THE ACCURACY OF THE GAUSSIAN APPROXIMATION TO THE SUM OF INDEPENDENT VARIATES

BY

ANDREW C. BERRY

1. Introduction

The sum of finitely many variates possesses, under familiar conditions, an almost Gaussian probability distribution. This already much discussed "central limit theorem"(1) in the theory of probability is the object of further investigation in the present paper. The cases of Liapounoff(2), Lindeberg(3), and Feller(4) will be reviewed. Numerical estimates for the degrees of approximation attained in these cases will be presented in the three theorems of §4. Theorem 3, the arithmetical refinement of the general theorem of Feller, constitutes our principal result. As the foregoing implies, we require throughout the paper that the given variates be totally independent. And we consider only one-dimensional variates.

The first three sections of the paper are devoted to the preparatory Theorem 1 in which the variates meet the further condition of possessing finite third order absolute moments. Let $X_1, X_2, \cdots, X_n$ be the given variates. For each k $(k=1, 2, \cdots, n)$ let $\mu_2(X_k)$ and $\mu_3(X_k)$ denote, respectively, the second and third order absolute moments of X_k about its mean (expected) value α_k. These moments are either both zero or both positive. The former case arises only when X_k is essentially constant, i.e., differs from its mean value at most in cases of total probability zero. To avoid trivialities we suppose that $\mu_2(X_k)>0$ for at least one k $(k=1, 2, \cdots, n)$. The non-negative square root of $\mu_2(X_k)$ is the standard deviation of X_k and will be denoted by σ_k. We call

$$\lambda(X_k) = \begin{cases} \mu_3(X_k)/\mu_2(X_k), & \text{if} \quad \mu_2(X_k) \neq 0, \\ 0, & \text{if} \quad \mu_2(X_k) = 0, \end{cases} \tag{1}$$

the *moment-ratio* of X_k. If X_k is essentially bounded(5) and its bound, meas-

Presented to the Society, February 24, 1940; received by the editors May 8, 1940.

(1) H. Cramér, *Random Variables and Probability Distributions* (Cambridge Tracts in Mathematics, no. 36), Cambridge University Press, 1937, pp. 56–64.

(2) A. Liapounoff, *Nouvelle forme du théorème sur la limite de probabilité*, Mémoires de l'Académie des Sciences de St-Petersbourg, (8), vol. 12 (1901). (Cramér, loc. cit., p. 60.)

(3) J. W. Lindeberg, *Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 15 (1922), pp. 211–225. (Cramér, loc. cit., Theorem 21.)

(4) W. Feller, *Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung*, Mathematische Zeitschrift, vol. 40 (1935), pp. 521–559. (Cramér, loc. cit., Theorem 22.)

(5) The author originally developed the early sections of this paper for the case of bounded variates, and is indebted to W. Feller who urged the study, in these sections, of the case of finite third order absolute moments.

ured from its mean value, is denoted by $l(X_k)$, then, as can be verified easily,

$$\lambda(X_k) \leqq l(X_k). \tag{2}$$

We set

$$\Lambda = \max \left\{\lambda(X_1), \lambda(X_2), \cdots, \lambda(X_n)\right\}. \tag{3}$$

The familiar inequality(6) that the square root of the second order absolute moment cannot exceed the cube root of the third order absolute moment readily yields the set of estimates:

$$\sigma_k \leqq \Lambda, \qquad (k = 1, 2, \cdots, n). \tag{4}$$

Consider, now, the variate sum

$$X = X_1 + X_2 + \cdots + X_n. \tag{5}$$

It has the mean value

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \tag{6}$$

and, in virtue of the independence of the given variates, the standard deviation

$$\sigma = (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)^{1/2}. \tag{7}$$

The formula which we are going to derive involves the two numbers Λ and σ, respectively the maximum moment-ratio of the *individual* variates and the standard deviation of their *sum*. Our assumption that the variates are not all constant implies that these numbers are both positive. We introduce their ratio

$$\epsilon = \Lambda/\sigma. \tag{8}$$

It is this ratio which serves as a convenient measure of the extent to which the sum X fails to be Gaussian.

The Gaussian (or Laplacean) distribution function

$$G(x) = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x} e^{-t^2/2}dt, \qquad (-\infty < x < \infty), \tag{9}$$

characterizes the law of probability obeyed by a normal variate of mean value 0 and standard deviation 1. For every x, $G(x)$ is the probability that such a variate assume a value less than x. A normal variate of mean value α and standard deviation σ has the distribution function $G((x-\alpha)/\sigma)$. Let the variate sum X have the distribution function $F(x)$. The least upper bound,

$$M = \sup_{-\infty < x < \infty} \left| F(x) - G((x - \alpha)/\sigma) \right|, \tag{10}$$

(6) Hardy-Littlewood-Pólya, *Inequalities*, Cambridge, 1934, p. 157.

of the modulus of the difference between $F(x)$ and the associated normal distribution function constitutes a precise measure of the "abnormality" of X. A theorem of Liapounoff, about which we shall have more to say in §4, implies

$$\lim_{\epsilon\to 0} M = 0. \tag{11}$$

Our first goal, an arithmetical refinement of (11), reads:

THEOREM 1. $M \leqq (1.88)\epsilon$.

