The Annals of Probability
1987, Vol. 15, No. 3, 920–931

EDGEWORTH EXPANSION FOR STUDENT'S *t* STATISTIC UNDER MINIMAL MOMENT CONDITIONS

BY PETER HALL[1]

University of Glasgow

Recent advances in the theory of Edgeworth expansion allow rigorous estimation of remainder terms, under explicit regularity conditions. However, even in the important case of Student's t statistic, the moment conditions required are unrealistically stringent. In this note we relax those conditions to the minimum needed to define terms in the Edgeworth expansion.

1. Introduction and results. The fundamental work of Bhattacharya and Ghosh (1978) makes rigorous the classical formalism of Edgeworth expansion. When applied to the case of Student's t statistic, these results allow an expansion with a remainder $o(n^{-k/2})$, provided the sampling distribution has finite $2(k+2)$th order moment. This sharpens the earlier work of Chung (1946), among others, but still falls short of what a statistician would regard as "minimal" conditions for Edgeworth expansion. Notice that terms before the remainder $o(n^{-k/2})$ depend only on moments up to the $(k+2)$th, and so the assumption of $2(k+2)$th moments appears to be "twice as stringent" as necessary.

The purpose of this note is to show rigorously how to relax moment assumptions to statistically "minimal" conditions, in the case of Student's t statistic. As prerequisites for an expansion with remainder $o(n^{-k/2})$, we need only finiteness of $(k+2)$th order moments plus nonsingularity of the sampling distribution. Edgeworth expansions of Student's statistic have more than half a century of history [see reviews by Wallace (1958), Bowman, Beauchamp and Shenton (1977) and Cressie (1980)], but there appears to be no previous derivation under minimal conditions. This is surprising, given the very extensive literature on the distribution of Student's statistic under nonnormality.

We pause to introduce necessary notation. Let $X, X_1, X_2, \dots$ be independent and identically distributed random variables, and set $\bar{X} \equiv n^{-1}\sum_{i=1}^n X_i$ and $\mu_k \equiv E(X^k)$, whenever this quantity is finite. The Studentized mean is given by $T_0 \equiv n^{1/2}(\bar{X}-\mu)(n^{-1}\sum_{i=1}^n X_i^2 - \bar{X}^2)^{-1/2}$ where $\mu = \mu_1$. Let P_i be the polynomial of degree $3i-1$ appearing in the formal Edgeworth expansion of the distribution of T_0:

$$(1.1) \qquad P(T_0 \le y) = \Phi(y) + \sum_{i=1}^{k} n^{-i/2} P_i(y)\phi(y) + o(n^{-k/2}),$$

where Φ, ϕ are, respectively, the standard normal distribution, density functions.

Received June 1985; revised March 1986.

[1] On leave from Australian National University.

AMS 1980 *subject classifications*. Primary 60E05; secondary 60F05.

Key words and phrases. Edgeworth expansion, moment conditions, t statistic.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability*. www.jstor.org

The coefficients in P_i are functions of $\mu_1, \ldots, \mu_{i+2}$; for example,

$$P_1(y) = \tfrac{1}{6}\tau(2y^2 + 1),$$

$$P_2(y) = -y\{\tfrac{1}{18}\tau^2(y^4 + 2y^2 - 3) - \tfrac{1}{12}\kappa(y^2 - 3) + \tfrac{1}{4}(y^2 + 3)\},$$

where $\tau \equiv E(X - \mu)^3(\operatorname{var} X)^{-3/2}$, $\kappa \equiv E(X - \mu)^4(\operatorname{var} X)^{-2} - 3$.

THEOREM. *Assume $k \geq 1$, $E|X|^{k+2} < \infty$ and the distribution of X is nonsingular. Then expansion* (1.1) *is available uniformly in* $-\infty < y < \infty$.

In particular, if $E|X|^3 < \infty$ and X is nonsingular, then

$$P(T_0 \leq y) = \Phi(y) + n^{-1/2}\tfrac{1}{6}\tau(2y^2 + 1)\phi(y) + o(n^{-1/2}) \tag{1.2}$$

uniformly in y. The work of Callaert and Veraverbeke (1981) gives $P(T_0 \leq y) = \Phi(y) + O(n^{-1/2})$, under the condition $E|X|^{4.5} < \infty$ but without the assumption of nonsingularity.

There is no difficulty in extending our techniques to obtain expansions of $E\{f(T_0)\}$ for functions f. However, statistical implications are clearer in the form (1.1).

2. Proof. Define $\operatorname{sgn}(x) = 1$ if $x > 0$, -1 if $x < 0$ and 0 if $x = 0$, and given $y \in (-\infty, \infty)$, let $S(y) \equiv \operatorname{sgn}(X - y)$, $p(y, x) \equiv P\{S(y) = 1 \,|\, |X - y| = x\}$,

$$\begin{aligned} Y(y) &\equiv (X - y) - E(X - y \,|\, |X - y|) \\ &= |X - y|[S(y) - \{2p(y, |X - y|) - 1\}], \\ \psi(y, t) &\equiv E[\exp\{itY(y)\} \,|\, |X - y|] \\ &= \exp[-it|X - y|\{2p(y, |X - y|) - 1\}] \\ &\quad \times [p(y, |X - y|)e^{it|X - y|} + \{1 - p(y, |X - y|)\}e^{-it|X - y|}]. \end{aligned}$$

Lemmas 2.1 and 2.2 establish that for some y, $Y(y)$ conditional on $|X - y|$ has a nondegenerate distribution with useful properties. In the integrand of (2.1), interpret the ratio $0/0$ as 0 whenever it appears.

