).

(iii) We must then go from the upper bound for $E\underline{h}_{a,\lambda}(W)$ and the lower bound for $E\bar{h}_{a,\lambda}(W)$ provided by steps (i) and (ii) to an approximation for $P\{W \leqq a\}$. The argument for this is similar to the concluding part of the argument in some versions of the proof of the Berry-Esseen theorem.

In order to carry out this program, we define for real a and (small) $\lambda > 0$,

$$(2.34)\qquad h_{a,\lambda}(x) = g\left(\frac{x-a}{\lambda}\right),$$

where

$$(2.35)\qquad g(y) = \begin{cases} 1 - \dfrac{1}{2(1-y)^2} & \text{if } y \leqq 0, \\[2ex] \dfrac{1}{2(1+y)^2} & \text{if } y \geqq 0. \end{cases}$$

We shall need the following.

LEMMA 2.2. *For all real y_1 and y_2,*

$$(2.36)\qquad |g(y_1) - g(y_2) - (y_1 - y_2)g'(y_2)| \leqq \frac{3(y_1 - y_2)^2}{1 + y_2^2},$$

and consequently for all real x_1 and x_2,

$$|h_{a,\lambda}(x_1) - h_{a,\lambda}(x_2) - (x_1 - x_2)h'(x_2)| \leqq \frac{3(x_1 - x_2)^2}{\lambda^2 + (x_2 - a)^2}. \tag{2.37}$$

PROOF OF LEMMA 2.2. For the proof of (2.36) we distinguish three cases.

Case 1. $|y_2| > 1$ and $y_1 y_2 \leqq 0$. Without essential loss of generality, we suppose $y_2 > 1$ and $y_1 \leqq 0$. Then

$$0 \leqq g(y_1) - g(y_2) \leqq \frac{2(y_1 - y_2)^2}{1 + y_2^2}, \tag{2.38}$$

and

$$0 \leqq (y_1 - y_2)g'(y_2) = \frac{y_2 - y_1}{(1 + y_2)^3} \leqq \frac{(y_2 - y_1)^2}{1 + y_2^2}. \tag{2.39}$$

It follows that, in this case, (2.36) holds even with 3 replaced by 2.

Case 2. $|y_2| \leqq 1$.

$$\begin{aligned} |g(y_1) - g(y_2) - (y_1 - y_2)g'(y_2)| &= \left| \int_{y_2}^{y_1} g''(y)(y_1 - y)\, dy \right| \\ &\leqq 3 \cdot \tfrac{1}{2}(y_1 - y_2)^2 \leqq \frac{3(y_1 - y_2)^2}{1 + y_2^2}, \end{aligned} \tag{2.40}$$

since

$$g''(y) = \begin{cases} -\dfrac{3}{(1 - y)^4} & \text{if} \quad y < 0, \\ \dfrac{3}{(1 + y)^4} & \text{if} \quad y > 0, \end{cases} \tag{2.41}$$

does not exceed 3.

Case 3. $y_1 y_2 \geqq 0$. Without essential loss of generality, we suppose $y_1 y_2 \geqq 0$. Then

$$\begin{aligned} &|g(y_1) - g(y_2) - (y_1 - y_2)g'(y_2)| \\ &= \frac{1}{2}\left| \frac{1}{(1 + y_1)^2} - \frac{1}{(1 + y_2)^2} + (y_1 - y_2)\frac{1}{(1 + y_2)^3} \right| \\ &= \frac{|(y_2 - y_1)^3[3 - y_2 - 2y_1]|}{2(1 + y_1)^2(1 + y_2)^3} \leqq \frac{3}{2}\frac{(y_1 - y_2)^2}{1 + y_2^2}. \end{aligned} \tag{2.42}$$

Thus, taking the three cases together, we have proved (2.36). Formula (2.37) follows from (2.36) by substitution of $(x_i - a)/\lambda$ for y_i.

Next let us start to approximate the left side of (2.7). By Lemma 2.2,

$$\begin{aligned} &E^{\mathscr{F}}\left[G \int_{W^*}^{W} h(z)\, dz \right] \\ &= E^{\mathscr{F}}\left\{ G \int_{W^*}^{W} \left[h(W) + (z - W)h'(W) + \frac{3\vartheta(z - W)^2}{\lambda^2 + (W - a)^2} \right] dz \right\} \end{aligned} \tag{2.43}$$

$$= h(W)E^{\mathscr{F}}[G(W - W^*)] - \tfrac{1}{2}h'(W)E^{\mathscr{F}}[G(W - W^*)^2] + \frac{\vartheta}{\lambda^2 + (W - a)^2} E^{\mathscr{F}}|G(W - W^*)^3|.$$

Here, as in the proof of Theorem 2.2, ϑ denotes a number or random variable, possibly depending on everything, but satisfying $|\vartheta| \leqq 1$. Two occurrences of ϑ, even in a single formula, need not denote the same number or random variable. For brevity, we have written h rather than $h_{a,\lambda}$.

