JOURNAL OF MULTIVARIATE ANALYSIS 27, 68–79 (1988)

On Moment Conditions for Valid Formal Edgeworth Expansions

RABI N. BHATTACHARYA*

Indiana University

AND

J. K. GHOSH

The Indian Statistical Institute, Calcutta, India

Communicated by the Editors

The validity of formal Edgeworth expansions for statistics which are functions of sample averages was established in R. N. Bhattacharya and J. K. Ghosh (1978, *Ann. Statist.* 6 434–451) under a moment condition which is sometimes too severe. In this article this moment condition is relaxed. Two examples of P. Hall (1983, *Ann. Probab.* 11 1028–1036; 1987, *Ann. Probab.* 15 920–931) are discussed in this context.

INTRODUCTION

The validity of formal Edgeworth expansions for classical statistics was established in Bhattacharya and Ghosh [2] under moment conditions which cannot be relaxed in general, but turn out to be too severe in some cases. Two such examples are considered in Hall [6, 7]. In these examples and many others the highest order of moments involved in the actual expansion is much smaller than the order of moments assumed finite in our earlier work [2], and special methods were used by Hall [6, 7] to relax this moment condition. Attempts to find minimal moment restrictions for the general case run into unexpected analytical difficulties.

Suppose that the statistic may be expressed as (or approximated by) $H(\bar{Z})$, where $\bar{Z} = (1/n) \sum_{j=1}^{n} Z_j$ is a mean of i.i.d. vectors and H is a smooth function in a neighborhood of $\mu = EZ_j$. If all the components of grad $H(\mu)$

Received February 16, 1988.

Research supported by NSF Grant DMS 8503358.

AMS 1980 subject classifications: Primary 60F05; Secondary 62E20.

Key words and phrases: Asymptotic expansion, student's statistic.

0047-259X/88 $3.00

are nonzero then one cannot significantly weaken the earlier moment assumptions. In this article we provide a relaxation of the moment condition in case grad $H(\mu)$ has some zero components, as is true in both examples of Hall. Apart from the method we present in detail here, another method using conditioning with respect to some coordinates of Z_j (namely coordinates $Z_j^{(i)}$ for which $(\partial H(z)/\partial z^{(i)})(\mu)=0$) is sketched as Remark 5 in Section 7. This last method generalizes some ideas of Hall [7] dealing with Student's statistic.

1. The Main Result

Many classical statistics are (or, may be approximated by statistics) of the form $H(\bar{Z})$, where $\bar{Z}=(1/n)\sum_1^n Z_j$ is a k-dimensional mean vector of sample characteristics and H is smooth in a neighborhood of $\mu=E\bar{Z}$. If grad $H(\mu)\neq 0$, and $E|Z_j|^2<\infty$, then the normalized statistic $W_n=\sqrt{n}\,(H(\bar{Z})-H(\mu))$ is asymptotically normal. This follows from the Taylor expansion

$$W_n=\sqrt{n}\,(\bar{Z}-\mu)\cdot \text{grad } H(\mu)+o_p(1)). \tag{1.1}$$

If $E|Z_j|^s<\infty$ for some integer $s\geqslant 3$ and H is s-times continuously differentiable in a neighborhood of μ, then one may approximate W_n better by

$$W_n''=n^{1/2}\left\{\sum_{i=1}^{k} l_i(\bar{Z}^{(i)}-\mu^{(i)})+\frac{1}{2!}\sum_{i_1,i_2=1}^{k} l_{i_1,i_2}(\bar{Z}^{(i_1)}-\mu^{(i_1)})(\bar{Z}^{(i_2)}-\mu^{(i_2)})\right.$$
$$\left.+\cdots+\frac{1}{(s-1)!}\sum_{i_1,\ldots,i_{s-1}=1}^{k} l_{i_1,\ldots,i_{s-1}}(\bar{Z}^{(i_1)}-\mu^{(i_1)})\cdots(\bar{Z}^{(i_{s-1})}-\mu^{(i_{s-1})})\right\}. \tag{1.2}$$

Here superscripts denote coordinates and $l_i=(D_iH)(\mu)$, $l_{i_1,i_2}=(D_{i_1}D_{i_2}H)(\mu)$, etc., with D_i denoting differentiation with respect to the ith coordinate. One may compute the jth cumulant $K_{j,n}$ of W_n'' algebraically $(1\leqslant j\leqslant s)$, and keep only terms up to order $O(n^{-(s-2)/2})$:

$$K_{j,n}=\tilde{K}_{j,n}+o(n^{-(s-2)/2}) \qquad (1\leqslant j\leqslant s), \tag{1.3}$$