LEMMA 2.1. *Let f be a density on $(-\infty, \infty)$. The set of y such that*

$$\int_0^\infty f(x + y)f(-x + y)\{f(x + y) + f(-x + y)\}^{-1}\,dx > 0 \tag{2.1}$$

has strictly positive measure.

PROOF. Fix $a > 0$, and define $g(x) \equiv \min\{a, f(x)\}$. The integrand in (2.1) dominates $(2a)^{-1}g(x + y)g(-x + y)$, and so if the lemma is false,

$$\int_{-\infty}^\infty g(x + y)g(-x + y)\,dx = 0, \tag{2.2}$$

for almost all y. Now integrate (2.2) over $-\infty < y < \infty$, obtaining $\tfrac{1}{2}(\int g)^2 = 0$. This contradicts the hypothesis that f is a density. □

LEMMA 2.2. *If X is nonsingular, then for some real number y,*

$$\sup_{|t|>\varepsilon} E\{|\psi(y,t)|\} < 1, \quad \textit{for all } \varepsilon > 0. \tag{2.3}$$

PROOF. We may write $dP(X \le x) = qf(x)\,dx + (1-q)\,dG(x)$, where $q > 0$, f is a density and G is a distribution. Let δ, δ' be, respectively, the left-hand side of (2.3), and the value δ would take if X had density f. Then $\delta \le q\delta' + 1 - q$, and so it suffices to consider the case where X has density f.

Choose y such that (2.1) holds, and then let a equal the left-hand side of (2.1). Then $g(x) \equiv a^{-1}f(x+y)f(-x+y)\{f(x+y)+f(-x+y)\}^{-1}$ is a density on $(0,\infty)$. By the Riemann–Lebesgue lemma,

$$\beta(t) \equiv \int_0^\infty \{1 - \cos(2tx)\}g(x)\,dx \to 1$$

as $|t| \to \infty$, and trivially $\beta(t) > 0$ for $t \ne 0$. Therefore $b_\varepsilon \equiv \inf_{|t|>\varepsilon}\beta(t) > 0$. If X has density f then $|X - y|$ has density $f(x+y) + f(-x+y)$, $x > 0$. Thus,

$$\begin{aligned}
&[E\{|\psi(y,t)|\}]^2 \\
&\quad = \left(E\left[1 - 4p(y,|X-y|)\{1 - p(y,|X-y|)\}\sin^2\{t(X-y)\}\right]^{1/2}\right)^2 \\
&\quad \le 1 - 2\int_0^\infty f(x+y)f(-x+y)\{f(x+y)+f(-x+y)\}^{-1} \\
&\qquad\qquad \times \{1 - \cos(2tx)\}\,dx.
\end{aligned}$$

Consequently, $\sup_{|t|>\varepsilon}E\{|\psi(y,t)|\} \le (1 - 2ab_\varepsilon)^{1/2} < 1$, as required. □

One consequence of (2.3) is

$$E\{Y^2(y)\} = E[E\{Y^2(y)||X-y|\}] > 0. \tag{2.4}$$

We may assume that (2.3) *holds with* $y = 0$. Notice that we are now not permitted to assume X has zero mean. Let $\mathscr{F}$ be the σ-field generated by $|X_1|,\ldots,|X_n|$; $S_j \equiv \operatorname{sgn}(X_j)$, $p_j \equiv p(0,|X_j|)$,

$$\begin{gathered}
Y_j \equiv X_j - E(X_j||X_j|) = \{S_j - (2p_j - 1)\}|X_j|, \\
\beta_{2j} \equiv E(Y_j^2||X_j|) = 4p_j(1-p_j)X_j^2, \qquad \beta_{kj} \equiv E(Y_j^k||X_j|), \\
\nu_2 \equiv E(\beta_{21}) > 0 \quad [\text{see (2.4)}], \qquad \nu_k \equiv E(\beta_{k1}), \\
\psi_j(t) \equiv E\{\exp(itY_j)||X_j|\}, \qquad s^2 \equiv \sum_{j=1}^n \beta_{2j}, \qquad T \equiv \sum_{j=1}^n Y_j.
\end{gathered}$$

Let $C_1, C_2, \ldots$ be constants depending on the distribution of X but not on n or other quantities, and let $D_1, D_2, \ldots$ be absolute constants. Conditional on $\mathscr{F}$, the variables $s^{-1}Y_j$ are independent with zero means and variances $s^{-2}\beta_{2j}$. [This type of conditioning argument has been used in the past; see, for example, Albers, Bickel and van Zwet (1976) and Bickel and van Zwet (1978).] Our first task is to derive an Edgeworth expansion of the distribution of $s^{-1}T$. To more clearly illustrate our argument we shall concentrate on the case $k = 1$, and establish (1.2). Result (2.7) below indicates the route taken if $k = 2$. It will become clear that there are no essential differences between treatments for different values of k, and at various points in our proof we shall indicate generalisations for arbitrary k.

