Edgeworth expansions for statistics which are functions of lattice and non-lattice variables

Gutti Jogesh Babu
Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA

Received March 1990
Revised August 1990

Abstract: Let \( G \) be a distribution function on \( \mathbb{R}^{k+1} \) such that the \((k+1)\)th marginal is lattice. Let \( \overline{Z}_n \) denote the sample mean of \( n \) independent observations from \( G \). For \( s \geq 3 \), the \( s \)-term Edgeworth expansions are obtained for a wide class of statistics which are smooth functions of \( \overline{Z}_n \). The result is then applied to a statistic similar to the Student’s \( t \)-statistic, where the scaling factor, the sample standard deviation is replaced by the more robust mean absolute deviation.

AMS 1980 Subject Classification: 62B20

Keywords: Cramér’s condition, mean absolute deviation, lattice distributions, Edgeworth expansions.

1. Introduction

Consider a \( k + 1 \) variate random vector \( Z = (X, Y) \), where \( X \) is a \( k \)-variate random vector and \( Y \) is a univariate random variable. Let \( Z_i = (X_i, Y_i) \), \( i = 1, \ldots, n \), be \( n \) independent copies of \( Z \). Let \( E(Z) = \nu \) and let \( \overline{Z}_n = (\overline{X}_n, \overline{Y}_n) \) denote the sample mean. Under certain assumptions on the characteristic function of \( Z \), the Edgeworth expansions (EE) for \( P(\sqrt{n} (\overline{Z}_n - \nu) \in A) \) for a fairly rich class of Borel measurable sets \( A \) in \( \mathbb{R}^{k+1} \) are well known. See Bhattacharya and Ranga Rao (1986). When \( Y \) has a lattice distribution, Babu and Singh (1989) have shown that the two term EE for \( T_n = \sqrt{n} (\overline{Z}_n - \nu) + (0, \eta)n^{-3/8} \) is the same as the usual two term expansion in the pure non-lattice case, where \( \eta \) is a univariate random variable having a smooth distribution. As a consequence, if a statistic is a smooth function of \( \overline{Z}_n \) and the lattice part \( \overline{Y}_n \) does not appear in the linear approximation of the function, then the two term formal EE for the statistic is valid. This result on \( T_n \) cannot be extended to obtain more than two terms in the regular EE. However, a modified \( s \)-term EE can be obtained for any \( s \geq 3 \), when enough moments are assumed. Even though this expansion is usually different from the standard expansions, the difference is visible only in the \((k+1)\)th variable. If \( H \) is a smooth function and if \( H \) is four times continuously differentiable, then the formal three term EE for \( \sqrt{n} (H(\overline{Z}_n) - H(\nu)) \) is valid if the \((k+1)\)th variable \( \overline{Y}_n \) does not appear in the linear as well as the quadratic term of the Taylor series expansion of \( H \). This result is applied to the statistic

\[
H_n = \sqrt{n} \left( \overline{W}_n - \mu \right) / M_n
\]

(1.1)

Research supported in part by NSA Grant MDA-904-90-H-1001 and by NSF Grant DMS-9007717.
where \( W_1, \ldots, W_n \) are i.i.d. random variables from a continuous population with mean \( \mu \), \( \overline{W}_n = (1/n) \sum_{i=1}^{n} W_i \) and

\[
M_n = \frac{1}{n} \sum_{i=1}^{n} |W_i - \overline{W}_n|
\]

is the mean absolute deviation.

In Section 2 it is shown that \( H_n \) can be expressed as a smooth function \( H \) of the mean of \( (W_i - \mu, |W_i - \mu|, I(W_i \leq \mu)) \) plus a negligible term. Further the lattice variables \( I(W_i \leq \mu) \) do not appear in the first two terms of the Taylor series approximation of \( H \). If the distribution of \( W_1 \) has density in a neighborhood of \( \mu \), then it is easy to see that the distribution of \( (W_1 - \mu, |W_1 - \mu|) \) satisfies Cramér's condition. Theorem 2 below is applicable and the formal three term EE for \( H_n \) is valid. The details are given in Section 3.