$\tilde{K}_{j,n})$ being a polynomial in $n^{-1/2}$ with coefficients determined by the moments of Z_j and the derivatives $l_i, k_{i_1,i_2},\ldots,l_{i_1,\ldots,i_{s-1}}$. One has $\tilde{K}_{1,n}=O(n^{-1/2})$, $\tilde{K}_{2,n}=\sigma^2+o(n^{-1/2})$, $\tilde{K}_{j,n}=O(n^{-(j-2)/2})$ $(j\geqslant 3)$, where

$$\sigma^2=\text{grad } H(\mu)\cdot V\,\text{grad } H(\mu),$$
$$V\equiv \text{cov } Z_j. \tag{1.4}$$

The characteristic function of W_n is now approximated by

$$\exp\left\{\sum_{j=1}^{s-2}\frac{(i\xi)^j}{j!}\bar{K}_{j,n}\right\}$$
$$=\exp\left\{-\frac{\sigma^2\xi^2}{2}\right\}\exp\left\{i\xi\bar{K}_{1,n}-\frac{\xi^2}{2}(\bar{K}_{2,n}-\sigma^2)+\sum_{j=3}^{s-2}\frac{(i\xi)^j}{j!}\bar{K}_{j,n}\right\}$$
$$=\exp\left\{-\frac{\sigma^2\xi^2}{2}\right\}\left[1+\sum_{j=1}^{s-2}n^{-j/2}\pi_j(i\xi)\right]$$
$$+\,o(n^{-(s-2)/2})=\hat{\psi}_{s,n}(\xi)+o(n^{-(s-2)/2}), \tag{1.5}$$

say. For the second equality in (1.5) one expands in powers of $n^{-1/2}$. Here $\pi_j(i\xi)$ is a polynomial (in $i\xi$) whose coefficients depend on the moments of Z_j and the derivatives of H at μ. Now $\hat{\psi}_{s,n}$ is the Fourier transform of the *density* $\psi_{s,n}$ *of the formal Edgeworth expansion* of the distribution of W_n, obtained by inversion:

$$\psi_{s,n}(x)=\left[1+\sum_{j=1}^{s-2}n^{-j/2}\pi_j\left(-\frac{d}{dx}\right)\right]\phi_{\sigma^2}(x),$$
$$\phi_{\sigma^2}(x)\equiv\frac{1}{\sqrt{2\pi\sigma^2}}e^{-x^2/2\sigma^2} \tag{1.6}$$

Suppose that the observations Y_j $(j=1,2,...)$ are i.i.d. m-dimensional with common distribution G and that

$$Z_j=(f_1(Y_j),f_2(Y_j),...,f_k(Y_j))=(Z_j^{(1)},Z_j^{(2)},...,Z_j^{(k)}), \tag{1.7}$$

where f_r $(1\leqslant r\leqslant k)$ are real-valued Borel measurable functions on $\mathbf{R}^m$. Let Q_1 denote the (common) distribution of $Z_j-\mu$. The following assumptions were made in Bhattacharya and Ghosh [2], Bhattacharya [1], to prove the validity of the formal expansion (1.6) (i.e., to establish $\text{Prob}(W_n\in B)=\int_B\psi_{s,n}(x)\,dx+o(n^{-(s-2)/2})$ uniformly for all Borel sets B):

(B$_1$) *H is $(s-1)$-times continuously differentiable in a neighborhood of μ.*

(B$_2$) *grad $H(\mu)\neq 0$.*

(B$_3$) *$E|f_r(Y_j)|^s<\infty$ for $1\leqslant r\leqslant k$.*

(B$_4$) *There exists a nonempty open subset U of $\mathbf{R}^m$ with the properties: (i) G has a nonzero absolutely continuous component (with respect to Lebesgue measure on $\mathbf{R}^m$) with a positive density on U; (ii) f_r $(1\leqslant r\leqslant k)$ are continuously differentiable on U; (iii) $1,f_1,...,f_k$ are linearly independent as elements of the vector space of real valued continuous functions on U.*

Let us now assume, instead of (B_1), (B_2), (B_3),

(B_1') *H is s-times continuously differentiable in a neighborhood of* μ.

(B_2') (i) $l_i \neq 0$ *for* $1 \leqslant i \leqslant k_1$; (ii) $l_i = 0$ *for* $k_1 < i \leqslant k$, *where* k_1 *is an integer satisfying* $1 \leqslant k_1 < k$.

(B_3') (i) $E|f_r(Y_j)|^s < \infty$ *for* $1 \leqslant r \leqslant k_1$; (ii) $E|f_r(Y_j)|^{s-1} < \infty$ *for* $k_1 < r \leqslant k$, *for some positive integer* $s \geqslant 3$.

Our main result relaxing earlier moment conditions is the following.