LEMMA 2.3. *Let $Z_1, \dots, Z_n$ be independent random variables with finite third moments, zero means and $\Sigma_{j=1}^n E(Z_j^2) = 1$. Set*

$$\chi_j(t) \equiv E(e^{itZ_j}) \quad \text{and} \quad \lambda_{kj}(t) \equiv E\left\{\exp(itZ_j) - \sum_{r=0}^{k+2} \frac{1}{r!}(itZ_j)^r\right\},$$

and choose l so large that

$$\sum_{j=1}^{n} E\{Z_j^2 I(|Z_j| > l)\} \le \tfrac{1}{8}. \tag{2.5}$$

Then whenever $|t| \le 1/12l$,

$$\begin{aligned}\left[\prod_{j=1}^{n}\chi_j(t) - \left\{1 + \sum_{j=1}^{n}\lambda_{1j}(t) + \tfrac{1}{6}(it)^3\sum_{j=1}^{n}E(Z_j^3)\right\}e^{-t^2/2}\right] \\ \le D_1|t|e^{-t^2/6}\left[\sum_{j=1}^{n}\{E(Z_j^2)\}^2 + \left\{\sum_{j=1}^{n}E(|Z_j|^3)\right\}^2\right].\end{aligned} \tag{2.6}$$

If each $E(Z_j^4) < \infty$ and if $|t| \le 1/12l$,

$$\begin{aligned}\Bigg[\prod_{j=1}^{n}\chi_j(t) - \Bigg(1 + \sum_{j=1}^{n}\lambda_{2j}(t) + \tfrac{1}{6}(it)^3\sum_{j=1}^{n}E(Z_j^3) \\ + (it)^4\sum_{j=1}^{n}\left[\tfrac{1}{24}E(Z_j^4) - \tfrac{1}{8}\{E(Z_j^2)\}^2\right] \\ + \tfrac{1}{72}(it)^6\left\{\sum_{j=1}^{n}E(Z_j^3)\right\}^2\Bigg)e^{-t^2/2}\Bigg] \\ \le D_2|t|e^{-t^2/6}\Bigg[\sum_{j=1}^{n}\{E(Z_j^2)\}^3 + \sum_{j=1}^{n}E(Z_j^2)E(|Z_j|^3) \\ + \left\{\sum_{j=1}^{n}E(|Z_j|^3)\right\}^2 + \left\{\sum_{j=1}^{n}E(Z_j^4)\right\}^2\Bigg].\end{aligned} \tag{2.7}$$

Inequality (2.6) may be derived from Lemma 2.1 of Hall (1982), and (2.7) may be proved in a similar manner.

Given $\delta \in (0, \frac{1}{2})$ and $\lambda > 0$, let $\mathscr{E}_1$ be the event:

$$|s^2 - n\nu_2| < \delta n\nu_2, \qquad \left|\sum_{j=1}^n X_j^2 - n\mu_2\right| \le \delta n\mu_2, \qquad |X_j| \le \delta n^{1/2}$$

for $1 \le j \le n$ and $\hat{l} \le \lambda s^{-1}$, where

$$\hat{l} \equiv \inf\left\{l > 0\colon s^{-2}\sum_{j=1}^n E\left[Y_j^2 I(|Y_j| > ls)\middle|\mathscr{F}\right] \le \tfrac{1}{8}\right\}.$$

We first prove that if $E|X|^3 < \infty$, then for large λ, $P(\tilde{\mathscr{E}}_1) = o(n^{-1/2})$. [In the case $E|X|^{k+2} < \infty$, $P(\tilde{\mathscr{E}}_1) = o(n^{-k/2})$, where $\tilde{\mathscr{E}}$ denotes the complement of $\mathscr{E}$.]

Use Theorem 28, page 286 of Petrov (1975), and $E|X|^3 < \infty$, to get

$$P\big(|s^2 - n\nu_2| > \delta n\nu_2\big) + P\left(\left|\sum_{j=1}^n X_j^2 - n\mu_2\right| > \delta n\mu_2\right) + nP\big(|X| > \delta n^{1/2}\big)$$
$$= o(n^{-1/2}),$$

for each $\delta > 0$. Let $\nu(\lambda) \equiv E\{Y_1^2 I(|Y_1| > \lambda)\}$, and choose λ so large that $\nu_2^{-1}\nu(\lambda) \le \frac{1}{24}$. Arguing as before, we see that the event

$$\mathscr{E}_2 \equiv \left\{|s^2 - n\nu_2| \le \tfrac{1}{2}n\nu_2\right\} \cap \left\{\left|\sum_{j=1}^n E\left[Y_j^2 I(|Y_j| > \lambda)\middle|\mathscr{F}\right] - n\nu(\lambda)\right| \le \tfrac{1}{2}n\nu(\lambda)\right\}$$

has probability $1 - o(n^{-1/2})$. On $\mathscr{E}_2$,

$$s^{-2}\sum_{j=1}^n E\left\{Y_j^2 I(|Y_j| > \lambda)\middle|\mathscr{F}\right\} \le (n\nu_2/2)^{-1}\{3n\nu(\lambda)/2\} \le \tfrac{1}{8},$$

and so $\hat{l} \le \lambda s^{-1}$. The estimate $P(\tilde{\mathscr{E}}_1) = o(n^{-1/2})$ follows from these results.