2. Edgeworth expansions

Let \( Z = (X, Y) \) be a random vector in \( \mathbb{R}^{k+1} \). Let \( \nu = E(Z) \) and let the dispersion \( \Sigma \) of \( Z \) be positive definite. Let \( Z_i = (X_i, Y_i) \) be independent copies of \( Z \). Edgeworth expansions for the distribution of \( \sqrt{n} Z_n \) are well known under Cramér's condition. See Theorem 20.1 of Bhattacharya and Ranga Rao (1986). Let \( E \|Z\|^s < \infty \) for some \( s \geq 3 \). Let \( \xi_{n,s} \) denote the formal \((s-1)\)-term EE of the distribution of \( \sqrt{n} (Z_n - \nu) \). Let \( \phi_\Sigma \) denote the density of a centered normal vector with positive definite dispersion \( \Sigma \). Typically

\[
\xi_{n,s}(x) = \left(1 + \sum_{r=1}^{s-2} n^{-r/2} p_r(x)\right) \phi_\Sigma(x),
\]

where \( p_r \) is a polynomial of degree \( r + 2 \), whose coefficients are functions of moments of \( Z \) of orders \( \alpha = (\alpha_1, \ldots, \alpha_{k+1}) \) with \( \alpha_1 + \cdots + \alpha_{k+1} \leq r + 2 \) and \( \alpha_i \geq 0 \).

To state the results of this section we need some additional notation. For any bounded function \( f \), \( x \in \mathbb{R}^{k+1} \), and \( \delta > 0 \), let

\[
\omega(f, \delta, x) = \sup \{|f(z) - f(x)|: \|x - y\| \leq \delta\}
\]

and

\[
\bar{\omega}(f, \delta, \Sigma) = \int \omega(f, \delta, x) \phi_\Sigma(x) \, dx.
\]

Let \( \eta \) be a symmetric random variable with

\[
E |\eta|^{3(s-1)} < \infty,
\]

whose characteristic function vanishes outside a compact interval. Further we assume that \( \eta \) is independent of the sequence \( \{Z_i\} \). Let

\[
\xi_{n,s}(u, \nu; \eta) = E(\xi_{n,s}(u, \nu - \eta n^{-3/8}))
\]

Finally let \( Q_n \) denote the distribution of

\[
(\sqrt{n} \overline{X}_n, \sqrt{n} \overline{Y}_n + \eta n^{-3/8}) - \sqrt{n} \nu.
\]

**Theorem 2.1.** Suppose that the distribution of \( X \) satisfies Cramér’s condition

\[
\lim_{\|\nu\| \to \infty} \sup |E(e^{\nu'X})| < 1,
\]

2
and \( Y \) has a lattice distribution. If \( E \| Z \|^s < \infty \) for some \( s \geq 3 \), then for any measurable \( f \) bounded by 1, we have
\[
\left| \int f \, dQ_n - \int f(x) \xi_n(x; \eta) \, dx \right| \leq c\bar{w}(f, 4 e^{-dn}, \Sigma) + o(n^{-(s-2)/2}),
\] (2.7)
where \( c \) and \( d \) are some positive constants independent of \( f \).

**Proof.** The proof essentially involves expansions of the derivatives of the characteristic function \( \varphi_n \) of \( Q_n \). See proof of Theorem 20.1 of Bhattacharya and Ranga Rao (1986). Let \( \gamma(v) = E(e^{iv\eta}) \) and \( \theta \) denote the characteristic function of \( Z - v \). Clearly,
\[
\varphi_n(u, v) = \theta \left( \frac{u}{\sqrt{n}} \log n, \frac{v}{\sqrt{n}} \right) \gamma\left( \frac{v}{\sqrt{n}} \right).
\]
The derivatives of \( \varphi_n \) can be obtained in the range \( |v| \leq \sqrt{n} / \log n \) and \( |u| \leq \epsilon \sqrt{n} \), for some fixed \( \epsilon > 0 \), as in the proof of Theorem 20.1 of Bhattacharya and Ranga Rao (1986). Under the conditions on \( X \) and \( Y \), it is clear that the dispersion \( \Sigma \) of \( Z \) is positive definite. Since the characteristic function \( \gamma \) of \( \eta \) is 0 outside a compact interval, the derivatives of \( \varphi_n(u, v) \) of orders \( \alpha = (\alpha_1, \ldots, \alpha_{k+1}) \) with \( \alpha_i \geq 0 \) and
\[
\alpha_1 + \cdots + \alpha_{k+1} \leq s
\]
vanish for \( |v| > \sqrt{n} / \log n \). Since
\[
\theta \left( \frac{u}{\sqrt{n}} \log n, \frac{v}{\sqrt{n}} \right) \leq \theta \left( \frac{u}{\sqrt{n}} \log n, \frac{v}{\sqrt{n}} \right) \gamma\left( \frac{v}{\sqrt{n}} \right)
\]
and for any \( \epsilon > 0 \),
\[
\rho_\epsilon = \sup \left\{ \left| E(e^{iv\eta}) \right| : |u| > \epsilon \right\} < 1,
\]
it follows that for all \( n > n_0(\epsilon) \),
\[
\sup \left\{ \left| \theta \left( \frac{u}{\sqrt{n}} \log n, \frac{v}{\sqrt{n}} \right) \right| : |u| > \epsilon \sqrt{n}, \ |v| \leq \sqrt{n} / \log n \right\} \leq \rho_\epsilon + (\log n)^{-1} < \delta_n < 1.
\]
Consequently, it follows that the derivatives of \( \varphi_n(u, v) \) are exponentially decaying in the range \( |v| \leq \sqrt{n} / \log n \) and \( e\sqrt{n} \leq |u| \).