We postpone the proof. The theorem shows that the ratios M/ϵ arising from admissible sets of variates constitute a bounded aggregate of real numbers. Let

$$C = \sup M/\epsilon. \tag{12}$$

Then, for each admissible set of variates, it will be true that $M \leqq C\epsilon$. In addition, corresponding to each positive $C' < C$ there will exist an admissible set of variates having $M > C'\epsilon$. By Theorem 1, $C \leqq 1.88$. We shall now prove the following

THEOREM 2. $C \geqq 1/(2\pi)^{1/2}$.

Proof. Let n be an odd positive integer. Let $X_1, X_2, \cdots, X_n$ be totally independent but similar. Indeed, let each X_k have only the two possible values $+1$ and -1, each with the associated probability $1/2$. We see that $\epsilon = 1/n^{1/2}$. The variate sum X has the $(n+1)$ possible values $\pm 1, \pm 3, \cdots, \pm n$. Its distribution function $F(x)$ is a step-function which is constant in each interval free of these possible values. By symmetry, $F(x) = 1/2$ throughout $-1 < x < 1$. Clearly,

$$M \geqq \lim_{x\to 1\ (x<1)} |F(x) - G(\epsilon x)| = G(\epsilon) - G(0).$$

If we assume $C < 1/(2\pi)^{1/2}$ we can determine n correspondingly large so as to force

$$\frac{M}{\epsilon} \geqq \frac{G(\epsilon) - G(0)}{\epsilon} > C.$$

This, because

$$\lim_{\epsilon\to 0} \frac{G(\epsilon) - G(0)}{\epsilon} = G'(0) = \frac{1}{(2\pi)^{1/2}} \cdot$$

But the variates under consideration form an admissible set and so have $M \leqq C\epsilon$. The contradiction establishes the theorem.

Remark. If, for a given $\eta > 0$, we denote by C_η the least upper bound of those ratios M/ϵ which arise from admissible sets of variates having $\epsilon < \eta$, it is clear that C_η is a monotone increasing function of η. In particular, $C_\eta \leqq C$

for all η. Since, in the proof of Theorem 2, we may replace a satisfactory n by any larger odd integer, we infer that $C_\eta \geqq 1/(2\pi)^{1/2}$ for all η.

Reduction to the case (R). We now take advantage of the fact that moments have been measured about mean values, the fact that Λ and σ enter Theorem 1 only in their ratio ϵ, and our claim that $C \leqq 1.88$. We shall say that we have the case (R) if the given variates meet the additional requirements:

$$\alpha_k = 0, \qquad (k = 1, 2, \cdots, n), \tag{13}$$

$$\sigma = 1, \tag{14}$$

$$\epsilon < \frac{1}{1.88}\cdot \tag{15}$$

The corresponding case of Theorem 1 we shall call Theorem 1(R). We are going to show that the general theorem is a corollary of its own special case. Noting that distribution functions assume only values between 0 and 1, hence that the inequality $M \leqq 1$ is always valid, we see that Theorem 1 is trivial when (15) is false. We may confine our attention, therefore, to the case (15). If, however, the original variates of Theorem 1 do not also satisfy (13) and (14), we introduce the associated variates

$$X_k' = \frac{1}{\sigma}(X_k - \alpha_k), \qquad (k = 1, 2, \cdots, n).$$

These have $M' = M$, $\epsilon' = \epsilon$, and satisfy in detail the hypotheses of Theorem 1(R), all of which can be demonstrated without difficulty. Since the inequalities $M' \leqq (1.88)\epsilon'$ and $M \leqq (1.88)\epsilon$ are equivalent, Theorem 1 is a consequence of Theorem 1(R).

Elementary properties of $F(x) - G(x)$. In the present case (R), the sum X is a reduced variate: its mean value is 0 and its standard deviation is 1. The Bienaymé-Tchebycheff(7) inequality for a reduced variate reads:

$$\Pr\{|X| \geqq x\} \leqq 1/x^2, \qquad (x > 0).$$

Interpreting this in terms of $F(x)$, and equally well for $G(x)$, we infer

$$|F(x) - G(x)| \leqq 1/x^2, \qquad (-\infty < x < \infty). \tag{16}$$

It follows that a sequence of points x on which the modulus $|F(x) - G(x)|$ tends to its least upper bound M forms a bounded set. An easily constructed argument establishes the existence of a finite point $x = b$ for which either

$$F(b+) - G(b) = M, \tag{17}$$

or

$$F(b-) - G(b) = -M. \tag{17'}$$

(7) Cramér, loc. cit., p. 21.