Now, with the positive constants B_1 and B_2 to be chosen later, define the event

$$Q = \{E^{\mathscr{F}}|G(W - W^*)^2| \leqq B_1\lambda \text{ and } E^{\mathscr{F}}|G(W - W^*)^3| \leqq B_2\lambda^2\} \tag{2.44}$$

and let Q^c be the complement of Q. Then

$$\begin{aligned}
EE^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right] &= E\left\{(\chi^Q + \chi^{Q^c})E^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right]\right\} \\
&= E\chi^Q\left\{h(W)E^{\mathscr{F}}[G(W - W^*)] - \tfrac{1}{2}h'(W)E^{\mathscr{F}}[G(W - W^*)^2] \right.\\
&\quad \left. + \frac{\vartheta}{\lambda^2 + (W - a)^2}E^{\mathscr{F}}|G(W - W^*)^3|\right\} + E\left\{\chi^{Q^c}E^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right]\right\} \\
&= E\chi^Q\left\{h(W)E^{\mathscr{F}}[G(W - W^*)] + \frac{\vartheta}{2}h'(W)B_1\lambda \right.\\
&\quad \left. + \frac{\vartheta B_2\lambda^2}{\lambda^2 + (W - a)^2}\right\} + E\left\{\chi^{Q^c}E^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right]\right\} \\
&= E\left\{h(W)E^{\mathscr{F}}[G(W - W^*)] + \frac{\vartheta}{2}h'(W)B_1\lambda + \frac{\vartheta B_2\lambda^2}{\lambda^2 + (W - a)^2}\right\} \\
&\quad + E\chi^{Q^c}\left\{E^{\mathscr{F}}\left[G\int_{W^*}^{W} h(z)\,dz\right] - h(W)E^{\mathscr{F}}[G(W - W^*)] \right.\\
&\quad \left. + \frac{\vartheta}{2}h'(W)B_1\lambda\right\} \\
&= Eh(W) + \vartheta E|E^{\mathscr{F}}[G(W - W^*)] - 1| \\
&\quad + \frac{\vartheta}{2}B_1\lambda Eh'(W) + \vartheta B_2\lambda^2 E\frac{1}{\lambda^2 + (W - a)^2} \\
&\quad + E\chi^{Q^c}\left\{E^{\mathscr{F}}\left[G\int_{W^*}^{W}\big(h(z) - h(W)\big)\,dz\right] + \frac{\vartheta}{2}h'(W)B_1\lambda\right\} \\
&= E\left\{h(W) + \frac{\vartheta}{2}B_1\lambda h'(W) + \frac{\vartheta B_2\lambda^2}{\lambda^2 + (W - a)^2}\right\} \\
&\quad + E\left\{|E^{\mathscr{F}}[G(W - W^*)] - 1| + \chi^{Q^c}\left[|G(W - W^*)| + \frac{B_1}{8}\right]\right\}.
\end{aligned} \tag{2.45}$$

Then, with $\underline{h}$ and $\bar{h}$ defined by

$$\underline{h}(W) = h(W) + \frac{B_1\lambda}{2} h'(w) - \frac{B_2\lambda^2}{\lambda^2 + (w - a)^2} \tag{2.46}$$

and

$$\bar{h}(w) = h(w) - \frac{B_1\lambda}{2} h'(w) + \frac{B_2\lambda^2}{\lambda^2 + (w - a)^2}, \tag{2.47}$$

formula (2.45) yields

$$E\underline{h}(W) - R_1 \leqq E\left[G \int_{W^*}^{W} h(z)\, dz\right] \leqq E\bar{h}(W) + R_1, \tag{2.48}$$

where

$$R_1 = E\left\{|E^{\mathscr{F}}[G(W - W^*] - 1| + \chi^{Q^c}[|G(W - W^*)| + \frac{B_1}{8}\right\}. \tag{2.49}$$

In order to use these bounds to derive Theorem 2.1 from (2.7), we need to obtain bounds for $\underline{h}$, $\bar{h}$, and R_1.

LEMMA 2.3. *If we take $B_1 = 1/8$ and $B_2 = 1/16$ in the definition* (2.44) *of Q, we have*

$$\underline{h}(w) \geqq \begin{cases} \dfrac{\lambda^2}{4[\lambda + (w - a)]^2} & \text{if } w \geqq a, \\ 1 - \dfrac{3\lambda^2}{4[\lambda - (w - a)]^2} & \text{if } w \leqq a, \end{cases} \tag{2.50}$$

and

$$\bar{h}(w) \leqq \begin{cases} 1 - \dfrac{\lambda^2}{4[\lambda - (w - a)]^2} & \text{if } w \leqq a, \\ \dfrac{3\lambda^2}{4[\lambda + (w - \mathrm{a})]^2} & \text{if } w \geqq a. \end{cases} \tag{2.51}$$