In view of (2.4) we have,
\[
P \left( | \eta | > n^{3/16} \right) \leq n^{-9(t-1)/16} E \left| \eta \right|^3(t-1) = O(n^{-(t-1)/2}).
\] (2.8)
Clearly for any \( m > 0 \),
\[
\int_{\|z\| > \log n} \left| \xi_{n,z}(z) \right| \, dz = O(n^{-m}).
\] (2.9)
Let \( g \) denote the density of \( \eta \). Now (2.8) and (2.9) yield, for any \( \epsilon > 0 \),
\[
\int [\omega(f, \epsilon, (x, y)) \left| \xi_{n,z}(x, y - wn^{-3/8}) \right| g(w) \, dw \, dx \, dy
\]
\[
\leq \int_{|w| \leq n^{-3/16}} \left( \| (x, y - wn^{-3/8}) \| \leq \log n \right) \omega(f, \epsilon, (x, y)) \left| \xi_{n,z}(x, y - wn^{-3/8}) \right| g(w) \, dw \, dx \, dy
\]
\[
+ O(n^{-(t-1)/2})
\]
\[
= O\left( \int_{\|z\| < 2 \log n} \omega(f, 2\epsilon, z) \varphi(z) \exp\left( \frac{1}{2} \left( \|z\Sigma^{-1}\|n^{-3/16}\right) \right) \, dz + O(n^{-(t-1)/2})
\]
\[
= O(n^{-(t-1)/2}) + O(\bar{w}(f, 2\epsilon, \Sigma)).
\] (2.10)

The rest of the proof is similar to that of Theorem 20.1 of Bhattacharya and Ranga Rao (1986). □
Using Theorem 2.1, we obtain a result similar to Theorem 2b of Bhattacharya and Ghosh (1978).

**Theorem 2.2.** Let $H$ be a real valued function on $\mathbb{R}^{k+1}$ and let all the derivatives of $H$ of orders $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$, with $\alpha_i \geq 0$ and $\alpha_1 + \cdots + \alpha_{k+1} \leq s$, are continuous in a neighborhood of $v$. Further we assume that the derivatives of $H$ of all orders $(\alpha_1, \ldots, \alpha_{k+1})$ with $\alpha_{k+1} \geq 1$ and $\sum_{i=1}^{k+1} \alpha_i = s - 2$ vanish at $v$.

Then under the conditions of Theorem 2.1, the formal $(s-1)$-term $EE$ is valid for $\sqrt{n}(H(Z_n) - H(v))$.

**Proof.** For $z = (z_1, \ldots, z_{k+1})$ and $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$, let

$$z^\alpha = \prod_{i=1}^{k+1} z_i^{\alpha_i}.$$  

Let

$$f_n(z) = \sqrt{n} \left( H(v + zn^{-1/2}) - H(v) \right).$$

The $(s-2)$-term Taylor series $T_s(z)$ of $f_n(z)$ does not depend on the $(k+1)$th coordinate $v$ of $z = (u, v)$. As a consequence, for any $|h| \leq n^{3/16}$,

$$f_n(u, v + hn^{-3/8}) = T_s(u, v) + n^{-(s-2)/2} \Sigma'(u, v + n^{-3/8}) l_\alpha + O(n^{-(s-1)/2} \log n)$$

$$= \xi_{n,s}(u, v) + n^{-(s-2)/2} \Sigma'(u, v) l_\alpha + o(n^{-(s-2)/2})$$

$$= f_n(u, v) + o(n^{-(s-2)/2}), \quad \text{(2.11)}$$

where $\Sigma'$ denotes sum over all $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$ satisfying $\sum_{i=1}^{k+1} \alpha_i = s - 1$ and $\alpha_i \geq 0$, and $l_\alpha$ denotes the $\alpha$th order derivative of $H$ at $v$. Observe that by (2.8), (2.9) (2.10) and (2.11),