Take $Z_j \equiv Y_j/s$, conditional on $\mathscr{F}$, in (2.6). Let E' denote expectation conditional on $\mathscr{F}$. On $\mathscr{E}_1$

$$\sum_{j=1}^n \left\{E'(Y_j/s)^2\right\}^2 + \left\{\sum_{j=1}^n E'\big(|Y_j/s|^3\big)\right\}^2$$
$$\le s^{-4}\sum_{j=1}^n X_j^4 + 64s^{-6}\left(\sum_{j=1}^n |X_j|^3\right)^2$$
$$\le s^{-4}\delta n^{1/2}\sum_{j=1}^n |X_j|^3 + 64s^{-6}\delta n^{1/2}n\mu_2(1+\delta)\sum_{j=1}^n |X_j|^3$$
$$\le C_1\delta n^{-3/2}\sum_{j=1}^n |X_j|^3,$$

$$\left|\sum_{j=1}^n E'\big(Y_j/s\big)^3 - n^{-1/2}\nu_2^{-3/2}\nu_3\right| \le C_2 n^{-3/2}\left\{\left|\sum_{j=1}^n (\beta_{3j} - \nu_3)\right| + \delta\sum_{j=1}^n |X_j|^3\right\},$$

$$\sum_{j=1}^n \left|\lambda_{1j}(t)\right| \le D_3 s^{-4}\sum_{j=1}^n |tX_j|^4 \le C_3 t^4\delta n^{-3/2}\sum_{j=1}^n |X_j|^3.$$

Substituting into (2.6) we obtain on $\mathscr{E}_1$

$$\left|\prod_{j=1}^{n} \chi_j(t) - \left\{1 + \tfrac{1}{6}(it)^3 n^{-1/2} \nu_2^{-3/2} \nu_3\right\} e^{-t^2/2}\right|$$
$$\leq C_4 |t| e^{-t^2/6} n^{-3/2} \left\{\left|\sum_{j=1}^{n} (\beta_{3j} - \nu_3)\right| + \delta \sum_{j=1}^{n} |X_j|^3\right\}.$$

Now apply the smoothing lemma for characteristic functions [Petrov (1975), Theorem 2, page 109] with $T = n$, obtaining on $\mathscr{E}_1$,

$$\sup_{-\infty<y<\infty} \left|P(T \leq sy|\mathscr{F}) - \left\{\Phi(y) + \tfrac{1}{6} n^{-1/2} \nu_2^{-3/2} \nu_3 (1 - y^2) \phi(y)\right\}\right|$$
$$\leq C_5 n^{-3/2} \left\{\left|\sum_{j=1}^{n} (\beta_{3j} - \nu_3)\right| + \delta \sum_{j=1}^{n} |X_j|^3 + n^{1/2}\right\}$$
$$+ C_5 \int_{s/12\lambda}^{n} \left\{\prod_{j=1}^{n} |\psi_j(t/s)|\right\} dt.$$

This inequality continues to hold if y is replaced by any $\mathscr{F}$-measurable random variable U_y indexed by y. Use that inequality on $\mathscr{E}_1$ and the trivial upper bound 2 (for large n) on $\tilde{\mathscr{E}}_1$, and take expectations:

$$\sup_{-\infty<y<\infty} \left|P(T \leq sU_y) - E\left\{\Phi(U_y) + \frac{1}{6} n^{-1/2} \nu_2^{-3/2} \nu_3 (1 - U_y^2) \phi(U_y)\right\}\right|$$
$$\leq C_5 \delta E(|X|^3) n^{-1/2} + C_5 n^{1/2} \int_{1/12\lambda}^{C_6 n^{3/2}} \{E|\psi_1(t)|\}^n dt + o(n^{-1/2}).$$

Use Lemma 2.2, and the fact that δ is arbitrarily small, to get

$$\sup_{-\infty<y<\infty} \left|P(T \leq sU_y) - E\left\{\Phi(U_y) + \tfrac{1}{6} n^{-1/2} \nu_2^{-3/2} \nu_3 (1 - U_y^2) \phi(U_y)\right\}\right| = o(n^{-1/2}). \tag{2.8}$$

(For general k, this expansion is carried to terms in $n^{-k/2}$.)

Define

$$u_1 \equiv -n^{-1/2} \nu_2^{-1/2} \mu y^2 (1 + n^{-1} y^2)^{-1},$$
$$u_2 \equiv \nu_2^{-1/2} (1 + n^{-1} y^2)^{-1} y \left\{(\mu_2 - \mu^2) + n^{-1} \mu_2 y^2\right\}^{1/2},$$
$$U_1 \equiv (1 + n^{-1} y^2) \left\{(\mu_2 - \mu^2) + n^{-1} \mu_2 y^2\right\}^{-1} n^{-1} \sum_{i=1}^{n} (X_i^2 - \mu_2)$$

if $\Sigma X_i^2 > n\mu^2$, $U_1 = -1$ otherwise,

$$U_2 \equiv -\nu_2^{-1/2} n^{-1/2} \sum_{i=1}^{n} \{(2p_i - 1)|X_i| - \mu\},$$
$$U_3 \equiv \nu_2^{-1} n^{-1} (s^2 - n\nu_2),$$
$$U_y \equiv \{u_1 + u_2 (1 + U_1)^{1/2} + U_2\} (1 + U_3)^{-1/2}.$$

