PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 90, Number 1, January 1984

REVERSING THE BERRY-ESSÉEN INEQUALITY

PETER HALL AND A. D. BARBOUR

ABSTRACT. We derive a lower bound to the rate of convergence in the central limit theorem. Our result is expressed in terms similar to those of the Berry-Esséen inequality, with the distance between two distributions on one side of the inequality and an easily calculated function of the summands on the other, related by a universal constant. The proof is based on Stein's method.

The Berry-Esséen inequality [**1**, **3**] is one of the most informative results on rates of convergence in the central limit theorem. It tells us how close the normal distribution is to the distribution of a sum of independent random variables. In its original form, it applies where the summands have finite third moments, but more general versions [**2**, **4**, **7**] permit this condition to be relaxed.

Our aim here is to present an inequality which describes not how close two distributions are, but how far they are apart. The Berry-Esséen inequality gives an upper bound to the rate of convergence in the central limit theorem, and our inequality gives a lower bound.

Let $Y_1, \ldots, Y_n$ be independent random variables with zero means and with variances $\sigma_i^2 = EY_i^2$ satisfying $\sum_{i=1}^n \sigma_i^2 = 1$. If the Y_i's form the nth row of a triangular array, and if Lindeberg's condition holds for the array, then the sum $\sum_{i=1}^n Y_i$ should be approximately normal $N(0, 1)$. Results of Rozovskii [**8**, **9**] and Hall [**5**, **6**, Theorems 2.2 and 2.4, pp. 25 and 46] suggest that the uniform distance,

$$\Delta \equiv \sup_w \left| P\left[\sum_{i=1}^n Y_i \leqslant w \right] - \Phi(w) \right|,$$

where Φ is the standard normal distribution function, is of order of magnitude equal to

$$\delta \equiv \sum_{i=1}^n E\{Y_i^2 I[|Y_i| > 1]\} + \sum_{i=1}^n E\{Y_i^4 I[|Y_i| \leqslant 1]\} + \left| \sum_{i=1}^n E\{Y_i^3 I[|Y_i| \leqslant 1]\} \right|,$$

to a first approximation. For example, when the summands are identically distributed, the ratio $(\delta + n^{-1/2})/(\Delta + n^{-1/2})$ is bounded away from zero and infinity as $n \to \infty$.

Received by the editors February 1, 1983 and, in revised form, May 20, 1983.
1980 *Mathematics Subject Classification*. Primary 60F05, 60G50; Secondary 60E15.
Key words and phrases. Berry-Esséen inequality, central limit theorem, lower bound, rate of convergence, sums of independent variables.

Note that $\delta \to 0$ under Lindeberg's condition. Therefore, an ideal lower bound might take the form $\delta \leqslant C\Delta$, for some constant C. However, such an inequality is impossible, since, if each Y_i is normal $N(0, \sigma_i^2)$, Δ equals zero and δ is roughly of order $\sum_{i=1}^n \sigma_i^4$. Instead, we prove the next best thing.

THEOREM. *There exists a universal constant C such that*

$$\delta \leqslant C\left(\Delta + \sum_{i=1}^n \sigma_i^4\right). \tag{1}$$

Note that, in the case where the Y_i's are identically distributed, $\sum_{i=1}^n \sigma_i^4 = 1/n$.

Several earlier attempts have been made to derive inequalities of this kind, but all have required extra conditions or have involved extra terms. See, for example, Theorem 1 of Rozovskii [**8**], but note that Rozovskii does not assume the summands to have finite mean and variance. Our proofs are very different from the Fourier arguments of earlier workers, and are based on Stein's method [**10**].

PROOF. Define $W = \sum_{i=1}^n Y_i$ and $W_i = W - Y_i$. Let $f(w)$ denote either $e^{-w^2/2}$ or $we^{-w^2/2}$, $-\infty < w < \infty$, and set $h(w) = f'(w) - wf(w)$. In both cases, $Eh(N) = 0$ if N is normal $N(0, 1)$, so that

$$|Eh(W)| = \left|\int_{-\infty}^{\infty} h'(w)\{P[W \leqslant w] - \Phi(w)\}\, dw\right| \leqslant C_1\Delta, \tag{2}$$

where we define

$$C_1 = \int_{-\infty}^{\infty} |h'(w)|dw, \quad C_2 = \sup_w |f'''(w)|, \quad C_3 = \int_{-\infty}^{\infty} |f''(w)|dw.$$