THEOREM. *Under the assumptions* (B_1'), (B_2'), (B_3'), (B_4) *one has*

$$\sup_{u \in \mathbf{R}^1} \left| \operatorname{Prob}(W_n \leqslant u) - \int_{-\infty}^{u} \psi_{s,n}(x)\,dx \right| = o(n^{-(s-2)/2}). \tag{1.8}$$

Proof. Recall the notation $W_n = \sqrt{n}(H(\bar{Z}) - H(\mu))$. Let

$$\begin{aligned} W_n' = &\sum_{1 \leqslant i \leqslant k_1} l_i \sqrt{n}\,(\bar{Z}^{(i)} - \mu^{(i)}) \\ &+ \frac{n^{-1/2}}{2!} \sum_{1 \leqslant i_1, i_2 \leqslant k} l_{i_1, i_2} \sqrt{n}\,(\bar{Z}^{(i_1)} - \mu^{(i_1)}) \sqrt{n}\,(\bar{Z}^{(i_2)} - \mu^{(i_2)}) \\ &+ \cdots + \frac{n^{-(s-1)/2}}{s!} \sum_{1 \leqslant i_1, i_2, \ldots, i_s \leqslant k} l_{i_1, i_2, \ldots, i_s} \sqrt{n}\,(\bar{Z}^{(i_1)} - \mu^{(i_1)}) \\ &\cdots \sqrt{n}\,(\bar{Z}^{(i_s)} - \mu^{(i_s)}). \end{aligned} \tag{1.9}$$

We first prove (1.8) with W_n replaced by W_n'. By Lemma 2.2 in Bhattacharya and Ghosh [2], Q^{*k} (i.e., the distribution of $\sum_1^k (Z_j - \mu)$) has a nonzero absolutely continuous component. Hence the distribution Q_n of $\sqrt{n}\,(\bar{Z} - \mu)$ has a nonzero absolutely continuous component for $n \geqslant k$. Write

$$\begin{aligned} h(z, \varepsilon) = &\sum_{1 \leqslant i \leqslant k_1} l_i z^{(i)} + \frac{\varepsilon}{2!} \sum_{1 \leqslant i_1, i_2 \leqslant k} l_{i_1, i_2} z^{(i_1)} z^{(i_2)} \\ &+ \cdots + \frac{\varepsilon^{s-1}}{s!} \sum_{1 \leqslant i_1, \ldots, i_s \leqslant k} l_{i_1, \ldots, i_s} z^{(i_1)} \cdots z^{(i_s)}, \\ h(z, 0) = &\sum_{1 \leqslant i \leqslant k_1} l_i z^{(i)}. \end{aligned} \tag{1.10}$$

Now it is shown in Bhattacharya and Ranga Rao [3] (see the proof of Theorem 19.5 and the remark on p. 207) that there exists a part q_n' of the density (component) of Q_n which has the properties

$$\sup_{B} \left| \int_B q_n'(z)\,dz - Q_n(B) \right| = o(n^{-(s-2)/2}) \qquad (B \text{ a Borel subset of } \mathbf{R}^k) \tag{1.11}$$

and

$$|q_n'(z)-\xi_{s-1,n}(z)|\leqslant c\delta_n n^{-(s-3)/2}(1+|z|^{s+k}), \qquad [z\in\mathbb{R}^k], \quad (1.12)$$

where $\xi_{s-1,n}(z)$ *is the density of the* $(s-2)$*-term Cramér–Edgeworth expansion of* Q_n, c is a positive constant, and $\delta_n\to 0$ as $n\to\infty$. Note that (1.11) holds under the assumptions (B_3'), (B_4); i.e., $E|Z_j|^{s-1}<\infty$ suffices. Indeed the right side in (1.11) is $o(n^{-m})$ for *every* positive integer m (see relations (19.73), (19.76), (19.77) in Bhattacharya and Ranga Rao [3]).

By (1.11) the following holds uniformly for all u:

$$\begin{aligned}
\text{Prob}(W_n'\leqslant u) &= \text{Prob}\left(\sum_{1\leqslant i\leqslant k_1} l_i\sqrt{n}\,(\bar{Z}^{(i)}-\mu^{(i)})\leqslant u\right)\\
&\quad+\text{Prob}\left(\{W_n'\leqslant u\}\Big\backslash\left\{\sum_{1\leqslant i\leqslant k_1} l_i\sqrt{n}\,(\bar{Z}^{(i)}-\mu^{(i)})\leqslant u\right\}\right)\\
&\quad-\text{Prob}\left(\left\{\sum_{1\leqslant i\leqslant k_1} l_i\sqrt{n}\,(\bar{Z}^{(i)}-\mu^{(i)})\leqslant u\right\}\Big\backslash\{W_n'\leqslant u\}\right)\\
&=\text{Prob}\left(\sum_{1\leqslant i\leqslant k_1} l_i\sqrt{n}\,(\bar{Z}^{(i)}-\mu^{(i)})\leqslant u\right)\\
&\quad+\int_{\{h(z,\varepsilon)\leqslant u\}\backslash\{h(z,0)\leqslant u\}} q_n'(z)\,dz\\
&\quad-\int_{\{h(z,0)\leqslant u\}\backslash\{h(z,\varepsilon)\leqslant u\}}+o(n^{-(s-2)/2}). \qquad (1.13)
\end{aligned}$$

