JOURNAL OF APPROXIMATION THEORY 13, 327–340 (1975)

On the Rate of Approximation in the Central Limit Theorem

P. L. BUTZER, L. HAHN,* AND U. WESTPHAL

Lehrstuhl A für Mathematik, Rheinisch-Westfälische Technische Hochschule Aachen, 51 Aachen, West Germany

DEDICATED TO PROFESSOR G. G. LORENTZ ON THE OCCASION OF HIS SIXTY-FIFTH BIRTHDAY

1. INTRODUCTION AND HISTORY

Let $(\Omega, \mathfrak{A}, Pr)$ be an arbitrary probability space with distribution function (d.f.) F_X of the real random variable (r.v.) $X\colon \Omega \to \mathbf{R}$, defined by $F_X(x) = Pr\{\omega \in \Omega\colon X(\omega) \leqslant x\}$, for every $x \in \mathbf{R}$. Let X^* be a normally distributed random variable with mean 0 and variance 1, i.e., X^* is a random variable with d.f. $F_{X^*}(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp(-u^2/2)\, du$.

A sequence $(X_n)_{n=1}^{\infty}$ of real r.v.'s with variance satisfying

$$0 < \operatorname{Var}(X_n) < +\infty, \qquad \text{for each } n \in \mathbf{N},$$

is said to satisfy the central limit theorem [2, p. 223] in case $(n \to \infty)$

$$F_{T_n}(x) \to F_{X^*}(x) \qquad (\text{for each } x \in \mathbf{R}), \tag{1.1}$$

where

$$T_n := s_n^{-1} \sum_{k=1}^{n} [X_k - E(X_k)], \qquad s_n{}^2 = \operatorname{Var}\left(\sum_{k=1}^{n} [X_k - E(X_k)]\right).$$

($F_{T_n}(x)$ denoting the d.f. of the normalized sum T_n, and the expectation $E(X) := \int_{\mathbf{R}} x\, dF_X(x)$). This theorem is actually satisfied provided the sequence of r.v. is independent (which is case below) and identically distributed.

Of the many versions equivalent to (1.1) let us recall two further ones needed below. One is in terms of the pointwise convergence of the corresponding characteristic functions, namely

$$\int_{\mathbf{R}} e^{iux} dF_{T_n}(x) \to \frac{1}{(2\pi)^{1/2}} \int_{\mathbf{R}} e^{iux} e^{-x^2/2}\, dx = e^{-u^2/2}, \tag{1.2}$$

* The research of Lothar Hahn was partially supported by DFG grant Ne 171/1.

mainly used to prove that a sequence of r.v. satisfies the central limit theorem. A second equivalence to (1.1) is

$$\int_{\mathbf{R}} f(x)\, dF_{T_n}(x) \to \int_{\mathbf{R}} f(x)\, dF_{X^*}(x) \qquad (n \to \infty), \tag{1.3}$$

for each $f \in C_B(\mathbf{R})$, $C_B(\mathbf{R})$ denoting the class of bounded, uniformly continuous functions defined on $\mathbf{R}$.

A sufficient condition for the validity of (1.1) is

$$\| V_{T_n} f - V_{X^*} f \| \to 0 \qquad (n \to \infty), \tag{1.4}$$

for each $f \in C_B{}^r(\mathbf{R})$ and some $r \in \mathbf{N}$, where $V_X : C_B(\mathbf{R}) \to C_B(\mathbf{R})$ is the linear operator defined by

$$V_X f(y) := \int_{\mathbf{R}} f(x + y)\, dF_X(x), \tag{1.5}$$

and $C_B{}^r(\mathbf{R}) = \{f \in C_B(\mathbf{R}): f^{(j)} \in C_B(\mathbf{R}),\ 1 \leqslant j \leqslant r\}$, $\|f\| = \sup_{y \in \mathbf{R}} |f(y)|$. The operator V_X was mainly introduced by H. F. Trotter [19] in order to present an elementary proof that a sequence $(X_n)_{n=1}^{\infty}$ of r.v. satisfies the central limit theorem; it was taken over in a modified form in the monograph [18] by A. Renyi (who, however, did not cite Trotter).

The study of the rate of convergence of $F_{T_n}(x)$ to $F_{X^*}(x)$ as $n \to \infty$ in the uniform norm, apparently initiated by A. Liapounov [11] in 1901, and carried out by H. Cramér [5] in 1937 and A. C. Berry [4] and C. G. Esseen [6] in the fourties, has been receiving considerable attention in recent years by V. M. Zolotarev [21], I. A. Ibragimov [9], V. Paulauskas [14], J. Banys, N. Kalinauskaiti and P. Vaitkus [1], V. V. Petrov [16], L. V. Osipov [12], L. V. Osipov and V. V. Petrov [13] and W. Feller [7].

There seem to be essentially two different types of results established so far, namely "large O" and "small o" approximation estimates.