Lemma 1. *There exists a number a such that one of the inequalities*

$$F(x+a) - G(x+a) \geqq \frac{\delta - x}{(2\pi)^{1/2}}, \tag{18}$$

$$F(x+a) - G(x+a) \leqq \frac{-\delta - x}{(2\pi)^{1/2}} \tag{18'}$$

holds throughout the interval $-\delta < x < \delta$ *where*

$$\delta = M \cdot \left(\frac{\pi}{2}\right)^{1/2}. \tag{19}$$

Proof. In case (17) write $a = b + \delta$. Then, for all $x > b$,

$$F(x) \geqq F(b+), \qquad G(x) \leqq G(b) + \frac{x-b}{(2\pi)^{1/2}},$$

the first because $F(x)$ is non-decreasing, the second since $G(x)$ is differentiable and $G'(x) \leqq 1/(2\pi)^{1/2}$. Whence,

$$F(x) - G(x) \geqq M - \frac{x-b}{(2\pi)^{1/2}}$$

for all $x > b$. This implies (18) for all $x > -\delta$, *a fortiori* for $-\delta < x < \delta$. In case (17′) a symmetric argument employing $a = b - \delta$ establishes (18′) in $-\delta < x < \delta$.

2. The proof of Theorem 1(R)

Our proof, which rests on the calculation to be presented in §3, utilizes the characteristic functions

$$\phi(t) = \int_{-\infty}^{\infty} e^{ixt} dF(x), \tag{20}$$

$$\psi(t) = \int_{-\infty}^{\infty} e^{ixt} dG(x) = e^{-t^2/2}, \qquad (-\infty < t < \infty), \tag{21}$$

of the distributions under discussion. In (21) the indicated evaluation is familiar. Since $G(x)$ is a differentiable function, the Stieltjes integral can be written as an ordinary integral:

$$\psi(t) = \int_{-\infty}^{\infty} e^{ixt} G'(x) dx = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} e^{-x^2/2} e^{ixt} dx.$$

Thus, $\psi(t)$ is the Fourier transform of $e^{-x^2/2}$. That this is $e^{-t^2/2}$ can be verified readily.

We are interested primarily in the difference

$$\phi(t) - \psi(t) = \int_{-\infty}^{\infty} e^{ixt} d\{F(x) - G(x)\}. \tag{22}$$

From (16) we infer that $F(x)-G(x)$ is absolutely integrable over $-\infty < x < \infty$ and tends to 0 when $x \to \pm\infty$. These facts justify the integration by parts which yields

$$\frac{\phi(t) - \psi(t)}{-it} = \int_{-\infty}^{\infty} \{F(x) - G(x)\} e^{ixt} dx. \tag{23}$$

It is known(8) that if the difference (22) is uniformly small in some finite interval about $t=0$, then M is small. This fact suggests the following procedure.

In (23) we replace x by $x+a$ (the a of Lemma 1) and obtain

$$\frac{\phi(t) - \psi(t)}{-it} \cdot e^{-iat} = \int_{-\infty}^{\infty} \{F(x + a) - G(x + a)\} e^{ixt} dx. \tag{24}$$

We confine t to the finite interval

$$-T \leqq t \leqq T, \tag{25}$$

by employing a weighting factor $w(t)$ which vanishes outside this interval. We choose

$$w(t) = T - |t|, \qquad \text{when } -T \leqq t \leqq T. \tag{26}$$

The Fourier transform of $w(t)$ is, except for a missing constant factor,

$$W(x) = \int_{-T}^{T} w(t) e^{ixt} dt = \frac{2(1 - \cos Tx)}{x^2}. \tag{27}$$

The Parseval theorem in the theory of Fourier transforms assures us of the validity of the equality

$$\int_{-\infty}^{\infty} W(x)\{F(x + a) - G(x + a)\} dx = \int_{-T}^{T} w(t) \frac{\phi(t) - \psi(t)}{-it} e^{-iat} dt. \tag{28}$$

This may be derived directly by multiplying (24) throughout by $w(t)$, integrating over (25), and inverting the order of integration in the resulting iterated integral. The last step is justified by the absolute integrability of the product $w(t)\{F(x+a)-G(x+a)\}$ over the strip $-T \leqq t \leqq T$, $-\infty < x < \infty$.

Since, as a brief inspection of (22) shows, the modulus $|\phi(t)-\psi(t)|$ is an even function of t, we can derive from the Parseval equality (28) the inequality

(8) P. Lévy, *Théorie de l'Addition des Variables Aléatoires* (Monographies des Probabilités), Paris, 1937, p. 49. Cramér, loc. cit., p. 29, Theorem 11.

$$\left|\int_{-\infty}^{\infty} \frac{1-\cos Tx}{x^2}\{F(x+a)-G(x+a)\}dx\right| \leqq \int_0^T (T-t)\frac{|\phi(t)-\psi(t)|}{t}dt. \tag{29}$$