But in view of (B_3')(i) (and (B_4)) one has, unformly for all u,

$$\begin{aligned}
&\text{Prob}\left(\sum_{1\leqslant i\leqslant k_1} l_i\sqrt{n}\,(\bar{Z}^{(i)}-\mu^{(i)})\leqslant u\right)\\
&\quad=\int_{\{z\in\mathbb{R}^{k_1}:\sum_1^{k_1} l_i z^{(i)}\leqslant u\}} {}^1\xi_{s,n}(z)\,dz+o(n^{-(s-2)/2}), \qquad (1.14)
\end{aligned}$$

where ${}^1\xi_{s,n}$ *is the density of the* $(s-1)$*-term Cramér–Edgeworth expansion of the distribution of* $\sqrt{n}\,(\bar{Z}^{(1)}-\mu^{(1)},\ldots,\bar{Z}^{(k_1)}-\mu^{(k_1)})$.

On the other hand,

$$\begin{aligned}
&\int_{\{h(z,\varepsilon)\leqslant u\}\backslash\{h(z,0)\leqslant u\}} q_n'(z)\,dz-\int_{\{h(z,0)\leqslant u\}\backslash\{h(z,\varepsilon)\leqslant u\}} q_n'(z)\,dz\\
&\quad=\int_{\{h(z,\varepsilon)\leqslant u\}\backslash\{h(z,0)\leqslant u\}}\xi_{s-1,n}(z)\,dz-\int_{\{h(z,0)\leqslant u\}\backslash\{h(z,\varepsilon)\leqslant u\}}\xi_{s-1,n}(z)\,dz+\eta_n,
\end{aligned} \qquad (1.15)$$

where, by (1.12),

$$\eta_n \leqslant \left(\int_{\{h(z,\varepsilon) \leqslant u\} \Delta \{h(z,0) \leqslant u\}} (1 + |z|^{s+k})^{-1} \, dz \right) c\delta_n n^{-(s-3)/2} \tag{1.16}$$

Here Δ denotes *symmetric difference*: $B \Delta C = (B \backslash C) \cup (C \backslash B)$. Note that for z in $\{|z| < 1/\varepsilon^{1/(s-1)}\}$ there are positive constants c_1, d_1 *such that*

$$h(z,\varepsilon) - c_1 \varepsilon \, |z||^2 - d_1 \varepsilon \leqslant h(z,0) \leqslant h(z,\varepsilon) + c_1 \varepsilon \, |z|^2 + d_1 \varepsilon. \tag{1.17}$$

Write, for given u satisfying $|u| < 2\,|l|/\varepsilon^{1/(s-1)} (|l|^2 = \sum_{1 \leqslant i \leqslant k_1} l_i^2)$,

$$A_\varepsilon = (\{h(z,\varepsilon) \leqslant u\} \; \Delta \{h(z,0) \leqslant u\} \cap \{|z| < 1/\varepsilon^{1/(s-1)}\}. \tag{1.18}$$

Then

$$\begin{aligned} &A_\varepsilon \subset A_{\varepsilon 1} \cup A_{\varepsilon 2}, \\ &A_{\varepsilon 1} = \{u - c_1 \varepsilon \, |z|^2 - d_1 \varepsilon \leqslant h(z,0) \leqslant u\} \cap \{|z| < 1/\varepsilon^{1/(s-1)}\}, \\ &A_{\varepsilon 2} = \{u < h(z,0) \leqslant u + c_1 \, |z|^2 + d_1 \varepsilon\} \cap \{|z| < 1/\varepsilon^{1/(s-1)}\}. \end{aligned} \tag{1.19}$$

Now make an orthogonal transformation $z \to y$ with $y^{(1)} = h(z,0)/|l| = \sum_{1 \leqslant i \leqslant k_1} l_i z^{(i)} / (\sum l_i^2)^{1/2}$. Then

$$\int_{A_{\varepsilon 1}} (1 + |z|^{s+k})^{-1} \, dz$$

$$= \int_{\{(u - c_1 \varepsilon |y|^2 - d_1 \varepsilon)/|l| \leqslant y^{(1)} \leqslant u/|l|\} \cap \{|y| < 1/\varepsilon^{1/(s-1)}\}} (1 + |y|^{s+k})^{-1} \, dy. \tag{1.20}$$