If the absolute third moment

$$\beta_3 := E(|X|^3) = \int_{\mathbf{R}} |x|^3\, dF_X(x),$$

is finite, then Berry and Esseen showed[1] that

$$\| F_{T_n} - F_{X^*} \| \leqslant C\beta_3/(n)^{1/2}, \tag{1.6}$$

if the sequence of real r.v. is independent and identically distributed. Concerning sharper estimates, the example of the lattice distributions shows that

[1] There are also investigations concerned with the best possible constant C. For example, $(2\pi)^{-1/2} < C \leqslant 0.82$. These investigations [20] do not interest us here.

the existence of higher absolute moments would not yield a better order of approximation than $n^{-1/2}$ (compare V.B. Gnedenko and A. N. Kolmogorov [8, p. 212]). However, if the d.f. $F_X(x)$ satisfies a condition of Cramér, namely

$$\limsup_{|u|\to\infty} \left| \int_{\mathbf{R}} e^{iux} dF_X(x) \right| < 1, \tag{1.7}$$

and if the pseudomoments (apparently first utilized by H. Bergström [3])

$$\mu(j) := \int_{\mathbf{R}} x^j d[F_X(x) - F_{X^*}(x)] = 0 \qquad (0 \leqslant j < r) \tag{1.8}$$

and if the rth absolute moment

$$\beta_r := E(|x|^r) < +\infty, \tag{1.9}$$

then Ibragimov [9] showed that for any *even* $r \geqslant 4$

$$\| F_{T_n} - F_{X^*} \| = O(n^{-(r-2)/2}) \qquad (n \to \infty). \tag{1.10}$$

The first question is whether it is possible to obtain an order $O(n^{-(r-2)/2})$ provided only conditions (1.8), (1.9) are satisfied, the crucial condition (1.7) being dropped. Here Paulauskas [14] (actually in the frame of more general investigations) only achieved the order $O(n^{-1/2})$ and not $O(n^{-(r-2)/2})$ as desired. However, if one would work in the equivalent convergence type (1.2), would it then be possible to show that

$$\int_{\mathbf{R}} f(x)\, d[F_{T_n}(x) - F_{X^*}(x)] = O(n^{-(r-2)/2})? \tag{1.11}$$

This will indeed be shown to be the case provided $f \in C_B^{r-1}(\mathbf{R})$ and the $(r-1)$th derivative $f^{(r-1)} \in \mathrm{Lip}\, 1$, conditions (1.8), (1.9) being satisfied (see Theorem 2). Here r may also be odd.

The next question is what happens when the sequence of r.v. is not identically distributed in which case very little seems to be known. If, instead of (1.8), one introduces the condition

$$\nu_i(j) := \int_{\mathbf{R}} x^j d[F_{X_i}(x) - F_{\sigma_i X^*}(x)] = 0 \qquad (0 \leqslant j < r;\ i \in \mathbf{N}), \tag{1.12}$$

with

$$\sigma_i^2 := \mathrm{Var}(X_i), \tag{1.13}$$

and, instead of (1.9), condition

$$\beta_{r,i} := E(|X_i|^r) < +\infty \qquad (i \in \mathbf{N}), \tag{1.14}$$

then it will be shown that a result of type (1.11) is possible; for the precise formulation see Theorem 1. (Note that if $(X_n)_{n=1}^\infty$ is identically distributed, then (1.8) implies (1.12) since $\sigma_i^2 = 1$, $i \in \mathbf{N}$, and $F_{X_i} = F_{X_j}$, $i, j \in \mathbf{N}$).

The second type of result is the small o-type theorem. If at first $(X_n)_{n=1}^\infty$ is identically distributed, then Esseen [6] showed that (see [8, p. 195] or [17, p. 180])

$$F_{T_n}(x) - F_{X^*}(x) = \frac{e^{-x^2/2}}{(2\pi)^{1/2}} \left[\frac{Q_1(x)}{n^{1/2}} + \frac{Q_2(x)}{n} + \cdots + \frac{Q_{r-2}(x)}{n^{(r-2)/2}} \right] + o\left(\frac{1}{n^{(r-2)/2}}\right),$$

uniformly in x provided conditions (1.7) and (1.9) are satisfied. Here the $Q_k(x)$ are rather intricate polynomials of degree $3k - 1$, determined indirectly (see V. V. Petrov [15]) with coefficients depending upon $\alpha_3/\sigma^3, \ldots, \alpha_{k+2}/\sigma^{k+2}$, where

$$\alpha_j = \int_{\mathbf{R}} x^j dF_X(x), \qquad \sigma^2 = \mathrm{Var}(X). \tag{1.15}$$

If condition (1.8) is satisfied not only for $0 \leqslant j < r$ but also for $j = r$, and (1.9) holds, then it can be shown that $Q_k(x) = 0$ for $1 \leqslant k \leqslant r - 2$, implying that

$$\| F_{T_n} - F_{X^*} \| = o\left(\frac{1}{n^{(r-2)/2}}\right) \qquad (n \to \infty).$$

Dropping the Cramér condition (1.7), the question arises as to what happens for the counterpart (1.11) with large-O replaced by small-o. This leads to Theorem 3.