$$\int_{f_n(u,v) \leq z} \xi_{n,s}(u, v; \eta) \, du \, dv = \int g(\omega) \left( \int_{f_n(u,v+wn^{-3/8}) \leq z} \xi_{n,s}(u, v) \, du \, dv \right) \, d\omega$$

$$= \int_{f_n(u,v) \leq z + o(n^{-(s-2)/2})} \xi_{n,s}(u, v) \, du \, dv + O \left( \int_{|q| > \log n} \phi_{\Sigma}(q) \, dq \right) + O \left( P \left( |\eta| > n^{3/16} \right) \right)$$

$$= \int_{f_n(q) \leq z} \xi_{n,s}(q) \, dq + \int_{|f_n(q) - z| = o(n^{-(s-2)/2})} \phi_{\Sigma}(q) \, dq + o(n^{-(s-2)/2}). \quad \text{(2.12)}$$

The rest of the proof is similar to that of Theorem 2b of Bhattacharya and Ghosh (1978). \[ \square \]

3. **An application**

Let $W_1, \ldots, W_n$ be independent random variables with a common continuous distribution $F$. Let $E(W_i) = \mu$. Instead of Student’s $t$-statistic consider the more robust statistic

$$H_n = \sqrt{n} \left( \bar{W}_n - \mu \right) / M_n, \quad \text{(3.1)}$$

where

$$M_n = \frac{1}{n} \sum_{i=1}^{n} |W_i - \bar{W}_n|. \quad \text{(3.2)}$$

Herrey (1965) derived the distribution of $H_n$ under normal assumptions using the independence of $\bar{W}_n$ and
If $F$ is not Gaussian, then $\bar{W}_n$ and $M_n$ are no longer independent. In this section we shall derive $EE$ for $H_n$ using Theorem 2.2.

**Theorem 3.1.** Suppose $F$ is continuously differentiable in a neighborhood of $\mu$ and the derivative $f$ of $F$ at $\mu$ is non-zero. Further assume that, for some $\beta > 0$, $c > 0$,

$$|f(\mu) - f(x)| \leq c |x - \mu|^{\beta}.$$

If $E(W_1^4) < \infty$, then the formal three term Edgeworth expansion for $H_n$ is valid. That is

$$n\left| P(H_n \leq x) - \Phi_{\sigma_2}(x) - \frac{\alpha}{\sqrt{n}} P_1(x/\sigma) \phi_{\sigma}(x) - \frac{\alpha^2}{n} P_2(x/\sigma) \phi_{\sigma}(x) \right| \to 0$$

uniformly in $x$, where

$$\alpha = (\text{Var} W_1)^{1/2}/(E |W_1 - \mu|),$$

and $\Phi_{\sigma}$ and $\phi_{\sigma}$ denote distribution and density of centered Gaussian variable with variance $\sigma^2$. Further $P_1$ and $P_2$ are the usual polynomials appearing in the three term $EE$. See (3.3) and (3.4).

**Remark.** The coefficients of the polynomials $P_1$ and $P_2$ can be obtained by computing the first four approximate cumulants of $H_n$. It can be shown that

$$P_1(x) = -k_1 + \frac{1}{6}k_3(1 - x^2) \quad \text{(3.3)}$$

and

$$P_2(x) = -\frac{1}{6}k_3x^2 + \frac{1}{6}(3x - x^3)k_4 + \frac{1}{6}(10x^3 - 15x - x^5)k_5^2, \quad \text{(3.4)}$$