Furthermore,

$$\begin{aligned} Eh(W) &= \sum_{i=1}^n E\{\sigma_i^2 f'(W_i + Y_i) - Y_i f(W_i + Y_i)\} \\ &= \sum_{i=1}^n E\big[\sigma_i^2\{f'(W_i + Y_i) - f'(W_i) - Y_i f''(W_i)\} \\ &\qquad - Y_i\{f(W_i + Y_i) - f(W_i) - Y_i f'(W_i)\}\big] \end{aligned} \tag{3}$$

and

$$\sum_{i=1}^n \sigma_i^2 |E\{f'(W_i + Y_i) - f'(W_i) - Y_i f''(W_i)\}| \leqslant \frac{1}{2} C_2 \sum_{i=1}^n \sigma_i^4. \tag{4}$$

Let W^0 and N denote, respectively, a copy of W and a standard normal variable, both independent of $Y_1, \ldots, Y_n$. Define

(5)

$$\begin{aligned} t_{1i} &\equiv E\big\{Y_i\big[\{f(W_i + Y_i) - f(W_i) - Y_i f'(W_i)\} \\ &\qquad - \{f(W^0 + Y_i) - f(W^0) - Y_i f'(W^0)\}\big]\big\} \\ &= E\left[Y_i^2 \int_0^1 E\{f'(W_i + tY_i) - f'(W^0 + tY_i) \mid Y_i\}\, dt\right] + \sigma_i^2 E\{f'(W^0) - f'(W_i)\}. \end{aligned}$$

For any constant β,

$$\begin{aligned}
&\left|E\{f'(W_i+\beta)-f'(W^0+\beta)\}\right| \\
&\quad=\left|\int_{-\infty}^{\infty} dP(W_i\leqslant w)\int_{-\infty}^{\infty}\{f'(w+\beta)+zf''(w+\beta)-f'(w+z+\beta)\}\,dP(Y_i\leqslant z)\right| \\
&\quad\leqslant\int_{-\infty}^{\infty} dP(W_i\leqslant w)\int_{-\infty}^{\infty}\tfrac{1}{2}z^2\Big\{\sup_v |f'''(v)|\Big\}\,dP(Y_i\leqslant z)=\tfrac{1}{2}C_2\sigma_i^2,
\end{aligned}$$

and so, from (5),

$$|t_{1i}|\leqslant C_2\sigma_i^4. \tag{6}$$

Similarly, we may derive the results

$$\begin{aligned}
t_{2i}&\equiv E\Big\{Y_i\big[\{f(W^0+Y_i)-f(W^0)-Y_if'(W^0)\} \\
&\qquad\qquad -\{f(N+Y_i)-f(N)-Y_if'(N)\}\big]\Big\} \\
&=E\left[Y_i^2\int_0^1 E\{f'(W^0+tY_i)-f'(N+tY_i)\,|\,Y_i\}\,dt\right]+\sigma_i^2E\{f'(N)-f'(W^0)\},
\end{aligned}$$

$$\left|E\{f'(W^0+\beta)-f'(N+\beta)\}\right|=\left|\int_{-\infty}^{\infty} f''(w+\beta)\{P[W\leqslant w]-\Phi(w)\}\,dw\right|\leqslant C_3\Delta,$$

and $|t_{2i}|\leqslant 2C_3\sigma_i^2\Delta$. It follows from this estimate, (2)–(4) and (6) that

$$\left|\sum_{i=1}^n A_i\right|\leqslant (C_1+2C_3)\Delta+\frac{3}{2}C_2\sum_{i=1}^n\sigma_i^4, \tag{7}$$

where $A_i\equiv E[Y_i\{f(N+Y_i)-f(N)-Y_if'(N)\}]=E\{Y_i^2\psi(Y_i)\}$ and $\psi(y)\equiv y^{-1}E\{f(N+y)-f(N)-yf'(N)\}$.