If $(X_n)_{n=1}^\infty$ is not identically distributed, Lindeberg gave a sufficient condition for (1.1) or (1.2) to hold. It is given by

$$L_n(\delta) := \frac{1}{s_n^2} \sum_{i=1}^{n} \int_{|x-\alpha_{1,i}| \geqslant \delta s_n} [x - \alpha_{1,i}]^2 \, dF_{X_i}(x) \to 0 \tag{1.16}$$

for $n \to \infty$ and every $\delta > 0$, where

$$\alpha_{j,i} = \int_{\mathbf{R}} x^j dF_{X_i}(x) \qquad (i, j \in \mathbf{N}). \tag{1.17}$$

Conversely, if Feller's condition is satisfied, namely

$$\lim_{n \to \infty} \max_{1 \leqslant i \leqslant n} (\sigma_i/s_n) = 0, \tag{1.18}$$

then Lindeberg's condition follows from (1.1) or (1.2), or equivalently from

$$\| F_{T_n} - F_{X^*} \| = o(1) \qquad (n \to \infty). \tag{1.19}$$

This led us to the result that (see Theorem 4)

$$\int_{\mathbf{R}} f(x)\, d[F_{T_n}(x) - F_{X^*}(x)] = o\left(s_n^{-r} \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r)\right) \tag{1.20}$$

where

$$\gamma_r = E(| X^* |^r), \tag{1.21}$$

provided (1.12) holds for $0 \leqslant j \leqslant r$, $i \in \mathbf{N}$ ($r \geqslant 2$) and Feller's condition as well as a generalized Lindeberg-type condition is satisfied, namely

$$L_n^r(\delta) := \frac{1}{t_n^r} \sum_{i=1}^{n} \int_{|x-\alpha_{1,i}| \geqslant \delta s_n} | x - \alpha_{1,i} |^r \, dF_{X_i}(x) \to 0 \tag{1.22}$$

for $n \to \infty$ and every $\delta > 0$, where

$$t_n^r := \sum_{i=1}^{n} E(| X_i - \alpha_{1,i} |^r).$$

(Note that in case $r = 2$ condition $L_n^2(\delta)$ reduces to (1.16), and (1.12) is satisfied for $j = 0, 1, 2$, with $\sum_{i=1}^{n} (\beta_{2,i} + \gamma_2 \sigma_i^2) = O(s_n^2)$. So (1.20) coincides with (1.19). However, condition $L_n^r(\delta)$ does not imply $L_n^{r-1}(\delta)$).

The four theorems announced will be established in Section 3. While Section 2 is concerned with some preliminary results, Section 4 is devoted to an application of our Theorem 2. Section 5 closes with concluding remarks on the norm chosen as well as with an open problem.

2. Preliminaries

We need to recall the definition of the modulus of continuity and Lipschitz classes. The former is defined for $f \in C_B(\mathbf{R})$, $\delta \geqslant 0$ by

$$\omega(f; \delta) = \sup_{|h| \leqslant \delta} \| f(x+h) - f(x) \|, \tag{2.1}$$

having the properties that $\omega(f; \delta)$ is a monotonely decreasing function of δ with $\omega(f; \delta) \to 0$ for $\delta \to 0+$, and

$$\omega(f; \lambda\delta) \leqslant (1 + \lambda)\, \omega(f; \delta) \qquad (\text{each } \lambda > 0). \tag{2.2}$$

A function $f \in C_B(\mathbf{R})$ is said to satisfy a Lipschitz condition of order α,

$0 < \alpha \leqslant 1$, in symbols $f \in \mathrm{Lip}\,\alpha$, if $\omega(f; \delta) = O(\delta^\alpha)$. It is obvious that $f' \in C_B(\mathbf{R})$ implies $f \in \mathrm{Lip}\,1$.

Concerning the operator V_X defined by (1.5), it is a contraction, i.e., $\| V_X f \| \leqslant \| f \|$ for all $f \in C_B(\mathbf{R})$. If $X_1, X_2, \ldots, X_n$ are independent r.v., then

$$V_{\sum_{i=1}^n X_i} = V_{X_1} \circ V_{X_2} \circ \cdots \circ V_{X_n}. \tag{2.3}$$

The operators V_{X_i} and V_{X_j} commute, and $V_{X_i} = V_{X_j}$ if $F_{X_i}(x) = F_{X_j}(x)$, i.e., if X_i and X_j are identically distributed.

If the $X_1, \ldots, X_n$ are moreover independent, and $c_i > 0$ for $i = 1, 2, \ldots, n$, then

$$V_{\sum_{i=1}^n c_i^{-1} X_i} = V_{c_1^{-1} X_1} \circ V_{c_2^{-1} X_2} \circ \cdots \circ V_{c_n^{-1} X_n}. \tag{2.4}$$

In particular, if $c_i = c > 0$, $i = 1, 2, \ldots, n$,

$$V_{\sum_{i=1}^n c^{-1} X_i} = (V_{c^{-1} X})^n, \tag{2.5}$$

where X represents some r.v. X_i.