Write $|y|^2 = (y^{(1)})^2 + \sum_2^k (y^{(i)})^2 = (y^{(1)})^2 + r^2$ and solve the quadratic equation (in $y^{(1)}$): $y^{(1)} = (u - c_1 \varepsilon (y^{(1)})^2 - c_1 \varepsilon r^2 - d_1 \varepsilon)//|l|$, to derive from (1.20) the inequality

$$\int_{A_{\varepsilon 1}} (1 + |z|^{s+k})^{-1} \, dz$$

$$\leqslant \int_{\{(u/|l|) - c_2 \varepsilon \leqslant y^{(1)} \leqslant u/|l|\} \cap \{|y| < 1/\varepsilon^{1/(s-1)}\}} (1 + |y|^{s+k})^{-1} \, dy \leqslant c_3 \varepsilon, \tag{1.21}$$

which holds for some positive constants c_2, c_3 and for all sufficiently small $\varepsilon > 0$. Similarly, one has

$$\int_{A_{\varepsilon 2}} (1 + |z|^{s+k})^{-1} \, dz \leqslant c_4 \varepsilon \tag{1.22}$$

for some positive constant c_4 and all sufficiently small $\varepsilon > 0$. Also,

$$\int_{\{|z| > 1/\varepsilon^{1/(s-1)}\}} (1+|z|^{s+k})^{-1}\, dz = \omega_k \int_{1/\varepsilon^{1/(s-1)}}^{\infty} x^{k-1}(1+x^{s+k})^{-1}\, dx$$

$$\leqslant \omega_k \int_{1/\varepsilon^{1/(s-1)}} \frac{1}{x^{s+1}}\, dx \leqslant c_5 \varepsilon, \qquad [0 < \varepsilon < 1], \tag{1.23}$$

where ω_k, c_5 are suitable positive constants.

Combining (1.16)–(1.23) one gets, with $\varepsilon = n^{-1/2}$,

$$\eta_n = o(n^{-(s-2)/2}), \tag{1.24}$$

uniformly for all u satisfying $|u| < 2\,|l|/\varepsilon^{1/(s-1)}$. For $u \geqslant 2\,|l|/\varepsilon^{1/(s-1)}$, $A_{\varepsilon 1}$ is empty for all sufficiently small ε (see (1.20)). For $u \leqslant -2\,|l|/\varepsilon^{1/(s-1)}$,

$$\int_{A_{\varepsilon 1}} (1+|z|^{s+k})^{-1}\, dz$$

$$\leqslant \int_{\{y^{(1)} \leqslant -2/\varepsilon^{1/(s-1)}\}} (1+|y|^{s+k})^{-1}\, dy$$

$$\leqslant c_6 \int_0^{\infty} r^{k-2} \left\{ \int_{\{y^{(1)} \leqslant -2/\varepsilon^{1/(s-1)}\}} (|y^{(1)}|+r)^{-s-k}\, dy^{(1)} \right\} dr$$

$$= \frac{c_6}{s+k-1} \int_0^{\infty} r^{k-2} \left(\frac{2}{\varepsilon^{1/(s-1)}} + r \right)^{-s-k+1} dr$$

$$\leqslant \frac{c_6}{s+k} \int_{2/\varepsilon^{1/(s-1)}}^{\infty} v^{-s-1}\, dv \leqslant c_7 \varepsilon, \tag{1.25}$$

for appropriate constants c_6, c_7. Similarly, one shows that

$$\int_{A_{\varepsilon 2}} (1+|z|^{s+k})^{-1}\, dz = O(\varepsilon) \qquad \text{as} \quad \varepsilon \downarrow 0, \tag{1.26}$$

in case $u \leqslant -2\,|l|/\varepsilon^{1/(s-1)}$. In exactly the same manner one shows that for $u \geqslant 2\,|l|/\varepsilon^{1/(s-1)}$, the integrals of $(1+|z|)^{-s-k}$ over $A_{\varepsilon 1}$ and $A_{\varepsilon 2}$ are $O(\varepsilon)$. Hence (1.24) holds uniformly for *all* u. Now use (1.24), (1.13)–(1.15) to get

$$\sup_{u \in \mathbb{R}^1} \Bigg| \operatorname{Prob}(W_n' \leqslant u) - \Bigg[\int_{\{z \in \mathbb{R}^k : \sum l_i z^{(i)} \leqslant u\}} {}^1\xi_{s,n}(z)\, dz$$

$$+ \int_{\{h(z,\varepsilon) \leqslant u\} \setminus \{h(z,0) \leqslant u\}} \xi_{s-1,n}(z)\, dz$$

$$- \int_{\{h(z,0) \leqslant u\} \setminus \{h(z,\varepsilon) \leqslant u\}} \xi_{s-1,n}(z)\, dz \Bigg] \Bigg| = o(n^{-(s-2)/2}). \tag{1.27}$$