Noting that $|F(x)-G(x)| \leqq M = \delta(2/\pi)^{1/2}$ for all x, we find

$$\left|\left(\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right)\frac{1-\cos Tx}{x^2}\{F(x+a)-G(x+a)\}dx\right| \leqq 2(2/\pi)^{1/2}\delta\int_{\delta}^{\infty}\frac{1-\cos Tx}{x^2}dx.$$

On the other hand, Lemma 1 yields, equally in its two cases,

$$\left|\int_{-\delta}^{\delta}\frac{1-\cos Tx}{x^2}\{F(x+a)-G(x+a)\}dx\right| \geqq (2/\pi)^{1/2}\delta\int_0^{\delta}\frac{1-\cos Tx}{x^2}dx.$$

These two estimates, the standard evaluation

$$\int_0^{\infty}\frac{1-\cos x}{x^2}dx = \frac{\pi}{2},$$

the triangle inequality

$$\left|\int_{-\infty}^{\infty}\right| \geqq \left|\int_{-\delta}^{\delta}\right| - \left|\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right|,$$

and certain obvious reductions enable us to deduce from (29) the following:

$$A(T\delta) \leqq \int_0^T (T-t)\frac{|\phi(t)-\psi(t)|}{t}dt, \tag{30}$$

where

$$A(u) = (2/\pi)^{1/2}\cdot u\cdot\left\{3\int_0^u \frac{1-\cos x}{x^2}dx - \pi\right\}. \tag{31}$$

In §3, for the particular choice $T=1.1/\epsilon$, we shall prove that the right member of (30) does not exceed

$$\frac{1.1}{6}\cdot\left(\frac{\pi}{2}\right)^{1/2}.$$

Thus,

$$A\left(\frac{1.1}{\epsilon}\delta\right) < 0.2298. \tag{32}$$

Granting this for the moment, we employ the Taylor inequality

$$\cos x \leqq 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!},$$

and make the calculation:

$$A(2.59) > 0.2298. \tag{33}$$

Now, at least when it is positive, $A(u)$ is an increasing function. It follows that

$$M = (2/\pi)^{1/2}\cdot\delta \leqq (2/\pi)^{1/2}\cdot\frac{(2.59)\epsilon}{1.1} < (1.88)\epsilon. \tag{34}$$

We have completed the proof of the fact that Theorem 1(R) rests on the calculation to be presented in the next section.

3. The calculation

The individual distribution functions. We must return to the individual variates $X_1, X_2, \cdots, X_n$ of Theorem 1(R). The corresponding distribution functions $F_1(x), F_2(x), \cdots, F_n(x)$ all increase monotonically and have the common limits 0, 1 when x extends respectively to $-\infty, +\infty$. The hypotheses of Theorem 1(R) require, for each k $(k=1, 2, \cdots, n)$,

$$\int_{-\infty}^{\infty} x dF_k(x) = 0, \qquad \int_{-\infty}^{\infty} x^2 dF_k(x) = \sigma_k^2,$$

$$\int_{-\infty}^{\infty} |x|^3 dF_k(x) = \mu_3(X_k) \leqq \epsilon\cdot\sigma_k^2.$$

Inequalities (4) here read: $\sigma_k \leqq \epsilon$. We shall feel free to use the foregoing facts without further explicit reference.

The individual characteristic functions. The variate X_k has the characteristic function

$$\phi_k(t) = \int_{-\infty}^{\infty} e^{ixt} dF_k(x) = 1 - \tfrac{1}{2}\sigma_k^2 t^2 + \cdots.$$

The constant term and, better, the two indicated terms of the formal Taylor expansion will be employed as approximations. We shall need estimates for the errors

$$u_k = \phi_k(t) - 1 = \int_{-\infty}^{\infty} (e^{ixt} - 1 - ixt) dF_k(x), \tag{35}$$

$$v_k = \phi_k(t) - 1 + \tfrac{1}{2}\sigma_k^2 t^2 = \int_{-\infty}^{\infty} (e^{ixt} - 1 - ixt + \tfrac{1}{2}x^2t^2) dF_k(x). \tag{36}$$

The elementary inequalities

$$\left| e^{ixt} - 1 - ixt \right| \leqq \tfrac{1}{2}x^2t^2, \qquad \left| e^{ixt} - 1 - ixt + \tfrac{1}{2}x^2t^2 \right| \leqq \tfrac{1}{6} \left| x \right|^3 \left| t \right|^3,$$

which are valid for all real x and t, at once yield

$$\left| u_k \right| \leqq \tfrac{1}{2}\sigma_k^2 t^2, \tag{37}$$

$$\left| v_k \right| \leqq \tfrac{1}{6}\epsilon \cdot \sigma_k^2 \left| t \right|^3. \tag{38}$$