Furthermore, if A and B are two contradiction endomorphisms of $C_B(\mathbf{R})$ which commute with each other, then we also make use of the inequality

$$\| A^n f - B^n f \| \leqslant n \| Af - Bf \| \qquad (f \in C_B(\mathbf{R});\ n \in \mathbf{N}). \tag{2.6}$$

More generally, if $A_1, A_2, \ldots, A_n$, $B_1, B_2, \ldots, B_n$ are endomorphisms of $C_B(\mathbf{R})$ consisting of commutative, linear contraction operators, then for any $f \in C_B(\mathbf{R})$,

$$\| A_1 A_2 \circ \cdots \circ A_n f - B_1 B_2 \circ \cdots \circ B_n f \| \leqslant \sum_{i=1}^n \| A_i f - B_i f \|. \tag{2.7}$$

Finally, if X is any r.v. with $E(|X|^r) < +\infty$, then $E(|X|^j) < +\infty$ for any $1 \leqslant j \leqslant r$, and

$$E(|X|^j) \leqslant 1 + E(|X|^r). \tag{2.8}$$

3. Main Results

THEOREM 1. *Let $(X_n)_{n=1}^\infty$ be a sequence of real independent r.v. (not necessarily identically distributed) such that*

$$\nu_i(j) = \int_{\mathbf{R}} x^j d[F_{X_i}(x) - F_{\sigma_i X^*}(x)] = 0 \qquad (0 \leqslant j < r;\ i \in \mathbf{N}), \tag{3.1}$$

$$\beta_{r,i} = E(|X_i|^r) < +\infty \qquad (i \in \mathbf{N}) \tag{3.2}$$

for some fixed $r \geqslant 3$, $r \in \mathbf{N}$. Then for any $f \in C_B^{r-1}(\mathbf{R})$

$$\| V_{T_n} f - V_{X^*} f \| \leqslant \frac{2}{(r-1)!\, s_n^{r-1}} \omega(f^{(r-1)}; s_n^{-1}) \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r + 1).$$

If in addition to the above hypotheses $f^{(r-1)} \in \operatorname{Lip} \alpha$, $0 < \alpha \leqslant 1$, then

$$\| V_{T_n} - V_{X^*} f \| = O\left[\frac{s_n^{-r+1-\alpha}}{(r-1)!} \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r + 1) \right].$$

THEOREM 2. *If $(X_n)_{n=1}^{\infty}$ is identically distributed, condition* (3.1) *being replaced by* (1.8) *with $\beta_r = E(|X|^r) < \infty$, then for any $f \in C_B^{r-1}(\mathbf{R})$*

$$\| V_{T_n} f - V_{X^*} f \| = O[n^{-(r-3)/2} \omega(f^{(r-1)}; n^{-1/2})]. \tag{3.3}$$

If in addition $f^{(r-1)} \in \operatorname{Lip} \alpha$, $0 < \alpha \leqslant 1$, then

$$\| V_{T_n} f - V_{X^*} f \| = O[n^{-(r-3+\alpha)/2}].$$

In particular, under the above hypotheses,

$$\int_{\mathbf{R}} f(x)\, d[F_{T_n}(x) - F_{X^*}(x)] = O[n^{-(r-3+\alpha)/2}]. \tag{3.4}$$

Proof of Theorem 1. First note that in view of (2.4)

$$V_{T_n} = V_{s_n^{-1} X_1} \circ V_{s_n^{-1} X_2} \circ \cdots \circ V_{s_n^{-1} X_n} \tag{3.5}$$

$$V_{X^*} = V_{\sigma_1 s_n^{-1} X^*} \circ V_{\sigma_2 s_n^{-1} X^*} \circ \cdots \circ V_{\sigma_n s_n^{-1} X^*}, \tag{3.6}$$

the latter holding since $\sum_{i=1}^{n} (\sigma_i s_n^{-1} X^*)$ is a normally distributed r.v. with mean zero and variance one.

Since $f \in C_B^{r-1}(\mathbf{R})$, one has by the Taylor series expansion

$$f(x+y) = \sum_{j=0}^{r-1} \frac{x^j}{j!} f^{(j)}(y) + \frac{x^{r-1}}{(r-1)!} [f^{(r-1)}(\eta) - f^{(r-1)}(y)],$$

where η is some number between y and $x + y$. Applying the operator $V_{s_n^{-1} X_i}$ to f, this yields

$$\begin{aligned} V_{s_n^{-1} X_i} f(y) &= \int_{\mathbf{R}} f(x+y)\, dF_{s_n^{-1} X_i}(x) = \int_{\mathbf{R}} f(x s_n^{-1} + y)\, dF_{X_i}(x) \\ &= \sum_{j=0}^{r-1} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(i)}(y) + \frac{s_n^{-(r-1)}}{(r-1)!} \int_{\mathbf{R}} x^{r-1} [f^{(r-1)}(\eta) - f^{(r-1)}(y)]\, dF_{X_i}(x), \end{aligned}$$

where η is now between y and $y + s_n^{-1}x$, and $\alpha_{j,i}$ is defined by (1.17). Since

