Probab. Th. Rel. Fields 72, 289–303 (1986)

Probability Theory and Related Fields

Asymptotic Expansions Based on Smooth Functions in the Central Limit Theorem

A.D. Barbour

Institut für Angewandte Mathematik, Universität Zürich, CH-8001 Zürich, Switzerland

Summary. Stein's method is used to derive asymptotic expansions for expectations of smooth functions of sums of independent random variables, together with Lyapounov estimates of the error in the approximation.

1. Introduction

When considering the error in the normal approximation to the partial sums of stationary sequences, Stein (1970) introduced a new technique, by means of which differences of the form $\mathbb{E}h(W)-\mathbb{E}h(\mathcal{N})$, for smooth functions h, could be directly estimated: here, $\mathcal{N}$ denotes a standard normal random variable, and W denotes the random variable whose distribution is being approximated. The technique has a structure which lends itself in principle to iterative application, by means of which asymptotic expansions could be obtained, but the possibility seems not to have been exploited, owing to the apparent complexity of the procedure. In this paper, a simplification is found, which enables asymptotic expansions for the expectations of smooth functions of sums of independent random variables to be derived, together with Lyapounov bounds on the approximation error, at the cost of some analytic argument, concerning the smoothness and rate of growth of solutions of Stein's ordinary differential equation. The principal tool is a lemma which, for a random variable X with exponentially decaying tails, makes explicit the way in which the difference between $\mathbb{E}h(X)$ and $\mathbb{E}h(\mathcal{N})$ depends on the cumulants of X of order greater than two.

Asymptotic expansions for $\mathbb{E}h(W)$, where h is smooth and W is a partial sum of independent random variables, were considered by Hsu (1945), von Bahr (1965) and Bhattacharya (1970), and have more recently been discussed in Hipp (1977) and in Götze and Hipp (1978). In the latter paper, asymptotic expansions are obtained by Fourier methods, under conditions which, although similar to those used here, are not equivalent, and their error estimates are more difficult to express: however, their results are proved for sums of independent random vectors in $\mathbb{R}^d$. The arguments used in this paper are understandably simpler.

An advantage of considering expansions only for expectations of smooth functions h of W is that, in contrast to expansions for distribution functions, there is no need to impose smoothness conditions on the distributions of the summands: the natural moment conditions are all that is required. However, the problem cannot be entirely avoided. Each extra term in the asymptotic expansion requires an extra derivative of the function h to exist, and the estimated error of the expansion depends on a Lipschitz measure of the smoothness of the highest required derivative of h.

2. Main Results

The essence of Stein's (1970) method is that, if h is any function for which $\mathbb{E}|h(\mathcal{N})| < \infty$, then, for any random variable X,

$$\mathbb{E}h(\mathcal{N}) - \mathbb{E}h(X) = \mathbb{E}\{Xg(X) - Dg(X)\},$$

where $g = \theta h$ is defined by

$$(\theta h)(x) = \int_{-\infty}^{x} e^{\frac{1}{2}(x^2-t^2)}\{h(t) - \mathbb{E}h(\mathcal{N})\}\,dt$$
$$= -\int_{x}^{\infty} e^{\frac{1}{2}(x^2-t^2)}\{h(t) - \mathbb{E}h(\mathcal{N})\}\,dt:$$

here, and subsequently, $\mathcal{N}$ denotes a standard normal random variable and $D_l f$ the l^{th} derivative of f. Note that g satisfies the differential equation

$$Dg(w) - wg(w) = h(w) - \mathbb{E}h(\mathcal{N}).$$

Thus the closeness of the distributions of X and $\mathcal{N}$ can be estimated, if an estimate of $\mathbb{E}\{Xg(X) - Dg(X)\}$ is available, for an appropriate class of functions g. The following lemma provides a starting point for such estimates.

Lemma 1. *Let g be an $l-1$ times differentiable function and X a random variable, and suppose that*

$$\mathbb{E}\{|X|^l u_{l-1}(g; X)\} < \infty,$$

where

$$u_k(g; x) = \sup_{|t| \leq x} |D_k g(t) - D_k g(0)|.$$

Let κ_r denote the r^{th} cumulant of X. Then

$$\mathbb{E}\{Xg(X)\} = \sum_{s=0}^{l-1} \frac{\kappa_{s+1}}{s!} \mathbb{E}\{D_s g(X)\} + \eta_{l-1}(g; X),$$

where

$$|\eta_k(g;X)| \leqq \sum_{s=0}^{k} \left\{ \frac{|\kappa_{s+1}|}{s!} \frac{\mathbb{E}|X^{k-s} u_k(g;X)|}{(k-s)!} \right\} + \frac{\mathbb{E}|X^{k+1} u_k(g;X)|}{k!}$$
$$\leqq d_k \mathbb{E}|X^{k+1} u_k(g;X)|;$$
$$d_k = \sum_{s=0}^{k} \frac{\alpha_{s+1}}{s!(k-s)!};$$

and the universal constants $(\alpha_s)_{s\geqq 1}$ *are defined by*

$$\alpha_1 = 2;\ \alpha_s = \sup_{Y\in L_s} \left\{ \frac{|\kappa_s|}{\mathbb{E}|Y-\mathbb{E}Y|^s} \right\}.$$

Remarks. 1. In particular, $d_1=3$, $d_2=5/2$, $d_3=5/3$.

2. If $\mathbb{E}X=0$ and $\mathbb{E}X^2=1$, the conclusion may be rewritten in the form

$$\mathbb{E}\{Xg(X)-Dg(X)\} = \sum_{s=2}^{l-1} \frac{\kappa_{s+1}}{s!} \mathbb{E}\{D_s g(X)\} + \eta_{l-1}(g;X). \tag{1}$$

In this case, since $\kappa_1=0$, we may take $\alpha_1=1$ and hence $d_1=d_2=2$, $d_3=3/2$.