Henceforth, we consider the functions $f_1(w)=we^{-w^2/2}$ and $f_2(w)=e^{-w^2/2}$ separately. Let C_{ij} denote the version of C_j computed for $f=f_i$, and define

$$D=\sup_{0<y\leqslant 1} y^{-2}\left|y^{-1}\big(1-e^{-y^2/4}\big)-y/4\right|<\frac{1}{32}.$$

When $f=f_1$, we have

$$-2^{3/2}\psi(y)=1-e^{-y^2/4}\geqslant(1-e^{-1/4})(y^2\wedge 1),$$

and so, by (7),

$$\begin{aligned}
&\sum_{i=1}^n E\big\{Y_i^2I\big[|Y_i|>1\big]\big\}+\sum_{i=1}^n E\big\{Y_i^4I\big[|Y_i|\leqslant 1\big]\big\} \\
&\qquad\leqslant 2^{3/2}(1-e^{-1/4})^{-1}\left\{(C_{11}+2C_{13})\Delta+\frac{3}{2}C_{12}\sum_{i=1}^n\sigma_i^4\right\}.
\end{aligned} \tag{8}$$

When $f=f_2$, we have $-2^{1/2}\psi(y)=y^{-1}(1-e^{-y^2/4})$, and so

$$\begin{aligned}
&|2^{1/2}\psi(y)+y/4|\leqslant Dy^2 &&\text{for } |y|\leqslant 1;\\
&|2^{1/2}\psi(y)|\leqslant 1-e^{-1/4} &&\text{for } |y|\geqslant 1.
\end{aligned}$$

Therefore, from (7), it follows that

$$(9)\quad \frac{1}{4}\left|\sum_{i=1}^{n} E\{Y_i^3 I[|Y_i| \leq 1]\}\right| \leq 2^{1/2}\left\{(C_{21} + 2C_{23})\Delta + \frac{3}{2}C_{22}\sum_{i=1}^{n}\sigma_i^4\right\}$$
$$+ (1 - e^{-1/4})\left[\sum_{i=1}^{n} E\{Y_i^2 I[|Y_i| > 1]\} + \sum_{i=1}^{n} E\{Y_i^4 I[|Y_i| \leq 1]\}\right],$$

since $D < (1 - e^{-1/4})$. From (8) and (9) we may deduce that

$$\sum_{i=1}^{n} E\{Y_i^2 I[|Y_i| > 1]\} + \sum_{i=1}^{n} E\{Y_i^4 I[|Y_i| \leq 1]\} + \left|\sum_{i=1}^{n} E\{Y_i^3 I[|Y_i| \leq 1]\}\right|$$
$$\leq \left[\{1 + 4(1 - e^{-1/4})\}2^{3/2}(1 - e^{-1/4})^{-1}(C_{11} + 2C_{13}) + 2^{5/2}(C_{21} + 2C_{23})\right]\Delta$$
$$+ \left[3 \cdot 2^{1/2}\{1 + 4(1 - e^{-1/4})\}(1 - e^{-1/4})^{-1}C_{12} + 3 \cdot 2^{3/2}C_{22}\right]\sum_{i=1}^{n}\sigma_i^4,$$

which produces the desired result. Explicit computation of the constants shows that the upper bound may be taken to be $392\Delta + 121\sum_{i=1}^{n}\sigma_i^4$.

ACKNOWLEDGEMENT. The second author wishes to express his gratitude for a Visiting Fellowship at the Australian National University, during which this work was undertaken.

REFERENCES

1. A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. Amer. Math. Soc. **49** (1941), 122–136.

2. A. Bikyalis, *Estimates of the remainder term in the central limit theorem*, Litovsk. Mat. Sb. **6** (1966), 323–346. (Russian)

3. C.-G. Esséen, *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law*, Acta Math. **77** (1945), 1–125.

4. W. Feller, *On the Berry-Esséen theorem*, Z. Wahrsch. Verw. Gebiete **10** (1968), 261–268.

5. P. Hall, *Characterizing the rate of convergence in the central limit theorem*, Ann. Probab. **8** (1980), 1037–1048.

6. ______, *Rates of convergence in the central limit theorem*, Pitman, London, 1982.

7. L. V. Osipov and V. V. Petrov, *On an estimate of the remainder term in the central limit theorem*, Theor. Probab. Appl. **12** (1967), 281–286.

8. L. V. Rozovskii, *A lower bound to the remainder term in the central limit theorem*, Mat. Zametki **24** (1978), 403–410. (Russian)

9. ______, *On the precision of an estimate of the remainder term in the central limit theorem*, Theor. Probab. Appl. **23** (1978), 712–730.

10. C. Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent variables*, Proc. Sixth Berkeley Sympos. Math. Statist. Prob., Vol. 2, Univ. of California Press, Berkeley, Calif., 1970.

DEPARTMENT OF STATISTICS, THE FACULTIES, AUSTRALIAN NATIONAL UNIVERSITY, G.P.O. BOX 4, CANBERRA ACT 2601, AUSTRALIA

GONVILLE AND CAIUS COLLEGE, CAMBRIDGE CB2 1TA, UNITED KINGDOM