We shall prove only (2.50) since (2.51) follows by symmetry or a similar proof. First we observe that, for $w \geqq 0$, with g defined by (2.35)

$$\begin{aligned} g(w) + B_1 g'(w) - \frac{B_2}{1 + w^2} &= \frac{1}{2(1 + w)^2} - \frac{B_1}{(1 + w)^3} - \frac{B_2}{1 + w^2} \\ &\geqq \frac{1}{2(1 + w)^2}(1 - 2B_1 - 4B_2) \geqq \frac{1}{4(1 + w)^2} \end{aligned} \tag{2.52}$$

and, for $w \leqq 0$,

$$\begin{aligned} &g(w) + B_1 g'(w) - \frac{B_2}{1 + w^2} \\ &= 1 - \frac{1}{2(1 - w)^2} - \frac{B_1}{(1 - w)^3} - \frac{B_2}{1 + w^2} \\ &\geqq 1 - \frac{1}{2(1 - w)^2}(1 + 2B_1 + 4B_2) \geqq 1 - \frac{3}{4(1 - w)^2}. \end{aligned} \tag{2.53}$$

Replacing w in (2.52) and (2.53) by $(w - a)/\lambda$ and using (2.46) and (2.14), we obtain (2.50).

To take care of R_1, defined by (2.49), we observe that, by Hölder's inequality and the fact that $(\chi^{Q^c})^{4/3} = \chi^{Q^c}$,

$$(2.54)\qquad \begin{aligned} &E\left\{\chi^{Q^c}\left[|G(W - W^*)| + \frac{B_1}{8}\right]\right\} \\ &\leqq (E\chi^{Q^c})^{3/4}(E[|G(W - W^*)| + \tfrac{1}{64}]^4)^{1/4} \\ &\leqq \tfrac{11}{10}\,(E\chi^{Q^c})^{3/4}\{E[G(W - W^*)]^4\}^{1/4}. \end{aligned}$$

We have also used the fact that, by (2.3), for $k \geqq 1$

$$(2.55)\qquad E|G(W - W^*)|^k \geqq |E[G(W - W^*)]|^k = 1,$$

and also $B_1 = 1/8$ and $(65/64)^4 \leqq 11/10$. By the definition (2.44) of Q

$$(2.56)\qquad \begin{aligned} E\chi^{Q^c} &= P(Q^c) \\ &= P\left\{E^{\mathscr{F}}|G(W - W^*)^2| > \frac{\lambda}{8} \text{ or } E^{\mathscr{F}}|G(W - W^*)^3| > \frac{\lambda^2}{16}\right\} \\ &\leqq P\left\{E^{\mathscr{F}}|G(W - W^*)^2| > \frac{\lambda}{8}\right\} + P\left\{E^{\mathscr{F}}|G(W - W^*)^3| > \frac{\lambda^2}{16}\right\} \\ &\leqq \frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^2|}{((\lambda/8) - E|G(W - W^*)^2|)^2} + \frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^3|}{((\lambda^2/16) - E|G(W - W^*)^3|)^2}, \end{aligned}$$

provided

$$(2.57)\qquad \frac{\lambda}{16} > E|G(W - W^*)^2|,$$

and

$$(2.58)\qquad \frac{\lambda^2}{16} > E|G(W - W^*)^3|.$$

We summarize these somewhat tedious calculations in

LEMMA 2.4. *Let h_*, h^*, and R_2 be defined by*

$$(2.59)\qquad h_*(w) = \begin{cases} \dfrac{\lambda^2}{4[\lambda + (w - a)]^2} & \text{if } w \geqq a, \\ 1 - \dfrac{3\lambda^2}{4[\lambda - (w - a)]^2} & \text{if } w \leqq a, \end{cases}$$

$$(2.60)\qquad h^*(w) = \begin{cases} \dfrac{3\lambda^2}{4[\lambda + (w - a)]^2} & \text{if } w \leqq a, \\ 1 - \dfrac{\lambda^2}{4[\lambda - (w - a)]^2} & \text{if } w \leqq a, \end{cases}$$

and

$$R_2 = \{\mathrm{Var}\, E^{\mathscr{F}}[G(W - W^*)]\}^{1/2} + \frac{11}{10}(E[G(W - W^*)]^4)^{1/4} \cdot \left\{ \frac{\mathrm{Var}\, E|G(W - W^*)^2|}{[(\lambda/8) - E|G(W - W^*)^2|]^2} + \frac{\mathrm{Var}\, E^{\mathscr{F}}|G(W - W^*)^3|}{[(\lambda^2/16) - E|G(W - W^*)^3|]^2} \right\}^{3/4}, \tag{2.61}$$

and suppose λ chosen to satisfy (2.57) and (2.58). Then

$$Eh_*(W) - R_2 \leqq E\left[G \int_{W^*}^{W} h(z)\, dz\right] \leqq Eh^*(W) + R_2. \tag{2.62}$$