Note that $P(\Sigma X_i^2 \le n\mu^2) = o(n^{-1/2})$. If $\Sigma X_i^2 > n\mu^2$, then $T_0 \le y$ if and only if $T \le sU_y$ (see the Appendix). The remainder of the proof of (1.2) consists of substituting this formula for U_y into (2.8), and checking that

$$\sup_{|y|\le \log n} \Big| E\big\{\Phi(U_y) + \tfrac{1}{6}n^{-1/2}\nu_2^{-3/2}\nu_3\big(1-U_y^2\big)\phi(U_y)\big\} - \big\{\Phi(y) + \tfrac{1}{6}n^{-1/2}\tau(2y^2+1)\phi(y)\big\}\Big| = o(n^{-1/2}). \tag{2.9}$$

[The case $|y| > \log n$ does not require attention; for example, if $y < -\log n$, then an Edgeworth expansion of $P\{n^{1/2}(\bar{X}-\mu) \le x\}$ shows that both sides of (1.2) are $o(n^{-1/2})$. If $E|X|^{k+2} < \infty$ and $y < -\log n$, both sides of (1.1) are $o(n^{-k/2})$.]

To make our proof of (2.9) easier to follow, we decompose it into five distinct steps. Many details are omitted. We assume the distribution of $(2p_1 - 1)|X_1|$ is nondegenerate; otherwise the proof is very easy.

Step (i). *Elimination of* U_3 *from* U_y. Let $\mathscr{E}_3$ be the event $|s^2 - n\nu_2| \le \delta n\nu_2$, where $\delta \in (0, \frac{1}{100})$, let $U_{y1} \equiv u_1 + u_2(1+U_1)^{1/2} + U_2$, and take $\psi(y)$ to be either $\Phi(y)$ or $(1-y^2)\phi(y)$. On $\mathscr{E}_3$, $|(1+U_3)^{-1/2} - (1-\frac{1}{2}U_3)| \le U_3^2$, and also $P(\tilde{\mathscr{E}}_3) = o(n^{-1/2})$. By treating separately the expectation on and off $\mathscr{E}_3$, we obtain

$$\sup_{|y|\le \log n} \Big| E\big\{\psi(U_y) - \psi(U_{y1}) + \tfrac{1}{2}U_3U_{y1}\psi'(U_{y1})\big\}\Big| \le C_7\delta^{1/2}n^{-1/2} + o(n^{-1/2}).$$

Since δ may be taken arbitrarily small,

$$\sup_{|y|\le \log n} \Big| E\big\{\psi(U_y) - \psi(U_{y1}) + \tfrac{1}{2}U_3U_{y1}\psi'(U_{y1})\big\}\Big| = o(n^{-1/2}). \tag{2.10}$$

Step (ii). *Expansion of* $E\{\psi(U_{y1})\}$. Let $X_i^* \equiv X_iI(|X_i| \le \delta n^{1/2})$, let $\mathscr{E}_4$ be the event $|\Sigma_{i=1}^n X_i^{*2} - n\mu_2| \le \delta n\mu_2$, define V_1, V_2, V_{y1} by replacing X_i by X_i^* in the definitions of U_1, U_2, U_{y1}, respectively, and set $V_{y2} \equiv u_1 + u_2 + V_2$. Then

$$\begin{aligned} &\Big|E\big\{\psi(V_{y1}) - \psi(V_{y2}) - \tfrac{1}{2}u_2V_1\psi'(V_{y2})\big\}\Big| \\ &\quad\le C_8\Big[y^2E\big\{V_1^2\big|\psi''(V_{y2}+yRV_1)\big|I(\mathscr{E}_4)\big\} + P(\tilde{\mathscr{E}}_4) \\ &\qquad + |y|E\big\{\big|V_1\psi'(V_{y2})\big|I(\tilde{\mathscr{E}}_4)\big\}\Big] \equiv t, \end{aligned} \tag{2.11}$$

where R is a random function of y satisfying $|R| \le C_9$ on $\mathscr{E}_4$. Choose $\delta_0 \in (0, \frac{1}{100})$ and $\varepsilon_1 > 0$ so small that whenever $\delta \in (0, \delta_0)$, we have $|V_{y2} + yRV_1| > \varepsilon_1|y|$ on the event $\mathscr{E}_4 \cap \{|V_2| \le \varepsilon_1|y|\}$. Then

$$\begin{aligned} t \le C_{10}\bigg(&\Big\{y^2 \sup_{|z|>\varepsilon_1|y|}|\psi''(z)|\Big\}E\big(V_1^2\big) + (\sup|\psi''|)y^2E\big\{V_1^2I(|V_2| > \varepsilon_1|y|)\big\} \\ &+ P\big(\tilde{\mathscr{E}}_4\big) + \Big[y^2E\big\{V_1^2\psi'(V_{y2})^2\big\}P\big(\tilde{\mathscr{E}}_4\big)\Big]^{1/2}\bigg). \end{aligned}$$

Choose $\varepsilon_2 > 0$ so small that $|V_2| \le \varepsilon_2 |y|$ implies $|V_{y2}| > \varepsilon_2 |y|$. Then

$$y^2 E\{V_1^2 \psi'(V_{y2})^2\} \le \Big\{y^2 \sup_{|z| > \varepsilon_2 |y|} \psi'(z)^2\Big\} E(V_1^2)$$
$$+ (\sup|\psi'|)^2 y^2 E\{V_1^2 I(|V_2| > \varepsilon_2 |y|)\}.$$