Proof. From Taylor's theorem,

$$\left| D_s g(x) - \sum_{r=0}^{l-1-s} \frac{x^r}{r!} D_{r+s} g(0) \right| \leqq \frac{|x|^{l-1-s}}{(l-1-s)!} u_{l-1}(g;x).$$

Hence

$$\left| \mathbb{E}(Xg(X)) - \sum_{r=0}^{l-1} \frac{\mathbb{E}(X^{r+1})}{r!} D_r g(0) \right| \leqq \frac{\mathbb{E}\{|X|^l u_{l-1}(g;X)\}}{(l-1)!}$$

and

$$\left| \mathbb{E}D_s g(X) - \sum_{r=0}^{l-1-s} \frac{\mathbb{E}(X^r)}{r!} D_{r+s} g(0) \right| \leqq \frac{\mathbb{E}\{|X|^{l-1-s} u_{l-1}(g;X)\}}{(l-1-s)!}, \qquad 0 \leqq s \leqq l-1.$$

But

$$\sum_{r=0}^{l-1} \frac{\mathbb{E}(X^{r+1})}{r!} D_r g(0) - \sum_{s=0}^{l-1} \frac{\kappa_{s+1}}{s!} \left\{ \sum_{r=0}^{l-1-s} \frac{\mathbb{E}(X^r)}{r!} D_{r+s} g(0) \right\}$$
$$= \sum_{m=0}^{l-1} D_m g(0) \left\{ \frac{\mathbb{E}(X^{m+1})}{m!} - \sum_{s=0}^{m} \frac{\kappa_{s+1}}{s!} \frac{\mathbb{E}(X^{m-s})}{(m-s)!} \right\} = 0.$$

The lemma now follows easily. □

In view of Remark 2, it can be seen that the closeness to zero of the cumulants of X of orders 3 to l would determine the closeness of X to the standard normal, provided that η_{l-1} was also constrained to be small. However, when $\mathbb{E}X=0$ and $\mathbb{E}X^2=1$, η_{l-1} has two terms which do not involve the higher order cumulants of X, namely

$$\frac{1}{(l-2)!}\mathbb{E}|X^{l-2}u_{l-1}(g;X)| \quad \text{and} \quad \frac{1}{(l-1)!}\mathbb{E}|X^{l}u_{l-1}(g;X)|. \tag{2}$$

It is tempting to suppose that these terms also could be estimated by cumulants of orders between three and $(l+2)$, say; but no such estimate can be obtained, since there is a distribution $\mathcal{Q}_l$, with the same moments as $\mathcal{N}$ of orders up to $(l+2)$, which has atoms at exactly $(l+3)$ points, and such an estimate would make the right hand side of (1) identically zero for $\mathcal{Q}_l$, which cannot be the case, since $\mathcal{Q}_l$ is not normally distributed. However, Remark 2 does show how closeness to the normal, expressed in terms of the expectations of smooth functions, can be established using the closeness of the higher order cumulants of X to zero. First, l has to be taken so large that the factorials in the denominators of the terms in (2) make their contributions to η_{l-1} small, and then the cumulants of orders 3 to l must also be small. The following corollary shows how this may be exploited.

Corollary 1. *Suppose X has a moment generating function with a non-zero radius of convergence, and let $g \in C_\infty$ satisfy*

$$\sup_x |D_k g(X)| \leqq C r^k, \qquad k \geqq 1, \tag{3}$$

for some $C>0$ and $r<R$, where R is the radius of convergence of the cumulant generating function of X. Then we have the identity

$$\mathbb{E}\{Xg(X)\} = \sum_{s \geqq 0} \frac{\kappa_{s+1}}{s!} \mathbb{E}D_s g(X). \tag{4}$$

In particular, if $\mathbb{E}X=0$ and $\mathbb{E}X^2=1$,

$$\mathbb{E}\{Xg(X) - Dg(X)\} = \sum_{s \geqq 2} \frac{\kappa_{s+1}}{s!} \mathbb{E}D_s g(X). \tag{5}$$

Remark. If (3) is satisfied with $r>R$, (4) holds with λX in place of X, for all $\lambda < R/r$.

Proof. The radius of convergence of m.g.f. (X) exceeds that of c.g.f. (X). Hence, from the first estimate of the lemma,

$$\begin{aligned} |\eta_k(g;X)| &\leqq C'\left\{\sum_{s=0}^{k} \frac{(s+1)}{(r')^{s+1}} \frac{(k-s+1)}{(r')^{k-s+1}} + \frac{(k+1)(k+2)}{(r')^{k+2}}\right\} r^{k+1} \\ &\leqq C' k^3 \left(\frac{r}{r'}\right)^{k+1}, \end{aligned}$$

for any $r<r'<R$ and for suitable constants C'. Thus $\lim_{k\to\infty} |\eta_k(g;X)| = 0$. □

For sums of independent random variables and smooth functions g, the error term in Lemma 1 can be expressed more informatively in Lyapounov form. For any function g, let

$$\mathscr{L}(g: p, \alpha) = \sup_{x \neq y} \{|g(x)-g(y)|/[|x-y|^\alpha(1+|x|^p+|y|^p)]\},$$

and define the function g_z by $g_z(x)=g(z+x)$. Let $(X_i)_{i=1}^N$ be independent random variables with zero mean, such that $\sum_{i=1}^N \mathbb{E}X_i^2=1$: set $W=\sum_{i=1}^N X_i$ and $W_i = W-X_i$.