Next we turn to the evaluation of the term $E\{Wf(W) - E[G\int_{W^*}^{W} zf(z)\, dz]\}$ occurring on the right side of (2.7). Here f is defined by (2.9) with $h = h_{a,\lambda}$ given by (2.34). However we shall first consider the case of $h = \chi^{(-\infty, b]}$ for some real b. Then f, defined by (2.9) is given by

$$f(w) = \begin{cases} \sqrt{2\pi}\, e^{w^2/2}\Phi(w)[1 - \Phi(b)] & \text{if } w \leqq b, \\ \sqrt{2\pi}\, e^{w^2/2}\Phi(b)[1 - \Phi(w)] & \text{if } w \geqq b. \end{cases} \tag{2.63}$$

We first prove

LEMMA 2.5. *For f defined by (2.63) we have*

$$|f(w)| \leqq 1, \tag{2.64}$$

$$|wf(w)| \leqq 1, \tag{2.65}$$

and

$$|f'(w)| \leqq 1. \tag{2.66}$$

PROOF OF LEMMA 2.5. Without essential loss of generality, we suppose $b \geqq 0$. We need the familiar fact that for $w \geqq 0$,

$$1 - \Phi(w) = \frac{1}{\sqrt{2\pi}} \int_w^\infty e^{-z^2/2}\, dz \leqq \frac{1}{\sqrt{2\pi}} \int_w^\infty \frac{z}{w} e^{-z^2/2}\, dz = \frac{e^{-w^2/2}}{w\sqrt{2\pi}}, \tag{2.67}$$

and, analogously, for $w \leqq 0$,

$$\Phi(w) \leqq \frac{e^{-w^2/2}}{|w|\sqrt{2\pi}}. \tag{2.68}$$

To prove (2.66), we consider three cases.

Case 1. $w \leqq 0 \leqq b$.

The first half of (2.63) yields (with $0 \leqq \vartheta \leqq 1$)

$$|f'(w)| = |[1 - \Phi(b)]\{\sqrt{2\pi}\, we^{w^2/2}\Phi(w) + 1\}| = |[1 - \Phi(b)]\{-\vartheta + 1\}| \leqq 1. \tag{2.69}$$

Case 2. $0 \leqq w \leqq b$.

Again, we use the first half of (2.63).

$$f'(w) = [1 - \Phi(b) + \Phi(w)\sqrt{2\pi}\, we^{w^2/2}[1 - \Phi(b)] \leqq [1 - \Phi(b) + \Phi(b)\sqrt{2\pi}\, we^{w^2/2}[1 - \Phi(w)] \leqq [1 - \Phi(b)] + \Phi(b) \leqq 1, \tag{2.70}$$

by (2.67). Since $f'(w)$ is obviously positive, (2.66) follows in this case.

Case 3. $0 \leqq b \leqq w$.

From the second half of (2.63) it follows that

$$f'(w) = \Phi(b)\{-1 + \sqrt{2\pi}\, we^{w^2/2}[1 - \Phi(w)]\}. \tag{2.71}$$

Because of (2.67), this is negative and

$$0 \geqq f'(w) \geqq -\Phi(b) \geqq -1. \tag{2.72}$$

This completes the proof of (2.66) for $h = \chi^{(-\infty,b]}$, that is, for f given by (2.63). Since the bound in (2.66) does not depend on b, the result also applies to any convex combination of the $\chi^{(-\infty,b]}$, in particular to $h = h_{a,\lambda}$. A similar remark will apply to (2.64) and (2.65). The proof of (2.65) follows easily by examination of (2.69), (2.70), and (2.71). Formula (2.64) can be proved by observing that f defined by (2.63) is nonnegative and achieves its maximum at b, and

$$f(b) = \sqrt{2\pi}\, e^{b^2/2}\Phi(b)[1 - \Phi(b)]. \tag{2.73}$$

Examination of tables of the normal distribution shows that this is everywhere less than one. This completes the proof of Lemma 2.5.

It follows that

$$\begin{aligned}
&E\left\{WF(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W} zf(z)\,dz\right]\right\} \qquad (2.74)\\
&= E\left\{Wf(W) - E^{\mathscr{F}}\left[G\int_{W^*}^{W}(f(W) + \vartheta(z - W))\,dz\right]\right\}\\
&= E\left\{Wf(W) - E^{\mathscr{F}}\left[f(W)G\frac{W^2 - W^{*2}}{2}\right.\right.\\
&\qquad\qquad \left.\left. + \vartheta G(|W| + |W - W^*|(W - W^*)\right]\right\}\\
&= E\left\{Wf(W) - E^{\mathscr{F}}\left[f(W)G(W - W^*)\left(W - \frac{W - W^*}{2}\right)\right.\right.\\
&\qquad\qquad \left.\left. + \vartheta G(|W| + |W - W^*|(W - W^*)\right]\right\}\\
&= E\left\{Wf(W)[1 - E^{\mathscr{F}}G(W - W^*)]\right.\\
&\qquad\qquad \left. + \vartheta E^{\mathscr{F}}[|G|(W - W^*)^2(|W| + \frac{f(w)}{2} + |G|\cdot|W - W^*|^3\right\}.
\end{aligned}$$