It is easily checked that $P(\tilde{\mathscr{E}}_4) = o(n^{-1/2})$ and $E(V_1^2) \le C_{11}\delta n^{-1/2} + o(n^{-1/2})$ uniformly in $|y| \le \log n$. Therefore if we prove that for each $\varepsilon > 0$,

$$\sup_{|y| \le \log n} y^2 E\{V_1^2 I(|V_2| > \varepsilon |y|)\} \le C_{12}(\varepsilon)\delta n^{-1/2} + o(n^{-1/2}), \tag{2.12}$$

then it will follow from the estimates below (2.11) and the result

$$\sup_{-\infty < y < \infty} \big|E\{\psi(U_{y1}) - \psi(V_{y1})\}\big| \le 2(\sup|\psi|) n P(|X| > \delta n^{1/2}) = o(n^{-1/2}),$$

that

$$\sup_{|y| \le \log n} \big|E\{\psi(U_{y1}) - \psi(V_{y2}) - \tfrac{1}{2} u_2 V_1 \psi'(V_{y2})\}\big| \le C_{13}\delta n^{-1/2} + o(n^{-1/2}). \tag{2.13}$$

Step (iii). *Proof of* (2.12). Notice that for $|y| \le \log n$,

$$V_1^2 \le C_{14}\Big\{n^{-1} \sum_{i=1}^{n} (X_i^{*2} - EX_i^{*2})\Big\}^2 + C_{14}\delta^{-2} n^{-1},$$

and so we need only prove that

$$y^2 n^{-2} E\Bigg[\Big\{\sum_{i=1}^{n} (X_i^{*2} - EX_i^{*2})\Big\}^2 I(|V_2| > \varepsilon |y|)\Bigg] \le C_{15}(\varepsilon)\delta n^{-1/2} + o(n^{-1/2})$$

uniformly in $|y| \le \log n$. The left-hand side equals

$$y^2 n^{-1} E\{(X_1^{*2} - EX_1^{*2})^2 P(|V_2| > \varepsilon |y| \,|\, X_1)\}$$
$$+ y^2 (1 - n^{-1}) E\{(X_1^{*2} - EX_1^{*2})(X_2^{*2}) P(|V_2| > \varepsilon |y| \,|\, X_1, X_2)\}.$$

Both terms are handled similarly. To indicate the method, we prove

$$y^2 E\{(X_1^{*2} - EX_1^{*2})(X_2^{*2} - EX_2^{*2}) P(V_2 < -\varepsilon |y| \,|\, X_1, X_2)\} \le C_{16}(\varepsilon)\delta n^{-1/2} + o(n^{-1/2}) \tag{2.14}$$

uniformly in $|y| \le \log n$. Let $\mu^* \equiv E(X_1^*)$, $\nu_k^* \equiv E\{(2p_1 - 1)|X_1^*| - \mu^*\}^k$ and

$$\Psi_n(z) \equiv 1 - \Phi(z) - \tfrac{1}{6} n^{-1/2} (\nu_2^*)^{-3/2} \nu_3^* (1 - z^2) \phi(z).$$

The classical derivation of Edgeworth expansions [Petrov (1975), Chapter 6] may be modified to yield

$$\xi_n \equiv \sup_{-\infty<z<\infty} (1+z^2)\left|P\left[\sum_{i=3}^{n}\{(2p_i-1)|X_i|-\mu^*\} > (n\nu_2^*)^{1/2}z\right] - \Psi_n(z)\right|$$
$$= o(n^{-1/2}).$$

This result, and the fact that $Z_2 < -\varepsilon|y|$ is equivalent to

$$(n\nu_2^*)^{-1/2}\sum_{i=3}^{n}\{(2p_i-1)|X_i^*|-\mu^*\} > \zeta_1 - \zeta_2,$$

where $\zeta_1 \equiv (\nu_2/\nu_2^*)^{1/2}\varepsilon|y| + (n/\nu_2^*)^{1/2}(\mu-\mu^*)$ and

$$\zeta_2 \equiv (n\nu_2^*)^{-1/2}\sum_{i=1}^{2}\{(2p_i-1)|X_i^*|-\mu^*\},$$

show that the left-hand side of (2.14) is dominated by

$$\begin{aligned}(2.15)\qquad & y^2\left|E\{(X_1^{*2}-EX_1^{*2})(X_2^{*2}-EX_2^{*2})\Psi_n(\zeta_1-\zeta_2)\}\right| \\ & +y^2\xi_n E\left|(X_1^{*2}-EX_1^{*2})(X_2^{*2}-EX_2^{*2})\{1+(\zeta_1-\zeta_2)^2\}^{-1}\right|.\end{aligned}$$

For some constant $\varepsilon_3>0$, $\zeta_1-\theta\zeta_2 \ge \varepsilon_3|y|-\varepsilon_3^{-1}$ uniformly in $0\le\theta\le1$, with probability one. For a *random* $\theta=\theta(y)\in(0,1)$, the first term in (2.15) equals

$$y^2\left|E\left[(X_1^{*2}-EX_1^{*2})(X_2^{*2}-EX_2^{*2})\{\Psi_n(\zeta_1)-\zeta_2\Psi_n'(\zeta_1)+\tfrac{1}{2}\zeta_2^2\Psi_n''(\zeta_1-\theta\zeta_2)\}\right]\right|$$
$$\le \tfrac{1}{2}\left\{y^2\sup_{z>\varepsilon_3|y|-\varepsilon_3^{-1}}|\Psi_n''(z)|\right\}E\left\{|(X_1^{*2}-EX_1^{*2})(X_2^{*2}-EX_2^{*2})|\zeta_2^2\right\}$$
$$\le C_{17}\delta n^{-1/2}+o(n^{-1/2}).$$