$$\left| \int_{\mathbf{R}} x^{r-1}[f^{(r-1)}(\eta) - f^{(r-1)}(y)]\, dF_{X_i}(x) \right|$$
$$\leqslant \int_{\mathbf{R}} | x |^{r-1} \omega(f^{(r-1)}; | \eta - y |)\, dF_{X_i}(x)$$
$$\leqslant \omega(f^{(r-1)}; s_n^{-1}) \int_{\mathbf{R}} | x |^{r-1} (1 + | x |)\, dF_{X_i}(x),$$

in view of (2.2) as $| \eta - y | \leqslant s_n^{-1} | x |$, it follows by (2.8) that

$$\left\| V_{s_n^{-1}X_i} f - \sum_{j=0}^{r-1} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(j)} \right\| \leqslant \frac{s_n^{-(r-1)}}{(r-1)!} (2\beta_{r,i} + 1)\, \omega(f^{(r-1)}; s_n^{-1}), \tag{3.7}$$

$\beta_{r,i}$ being defined by (1.14).

Analogously, since

$$\alpha_{j,i} = E(X_i^j) = E((\sigma_i X^*)^j) \qquad (0 \leqslant j < r;\ i \in \mathbf{N})$$

by (3.1), and, in view of (1.21) that

$$\gamma_r \sigma_i^r = \sigma_i^r E(| X^* |^r) = E(| \sigma_i X^* |^r),$$

one has, again by (2.2), that

$$\left\| V_{\sigma_i s_n^{-1} X^*} f - \sum_{j=0}^{r-1} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(j)} \right\| \leqslant \frac{s_n^{-(r-1)}}{(r-1)!} (2\gamma_r \sigma_i^r + 1)\, \omega(f^{(r-1)}; s_n^{-1}). \tag{3.8}$$

Combining the estimates (3.7) and (3.8) one has for each $i = 1, 2, \ldots, n$

$$\| V_{s_n^{-1}X_i} f - V_{\sigma_i s_n^{-1} X^*} f \| \leqslant \frac{s_n^{-(r-1)}}{(r-1)!} (2\beta_{r,i} + 2\gamma_r \sigma_i^r + 2)\, \omega(f^{(r-1)}; s_n^{-1}).$$

By (3.1), (3.2) and (2.7) this implies

$$\| V_{T_n} f - V_{X^*} f \| \leqslant \sum_{i=1}^{n} \| V_{s_n^{-1}X_i} f - V_{\sigma_i s_n^{-1} X^*} f \|$$
$$\leqslant \frac{2 s_n^{-(r-1)}}{(r-1)!}\, \omega(f^{(r-1)}; s_n^{-1}) \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r + 1),$$

completing the proof of Theorem 1.

Concerning the proof of Theorem 2, if $(X_n)_{n=1}^{\infty}$ is identically distributed, then $\sigma_i^2 = 1$, $\beta_{r,i} = \beta_r$, $i = 1, 2, \ldots, n$, $s_n^2 = \sum_{i=1}^n \sigma_i^2 = n$ or $s_n^{r-1} = n^{(r-1)/2}$, and $\sum_{i=1}^n (\beta_{r,i} + \gamma_r \sigma_i^r + 1) = O(n)$.

THEOREM 3. *Let $(X_n)_{n=1}^{\infty}$ be a sequence of real, independent and identically distributed r.v. such that*

$$\mu(j) = \int_{\mathbf{R}} x^j d[F_X(x) - F_{X^*}(x)] = 0 \qquad (0 \leqslant j \leqslant r)$$

and $\beta_r = E(|X|^r) < +\infty$ *for some fixed* $r \geqslant 2$, $r \in N$. *Then for any* $f \in C_B^r(\mathbf{R})$

$$\| V_{T_n} f - V_{X^*} f \| = o(n^{-(r-2)/2}) \qquad (n \to \infty). \tag{3.9}$$

In particular, for any $f \in C_B^r(\mathbf{R})$

$$\int_{\mathbf{R}} f(x)\, d[F_{T_n}(x) - F_{X^*}(x)] = o(n^{-(r-2)/2}) \qquad (n \to \infty).$$

Proof. Since $f \in C_B^r(\mathbf{R})$, we may apply the operator $V_{s_n^{-1}X}$ to the Taylor series expansion of f of order r (instead of order $r - 1$ as in proof of Thm. 1) to yield

$$V_{s_n^{-1}X} f(y) = \sum_{j=0}^{r} \frac{s_n^{-j}}{j!} \alpha_j f^{(j)}(y) + \frac{s_n^{-r}}{r!} \int_{\mathbf{R}} x^r [f^{(r)}(\eta) - f^{(r)}(y)]\, dF_X(x),$$

where η is some number between y and $y + s_n^{-1}x$. Since $f \in C_B^r(\mathbf{R})$, to each $\epsilon > 0$ there exists $\delta > 0$ such that $|\eta - y| < \delta$ implies $|f^{(r)}(\eta) - f^{(r)}(y)| < \epsilon$. For this δ we split up the above integral into

$$\int_{|x|<\delta s_n} x^r [f^{(r)}(\eta) - f^{(r)}(y)]\, dF_X(x) + \int_{|x|\geqslant \delta s_n} x^r [f^{(r)}(\eta) - f^{(r)}(y)]\, dF_X(x)$$
$$= I_1 + I_2, \quad \text{say.}$$