The reduction of the above integrals is now carried out exactly as in Bhattacharya and Ghosh [2] to yield

$$\sup_{u \in \mathbb{R}^1} \left| \operatorname{Prob}(W_n' \leqslant u) - \int_{-\infty}^{u} \psi_{s,n}(x)\, dx \right| = o(n^{-(s-2)/2}). \tag{1.28}$$

Finally note that there exists a constant c_8 such that

$$|W_n - W_n'| \leqslant C_8 n^{-s/2} / \sqrt{n}\,(\bar{Z} - \mu)|^{s+1}. \tag{1.29}$$

Now, by Corollary 17.12 in Bhattacharya and Ranga Rao [3] one has, for every $\varepsilon > 0$,

$$\begin{aligned} \operatorname{Prob}(\sqrt{n}\,|\bar{Z} - \mu| > \varepsilon n^{1/(s+1)}) &= o(n^{-(s-3)/2} n^{-(s-1)/(s+1)}) \\ &= o(n^{-(s-2)/2}) \qquad (s \geqslant 3). \end{aligned} \tag{1.30}$$

Since $\psi_{s,n}$ is bounded (uniformly in n), (1.28)–(1.30) imply (1.8). ∎

Remark 1. The proof essentially shows that one may replace the assumption (B_3') by (B_3''): $E|Z_1^{(i)}|^{s-r} < \infty$ *for all i which appear in the expression* (1.9) *for the first time in the sum* $n^{-r/2} \sum l_{i_1, \ldots, i_{r+1}} \sqrt{n}\,(\bar{Z}^{(i_1)} - \mu^{(i_1)}) \cdots \sqrt{n}\,(\bar{Z}^{(i_{r+1})} - \mu^{(i_{r+1})})$ $(0 \leqslant r \leqslant s-2)$.

Remark 2. The proof goes over to the case of vector-valued statistics $\sqrt{n}\,(H(\bar{Z}) - H(\mu))$ (or, more generally, vector-valued statistics which may be adequately approximated, coordinate wise, in the form (1.9)).

Remark 3. In Bhattacharya and Ghosh [2], (also see Bhattacharya [1]) it is proved under the assumptions (B_1)–(B_4) that

$$\sup_{B} \left| \operatorname{Prob}(W_n \in B) - \int_B \psi_{s,n}(x)\, dx \right| = o(n^{-(s-2)/2}), \tag{1.31}$$

where the supremum is over the class of *all Borel subsets B of* $\mathbb{R}^1$. Our proof above, under the moment relaxation (B_3') (or (B_3'')), only provides an approximation of the *distribution function*. Although this proof may be extended to carry over to the case of probabilities of sets with *smooth boundaries* (e.g., Borel measurable convex sets), it does not yield (1.31). We do not know if (1.31) is valid under the hupothesis of the present theorem. (Of course, (1.31) holds in this case if the right side is replaced by $o(n^{-(s-3)/2})$.)

Remark 4. An entirely analogous result holds for statistics $H(\bar{Z})$ for which $l_i = 0$ for all i, while $l_{i_1, i_2} \neq 0$ for some i_1, i_2. Thus for statistics $n(H(\bar{Z}) - H(\mu))$ arising in testing statistical hypotheses (See Chandra and Ghosh [4]) moment conditions may be relaxed for those coordinates which do not appear in the principal term of the Taylor expansion around μ.

Remark 5 (Conditioning argument). We write $Z_j' = (Z_j^{(1)}, ..., Z_j^{(k_1)})$, $Z_j'' = (Z_j^{(k_1+1)}, ..., Z_j^{(k)})$, $EZ_j' = \mu'$, $EZ_j'' = \mu''$. Under (B_4), $(\sum_1^k Z_j', \sum_1^k Z_j'')$ has a joint density and, therefore, $\sum_1^k Z_j'$ has a conditional density given $\sum_1^k Z_j''$. Dividing up $\sum_1^n Z_j'$, $\sum_1^n Z_j''$ into consecutive blocks of k summands each, one may first obtain an asymptotic expansion of the conditional distribution of the first sum (centered around its conditional expectation) given block sums of Z_j''. The successive block sums of Z_j' are still *independent* under this conditioning, but *not identically distributed*. However, by restricting $\bar{Z}''$ close to μ'' (the complementary event having small probability), one may often justify an asmptotic expansion of the above conditional distribution (see, e.g., Bhattacharya and Ranga Rao [3, Theorem (9.3)]). Under this conditioning regard $H(\bar{Z})$ as a function of $\bar{Z}'$ with (block sums of) Z_j'' as parameters, center $H(\bar{Z})$ around its conditional expectation, rewrite $\sqrt{n}\,(H(\bar{Z}) - H(\mu))$ in terms of this new centering, and proceed as in Bhattacharya and Ghosh [2] to obtain an asymptotic expansion of its conditional distribution. Finally expand the expectation of this expansion, this time dealing with (sample) means of i.i.d. summands. Such a procedure sometimes also succeeds in relaxing moment conditions. See Hall [7] for a similar procedure applied to the *Student's statistic*. Clearly, for the expansion of the conditional distribution of the statistic up to an error $o(n^{-(s-2)/2})$ one only needs $E\,|Z_j'|^s < \infty$, together with an appropriate moment condition on Z_j'' to ensure that $\bar{Z}''$ remains sufficiently close to μ'' with probability $1 - o(n^{-(s-2)/2})$. However, higher moments may be needed in carrying out the expansion of the expectation of the conditional expansion mentioned above. See Example 2 in Section 2 for an additional comment on this.