Lemma 2. *Let g be $k-1$ times differentiable, for some $k \geqq 2$, and suppose that*

$$G = \mathscr{L}(D_{k-1}g; p, \alpha) < \infty,$$

for some $p \geqq 0$, $0 \leqq \alpha \leqq 1$. Then, if $\mathbb{E}|X_i|^{k+p+\alpha} < \infty$, $1 \leqq i \leqq N$,

$$\mathbb{E}\{Wg(W)-Dg(W)\} = \sum_{s=1}^{k-2} \frac{\kappa_{s+2}(W)}{(s+1)!} \mathbb{E}\{D_{s+1}g(W)\} + e_{k-1}, \tag{6}$$

where

$$|e_{k-1}| \leqq d_{k-1} \sum_{i=1}^N \mathbb{E}|X_i^k u_{k-1}(g_{W_i}; X_i)|$$
$$\leqq (2^p+2)\,2^p G d_{k-1} \sum_{i=1}^N \mathbb{E}\{|X_i|^{k+\alpha}(1+|X_i|^p+\mathbb{E}|W|^p)\}.$$

In particular, for $p \leqq 2$,

$$|e_{k-1}| \leqq 48 G d_{k-1} \sum_{i=1}^N \mathbb{E}\{|X_i|^{k+\alpha}(1+|X_i|^p)\}. \tag{7}$$

Proof. It is immediate that

$$\mathbb{E}\{Wg(W)-Dg(W)\} = \sum_{i=1}^N \mathbb{E}\{X_i g_{W_i}(X_i) - Dg_{W_i}(X_i)\,\mathbb{E}X_i^2\}.$$

Now

$$u_{k-1}(g_{W_i}; X_i) \leqq G|X_i|^\alpha\{1+|W_i|^p+(|W_i|+|X_i|)^p\},$$

and so, since W_i and X_i are independent,

$$\mathbb{E}\{|X_i|^k u_{k-1}(g_{W_i}; X_i)\} \leqq G\mathbb{E}\{|X_i|^{k+\alpha}(1+(2^p+1)\,\mathbb{E}|W_i|^p+2^p|X_i|^p)\} < \infty.$$

Hence, from Lemma 1,

$$\mathbb{E}\{Wg(W)-Dg(W)\} = \sum_{s=2}^{k-1} \frac{\kappa_{s+1}(W)}{s!} \mathbb{E}\{D_s g(W)\} + \sum_{i=1}^N \mathbb{E}\eta_{k-1}(g_{W_i}; X_i);$$

the lemma now follows by noting that

$$\mathbb{E}|W_i|^p \leqq 2^p\{\mathbb{E}|W|^p + \mathbb{E}|X_i|^p\}. \quad \square$$

Now, if $X_i = Y_i/\sqrt{N}$, where the $(Y_i)_{i=1}^N$ are i.i.d. with mean zero and variance 1, the cumulants of W satisfy

$$\kappa_{s+2}(W) = N^{-s/2}\kappa_{s+2}(Y_1), \quad s \geqq 0,$$

so that (6) represents an asymptotic expansion for the quantity

$$\mathbb{E}\{Wg(W) - Dg(W)\}$$

in descending powers of $N^{\frac{1}{2}}$. Thus, by Stein's argument, Eq. (6), with $g = \theta h$, is an asymptotic expansion for the discrepancy $\mathbb{E}h(\mathcal{N}) - \mathbb{E}h(W)$. However, the expansion involves expectations taken with respect to the distribution of W, rather than the standard Normal distribution, and transforming (6) into a true asymptotic expansion for $\mathbb{E}h(\mathcal{N}) - \mathbb{E}h(W)$ involves applying (6) to its own right hand side - converting each W-expectation to an $\mathcal{N}$-expectation with further errors - and then applying (6) to the new set of W-expectations, and so on until the required order is reached. The result is as follows.

Theorem. *In the setting of Lemma* 2, *suppose that, for some* $2 \leqq k \leqq K$ *and for some* $0 \leqq \alpha \leqq 1$,

(i) *h is $k-2$ times differentiable, and*

$$H_1 = \mathcal{L}(D_{k-2}h; p, \alpha) < \infty$$

or

(ii) *h is $k-1$ times differentiable, and*

$$H_2 = \mathcal{L}(D_{k-1}h; p+1, \alpha) < \infty.$$

Then, if $\mathbb{E}|X_i|^{k+p+\alpha} < \infty$, $1 \leqq i \leqq N$, *it follows that*

$$\mathbb{E}h(W) = Eh(\mathcal{N}) + \sum_{(k-2)} (-1)^r \prod_{j=1}^r \left\{\frac{\kappa_{s_j+2}(W)}{(s_j+1)!}\right\} \mathbb{E}\left\{\prod_{j=1}^r (D_{s_j+1}\theta) h(\mathcal{N})\right\} + \eta, \quad (8)$$

where $\sum_{(k-2)}$ *denotes the sum over* $\left\{r \geqq 1;\ s_j \geqq 1,\ 1 \leqq j \leqq r\colon \sum_{j=1}^r s_j \leqq k-2\right\}$,

$$|\eta| \leqq C_j H_j \sum_{i=1}^N \mathbb{E}\{|X_i|^{k+\alpha} + |X_i|^{k+p+\alpha}\}$$

($j=1$ *for case* (i), $j=2$ *for case* (ii)), and $C_j = C_j(p, k, \alpha)$ *are universal constants.*

Remark. The expansion of Theorem 1 is not obviously the same as the expectation of h with respect to the usual signed measure given by the Edgeworth expansion to $k-2$ terms. However, the equivalence may be proved as follows. First, if $|D_s g(x)| \leqq G\{1 + |x|^p\}$ for some $G, p > 0$, integration by parts shows that

$$\mathbb{E}\{H_m(\mathcal{N})\, D_s g(\mathcal{N})\} = \mathbb{E}\{H_{m+s}(\mathcal{N})\, g(\mathcal{N})\},$$

where $(H_m)_{m \geqq 1}$ are the Hermite polynomials, and

$$\mathbb{E}\{H_r(\mathcal{N})\,\theta g(\mathcal{N})\} = -\mathbb{E}\left\{(g(\mathcal{N}) - \bar{g}) \int_0^{\mathcal{N}} H_r(x)\,dx\right\}.$$