Using Lemma 2.5, we obtain the bound

$$\left|E\left\{Wf(W) - E\left[G\int_{W^*}^{W} zf(z)\,dz\right]\right\}\right| \tag{2.75}$$
$$\leq E|1 - E^{\mathscr{F}}[G(W - W^*)]| + E[|G|\cdot|W - W^*|^3]$$
$$+ [E\{E^{\mathscr{F}}[|G|(W - W^*)^2]\}^2 E(|W| + \tfrac{1}{2})^2]^{1/2}.$$

Theorem 2.1 will follow from the lemmas we have proved together with

LEMMA 2.6. *Let R_3 be a nonnegative number and W, Y, Z independent random variables with* c.d.f.'s *F, G, H respectively, such that for all real a,*

$$P\{W + Z \leq a\} - R_3 \leq \Phi(a) \leq P\{W + Y \leq a\} + R_3. \tag{2.76}$$

Then, for all real a,

$$|P\{W \leq a\} - \Phi(a)| \leq 3\left(R_3 + \int_{-\infty}^{0} |y|\,dG(y) + \int_{0}^{\infty} z dH(z)\right). \tag{2.77}$$

I omit the proof of Lemma 2.6 because, except for the numerical constants it is essentially an inferior version of part of the usual proof of the theorem of Berry and Esséen. See for example Loève ([4], p. 283, Proposition 20.3b).

To prove Theorem 2.1, we need only put all of the preceding together. The distributions G and H of Lemma 2.6 are defined by

$$G(y) = h^*(a - y) \tag{2.78}$$

and

$$H(z) = h_*(a - z) \tag{2.79}$$

with h_* and h^* given by (2.59) and (2.60). Then

$$\int_{-\infty}^{0} |y|\,dG(y) = \int_{0}^{\infty} ud\left(-\frac{3\lambda^2}{4[\lambda + u]^2}\right) = \frac{3\lambda}{4} \tag{2.80}$$

and

$$\int_{0}^{\infty} zdG(z) = \frac{3\lambda}{4}. \tag{2.81}$$

Also

$$P\{W + Y \leq a\} = \int G(a - x)\,dF(x) = EG(a - W) = Eh^*(W), \tag{2.82}$$

and

$$P\{W + Z \leq a\} = Eh_*(W). \tag{2.83}$$

But, by (2.7), (2.62), and (2.75), letting

$$(2.84)\quad R_3 = 2\{\operatorname{Var} E^{\mathscr{F}}[G(W - W^*)]\}^{1/2} + E|G(W - W^*)^3| + \{E(|W| + \tfrac{1}{2})^2 E\{E^{\mathscr{F}}|G(W - W^*)^2|\}^2\}^{1/2} + \frac{11}{10}(E[G(W - W^*)]^4)^{1/4}\left\{\frac{\operatorname{Var} E|G(W - W^*)^2|}{[(\lambda/8) - E|G(W - W^*)^2|]^2} + \frac{\operatorname{Var} E|G(W - W^*)^3|}{[(\lambda^2/16) - E|G(W - W^*)^3|]^2}\right\}^{3/4},$$

we have

$$(2.85)\quad \begin{aligned} P\{W + Z \leqq a\} - R_3 &= Eh_*(W) - R_3 \\ &\leqq \Phi(a) \leqq Eh^*(W) + R_3 = P\{W + Y \leqq a\} + R_3. \end{aligned}$$

Then Lemma 2.6 and (2.93) and (2.94) imply that

$$(2.86)\quad |P\{W \leqq a\} - \Phi(a)| \leqq 3\left(R_3 + \frac{3\lambda}{2}\right).$$

We recall that λ must be chosen so as to satisfy (2.57) and (2.58). If we take

$$(2.87)\quad \lambda = \max\{12E|G(W - W^*)^2|, 5[E|G(W - W^*)^3|]^{1/2}\}$$

we obtain (2.4) with R given by (2.6).