2. Examples

Example 1 (Hall [6]). Let Y_j $(j = 1, 2, ...)$ be a sequence of i.i.d. radom variables having zero mean, unit standard deviation and a nonzero third moment μ_3, say $\mu_3 > 0$. One may expect that the $100(1-\alpha)\%$ point of the distribution of $\sqrt{n}\,\bar{Y} = (Y_1 + \cdots + Y_n)/n^{1/2}$ is better approximated (than the $100(1-\alpha)\%$ point $z = z(\alpha)$ of the standard normal) by that of the normalized chisquare χ_N^2 having N degrees of freedom, where N is chosen so that the third moment (namely, $(8/N)^{1/2}$) of $T_N \equiv (2N)^{-1/2}(\chi_N^2 - N)$ equals that of $\sqrt{n}\,\bar{Y}$ (namely, $\mu_3/n^{1/2}$); i.e.,

$$N = 8n/\mu_3^2. \tag{2.1}$$

One may use the gamma tables to find $z_N = z_N(\alpha)$ such that

$$\text{Prob}(T_N \leqslant z_N) = 1 - \alpha. \tag{2.2}$$

Hall [6] shows that z_N is indeed a better approximation of the $100(1-\alpha)\%$ point for $\sqrt{n}\,\bar{Y}$ than usual estimates, under Cramér's condition as well as in the lattice case. In case μ_3 is unknown, replace it by the sample third moment $\hat{\mu}_3$ and write

$$\hat{N} = 8n/\hat{\mu}_3^2. \tag{2.3}$$

Hall [6, Theorem 5] provides an asymptotic expansion of $\text{Prob}(\sqrt{n}\,\bar{Y} \leq z_{\hat{N}})$ up to order $o(n^{-1})$, uniformly for $\alpha \in [\varepsilon, 1-\varepsilon]$ for every $\varepsilon > 0$, under the assumptions (i) $EY_1^6 < \infty$ and (ii) (Y_1, Y_1^3) satisfies Cramér's condition. He correctly points out that this expansion may be derived from Bhattacharya and Ghosh [2] only if (i) is strengthened to (i)′ $EY_1^{12} < \infty$. Let us show that our present results may be used to derive Hall's expansion under the conditions (i) $EY_1^6 < \infty$ and (ii)″ (B_4) *holds with* $m=1$, $k=2$; $f_1(y)=y$, $f_2(y)=y^3$.

By Lemma 1 of Hall [6], obtained by equating the asymptotic expansion of $\text{Prob}(T_N \leq y)$ with $1-\alpha$, one has

$$z_N = z + N^{-1/2}P_1(z) + N^{-1}P_2(z) + o(N^{-1}), \tag{2.4}$$

uniformly for $\alpha \in [\varepsilon, 1-\varepsilon]$ (for every fixed positive ε). Here P_1, f_2 are polynomials. Thus it is enough to expand $\text{Prob}(\sqrt{n}\,\bar{Y} \leq z')$, where

$$\begin{aligned} z' &= z + \hat{N}^{-1/2}P_1(z) + \hat{N}^{-1}P_2(z) \\ &= z + \frac{\hat{\mu}_3}{\sqrt{8n}}P_1(z) + \frac{\hat{\mu}_3^2}{8n}P_2(z) \\ &= z + \frac{\mu_3 P_1(z)}{\sqrt{8n}} + \frac{\mu_3^2 P_2(z)}{8n} \\ &\quad + n^{-1}\left\{\sqrt{n}\,(\hat{\mu}_3 - \mu_3)\left(\frac{P_1(z)}{\sqrt{8}} + \frac{2\mu_3}{8n^{1/2}}P_2(z)\right)\right\} \\ &\quad + n^{-2}(\sqrt{n}\,(\hat{\mu}_3 - \mu_3))^2\frac{P_2(z)}{8}. \end{aligned} \tag{2.5}$$