We now confine our attention to the interval

$$0 \leqq t < \frac{2^{1/2}}{\epsilon}, \tag{39}$$

in which, by virtue of (37), we may be certain that $|u_k| < 1$. In this interval, then, $\phi_k(t)$ is never 0 and we may define

$$\log \phi_k(t) = \log (1 + u_k) = \int_0^{u_k} \frac{dz}{1+z}, \tag{40}$$

for example by employing as the path of integration in the complex plane the straight line segment joining $z=0$ to $z=u_k$. For the difference

$$r_k = \log (1 + u_k) - u_k = - \int_0^{u_k} \frac{zdz}{1+z},$$

we obtain the estimate

$$\begin{aligned} \left| r_k \right| &\leqq \int_0^{|u_k|} \frac{xdx}{1-x} \leqq \int_0^{\sigma_k^2 t^2/2} \frac{xdx}{1-x} = \sigma_k^4 \int_0^{t^2/2} \frac{ydy}{1-\sigma_k^2 y} \\ &\leqq \epsilon^2 \sigma_k^2 \int_0^{t^2/2} \frac{ydy}{1-\epsilon^2 y} = - \sigma_k^2 t^2 \left\{ \frac{\log (1 - \frac{1}{2}\epsilon^2 t^2) + \frac{1}{2}\epsilon^2 t^2}{\epsilon^2 t^2} \right\}. \end{aligned}$$

If we write

$$h(x) = \frac{x}{6} - \frac{1}{x^2}\left\{ \log\left(1 - \frac{x^2}{2}\right) + \frac{x^2}{2} \right\}, \qquad (0 \leqq x < 2^{1/2}), \tag{41}$$

and observe that

$$\log \phi_k(t) = - \tfrac{1}{2}\sigma_k^2 t^2 + v_k + r_k,$$

we find that we have proved

LEMMA 2. *Throughout the interval* (39),

$$\log \phi_k(t) = - \tfrac{1}{2}\sigma_k^2 t^2 + \sigma_k^2 \Delta_k,$$

where, for each k $(k = 1, 2, \cdots, n)$, $\left| \Delta_k \right| \leqq t^2 \cdot h(\epsilon t)$.

The characteristic function of the variate sum. We now encounter the essential reason for the hypothesis that the given variates be totally independent. This hypothesis guarantees the relation

$$\phi(t) = \phi_1(t)\phi_2(t) \cdots \phi_n(t). \tag{42}$$

This implication is well known(9). With a suitable determination of the logarithm of the product, we can now derive from Lemma 2

LEMMA 3. *Throughout the interval* (39),

$$\log \phi(t) = -\tfrac{1}{2}t^2 + \Delta,$$

where

$$|\Delta| \leqq t^2 \cdot h(\epsilon t).$$

Recalling that $\psi(t) = e^{-t^2/2}$, and observing that

$$\phi(t) - \psi(t) = (e^{\Delta} - 1)\psi(t),$$
$$|e^{\Delta} - 1| \leqq e^{|\Delta|} - 1,$$

we obtain from Lemma 3 the final

LEMMA 4. *Throughout the interval* (39),

$$|\phi(t) - \psi(t)| \leqq \{e^{t^2 \cdot h(\epsilon t)} - 1\}e^{-t^2/2}.$$

The integral B. We are now in position to calculate an upper bound for the integral which constitutes the right member of the inequality (30). For the choice $T = 1.1/\epsilon$ it is clear from Lemma 4 that this integral is dominated in magnitude by the integral

$$B = \int_0^{1.1/\epsilon} \left(\frac{1.1}{\epsilon} - t\right) \frac{e^{t^2 \cdot h(\epsilon t)} - 1}{t} e^{-t^2/2} dt. \tag{43}$$

We shall divide the calculation into three major parts by writing

$$B = B_1 + B_2 + B_3, \tag{44}$$

where, in terms of the abbreviation $c = 0.75$,

$$\begin{aligned} B_1 &= \frac{\epsilon}{6} \int_0^{c/\epsilon} \left(\frac{1.1}{\epsilon} - t\right) t^2 e^{-t^2/2} dt, \\ B_2 &= \int_0^{c/\epsilon} \left(\frac{1.1}{\epsilon} - t\right) \frac{e^{t^2 \cdot h(\epsilon t)} - 1 - \frac{1}{6}\epsilon t^3}{t} e^{-t^2/2} dt, \\ B_3 &= \int_{c/\epsilon}^{1.1/\epsilon} \left(\frac{1.1}{\epsilon} - t\right) \frac{e^{t^2 \cdot h(\epsilon t)} - 1}{t} e^{-t^2/2} dt. \end{aligned} \tag{45}$$

(9) Cramér, loc. cit., p. 36.