Hence

$$\begin{aligned}
&\mathbb{E}\{H_m(\mathcal{N})D_{s+1}\,\theta h(\mathcal{N})\}\\
&\quad=-\mathbb{E}\left\{(h(\mathcal{N})-\bar{h})\int_0^{\mathcal{N}} H_{m+s+1}(x)\,dx\right\}\\
&\quad=-\mathbb{E}\left\{h(\mathcal{N})\left[\int_0^{\mathcal{N}} H_{m+s+1}(x)\,dx-\mathbb{E}\left(\int_0^{\mathcal{N}} H_{m+s+1}(x)\,dx\right)\right]\right\}\\
&\quad=-\left(\frac{1}{m+s+2}\right)\mathbb{E}\{H_{m+s+2}(\mathcal{N})\,h(\mathcal{N})\}.
\end{aligned}$$

Thus, iteratively,

$$\mathbb{E}\left\{\prod_{j=1}^{r}(D_{s_j+1}\,\theta)\,h(\mathcal{N})\right\}=(-1)^r\prod_{j=1}^{r}\left(\sum_{i=j}^{r}(s_i+2)\right)^{-1}\mathbb{E}\{H_{v+2r}(\mathcal{N})\,h(\mathcal{N})\},$$

where $v=\sum_{j=1}^{r}s_j$. The equivalence of (8) and the integral of h with respect to the Edgeworth density, expressed as usual in terms of Hermite polynomials, now follows by a combinatorial argument.

Proof. In order to prove the theorem, we use several analytical lemmas, relating estimates of the derivatives of a function h to those of θh. These are stated and proved in the next section.

Iteration of (6) leads to the statement of Theorem 1, with

$$\begin{aligned}
|\eta|\leqq{}&\sum_{(k-2)}{}'\left\{\prod_{j=1}^{r-1}\frac{|\kappa_{s_j+2}(W)|}{(s_j+1)!}\right\}\\
&\cdot\left[\frac{|\kappa_{s_r+2}(W)|}{(s_r+1)!}\,d_1\sum_{i=1}^{N}\mathbb{E}\left|X_i^2u_1\left(\theta\prod_{j=1}^{r}(D_{s_j+1}\,\theta)h_{W_i};X_i\right)\right|\right.\\
&\left.+d_{s_r+1}\sum_{i=1}^{N}E\left|X_i^{s_r+2}u_{s_r+1}\left(\theta\prod_{j=1}^{r-1}(D_{s_j+1}\,\theta)\,h_{W_i};X_i\right)\right|\right],\qquad k\geqq 3;\\
|\eta|\leqq{}& d_1\sum_{i=1}^{N}\mathbb{E}|X_i^2u_1(\theta h_{W_i};X_i)|,\qquad k=2;
\end{aligned}\tag{9}$$

where $\sum_{(k-2)}'$ denotes the sum over $\left\{r\geqq 1;\ s_j\geqq 1,\ 1\leqq j\leqq r\colon\ \sum_{j=1}^{r}s_j=k-2\right\}$. Now, as in the proof of Lemma 2,

$$\begin{aligned}
&u_{v_t+1}\left(\theta\prod_{j=1}^{t-1}(D_{v_j+1}\,\theta)\,h_w;x\right)\\
&\qquad\leqq|x|^{\alpha}\{1+(2^p+1)|w|^p+2^p|x|^p\}\,\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,h;p,\alpha\right),
\end{aligned}\tag{10}$$

and each term in (9) involves a quantity of the form estimated in (10), with $\sum_{j=1}^{t}v_j=k-2$.

Under condition (i), apply (44) of Lemma 6 with $g=h$, $q=k-1$ and $r=p$, obtaining, whenever $\sum_{j=1}^{t} v_j = k-2$,

$$\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)h;p,\alpha\right) \leqq KH_1,$$

for a universal constant $K=K(\alpha,p,t,(v_j)_{j=1}^{t})$. Thus, as in the proof of Lemma 2, for $\sum_{j=1}^{t} v_j = k-2$,

$$\begin{aligned}\mathbb{E}&\left|X_i^{v_0+2} u_{v_t+1}\left(\theta\prod_{j=1}^{t-1}(D_{v_j+1}\,\theta)h_{W_i};X_i\right)\right| \\ &\leqq 2^p(2^p+2)KH_1\,\mathbb{E}\{|X_i|^{v_0+2+\alpha}(1+|X_i|^p+\mathbb{E}|W|^p)\}. \qquad (11)\end{aligned}$$

Note also that

$$|\kappa_{s+2}(W)| = \left|\sum_{i=1}^{N}\kappa_{s+2}(X_i)\right| \leqq \alpha_{s+2}\sum_{i=1}^{N}\mathbb{E}|X_i|^{s+2}, \qquad (12)$$

and that because, for any random variable Y, the function $\log \mathbb{E}|Y|^s$ of $s \geqq 1$ is convex,

$$\begin{aligned}\sum_{i=1}^{N}\mathbb{E}|X_i|^{s+2}\sum_{j=1}^{N}\mathbb{E}|X_j|^{s'+2} &\leqq \sum_{i=1}^{N}\mathbb{E}|X_i|^{s+s'+2}\sum_{j=1}^{N}\mathbb{E}|X_j|^2 \\ &= \sum_{i=1}^{N}\mathbb{E}|X_i|^{s+s'+2}. \qquad (13)\end{aligned}$$

Taking (11), (12) and (13) into (9), we conclude easily that, under condition (i),

$$|\eta| \leqq C_1' H_1 \sum_{i=1}^{N}\mathbb{E}\{|X_i|^{k+\alpha}(1+\mathbb{E}|W|^p+|X_i|^p)\}. \qquad (14)$$

Under condition (ii), apply (45) of Lemma 6 instead of (44) to estimate $\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)h;p,\alpha\right)$, obtaining

$$|\eta| \leqq C_2' H_2 \sum_{i=1}^{N}\mathbb{E}\{|X_i|^{k+\alpha}(1+\mathbb{E}|W|^p+|X_i|^p)\}. \qquad (15)$$

Estimates (14) and (15) establish the theorem for $p \leqq 2$.