Expressing $\sqrt{n}\,\bar{Y} \leq z'$ in the form (1.9), one may now apply Remark 1 with $s=4$. Note that $\sqrt{n}\,(\bar{Z}^{(2)} - \mu^{(2)}) = \sqrt{n}\,(\hat{\mu}_3 - \mu_3)$ appears the first time with coefficient n^{-1}, so that (B_3'') becomes

$$EY_1^4 < \infty, \qquad E\,|Y_1^3|^2 \equiv EY_1^6 < \infty. \tag{2.6}$$

We have taken $\hat{\mu}_3 = n^{-1}\sum_{j=1}^n Y_j^3$ above. One may modify the calculations a little in case $\hat{\mu}_3 = n^{-1}\sum_{j=1}^n (Y_j - \bar{Y})^3$, to prove that (2.6) suffices along with (B_4) (with $k=3$, $f_i(y) = y^i$ for $i = 1, 2, 3$).

The expansion of Prob$(\sqrt{n}\,\bar{Y} \leq z')$ in terms up to order n^{-1} involves EY_1^4 (see Hall [6, p. 1032]). It may be shown by complicated algebra that the coefficient of $n^{-3/2}$ in the formal expansion involves EY_1^6. Also, looking at (2.5) one would not expect a valid asymptotic expansion with error $o(n^{-1})$ unless $\sqrt{n}(\hat{\mu}_3 - \mu_3)$ converges in distribution. Thus it is unlikely that the desired expansion holds in general under the condition $E|Y_1|^r < \infty$ for some $r < 6$.

EXAMPLE 2 (Studentized statistics). Consider the Student's statistic $t = \bar{Y}/\hat{\sigma}$, where $\hat{\sigma}^2 = (1/n)\sum_{j=1}^{n} Y_j^2 - \bar{Y}^2$. Here $m = 1$, $k = 2$; $Z_j^{(1)} = Y_j$, $Z_j^{(2)} = Y_j^2$, $EY_j = 0$. According to the theorem in Section 1, under (B_4) the distribution of $n^{1/2}t$ has an asymptotic expansion with error $o(n^{-(s-2)/2})$ if

$$EY_j^{2(s-1)} < \infty, \tag{2.7}$$

instead of the earlier requirement: $EY_j^{2s} < \infty$. Thus for an error $o(n^{-1/2})$ one needs finite fourth moments. By a conditioning argument, similar to the one sketched in Remark 5, Hall [7] proves that for an error $o(n^{-1/2})$, $E|Y_1^3| < \infty$ is enough. He also shows that for a higher order expansion of the conditional distribution of t, given $\{Y_j^2, 1 \leq j \leq n\}$, $E|Y_j|^s < \infty$ suffices; but we are unable to obtain the appropriate expansion of the expectation of the conditional expansion under this moment condition.

Consider now the asmptotic expansion of the Studentized sample moment $\hat{\mu}_r = n^{-1}\sum_{j=1}^{n} Y_j^r$ (r is a positive integer). The studentized statistic is $T = (\hat{\mu}_r - \mu_r)/\hat{\sigma}_r$, where $\hat{\sigma}_r^2$ is obtained by replacing population moments by sample moments in the expression var$(\hat{\mu}_r)$ calculated at least approximately keeping the principal terms (i.e., terms of order n^{-1}). For an expansion with an error term $o(n^{-(s-2)/2})$, the theorem in Section 1 requires $E|Y_j|^{2r(s-1)} < \infty$ instead of the older moment condition $E|Y_j|^{2rs} < \infty$.

ACKNOWLEDGMENTS

Professor W. van Zwet has kindly brought to our attention the articles by Chibishov [5] in which moment conditions are relaxed much further for *polynomial statistics*. It is not clear to us if Chibishov's results lead in general to better moment conditions for nonpolynomial statistics. Also our method is different and much simpler than that of Chibishov.

REFERENCES

[1] BHATTACHARYA, R. N. (1985). Some recent results on Cramér–Edgeworth expansions with applications. *Multivariate Analysis-VI* (P. R. Krishnaiah, Ed.) pp. 57–75. Elsevier, New York.

[2] BHATTACHARYA, R. N., AND GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* 6 434–451.

[3] BHATTACHARYA, R. N., AND RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. Krieger, Melbourne, Fl.

[4] CHANDRA, T. K., AND GHOSH, J. K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables. *Sankhyā Ser. A* 41 22–47.

[5] CHIVISHOV, D. M. (1980, 1981). An asymptotic expansion for the distribution of a statistic admitting a stochastic expansion I, II. *Theory Probab. Appl.* 15 732–744; 16 1–12.

[6] HALL, P. (1983). Chi squared approximations to the distribution of a sum of independent random variables. *Ann. Probab.* 11 1028–1036.

[7] HALL, P. (1987). Edgeworth expansion for student's t statistic under minimal moment conditions. *Ann. Probab.* 15 920–931.