The term B_1. With the aid of a few integrations by parts, and standard evaluations, we find we can write B_1 in the form

$$B_1 = \frac{1.1}{6}\left(\frac{\pi}{2}\right)^{1/2} - \frac{\epsilon}{3} - \frac{1}{6}\int_{c/\epsilon}^{\infty}\left\{(1.1 - c)t^2 + c - \frac{2c}{t^2}\right\}e^{-t^2/2}dt.$$

The factor in braces in the integrand is an increasing function of t. At precisely this point we use the hypothesis (15). This implies that $t > (1.88)c$ in the mentioned factor. We infer that this factor is positive. Hence,

$$B_1 \leqq \frac{1.1}{6}\left(\frac{\pi}{2}\right)^{1/2} - \frac{\epsilon}{3}. \tag{46}$$

The term B_2. We begin with the observation that, for each fixed positive $t \leqq c/\epsilon$,

$$\frac{1}{\epsilon^2}\left(e^{t^2\cdot h(\epsilon t)} - 1 - \frac{\epsilon t^3}{6}\right)$$

is an increasing function of ϵ throughout $0 < \epsilon \leqq c/t$. This can be proved, for example, by differentiating with respect to ϵ and employing $1 \leqq e^{t^2\cdot h(\epsilon t)}$ and $\frac{1}{6} = h'(0) \leqq h'(\epsilon t)$, to show that this derivative is positive. It follows that

$$B_2 \leqq B_2' = \frac{\epsilon^2}{c^2}\int_0^{c/\epsilon}\left(\frac{1.1}{\epsilon} - t\right)t\left\{e^{t^2\cdot h(c)} - 1 - \frac{ct^2}{6}\right\}e^{-t^2/2}dt.$$

Let us increase the upper limit of integration from c/ϵ to $1.1/\epsilon$ thus replacing B_2' by a still larger number, say B_2''. An integration by parts then shows that

$$B_2'' = \frac{(1.1)\epsilon}{c^2}\left\{\frac{1}{[1 - 2h(c)]^{1/2}} - 1 - \frac{c}{3}\right\} - \frac{\epsilon}{c}\int_0^{1.1/\epsilon}\left\{\frac{e^{t^2\cdot h(c)}}{[1 - 2h(c)]^{1/2}} - 1 - \frac{c}{3} - \frac{ct^2}{6}\right\}e^{-t^2/2}dt.$$

The computation $0.2121 < h(c) < 0.2122$ proves $c[1 - 2h(c)]^{1/2} < 6h(c)$ and so demonstrates that the factor in braces in the last integrand is an increasing function of t^2. Since this factor is positive at $t = 0$, it follows that we may neglect the integral. Thus,

$$B_2 \leqq (0.134)\epsilon. \tag{47}$$

The term B_3. From the numerator in the middle factor of the integrand in B_3 we reject the term -1. The resulting larger integral we call B_3'. This can be written

$$B_3' = \frac{1}{\epsilon}\int_c^{1.1}\left(\frac{1.1}{u} - 1\right)e^{-H(u)/\epsilon^2}du,$$

where

$$H(u) = u^2\{\tfrac{1}{2} - h(u)\}.$$

It can be seen that $H(u) > 0$ in the interval of integration. This permits us to employ the elementary inequality

$$x \cdot e^{-x} \leqq 1/e, \qquad (0 < x),$$

and find

$$B_3' \leqq \frac{\epsilon}{e} \int_c^{1.1} \left(\frac{1.1}{u} - 1\right) \frac{du}{H(u)} \cdot$$

An examination of the derivative

$$H'(u) = u \cdot \left(1 - \frac{u}{2} - \frac{u^2}{2 - u^2}\right)$$

shows that $H(u)$ increases in $c \leqq u \leqq u'$ and decreases in $u' \leqq u \leqq 1.1$ where u' is the unique root of $H'(u)$ in the interval of integration. We find $0.85 < u' < 0.851$. Since $H(1) < H(0.75)$ we infer that

$$H(u) \geqq \begin{cases} H(1) & > 0.14010 \text{ when } 0.75 \leqq u \leqq 1, \\ H(1.05) & > 0.10814 \text{ when } 1 \leqq u \leqq 1.05, \\ H(1.1) & > 0.05674 \text{ when } 1.05 \leqq u \leqq 1.1. \end{cases}$$

By dividing the interval of integration into the three indicated parts and by replacing the reciprocal of $H(u)$ in each part by its upper bound in that part, we obtain the estimate

$$B_3 \leqq (0.195)\epsilon. \tag{48}$$

Summary of calculation. We have proved

$$B \leqq \frac{1.1}{6}\left(\frac{\pi}{2}\right)^{1/2} + \epsilon(-\tfrac{1}{3} + 0.134 + 0.195) < \frac{1.1}{6}\left(\frac{\pi}{2}\right)^{1/2}. \tag{49}$$

This shows that (32) is a valid consequence of (30) and so completes the argument establishing Theorem 1(R) and its dependent extension Theorem 1.