For I_1 one has $|\eta - y| \leqslant s_n^{-1} |x| < \delta$, implying

$$|I_1| \leqslant \epsilon \int_{|x|<\delta s_n} |x|^r\, dF_X(x) \leqslant \epsilon \beta_r .$$

For I_2 one has $|f^{(r)}(\eta) - f^{(r)}(y)| \leqslant 2 \| f^{(r)} \|$, giving

$$|I_2| \leqslant 2 \| f^{(r)} \| \int_{|x|\geqslant \delta s_n} |x|^r\, dF_X(x).$$

Since β_r is finite, $\int_{|x|\geqslant k} |x|^r\, dF_X(x) \to 0$ for $k \to \infty$. Therefore for n sufficiently large $|I_2| < \epsilon$ since $s_n = (n)^{1/2}$.

Thus one has for n sufficiently large

$$\left\| V_{s_n^{-1}X} f - \sum_{j=0}^{r} \frac{s_n^{-j}}{j!} \alpha_j f^{(j)} \right\| = o(s_n^{-r}) \qquad (n \to \infty).$$

Since $E(X^{*j}) = \alpha_j$, $0 \leqslant j \leqslant r$ and $E(|X^*|^r) = \gamma_r$ is finite, one has a corresponding estimate for the r.v. X^*, and so by the triangle inequality

$$\| V_{s_n^{-1}X}f - V_{s_n^{-1}X^*}f \| = o(s_n^{-r}) \qquad (n \to \infty).$$

In view of (2.6), (3.5) and (3.6) this gives

$$\| V_{T_n}f - V_{X^*}f \| \leqslant n \| V_{s_n^{-1}X}f - V_{s_n^{-1}X^*}f \| = o(n^{-(r-2)/2}),$$

and so the desired estimate (3.9).

THEOREM 4. *Let $(X_n)_{n=1}^{\infty}$ be a sequence of real, independent r.v.* (*not necessarily identically distributed*) *such that* (3.1) *and* (3.2) *hold for some* $r \geqslant 2$. *Assume further that the generalized Lindeberg condition* (1.22) *of order* r *as well as Feller's condition* (1.18) *be satisfied. Then for any* $f \in C_B^r(\mathbf{R})$

$$\| V_{T_n}f - V_{X^*}f \| = o\left[\frac{s_n^{-r}}{r!} \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r)\right] \qquad (n \to \infty).$$

Proof. On account of (3.5), (3.6) and (2.7) it suffices to show that

$$\sum_{i=1}^{n} \| V_{s_n^{-1}X_i}f - V_{s_n^{-1}\sigma_i X^*}f \| = o\left[s_n^{-r} \sum_{i=1}^{n} (\beta_{r,i} + \gamma_r \sigma_i^r)\right] \qquad (n \to \infty)$$

for each $f \in C_B^r(\mathbf{R})$.

Analogously as in the beginning of the proof of Theorem 3 we have

$$V_{s_n^{-1}X_i}f(y) - \sum_{j=0}^{r} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(j)}(y)$$

$$= \frac{s_n^{-r}}{r!} \left(\int_{|x|<\delta s_n} + \int_{|x|\geqslant \delta s_n}\right)[f^{(r)}(\eta) - f^{(r)}(y)]\, x^r dF_{X_i}(x) = I_1 + I_2, \quad \text{say,}$$

where η is some number between y and $y + s_n^{-1}x$, and δ is chosen as in the proof of Theorem 3.

Since $|\eta - y| \leqslant s_n^{-1} |x| < \delta$, $|I_1| \leqslant \epsilon\beta_{r,i}/(r! s_n^r)$, and one has

$$\left\| V_{s_n^{-1}X_i}f - \sum_{j=0}^{r} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(j)} \right\| \leqslant \epsilon\beta_{r,i} \frac{s_n^{-r}}{r!} + 2 \| f^{(r)} \| \frac{s_n^{-r}}{r!} \int_{|x|\geqslant \delta s_n} |x|^r dF_{X_i}(x).$$

The counterinequality for the r.v. X^* reads on account of (3.1)

$$\left\| V_{s_n^{-1}\sigma_i X^*}f - \sum_{j=0}^{r} \frac{s_n^{-j}}{j!} \alpha_{j,i} f^{(j)} \right\|$$

$$\leqslant \epsilon \frac{\gamma_r}{r!} \sigma_i^r s_n^{-r} + 2 \| f^{(r)} \| \frac{s_n^{r}}{r!} \int_{|x|\geqslant \delta s_n} |x|^r dF_{\sigma_i X^*}(x).$$