To complete the proof in the case $p>2$, take $h(w)=|w|^p$. Then, writing $p=r+\alpha$, where $r=[p]$, it follows that h is $r-1$ times differentiable, and satisfies condition (ii) with $k=r$ and $p=0$, since $\mathscr{L}(D_{r-1}h;1,\alpha)=H<\infty$. It thus follows from (15) that, for this function h,

$$|\eta| \leqq 3\,C_2' H \sum_{i=1}^{N} E|X_i|^p.$$

Hence, from (8), (12) and (13),

$$\mathbb{E}|W|^p \leqq \lambda_{p0} + \sum_{m=1}^{r-2} \lambda_{pm} \sum_{i=1}^{N} \mathbb{E}|X_i|^{2+m} + \lambda_{p,r-1} \sum_{i=1}^{N} \mathbb{E}|X_i|^p, \tag{16}$$

for universal constants $(\lambda_{pj})_{j=0}^{r-1}$. But now, from (13), for $t \geqq 2$,

$$\sum_{i=1}^{N} \mathbb{E}|X_i|^t \sum_{i=1}^{N} \mathbb{E}|X_i|^{2+m} \leqq \sum_{i=1}^{N} \mathbb{E}|X_i|^{t+m}. \tag{17}$$

Using (16) and (17) with $t = k+\alpha$, it follows that

$$\sum_{i=1}^{N} \mathbb{E}\{|X_i|^{k+\alpha}(1+\mathbb{E}|W|^p + |X_i|^p)\}$$
$$\leqq K' \left\{ \sum_{m=0}^{[p]} \sum_{i=1}^{N} \mathbb{E}|X_i|^{k+\alpha+m} + \sum_{i=1}^{N} \mathbb{E}|X_i|^{k+\alpha+p} \right\}, \tag{18}$$

and the convexity of $\log \mathbb{E}|Y|^t$ implies that this in turn is no greater than

$$K'' \sum_{i=0}^{N} \mathbb{E}\{|X_i|^{k+\alpha}(1+|X_i|^p)\}.$$

The theorem now follows from (14) and (15). □

Remarks 1. The difference between conditions (i) and (ii) lies in the trade-off between growth rate and smoothness of the functions h. Using condition (ii), functions $h(x)$ growing as fast as possible can be considered: that is, if $\mathbb{E}|X_i|^t < \infty$, $1 \leqq i \leqq N$, functions h such that $h(x) = O(|x|^t)$ are feasible, whereas, using condition (i), the maximum growth rate is $O(|x|^{t-2})$. On the other hand, functions h with an l-th derivative which is Lipschitz continuous with index α give an expansion up to $l_\wedge [t-2]$ terms using condition (i), but only up to $(l-1)_\wedge [t-2]$ terms using condition (ii).

2. Taking $h(x) = |x|^r$ for $r > 2$, let $k = [r]$, the greatest integer strictly smaller than r, and set $\alpha = r - k$. Then $\mathscr{L}(D_{k-1} h; 1, \alpha)$ is finite, and the theorem expresses $\mathbb{E}|W|^r$ in terms of an asymptotic expansion, involving the cumulants of W up to order k, and an error term no greater than a universal constant times $\sum_{i=1}^{N} \mathbb{E}|X_i|^r$. This, as a weak consequence, gives an inequality of Marcinkiewicz-Zygmund (1937) form,

$$\mathbb{E}\left|\sum_{i=1}^{N} X_i\right|^r \leqq C_r \left\{ \sum_{i=1}^{N} \mathbb{E}|X_i|^r + \left(\sum_{i=1}^{N} \mathbb{E}|X_i|^2\right)^{r/2} \right\}.$$

3. Subsidiary Analytic Results

Lemma 3. *Let h be $l-1$ times differentiable, with $D_s h(0) = 0$, $0 \leqq s \leqq l-1$, and satisfy*

$$H = \mathscr{L}(D_{l-1} h; L-l, \alpha) < \infty \tag{19}$$

for some $0 \leqq \alpha \leqq 1$ *and* $L \geqq l$. *Then*

$$|D_s \theta h(x)| \leqq K_0 H \begin{cases} 1 & |x| \leqq 1 \\ |x|^{L+\alpha-s-2} & |x| > 1 \end{cases}$$

for $0 \leqq s \leqq l-1$, *for universal constants* $K_0 = K_0(L, l, s, \alpha)$.

Proof. It follows from (19), by integration, that

$$\begin{aligned} |D_s h(x)| &\leqq \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+l-s)} H |x|^{\alpha+l-1-s}\{1+|x|^{L-l}\} \\ &\leqq \frac{2\Gamma(\alpha+1)}{\Gamma(\alpha+l-s)} H\{1+|x|^{L-1-s+\alpha}\}, \quad 0 \leqq s \leqq l-1. \end{aligned} \tag{20}$$

Thus also $\bar{h} = \mathbb{E} h(\mathcal{N})$ satisfies

$$|\bar{h}| \leqq \frac{2\Gamma(\alpha+1)}{\Gamma(\alpha+l)} H \mathbb{E}\{1+|\mathcal{N}|^{L+\alpha-1}\}, \tag{21}$$

where $\mathcal{N}$ is a standard normal random variable.