4. Generalizations

Let $X_1, X_2, \cdots, X_n$ be totally independent variates with the respective distribution functions $F_1(x), F_2(x), \cdots, F_n(x)$. Let $s > 0$ and $a_1, a_2, \cdots, a_n$ be any real numbers. We wish to compare the distribution function $F(x)$ of the variate sum $X = X_1 + X_2 + \cdots + X_n$ with the distribution function $G((x-a)/s)$ of a normal variate of mean value $a = a_1 + a_2 + \cdots + a_n$ and standard deviation s. We put

$$M = \sup_{-\infty<x<\infty} |F(x) - G((x-a)/s)|.$$

For a given $\epsilon > 0$ we introduce the three quantities

$$\epsilon_0 = \sum_{k=1}^{n} \Pr\{|X_k - a_k| > \epsilon s\},$$

$$\epsilon_1 = \frac{1}{s} \cdot \sum_{k=1}^{n} \left| \int_{a_k-\epsilon s}^{a_k+\epsilon s} (x_k - a_k) dF_k(x) \right|,$$

$$\epsilon_2 = \left| 1 - \frac{1}{s^2} \cdot \sum_{k=1}^{n} \int_{a_k-\epsilon s}^{a_k+\epsilon s} (x_k - a_k)^2 dF_k(x) \right|.$$

Feller (loc. cit.) has proved that if these three numbers are small then M is small. We shall now establish

THEOREM 3. *If* $\epsilon_0 \leqq \epsilon$, $\epsilon_1 \leqq \epsilon$, $\epsilon_2 \leqq \epsilon$, *then*

$$M \leqq (5.8)\epsilon.$$

More generally, whenever $\epsilon_1^2 + \epsilon_2 < 1$, *we have*

$$M \leqq \frac{C(\epsilon + \epsilon_1)}{(1 - \epsilon_1^2 - \epsilon_2)^{1/2}} + \epsilon_0 + \frac{\epsilon_1}{(2\pi)^{1/2}} + \frac{1}{(2\pi e)^{1/2}} \log \frac{1}{(1 - \epsilon_1^2 - \epsilon_2)^{1/2}},$$

where C ($\leqq 1.88$) *is the constant furnished by Theorem* 1.

Proof. Without loss of generality we may confine our attention to the special case

(R) $$s = 1, \quad a_1 = a_2 = \cdots = a_n = 0.$$

(We have but to introduce $X_k' = (1/s)(X_k - a_k)$ and to set $s' = 1$, $a_1' = a_2' = \cdots = a_n' = 0$, $\epsilon' = \epsilon$ in order to discover, first, that $M' = M$, $\epsilon_0' = \epsilon_0$, $\epsilon_1' = \epsilon_1$, $\epsilon_2' = \epsilon_2$ and, therefore, that the theorem is equivalent to its special case.)

We use a familiar device([10]). We approximate the given variates by the associated bounded variates

$$\overline{X}_k = \begin{cases} X_k, & \text{if } |X_k| \leqq \epsilon, \\ 0, & \text{if } |X_k| > \epsilon, \end{cases} \qquad (k = 1, 2, \cdots, n).$$

To these totally independent variates we are going to apply Theorem 1. Now, $\overline{X}_k$ has the mean value

$$\bar{\alpha}_k = \int_{-\epsilon}^{\epsilon} x dF_k(x), \qquad (k = 1, 2, \cdots, n).$$

This is easily verified. Next, since

([10]) Lévy, loc. cit., pp. 104–110.

$$\sum_{k=1}^{n} |\bar{\alpha}_k| = \epsilon_1, \tag{50}$$

we see that the bound of $\overline{X}_k$, measured from the mean value, does not exceed $\epsilon+\epsilon_1$. Recalling (2) we find

$$\overline{\Lambda} \leqq \epsilon + \epsilon_1, \tag{51}$$

where $\overline{\Lambda}$ is the maximum moment-ratio of the variates $\overline{X}_k$. The sum, $\overline{X}$, of these variates has the standard deviation $\bar{\sigma}$ given by

$$\bar{\sigma}^2 = \sum_{k=1}^{n} \int_{-\epsilon}^{\epsilon} x^2 dF_k(x) - \sum_{k=1}^{n} \bar{\alpha}_k^2.$$

The second term of this difference is non-negative and, by (50), not greater than ϵ_1^2. Thus,

$$(1 - \epsilon_1^2 - \epsilon_2)^{1/2} \leqq \bar{\sigma} \leqq (1 + \epsilon_2)^{1/2}. \tag{52}$$

By Theorem 1,

$$\left| \overline{F}(x) - G\left(\frac{x - \bar{\alpha}}{\bar{\sigma}}\right) \right| \leqq \frac{C(\epsilon + \epsilon_1)}{(1 - \epsilon_1^2 - \epsilon_2)^{1/2}}, \qquad (-\infty < x < \infty), \tag{53}$$

where $\overline{F}(x)$ denotes the distribution function of $\overline{X}$, and $\bar{\alpha}$ its mean value.