Combining both inequalities gives

$$\| V_{s_n^{-1}X_i}f - V_{s_n^{-1}\sigma_i X^*}f \| \leqslant \epsilon \frac{s_n^{-r}}{r!} (\beta_{r,i} + \gamma_r \sigma_i^r) + 2 \| f^{(r)} \| \frac{s_n^{-r}}{r!} (\xi_{n,i} + \zeta_{n,i}),$$

where

$$\xi_{n,i} = \int_{|x| \geqslant \delta s_n} | x |^r \, dF_{X_i}(x),$$

$$\zeta_{n,i} = \int_{|x| \geqslant \delta s_n} | x |^r \, F_{\sigma_i X^*}(x).$$

Multiplying by $r! s_n^r$, summing over the i's and dividing by $\sum_{i=1}^n (\beta_{r,i} + \gamma_r \sigma_i^r)$, one has

$$\left(r!\, s_n^r \Big/ \sum_{i=1}^n (\beta_{r,i} + \gamma_r \sigma_i^r)\right) \sum_{i=1}^n \| V_{s_n^{-1}X_i}f - V_{s_n^{-1}\sigma_i X^*}f \|$$

$$\leqslant \epsilon + \| f^{(r)} \| \left\{ \left(\sum_{i=1}^n \xi_{n,i} \Big/ \sum_{i=1}^n \beta_{r,i} \right) + \left(\sum_{i=1}^n \zeta_{n,i} \Big/ \gamma_r \sum_{i=1}^n \sigma_i^r \right) \right\}$$

since $\beta_{r,i} > 0$, $\sigma_i > 0$, $\gamma_r > 0$.

Since in view of Lindeberg's condition (1.22) the first term in the curly brackets tends to zero for $n \to \infty$ (noting that $\sum_{i=1}^n \beta_{r,i} = t_n^r$), one need only show that $u_n \sum_{i=1}^n \zeta_{n,i} \to 0$ for $n \to \infty$, where

$$u_n^{-1} = \sum_{i=1}^n \sigma_i^r.$$

Indeed, let j be the index such that $\sigma_j = \max_{1 \leqslant i \leqslant n} \sigma_i$. Then

$$u_n \sum_{i=1}^n \int_{|x| \geqslant \delta s_n} | x |^r \, dF_{\sigma_i X^*}(x) = u_n \sum_{i=1}^n \sigma_i^r \int_{|x| \geqslant (\delta s_n)/\sigma_j} | x |^r \, dF_{X^*}(x)$$

$$= \int_{|x| \geqslant (\delta s_n)/\sigma_j} | x |^r \, dF_{X^*}(x).$$

Applying Feller's condition (i.e., $s_n/\sigma_j \to \infty$ for $n \to \infty$) and noting that $E(| X^* |^r) < +\infty$, the theorem follows.

4. An Application

Let us consider an application of Theorem 2 to a particular sequence of identically distributed r.v. $(X_n)_{n=1}^\infty$, kindly suggested to us by Professor

Kaerkes, Aachen. These may be introduced via the d.f's F_{X_n} of the X_n which are equal to another for all $n \in \mathbf{N}$, i.e.,

$$F_X(x) = \begin{cases} 0 & \text{for} \quad x < -\sqrt{3}, \\ \frac{1}{6} & \text{for} \quad -\sqrt{3} \leqslant x < 0, \\ \frac{5}{6} & \text{for} \quad 0 \leqslant x < \sqrt{3}, \\ 1 & \text{for} \quad \sqrt{3} \leqslant x. \end{cases}$$

The hypotheses of Theorem 2 are satisfied with $r = 6$. Indeed, $E(X^j) = 0$ for j odd, and $E(X^j) = 1, 1, 3, 9$ for $j = 0, 2, 4, 6$, respectively, implying $\mu(j) = 0$ for $0 \leqslant j < 6$. Hence (3.4) takes on the form

$$\int_{\mathbf{R}} f(x)\, d[F_{T_n}(x) - F_{X^*}(x)] = O(n^{-(3+\alpha)/2})$$

for each $f^{(5)} \in \operatorname{Lip} \alpha$, $0 < \alpha \leqslant 1$.

In the particular instance that $f(x) = e^{iux}$, fixed $u \in \mathbf{R}$, one has the estimate

$$\frac{[2 + \cos(3u^2/n)^{1/2}]^n}{3^n} - e^{-u^2/2} = O_u(n^{-2}),$$

the large O depending on u, indeed

$$\int_{\mathbf{R}} e^{iux} dF_{T_n}(x) = \left[\int_{\mathbf{R}} e^{iux/(n)^{1/2}} dF_X(x)\right]^n$$

and

$$\int_{\mathbf{R}} e^{iux} dF_X(x) = \tfrac{1}{3}[2 + \cos u\sqrt{3}]. \tag{4.1}$$

It is important to note that (4.1) reveals that the condition (1.7) of Cramér is not satisfied for this example. More generally (see [10 p. 26]), Cramér's condition is not satisfied if the identically distributed r.v. X_n have lattice distribution.

5. Concluding Remarks

The original question of this paper was to examine conditions upon a sequence of ident. distributed r.v. such that the approximation (1.10) holds. However, in attempting to reach this goal we found conditions yielding (1.11). As noted, we needed one condition less than what Ibragimov [9] needed to establish (1.10) in the case of even r, namely condition (1.7).