The second term in (2.15) is dominated by

$$C_{17}\xi_n\left(E|X_1^{*2}-EX_1^{*2}|\right)^2\sup_{y>0}y^2\left\{1+\left(\varepsilon_3 y-\varepsilon_3^{-1}\right)^2\right\}^{-1} = o(n^{-1/2}).$$

Result (2.14) follows on combining these estimates.

Step (iv). *Refinement of* (2.13). Let

$$u_3 \equiv (1+n^{-1}y^2)\{(\mu_2-\mu^2)+n^{-1}\mu_2y^2\}^{-1},$$
$$V_{y3} \equiv u_1+u_2-(n\nu_2)^{-1/2}\sum_{i=2}^{n}\{(2p_i-1)|X_i^*|-\mu\},$$
$$\Delta \equiv -(n\nu_2)^{-1/2}\{(2p_1-1)|X_1^*|-\mu\}.$$

Then $V_{y2}=V_{y3}+\Delta$, and since V_{y3} is independent of X_1^*,

$$\begin{aligned}E\{V_1\psi'(V_{y2})\} &= u_3E\{(X_1^{*2}-\mu_2)\psi'(V_{y2})\} \\ (2.16)\qquad &= -(n\nu_2)^{-1/2}u_3E\left[(X_1^2-\mu_2)\{(2p_1-1)|X_1|-\mu\}\right] \\ &\quad\times E\{\psi''(V_{y3})\}+r_1(y),\end{aligned}$$

where $|yr_1(y)| \le C_{18}\delta n^{-1/2}+o(n^{-1/2})$ uniformly in $|y|\le\log n$. To simplify

$E\{\psi''(V_{y3})\}$, let $V_{y4} \equiv (n\nu_2)^{-1/2}\Sigma_{i=2}^n\{(2p_i - 1)|X_i^*| - \mu\}$ and $V_{y3} = y_1 + y_2 - V_{y4}$. A nonuniform Berry–Esseen bound [Petrov (1975), page 125] gives

$$\eta_n \equiv \sup_{-\infty<z<\infty} (1 + |z|^3)\left|P(V_{y4} \le z) - \Phi(z)\right| = O(n^{-1/2}).$$

Therefore

$$E\{\psi''(V_{y3})\} = \int_{-\infty}^{\infty} \psi''(y_1 + y_2 - z)\phi(z)\,dz + r_2(y),$$

$$|r_2(y)| \le \eta_n \int_{-\infty}^{\infty} |\psi'''(y_1 + y_2 - z)|(1 + |z|^3)^{-1}\,dz.$$

Substituting into (2.16) we conclude that for a polynomial π_1,

$$(2.17)\qquad \sup_{|y| \le \log n} \left|\tfrac{1}{2}u_2 E\{V_1\psi'(V_{y2})\} - n^{-1/2}\pi_1(y)\phi(y)\right| \le C_{19}\delta n^{-1/2} + o(n^{-1/2}).$$

A similar but simpler argument gives an expansion of $E\{\psi(V_{y2})\}$. In the cases $\psi(y) \equiv \Phi(y)$, $\psi(y) \equiv (1 - y^2)\phi(y)$, we have, respectively,

$$\sup_{|y| \le \log n} \left|E\{\psi(V_{y2})\} - \Phi(y) - n^{-1/2}\pi_2(y)\phi(y)\right| \le C_{20}\delta n^{-1/2} + o(n^{-1/2}),$$

$$\sup_{|y| \le \log n} \left|E\{\psi(V_{y2})\} - \pi_3(y)\phi(y)\right| \le C_{20}\delta n^{-1/2} + o(n^{-1/2}),$$

for polynomials π_2, π_3. Combining with (2.13) and (2.17), and noting that δ is arbitrarily small, we obtain for polynomials π_4, π_5,

$$(2.18)\qquad \sup_{|y| \le \log n} \left|E\{\Phi(U_{y1})\} - \Phi(y) - n^{-1/2}\pi_4(y)\phi(y)\right| = o(n^{-1/2}),$$

$$(2.19)\qquad \sup_{|y| \le \log n} \left|E\{(1 - U_{y1}^2)\phi(U_{y1})\} - \pi_5(y)\phi(y)\right| = o(1).$$

Step (v). *Completion.* Modifying arguments in Steps (ii)–(iv), we obtain in the cases $\psi(y) \equiv \Phi(y)$, $\psi(y) \equiv (1 - y^2)\phi(y)$, respectively,

$$\sup_{|y| \le \log n} \left|E\{U_3 U_{y1}\psi'(U_{y1})\} - n^{-1/2}\pi_6(y)\phi(y)\right| = o(n^{-1/2}),$$

$$\sup_{|y| \le \log n} \left|E\{U_3 U_{y1}\psi'(U_{y1})\} - \pi_7(y)\phi(y)\right| = o(1).$$