We consider only the case $x \geqq 0$, starting from the equation

$$\theta h(x) = -\int_x^\infty e^{\frac{1}{2}(x^2-t^2)}\{h(t)-\bar{h}\}\,dt: \tag{22}$$

for $x<0$, the argument is similar, based on the equation

$$\theta h(x) = \int_{-\infty}^x e^{\frac{1}{2}(x^2-t^2)}\{h(t)-\bar{h}\}\,dt. \tag{23}$$

Changing variable in (22) gives

$$\theta h(x) = -\int_0^\infty e^{-xz-z^2/2}\{h(z+x)-\bar{h}\}\,dz,$$

from which it follows easily that

$$D_s \theta h(x) = -\int_0^\infty \sum_{m=0}^s \binom{s}{m}(-z)^m D_{s-m}\{h(z+x)-\bar{h}\}\, e^{-xz-z^2/2}\,dz, \qquad 0 \leqq s \leqq l-1. \tag{24}$$

Now direct calculation shows that, for $m \in \mathbb{Z}^+$,

$$\int_0^\infty z^m e^{-zx-z^2/2}\,dz \leqq \min\{\sqrt{\pi/2}\,\mathbb{E}|\mathcal{N}|^m; m!\,x^{-m-1}\}, \tag{25}$$

and that, for $m, n \in \mathbb{Z}^+$ and $0 \leqq \beta \leqq 1$,

$$\int_0^\infty z^m (z+x)^{n+\beta} e^{-zx-z^2/2}\,dz \leqq \begin{cases} \sqrt{\pi/2} \sum_{j=0}^n \binom{n}{j} \mathbb{E}\{|\mathcal{N}|^{m+j}(|\mathcal{N}|+1)^\beta\}, & 0 \leqq x \leqq 1 \\ x^{n+\beta-m-1} \sum_{j=0}^{n+1} \binom{n+1}{j}(m+j)!, & x > 1. \end{cases} \tag{26}$$

Hence easily, from (20), (21) and (24), the statement of the lemma follows. □

Lemma 4. *Suppose that the conditions of Lemma 3 hold. Define*

$$\varepsilon_1(x, y)=|xD_{l-1}\,\theta h(x)-yD_{l-1}\,\theta h(y)|/[|x-y|^\alpha(1+|x|^{L-l}+|y|^{L-l})];$$
$$\varepsilon_2(x, y)=|D_{l-1}\,\theta h(x)-D_{l-1}\,\theta h(y)|/[|x-y|^\alpha(1+|x|^{L-l-1}+|y|^{L-l-1})].$$

Then

$$\sup_{x\neq y}\varepsilon_i(x, y)\leqq K_i H, \qquad i=1,2 \tag{27}$$

for universal constants $K_i=K_i(L, l, \alpha)$, $i=1,2$.

Proof. We estimate ε_1 first: without loss of generality, take $|x|\geqq|y|$ throughout. All constants of the form K'_j, $j\geqq 1$, are universal.

If $|x-y|\geqq\frac{1}{2}|x|$, it follows that

$$\varepsilon_1(x, y)\leqq\{|xD_{l-1}\,\theta h(x)|+|yD_{l-1}\,\theta h(y)|\}/[(\tfrac{1}{2}|x|)^\alpha(1+|x|^{L-l})]. \tag{28}$$

Note that the function Ψ defined by

$$\Psi(t)=\begin{cases} t, & 0\leqq t\leqq 1\\ t^{L-l+\alpha}, & t>1\end{cases}$$

is non-decreasing in t whenever $L\geqq l$ and $\alpha\geqq 1$. Hence, using Lemma 3,

$$|xD_{l-1}\,\theta h(x)|+|yD_{l-1}\,\theta h(y)|\leqq 2K_0(L, l, l-1, \alpha)\,H\Psi(|x|),$$

and it now follows easily from (28) that, for $|x-y|\geqq\frac{1}{2}|x|$,

$$\varepsilon_1(x, y)\leqq 2^{\alpha+1}K_0(L, l, l-1, \alpha)\,H. \tag{29}$$

For $|x-y|\leqq\frac{1}{2}|x|$, take, without loss of generality, $0<\frac{x}{2}\leqq y<x$. Then (24) yields

$$\begin{aligned}&|xD_{l-1}\,\theta h(x)-yD_{l-1}\,\theta h(y)|\\ &\quad\leqq\sum_{m=0}^{l-1}\binom{l-1}{m}\int_0^\infty z^m e^{-z^2/2}|xe^{-xz}D_{l-1-m}h(z+x)-ye^{-yz}D_{l-1-m}h(z+y)|\,dz\\ &\qquad+|\bar{h}|\int_0^\infty z^{l-1}e^{-z^2/2}|xe^{-xz}-ye^{-yz}|\,dz.\end{aligned} \tag{30}$$

Now, from the mean value theorem,

$$\begin{aligned}|xe^{-xz}-ye^{-yz}|&\leqq(x-y)\,e^{-yz}(1+xz)\\ &\leqq(x-y)\,e^{-xz/2}(1+xz),\end{aligned}$$

and so, from (20), (25) and (26), for $0\leqq m\leqq l-1$,

$$\begin{aligned}&\int_0^\infty z^m e^{-z^2/2}|xe^{-xz}-ye^{-yz}|\,|D_{l-1-m}h(z+x)|\,dz\\ &\quad\leqq\frac{2\Gamma(\alpha+1)}{\Gamma(\alpha+m+1)}H\int_0^\infty z^m e^{-z^2/2-xz/2}(x-y)(1+xz)(1+(z+x)^{L-l+\alpha+m})\,dz\\ &\quad\leqq K'_1 H(x-y)\begin{cases}1, & x\leqq 1\\ x^{L-l+\alpha-1}, & x>1,\end{cases}\end{aligned} \tag{31}$$

for a universal constant $K_1'=K_1'(L,l,m,\alpha)$. From (21), a similar estimate holds easily for the term involving $\bar{h}$. It thus remains to estimate

$$\int_0^\infty z^m e^{-z^2/2} y e^{-yz} |D_{l-1-m} h(z+x) - D_{l-1-m} h(z+y)|\, dz, \qquad 0 \leqq m \leqq l-1. \tag{32}$$