3. The sum of a weakly dependent stationary sequence

Let $X_1, X_2, \cdots$ be a stationary sequence of random variables. In formulas (3.17), (3.24), (3.26), and (3.30) below, combined with Theorem 2.1, we obtain a bound for the difference between the distribution of $\Sigma_1^n X_i$ and a normal distribution under mildly restrictive conditions, including

$$(3.1)\quad EX_i = 0, \qquad EX_i^2 = 1,$$

$$(3.2)\quad \beta = EX_i^8 < \infty,$$

and

$$(3.3)\quad 0 < C = \lim_{n \to \infty} \frac{1}{n} \operatorname{Var} \sum_1^n X_i < \infty.$$

In Corollary 3.1 we consider the particular case of an m-dependent sequence, proving that the difference is bounded by $An^{-1/2}$, where A depends on the properties of the process in a fairly complicated way. In other cases the results obtained are not as satisfactory. In Corollary 3.2, we see that, in the case where, roughly speaking, dependence decreases exponentially with time, the difference is bounded by $An^{-1/2}(\log n)^2$. It seems likely that the correct order of magnitude of the difference is at worst $n^{-1/2}$ or perhaps $n^{-1/2}(\log n)^{1/2}$.

For completeness, we start by recalling the definition of a stationary sequence.

DEFINITION 3.1. *A sequence $X_1, X_2, \cdots$ of random variables is said to be stationary if, for every pair t, j of natural numbers, the sequence $X_{t+1}, \cdots, X_{t+j}$ has the same distribution as $X_1, \cdots, X_j$.*

We consider a stationary sequence $X_1, X_2, \cdots$ of random variables satisfying conditions (3.1) and (3.2) above. Let

$$\rho_i = EX_tX_{t+i}, \tag{3.4}$$

and let $\{\alpha_k\}k = 1, 2, \cdots$ be a sequence such that, if A and B are any two finite sets of natural numbers for which

$$\inf_{i\in A, j\in B} |i - j| \geqq k. \tag{3.5}$$

and Y and Z are random variables with finite variance depending only on the $\{X_i\}_{i\in A}$ and the $\{X_j\}_{j\in B}$ respectively, then

$$|\mathrm{Corr}\,(Y, Z)| \leqq \alpha_k. \tag{3.6}$$

Let m be an arbitrary natural number less than n, which will be chosen appropriately in Corollaries 3.1 and 3.2. Ordinarily m will be much smaller than n. Let $\mathscr{B}$ be the σ-algebra of events generated by $X_1, \cdots, X_n$ and a random variable I uniformly distributed over $\{1, \cdots, n\}$ independent of $X_1, \cdots, X_n$. Let $\mathscr{F}$ be the σ-algebra generated by $X_1, \cdots, X_n$ and $\mathscr{C}$ the smallest σ-algebra containing all events of the form

$$\{I = i \text{ and for all } j \text{ such that } |j - i| > m,\ X_j \leqq a_j\}, \tag{3.7}$$

where the a_j are real numbers. Here and in the rest of this section, indices on the X, such as the j in X_j, are assumed to be restricted to $\{1, \cdots, n\}$. Let

$$G = \frac{n}{\sqrt{\delta_n}} X_I, \tag{3.8}$$

where

$$\delta_n = ny_1 - y_2, \tag{3.9}$$

$$y_1 = 1 + 2\sum_{i=1}^{m} \rho_i, \tag{3.10}$$

and

$$y_2 = 2\sum_{i=1}^{m} i\rho_i. \tag{3.11}$$

Also let

$$W = E^{\mathscr{F}}G = \frac{1}{\sqrt{\delta_n}} \sum_{i=1}^{n} X_i, \tag{3.12}$$

and

$$W^* = \frac{1}{\sqrt{\delta_n}} \sum_{|j-I|>m} X_j. \tag{3.13}$$

Then W^* is $\mathscr{C}$-measurable and

$$W - W^* = \frac{1}{\sqrt{\delta_n}} \sum_{|j-I|\leqq m} X_j. \tag{3.14}$$

To obtain a bound for the difference between the distribution of W and a unit normal distribution we must evaluate the remainder R given by (2.6) in this case. The last term in (2.6) is taken care of by observing that

$$E|E^{\mathscr{C}}G| \leqq \alpha_{m+1}\sqrt{EG^2}. \tag{3.15}$$

For, given $I = i$, the random variables G and $E^{\mathscr{C}}G$ are determined by X with indices differing by at least $m + 1$ so that, by hypothesis (3.6), their correlation is at most α_{m+1}. Thus

$$\begin{aligned} E(E^{\mathscr{C}}G)^2 &= E[G(E^{\mathscr{C}}G)] = EE^I[G(E^{\mathscr{C}}G)] \leqq E\alpha_{m+1}[E^I G^2(E^{\mathscr{C}}G)^2]^{1/2} \\ &\leqq \alpha_{m+1}[EG^2E(E^{\mathscr{C}}G)^2]^{1/2}. \end{aligned} \tag{3.16}$$

Consequently

$$E|E^{\mathscr{C}}G| \leqq [E(E^{\mathscr{C}}G)^2]^{1/2} \leqq \alpha_{m+1}\sqrt{EG^2} = \alpha_{m+1}\frac{n}{\sqrt{\delta_n}}. \tag{3.17}$$