Moreover, our proof proceeded in the "original" function space and so was rather elementary, essentially only making use of the Taylor series expansion. Ibragimov's proof, as well as all other proofs yielding estimates on the rate of convergence—as far as the authors are aware—always proceed via the "transformed" function space, in other words are carried out by means of characteristic functions, thus Fourier transforms. Moreover, these proofs are rather long and use intricate estimates. See also Feller [22, p. 487].

The question still remains whether it is possible to show that (1.11) implies (1.10) under some additional condition such as (1.7). More precisely, does

$$\sup_{y\in\mathbf{R}} \left| \int_{\mathbf{R}} f(x+y)\, d[F_{T_n}(x) - F_{X^*}(x)] \right| = O(n^{-(r-2)/2}), \tag{5.1}$$

for $f \in C_B{}^r(\mathbf{R})$ imply that

$$\sup_{x\in\mathbf{R}} \left| F_{T_n}(x) - F_{X^*}(x) \right| = O(n^{-(r-2)/2}), \tag{5.2}$$

under (1.8) together with some further condition?

Note that if (5.2) holds, then in the case of even $r \geqslant 4$ an inverse result of Ibragimov [9] implies that (without use of (1.7)) $\beta_r < +\infty$ and $\mu(j) = 0$ for $0 \leqslant j < r$ in (1.8), which in turn implies (5.1) by our direct Theorem 2. Recall that in our application of Section 4 the crucial condition (1.7) is not satisfied.

The matter described above could also be discussed in the case of non-identically distributed r.v.'s. But this should be much more difficult as there seem to be no comparable results known.

References

1. J. Banys, N. Kalinauskaiti and P. Vaitkus, The rate of convergence to stable distributions in a local theorem (Russian), *Litovsk. Mat. Sb.* **11** (1971), 511–516.
2. H. Bauer, "Wahrscheinlichkeitstheorie und Maßtheorie," De Gruyter, Berlin, 1968.
3. H. Bergström, On distribution functions with a limiting stable distribution function, *Ark. Mat.* **2** (1953), 463–474.
4. A. C. Berry, The accuracy of the Gaussian approximation to the sum of independent variables, *Trans. Amer. Math. Soc.* **49** (1941), 122–136.
5. H. Cramér, "Random Variables and Probability Distributions," (Third edition), University Press, Cambridge, 1970.
6. C. G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law, *Acta Math.* **77** (1945), 1–125.
7. W. Feller, On the Berry-Esseen theorem, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **10** (1968), 261–268.
8. B. V. Gnedenko and A. N. Kolmogorov, "Limit Distributions for Sums of Independent Random Variables," Addison-Wesley, Reading, MA, 1954.

9. I. A. Ibragimov, The rate of convergence to the normal distribution. (English Translation), *Dokl. Akad. Nauk SSSR* **161** (1965), 580–582.
10. I. A. Ibragimov and Yu. V. Linnik, "Independent and Stationary Sequences of Random Variables," Wolters- Noordhoff, Groningen 1971.
11. A. M. Liapounov, Nouvelle forme du théorème sur la limite de théorie des probabilités, *Mem. Acad. Sci. St. Pétersbourg* **12** (1901), 1–24.
12. L. V. Osipov, A refinement of Lindeberg's theorem (Russian), *Teor. Verojatnost. i. Primenen.* **11** (1966), 339–342.
13. L. V. Osipov and V. V. Petrov, On the estimation of the remainder term in the central limit theorem (Russian), *Teor. Verojatnost. i. Primenen.* **12** (1967), 468–475.
14. V. I. Paulauskas, One theorem on the rate of convergence in the central limit theorem (Russian), *Litovsk. Mat. Sb.* **11** (1971), 173–179.
15. V. V. Petrov, On some polynomials occurring in probability theory (Russian), *Vestnik Leningrad. Univ.* **19** (1962), 150–153.
16. V. V. Petrov, An estimate of the deviation of the distribution of the sum of independent random variables from the normal law (Russian), *Dokl. Akad. Nauk SSSR* **160** (1965), 1013–1015.
17. Yu. V. Prohorov and Yu. A. Rozanov, "Probability Theory," Springer, Berlin, Heidelberg, New York, 1969.
18. A. Renyi, "Foundations of Probability," Holden-Day, San Francisco CA, 1970.
19. H. F. Trotter, An elementary proof of the central limit theorem, *Arch. Math. (Basel)* **10** (1959), 226–234.
20. V. M. Zolotarev, Some inequalities of probability theory and their application to the refinement of A. M. Liapunov's theorem (Russian), *Dokl. Akad. Nauk SSSR* **177** (1967), 501–504.
21. V. M. Zolotarev, On the accuracy of approximation in the central limit theorem, *Dokl. Akad. Nauk SSSR* **203** (1972), 326–328.
22. W. Feller, "An Introduction to Probability Theory and its Applications," Vol. II, John Wiley and Sons, New York, London, Sydney, 1966.

Printed in Belgium