where $Y_i$, $U_i$ and $V_i$ are defined in the proof of Theorem 3.1,

$$k_1 = -E(U_iY_i),$$

$$k_2 = 6E(Y_iV_i) + 6(E(U_iY_i))^2 - 2E(U_iY_i^2) + 3E(U_i^2),$$

$$k_3 = E(Y_i^3) - 6E(U_iY_i)$$

and

$$k_4 = E(Y_i^4) - 3 - 16E(U_iY_i)E(Y_i^3) - 12E(U_iY_i^2) + 24E(V_iY_i) + 12E(U_i^3) + 84(E(U_iY_i))^2.$$
where for some $K > 0$ and $\epsilon > 0$,
\[
P\left( |R_{n1}| > Kn^{-1-\epsilon} \right) = o(n^{-1}).
\]
Consequently,
\[
H_n = S_n + R_{n2},
\]
where
\[
S_n = \sqrt{n} \sigma \bar{Y}_n \left[ 1 - \bar{U}_n + \bar{V}_n \bar{\psi}_n + (\bar{U}_n)^2 \right],
\]
\[
U_i = \sigma \left( |Y_i| - \sigma^{-1} + (2F(\mu) - 1)Y_i \right), \quad V_i = \sigma \left( 2F(\mu) - 2I(Y_i \leq 0) - \tau F(\mu) Y_i \right).
\]
and for some $K > 0$ and $\epsilon > 0$,
\[
P\left( |R_{n2}| > Kn^{-1-\epsilon} \right) = o(n^{-1}).
\]
Clearly,
\[
S_n = \sqrt{n} \left( H(Z_n) - H(E(Z_1)) \right),
\]
where
\[
Z_i = \left( Y_i, |Y_i| - \sigma^{-1}, I(Y_i \leq 0) - F(\mu) \right)
\]
and
\[
H(x, y, z) = \sigma x \left[ 1 - \sigma y - \sigma x(2F(\mu) - 1) - x(2z + \tau F(\mu)) + \sigma^2 \left( y + x(2F(\mu) - 1) \right)^2 \right].
\]
Note that
\[
E(Z_i) = (0, 0, 0), \quad H(0, 0, 0) = 0,
\]
and that $H$ is a smooth function satisfying the conditions of Theorem 2.2 with $s = 4$. As the distribution of $W_i$ is assumed to have a density in a neighbourhood of $\mu$, it follows that the distribution of $X_1 = (Y_1 / |Y_1|)$ satisfies the Cramér’s condition (2.6). So the formal three term EE is valid for the distribution of $H_n$. $\square$

Appendix

Let $F_n$ denote the empirical distribution of $W_1, \ldots, W_n$.

**Lemma.** Under the conditions of Theorem 3.1,
\[
M_n = \frac{1}{n} \sum_{i=1}^{n} |W_i - \mu| - 2(F(\mu) - F_n(\mu))(\bar{W}_n - \mu)
\]
\[
+ (\bar{W}_n - \mu)(2F(\mu) - 1) + (\bar{W}_n - \mu)^2 f(\mu) + O(|\bar{W}_n - \mu|^{2+\beta}) + (\bar{W}_n - \mu)R_n \tag{A1}
\]
and
\[
P\left( \sqrt{n} |\bar{W}_n - \mu| > \log n \right) = o(n^{-1}), \tag{A2}
\]
where for some $K > 0$,
\[
P\left( \sqrt{n} |\bar{W}_n - \mu| \leq \log n, |R_n| > K(\log n)n^{-3/4} \right) + O(n^{-2}).
\]
Proof. (A2) is an immediate consequence of Theorem 2 of Michel (1976). Also see Theorem 17.11 of Bhattacharya and Ranga Rao (1986) for a more general and sharper moderate deviation estimate. To prove (A1), we note that for any real \( a \) and \( x \),

\[
|x| - |x - a| = \int_0^a (1 - 2I(x \leq y)) \, dy.
\]

It follows that

\[
M_n - \frac{1}{n} \sum_{i=1}^n |W_i - \mu| = \int_{-\infty}^{\infty} (2F_n(\mu + y) - 1) \, dy
\]

\[
= (W_n - \mu)(2F_n(\mu) - 1) + 2\int_0^{(W_n - \mu)} (F_n(\mu + y) - F_n(\mu)) \, dy
\]

\[
= (W_n - \mu)(2F_n(\mu) - 1) + 2\int_0^{(W_n - \mu)} (F(\mu + y) - F(\mu)) \, dy
\]

\[
+ 2\int_0^{(W_n - \mu)} [F_n(\mu + y) - F(\mu + y) - F_n(\mu) + F(\mu)] \, dy. \quad (A3)
\]

Further

\[
2\int_0^{(W_n - \mu)} (F(\mu + y) - F(\mu)) \, dy = f(\mu)(W_n - \mu)^2 + O\left(\frac{1}{n}\right). \quad (A4)
\]

Using the arguments similar to Bahadur representation of quantiles (see Bahadur, 1966) we obtain for some \( K > 0 \),

\[
P\left( R'_n > K(\log n) n^{-3/4} \right) = o(n^{-1}) \quad (A5)
\]

where

\[
R'_n = \sup \left\{ \left| F_n(\mu + y) - F(\mu + y) - F_n(\mu) + F(\mu) \right| : |y| \sqrt{n} \leq \log n \right\}.
\]

It should be mentioned here that Bahadur (1966) obtained a sharper estimate under the assumption \( \beta = 1 \). For our purpose (A5) is enough, which can be proved assuming only \( \beta > 0 \). The representation (A1) now follows from (A3), (A4) and (A5). This completes the proof. \( \square \)

References