Next we consider the second, third and fifth terms in (2.6). We have

$$EG^8 = \frac{n^8}{\delta_n^4} EX_I^8 = \frac{n^8}{\delta_n^4}\beta, \tag{3.18}$$

and

$$E(W - W^*)^8 = \frac{1}{\delta_n^4} E\Big(\sum_{|j-I|\leqq m} X_j\Big)^8 \leqq \frac{1}{\delta_n^4}(2m+1)^8\beta. \tag{3.19}$$

Thus, by Hölder's inequality, for $0 \leqq k, \ell$ and $k + \ell \leqq 8$,

$$E|G|^k|W - W^*|^\ell \leqq n^k(2m+1)^\ell\left(\frac{\beta}{\delta_n^4}\right)^{(k+\ell)/8} \tag{3.20}$$

In particular,

$$E|G(W - W^*)^2| \leqq \frac{n}{\delta_n^{3/2}}(2m+1)^2\beta^{3/8}, \tag{3.21}$$

$$E|G(W - W^*)^3| \leqq \frac{n}{\delta_n^2}(2m+1)^3\beta^{1/2}, \tag{3.22}$$

and

$$E[E^{\mathscr{F}}|G(W - W^*)^2|^2] \leqq EG^2(W - W^*)^4 \leqq \frac{n^2}{\delta_n^3}(2m+1)^4\beta^{3/4}. \tag{3.23}$$

The sum of the second, third and fifth terms in (2.6) is then bounded by

$$
\begin{aligned}
(3.24)\quad & 3E|G(W - W^*)^3| + 3\{E(|W| + \tfrac{1}{2})^2 E[E^{\mathscr{F}}|G(W - W^*)^2|]^2\}^{1/2} \\
& \qquad + \max\{54E|G(W - W^*)^2, 23[E|G(W - W^*)^3|]^{1/2}\} \\
& \leqq \frac{3n}{\delta_n^2}(2m + 1)^3\beta^{1/2} + 9\frac{n}{\delta_n^{3/2}}(2m + 1)^2\beta^{3/8} \\
& \qquad + \max\left\{\frac{54n}{\delta_n^{3/2}}(2m + 1)^2\beta^{3/8}, 23\left[\frac{n}{\delta_n^2}(2m + 1)^3\beta^{1/2}\right]^{1/2}\right\},
\end{aligned}
$$

provided we assume that

$$
(3.25)\qquad EW^2 \leqq 4.
$$

Now let us look at the first term in formula (2.6). We have

$$
\begin{aligned}
(3.26)\quad & \operatorname{Var} E^{\mathscr{F}}[G(W - W^*)] \\
& = \operatorname{Var}\frac{1}{\delta_n}\sum_{i=1}^{n} X_i \sum_{|j-i|\leqq m} X_j \\
& = \frac{1}{\delta_n^2}\sum_{i,i'}\operatorname{Cov}\left(X_i \sum_{|j-i|\leqq m} X_j, X_{i'} \sum_{|j'-i'|\leqq m} X_{j'}\right) \\
& \leqq \frac{1}{\delta_n^2}\sum_{i,i'}\alpha_{(|i-i'|-2m)} + E\left[X_{m+1}^2\left(\sum_{j=1}^{2m+1}|X_j|\right)^2\right] \\
& \leqq \frac{(2m + 1)^2}{\delta_n^2}\beta^{1/2}n\left((2m + 1) + \sum_{i=1}^{n}\alpha_i\right).
\end{aligned}
$$

The bounding of the fourth term in (3.26) is similar to this but more tedious. First we observe that

$$
\begin{aligned}
(3.27)\quad & \operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^2| \\
& = \operatorname{Var}\frac{1}{\delta_n^{3/2}}\sum_{i=1}^{n}|X_i|\left(\sum_{|j-i|\leqq m} X_j\right)^2 \\
& = \frac{1}{\delta_n^3}\sum_{i,i'}\operatorname{Cov}\left(X_i\left(\sum_{|j-i|\leqq m} X_j\right)^2, X_{i'}\left(\sum_{|j'-i'|\leqq m} X_{j'}\right)^2\right) \\
& \leqq \frac{1}{\delta_n^3}\sum_{i,i'}\alpha_{(|i-i'|-2m)} + \frac{1}{2}\left[EX_i^2\left(\sum_{|j-i|\leqq m} X_j\right)^4 + EX_{i'}^2\left(\sum_{|j'-i'|\leqq m} X_{j'}\right)^7\right] \\
& \leqq \frac{n}{\delta_n^3}\left((2m + 1) + \sum_{1}^{n}\alpha_i\right)E\left[X_I\left(\sum_{|j-I|\leqq m} X_j\right)^2\right]^2 \\
& = \frac{1}{n}\left[(2m + 1) + \sum_{1}^{n}\alpha_i\right]E[G(W - W^*)^2]^2,
\end{aligned}
$$

and similarly

$$\text{(3.28)}\qquad \operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^3| \leqq \frac{1}{n}\left[(2m + 1) + \sum_1^n \alpha_i\right] E[G(W - W^*)^3]^2.$$