For $1 \leqq m \leqq l-1$, using (20),

$$|D_{l-1-m} h(z+x) - D_{l-1-m} h(z+y)| \leqq \frac{2(x-y)\Gamma(\alpha+1)}{\Gamma(\alpha+m)} H\{1+(z+x)^{L-l-1+m+\alpha}\},$$

and hence, using (25) and (26), (32) is bounded by

$$K_2' H(x-y)\begin{cases} 1, & x \leqq 1 \\ x^{L-l+\alpha-1}, & x>1, \end{cases} \tag{33}$$

for $K_2'=K_2'(L,l,m,\alpha)$, and it remains to treat the case $m=0$. Here, from (19), (25) and (26),

$$\begin{aligned} &\int_0^\infty y e^{-yz-z^2/2} |D_{l-1} h(z+x) - D_{l-1} h(z+y)|\, dz \\ &\quad \leqq H \int_0^\infty y e^{-yz-z^2/2} (x-y)^\alpha \{1+2(x+z)^{L-l}\}\, dz \\ &\quad \leqq K_3' H x (x-y)^\alpha \begin{cases} 1, & 0<x \leqq 1 \\ x^{L-l-1}, & x>1, \end{cases} \end{aligned} \tag{34}$$

for $K_3'=K_3'(L,l,\alpha)$.

Hence, using (31), (33) and (34), we have, for $0<\frac{x}{2} \leqq y \leqq x \leqq 1$,

$$\varepsilon_1(x,y) \leqq \frac{K_4' H\{x(x-y)^\alpha + (x-y)\}}{(x-y)^\alpha} \leqq 2K_4' H; \tag{35}$$

for $0<\frac{x}{2} \leqq y \leqq x$ and $x>1$, we have

$$\begin{aligned} \varepsilon_1(x,y) &\leqq \frac{K_5' H\{x^{L-l}(x-y)^\alpha + (x-y) x^{L-l+\alpha-1}\}}{(x-y)^\alpha x^{L-l}} \\ &= K_5' H\left\{1+\left(1-\frac{y}{x}\right)^{1-\alpha}\right\} \leqq 2K_5' H. \end{aligned} \tag{36}$$

Inequalities (29), (35) and (36) complete the estimate of ε_1.

The estimate of ε_2 is rather similar. For $|x-y| \leqq \frac{1}{2}|x|$, proceed in a manner similar to that used for ε_1, yielding, for $0<\frac{x}{2} \leqq y<x$,

$$\begin{aligned} &|D_{l-1}\,\theta h(x) - D_{l-1}\,\theta h(y)| \\ &\quad \leqq K_6' H \begin{cases} (x-y)+(x-y)^\alpha, & 0 \leqq x \leqq 1 \\ (x-y) x^{L-l+\alpha-2} + (x-y)^\alpha x^{L-l-1}, & x>1, \end{cases} \end{aligned}$$

and hence

$$\varepsilon_2(x, y) \leqq 2K_6' H, \qquad |x-y| \leqq \tfrac{1}{2}|x|. \tag{37}$$

For $|x-y| \geqq \frac{1}{2}|x|$ and $0 \leqq |y| \leqq |x| \leqq 1$, make estimates in a similar manner, replacing e^{-yz} by e^z instead of $e^{-xz/2}$. All the integrals converge because of the factor $e^{-z^2/2}$, yielding easily, in this range

$$|D_{l-1}\,\theta h(x) - D_{l-1}\,\theta h(y)| \leqq K_7' H((x-y) + (x-y)^\alpha),$$

from which

$$\varepsilon_2(x, y) \leqq 2K_7' H, \qquad 0 \leqq |y| \leqq |x| \leqq 1. \tag{38}$$

Finally, for $|x-y| \geqq \frac{1}{2}|x|$ and $|x|>1$, use the estimates of Lemma 3 directly:

$$\frac{|D_{l-1}\,\theta h(x)|}{|x-y|^\alpha[1+|x|^{L-l-1}+|y|^{L-l-1}]} \leqq K_8' H \frac{|x|^{L+\alpha-l-1}}{|x|^\alpha |x|^{L-l-1}} = K_8' H,$$

and

$$\frac{|D_{l-1}\,\theta h(y)|}{|x-y|^\alpha[1+|x|^{L-l-1}+|y|^{L-l-1}]} \leqq K_9' H \begin{cases} 1, & |y| \leqq 1 \\ \dfrac{1+|x|^{L+\alpha-l-1}}{|x|^\alpha[1+|x|^{L-l+1}]}, & |y|>1, \end{cases}$$

and so

$$\varepsilon_2(x, y) \leqq K_{10}' H, \qquad |x-y| \geqq \tfrac{1}{2}|x| \geqq \tfrac{1}{2}. \tag{39}$$

The second part of the lemma follows from (37), (38) and (39). □

Lemma 5. *Under the conditions of Lemma* 3, *θh is l times differentiable, and satisfies*

$$\sup_x \{|D_s\,\theta h(x)|/[1+|x|^{L+\alpha-s}]\} \leqq K_4(L, l, s, \alpha) H, \; 0 \leqq s \leqq l; \tag{40}$$

furthermore,

$$\mathscr{L}(D_l\,\theta h; L-l, \alpha) \leqq K_4(L, l, l+\alpha, \alpha) H. \tag{41}$$

The constants $K_4(L, l, s, \alpha)$ are universal.