Next, we observe that $\overline{X}$ differs from X at most in those cases in which for at least one k, $|X_k| > \epsilon$. These cases have a total probability of occurrence not greater than ϵ_0. Hence,

$$|F(x) - \overline{F}(x)| \leqq \epsilon_0, \qquad (-\infty < x < \infty). \tag{54}$$

If we combine (53) and (54) we obtain an inequality sharper, in some respects, than that announced in the theorem. But we must pass from $G((x-\bar{\alpha})/\bar{\sigma})$ to the reduced normal distribution function $G(x)$ desired in the present case (R). Since, by (50), $|\bar{\alpha}| \leqq \epsilon_1$, and since $e^{-t^2/2} \leqq 1$, we have

$$|G(x) - G(x - \bar{\alpha})| \leqq \frac{\epsilon_1}{(2\pi)^{1/2}}, \qquad (-\infty < x < \infty). \tag{55}$$

And the elementary inequality $|te^{-t^2/2}| \leqq 1/e^{1/2}$ yields, for all x,

$$\left| G(x - \bar{\alpha}) - G\left(\frac{x - \bar{\alpha}}{\bar{\sigma}}\right) \right| \leqq \frac{|\log \bar{\sigma}|}{(2\pi e)^{1/2}} \leqq \frac{1}{(2\pi e)^{1/2}} \log \frac{1}{(1 - \epsilon_1^2 - \epsilon_2)^{1/2}}. \tag{56}$$

The general inequality of the theorem is an immediate consequence of the inequalities (53)–(56).

Finally, we consider the special case $\epsilon_0 \leqq \epsilon$, $\epsilon_1 \leqq \epsilon$, $\epsilon_2 \leqq \epsilon$. Since it is always true that $M \leqq 1$ we may assume that $\epsilon < 1/5.8$, the desired inequality $M \leqq (5.8)\epsilon$ being trivial in the contrary case. Thus, the condition $\epsilon_1^2 + \epsilon_2 < 1$

is met generously and we may apply the already established general inequality. This yields

$$\frac{M}{\epsilon} \leqq \frac{3.76}{(1-\epsilon^2-\epsilon)^{1/2}} + 1 + \frac{1}{(2\pi)^{1/2}} + \frac{1}{\epsilon(2\pi e)^{1/2}} \log \frac{1}{(1-\epsilon^2-\epsilon)^{1/2}} \cdot$$

It is easily seen that the right member is an increasing function of ϵ in $0<\epsilon<1/5.8$ and is less than 5.8 when $\epsilon=1/5.8$. This proves the theorem.

The case (L). Let the standard deviation of each X_k be finite. Let $s=\sigma$, $a_1=\alpha_1$, $a_2=\alpha_2$, $\cdots$, $a_n=\alpha_n$, where the α's are the mean values of the individual variates and σ is the standard deviation of their sum. If these conditions are met we shall say we have the case (L). In this case we can express ϵ_1 and ϵ_2 in the convenient forms:

$$\epsilon_1 = \frac{1}{\sigma} \cdot \sum_{k=1}^{n} \left| \left(\int_{-\infty}^{\alpha_k-\epsilon\sigma} + \int_{\alpha_k+\epsilon\sigma}^{\infty} \right) (x-\alpha_k) dF_k(x) \right|,$$

$$\epsilon_2 = \frac{1}{\sigma^2} \cdot \sum_{k=1}^{n} \left(\int_{-\infty}^{\alpha_k-\epsilon\sigma} + \int_{\alpha_k+\epsilon\sigma}^{\infty} \right) (x-\alpha_k)^2 dF_k(x).$$

These imply

$$\epsilon_2 \geqq \epsilon^2 \epsilon_0, \qquad \epsilon_2 \geqq \epsilon \cdot \epsilon_1. \tag{57}$$

Lindeberg (loc. cit.) has proved that if ϵ_2 is small then M is small. We shall now establish

THEOREM 4. *In case* (L) *if* $\epsilon_2 \leqq \epsilon^3$, *then* $M \leqq (3.6)\epsilon$.

Proof. We shall assume $\epsilon<1/3.6$ since the theorem is trivial in the contrary case. By (57), we have $\epsilon_0 \leqq \epsilon$, $\epsilon_1 \leqq \epsilon^2$. The general inequality of Theorem 3 yields

$$\frac{M}{\epsilon} \leqq \frac{(1.88)(1+\epsilon)}{(1-\epsilon^4-\epsilon^3)^{1/2}} + 1 + \frac{1}{(2\pi)^{1/2}} + \frac{1}{\epsilon(2\pi e)^{1/2}} \log \frac{1}{(1-\epsilon^4-\epsilon^3)^{1/2}} \cdot$$

The right member is an increasing function of ϵ in $0<\epsilon<1/3.6$ and has a value <3.6 when $\epsilon=1/3.6$. This proves the theorem.

Finally, we consider the subcase in which for some (not necessarily integral) $m>2$, each X_k has a finite absolute moment (about its mean value) of order m: $\mu_m(X_k)$. We put $\eta=(1/\sigma^m)\sum_{k=1}^{n}\mu_m(X_k)$. Liapounoff (loc. cit.) proved M small when η is small.

THEOREM 5. $M \leqq (3.6)\eta^{1/(m+1)}$.

Proof. We define $\epsilon=\eta^{1/(m+1)}$. For this ϵ it is easy to prove $\epsilon_2 \leqq \epsilon^3$. The present theorem thus follows from Theorem 4.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.