Also, as a special case of (3.20),

$$\text{(3.29)}\qquad (E[G(W - W^*)]^4)^{1/4} \leqq \frac{(2m + 1)n}{\delta_n}\beta^{1/4}.$$

Thus we obtain for the fourth term in formula (2.6)

$$\text{(3.30)}\qquad (E[G(W - W^*)]^4)^{1/4}\left\{\frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^2|}{(E|G(W - W^*)^2)^2} + \frac{\operatorname{Var} E^{\mathscr{F}}|G(W - W^*)^3|}{(E|G(W - W^*)^3|)^2}\right\}^{3/4}$$
$$\leqq \frac{(2m + 1)n}{\delta_n}\beta^{1/4}\left[\frac{1}{n}\left((2m + 1) + \sum_1^n \alpha_i\right)\right]^{3/4}$$
$$\left\{\frac{E[X_I(\sum_{|j-I|\leqq m} X_j)^2]^2}{(E|X_I(\sum_{|j-I|\leqq m} X_j)^2|)^2} + \frac{E[X_I(\sum_{|j-I|\leqq m} X_j)^3]^2}{(E|X_i \sum_{|j-I|\leqq m} X_j)^3|)^2}\right\}^{3/4}.$$

Bounds for all the terms in (2.6) have now been given in formulas (3.17), (3.24), (3.26), and (3.30), under the conditions described at the beginning of this section.

DEFINITION 3.2. *A sequence $X_1, X_2, \cdots$ of random variables is m-dependent where m is a nonnegative integer if for any two subsets $A, B \subset \{1, 2, \cdots\}$ for which* (3.5) *holds with $k = m + 1$, the sets of random variables $\{X_i\}_{i\in A}$ and $\{X_j\}_{j\in B}$ are independent.*

An independent sequence is 0-dependent according to this definition.

COROLLARY 3.1. *If $X_1, X_2, \cdots$ is a stationary m-dependent sequence of random variables satisfying* (3.1), (3.2), *and* (3.3), *there exists a constant A (depending on the distribution of the sequence $X_1, X_2, \cdots$ but not on n) such that for all n and all real a*

$$\text{(3.31)}\qquad \left|P\left\{\frac{\sum_1^n X_i}{\sqrt{nC}} \leqq a\right\} - \Phi(a)\right| \leqq An^{-1/2}.$$

The proof is easily obtained by examining (3.17), (3.24), (3.26), and (3.30), observing that $\alpha_k = 0$ for $k \geqq m + 1$ and that, by (3.9) and assumption (3.3), δ_n is of the exact order of n.

COROLLARY 3.2. *If $X_1, X_2, \cdots$ is a stationary sequence of random variables satisfying* (3.1), (3.2), (3.3), *and* (3.6) *and if there exists a positive number λ such that for all sufficiently large k,*

$$\alpha_k \leqq e^{-\lambda k} \tag{3.32}$$

then there exists a constant A (depending on the distribution of $X_1, X_2, \cdots$ but not on n) such that

$$\left| P\left\{ \frac{\sum_1^n X_i}{\left(\operatorname{Var} \sum_1^n X_i\right)^{1/2}} \leqq a \right\} - \Phi(a) \right| \leqq A n^{-1/2}. \tag{3.33}$$

I must admit that I have not written out every detail in the proof of this. However, in applying the bounds (3.17), (3.24), (3.26), and (3.30), let us take

$$m = [K \log n] \tag{3.34}$$

with K an appropriate positive constant. Then, for sufficiently large K, the bound (3.17) is $O(n^{-1/2})$ (or any other power of n), (3.24) is $O(n^{-1/2} \log^2 n)$, (3.26) is $O(n^{-1} \log^3 n)$, and (3.30) (except for the final factor) is $O(n^{-3/4} \log^2 n)$. It seems clear that this final factor cannot be large enough to destroy the bound.

REFERENCES

[1] A. C. BERRY, "The accuracy of the Gaussian approximation to the sum of independent variates," *Trans. Amer. Math. Soc.*, Vol. 49 (1941), pp. 122–136.

[2] C-G. ESSÉEN, "Fourier analysis of distribution functions, a mathematical study of the Laplace-Gaussian law," *Acta Math.*, Vol. 77 (1945), pp. 1–125.

[3] I. A. IBRAGIMOV and YU. V. LINNIK, *Independent and Stationarily Connected Random Variables*, Moscow, Nauka, 1965. (In Russian.)

[4] M. LOÈVE, *Probability Theory*, New York, Van Nostrand, 1963 (3rd ed.).

[5] W. PHILIPP, "The remainder in the central limit theorem for mixing stochastic processes," *Ann. Math. Statist.*, Vol. 40 (1969), pp. 601–609.