Proof. The inequalities in (40) for $0 \leqq s \leqq l-1$ are already implied by Lemma 3. Then, since

$$D\theta h(x) = x\theta h(x) + h(x),$$

it follows, differentiating $l-1$ times, that

$$D_l\,\theta h(x) = x D_{l-1}\,\theta h(x) + (l-1) D_{l-2}\,\theta h(x) + D_{l-1}\,h(x). \tag{42}$$

Inequality (40) for $s=l$ now follows from (20) and Lemma 3. For the last part, (42) implies that

$$|D_l\,\theta h(x)-D_l\,\theta h(y)|/[|x-y|^\alpha(1+|x|^{L-l}+|y|^{L-l})]$$
$$\leqq\{|xD_{l-1}\,\theta h(x)-yD_{l-1}\,\theta h(y)|+(l-1)|D_{l-2}\,\theta h(x)-D_{l-2}\,\theta h(y)|$$
$$+|D_{l-1}\,h(x)-D_{l-1}\,h(y)|\}/[|x-y|^\alpha(1+|x|^{L-l}+|y|^{L-l})]. \tag{43}$$

The first term in (43) is estimated using Lemma 4, and the last term by (19). For the second,

$$|D_{l-2}\,\theta h(x)-D_{l-2}\,\theta h(y)|\leqq|x-y|\,|D_{l-1}\,\theta h(u)|,$$

for some u between x and y, and Lemma 3 now completes the argument. □

Lemma 6. *Let g be any $(q-1)$ times differentiable function such that $G_1=\mathscr{L}(D_{q-1}g;\,r,\alpha)<\infty$. Then, for any $t\geqq 1$ and $v_j\geqq 1$, $1\leqq j\leqq t$, such that $\sum_{j=1}^{t} v_j=q-1$,*

$$\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g;r,\alpha\right)\leqq G_1\prod_{j=1}^{t}K_4(r+q_j,q_j,q_j+\alpha,\alpha), \tag{44}$$

where $q_j=1+\sum_{i=j}^{t} v_i$. If g is q times differentiable and $G_2=\mathscr{L}(D_q g;r+1,\alpha)<\infty$, then

$$\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g;r,\alpha\right)$$
$$\leqq G_2K_2(v_t+r+3,v_t+2,\alpha)\prod_{j=1}^{t-1}K_4(r+q_j+2,q_j+1,q_j+1+\alpha,\alpha). \tag{45}$$

Proof. We prove (44) by induction on t. Note first that since, for the Hermite polynomials, we have

$$\theta H_m=\begin{cases}0 & m=0\\ -H_{m-1} & m\geqq 1,\end{cases}$$

any polynomial π of degree m is converted by θ to a polynomial of degree $m-1$. Hence, if $\deg(\pi)\leqq 2t+\sum_{j=1}^{t} v_j$,

$$\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\{g+\pi\}=\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g.$$

For $t=1$, we have $v_1=q-1$. Take

$$\pi(x)=-\sum_{j=0}^{q-1}x^jD_jg(0)/j!, \tag{46}$$

and apply Lemma 5 to $(g+\pi)$ with $l=q$: it follows that $D_{v_1+1}\,\theta g$ exists, and that

$$\mathscr{L}(D_{v_1+1}\,\theta g;r,\alpha)\leqq K_4(r+q,q,q+\alpha,\alpha)\,G_1.$$

This establishes (44) for $t=1$, and for any g and q such that g is $(q-1)$ times differentiable. Now assume that (44) is true for products with up to $t-1$ factors, and for any g and q such that g is $(q-1)$ times differentiable. We analyse $\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g; r, \alpha\right)$ as $\mathscr{L}\left(\prod_{j=2}^{t}(D_{v_j+1}\,\theta)(D_{v_1+1}\,\theta(g+\pi)); r, \alpha\right)$, where π is defined as in (46) and, as above, from Lemma 5,

$$\mathscr{L}(D_{q-v_1-1}\{D_{v_1+1}\,\theta(g+\pi)\}; r, \alpha) \leqq G_1\,K_4(r+q_1, q_1, q_1+\alpha, \alpha).$$

By the induction hypothesis, this implies that

$$\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g; r, \alpha\right) \leqq G_1\,K_4(r+q_1, q_1, q_1+\alpha, \alpha)\prod_{j=2}^{t}K_4(r+q_j, q_j, q_j+\alpha, \alpha)$$

as required.

To establish (45), first use (44) with $q+1$ for q and $r+1$ for r, to deduce that

$$\mathscr{L}(D_{v_t+1}\prod_{j=1}^{t-1}(D_{v_j+1}\,\theta)\,g; r+1, \alpha) \leqq G_2\prod_{j=1}^{t-1}K_4(r+q_j+2, q_j+1, q_j+1+\alpha, \alpha).$$

Applying (27) of Lemma 4 to $\prod_{j=1}^{t-1}(D_{v_j+1}\theta)\,g+\pi^*$ with $i=2$, $l=v_t+2$ and $L-l=r+1$, where

$$\pi^*(x) = -\sum_{j=0}^{v_t+1}x^j D_j\left\{\prod_{j=1}^{t-1}(D_{v_j+1}\,\theta)\,g\right\}(0)/j!,$$

it follows that

$$\mathscr{L}\left(\prod_{j=1}^{t}(D_{v_j+1}\,\theta)\,g; r, \alpha\right) \leqq G_2\,K_2(v_t+r+3, v_t+2, \alpha)\prod_{j=1}^{t-1}K_4(r+q_j+2, q_j+1, q_j+1+\alpha, \alpha),$$

as required. □

References

von Bahr, B.: On the convergence of moments in the central limit theorem. Ann. Math. Statist. **36**, 808–818 (1965)

Bhattacharya, R.N.: Recent results on refinements of the central limit theorem. Proc. VIth Berkeley Sympos. Math. Statist. Probab. **2**, 453–484 (1970)

Götze, F., Hipp, C.: Asymptotic expansions in the central limit theorem under moment conditions. Z. Wahrscheinlichkeitstheor. Verw. Geb. **42**, 67–87 (1978)

Hipp, C.: Edgeworth expansions for integrals of smooth functions. Ann. Probab. **5**, 1004–1011 (1977)

Hsu, P.L.: The approximate distribution of the mean and variance of a sample of independent variables. Ann. Math. Statist. **16**, 1–29 (1945)

Marcinkiewicz, J., Zygmund, A.: Sur les fonctions indépendantes. Fundam. Math. **29**, 60–90 (1937)

Stein, C.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proc. VIth Berkeley Sympos. Math. Statist. Probab. **2**, 583–602 (1970)

Received November 25, 1984