Combining with (2.10), (2.18) and (2.19), we get

$$\sup_{|y| \le \log n} \left|E\{\Phi(U_y)\} - \Phi(y) - n^{-1/2}\{\pi_4(y) - \tfrac{1}{2}\pi_6(y)\}\phi(y)\right| = o(n^{-1/2}),$$

$$\sup_{|y| \le \log n} \left|E\{(1 - U_y^2)\phi(U_y)\} - \{\pi_5(y) - \tfrac{1}{2}\pi_7(y)\}\phi(y)\right| = o(1).$$

Result (2.9) is now immediate, provided

$$(2.20)\qquad \pi_4(y) - \tfrac{1}{2}\pi_6(y) + \tfrac{1}{6}\nu_2^{-3/2}\nu_3\{\pi_5(y) - \tfrac{1}{2}\pi_7(y)\} \equiv \tfrac{1}{6}\tau(2y^2 + 1).$$

Tracing back through arguments prior to (2.9), we see that (1.2) holds *if and only* *if* (2.20) is true. Now, coefficients of $\pi_4, \dots, \pi_7$ depend only on the first three moments of X and $(2p_1 - 1)|X_1|$. Expansion (1.2) is known to be correct when $E(X^6) < \infty$, and this case covers all possible values of the aforementioned moments. Therefore (2.20) holds whenever $E|X|^3 < \infty$.

APPENDIX

Proof that if $\sum X_i^2 > n\mu^2$ then $T_0 \le y$ if and only if $T \le sU_y$.

Step (i). *Simplification of* $T_0 \le y$. Let $D \equiv \bar{X} - \mu$. Then

$$T_0 = n^{1/2}(\bar{X} - \mu)\left(n^{-1}\sum_1^n X_i^2 - \bar{X}^2\right)^{-1/2}$$
$$= n^{1/2}D\left\{n^{-1}\sum_1^n X_i^2 - \mu^2 - \left(D^2 + 2\mu D\right)\right\}^{-1/2},$$

and so the equation $T_0^2 = y^2$ is equivalent to

$$Q(D) \equiv D^2(1 + n^{-1}y^2) + 2n^{-1}\mu y^2 D - n^{-1}y^2\left(n^{-1}\sum_1^n X_i^2 - \mu^2\right) = 0.$$

This has solutions

$$D = -\left(1 + n^{-1}y^2\right)^{-1}\left[n^{-1}\mu y^2 \pm n^{-1/2}y\left\{(1 + n^{-1}y^2)n^{-1}\sum_1^n X_i^2 - \mu^2\right\}^{1/2}\right]$$
$$= D_1, D_2,$$

say, where $D_1 \le D_2$. If $y \le 0$ then the inequality $T_0 \le y$ is equivalent to $(D \le 0) \cap (Q(D) \ge 0)$, and so to $D \le D_1$, i.e., to

$$D \le -\left(1 + n^{-1}y^2\right)^{-1}\left[n^{-1}\mu y^2 - n^{-1/2}y\left\{(1 + n^{-1}y^2)n^{-1}\sum_1^n X_i^2 - \mu^2\right\}^{1/2}\right]. \tag{A.1}$$

If $y > 0$, then $T_0 \le y$ is equivalent to $(D \le 0) \cup \{(D > 0) \cap (Q(D) \le 0)\}$, and so to $D \le D_2$, which again reduces to inequality (A.1).

Therefore $T_0 \le y$ is equivalent to (A.1).

Step (ii). *Simplification of* $T \le sU_y$. Note that $1 + U_3 = s^2/n\nu_2$, and so $T \le sU_y$ is equivalent to

$$T \le (n\nu_2)^{1/2}\left\{u_1 + u_2(1 + U_1)^{1/2} + U_2\right\},$$

which we write as

$$n^{-1}\left\{T - (n\nu_2)^{1/2}U_2\right\} \le (\nu_2/n)^{1/2}\left\{u_1 + u_2(1 + U_1)^{1/2}\right\}. \tag{A.2}$$

A little algebra shows that the left-hand side of (A.2) is just D, while the right-hand side is identical to the right-hand side of (A.1). Therefore inequalities (A.1) and (A.2) are identical.

Combining steps (i) and (ii), we see that $T_0 \leq y$ if and only if $T \leq sU_y$, as had to be proved.

Acknowledgments. During the course of this work I have benefitted from stimulating conversations with R. V. Erickson and D. L. Wallace. The referee's comments have also proved helpful.

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 435–451.

BICKEL, P. J. and VAN ZWET, W. R. (1978). Asymptotic expansions for the power of distributionfree tests in the two-sample problem. *Ann. Statist.* **6** 937–1004.

BOWMAN, K. O., BEAUCHAMP, J. J. and SHENTON, L. R. (1977). The distribution of the t-statistic under non-normality. *Internat. Statist. Rev.* **45** 233–242.

CALLAERT, H. and VERAVERBEKE, N. (1981). The order of the normal approximation for a Studentized U-statistic. *Ann. Statist.* **9** 194–200.

CHUNG, K. L. (1946). The approximate distribution of Student's statistic. *Ann. Math. Statist.* **17** 442–465.

CRESSIE, N. (1980). Relaxing assumptions in the one-sample t-test. *Austral. J. Statist.* **22** 143–153.

HALL, P. (1982). *Rates of Convergence in the Central Limit Theorem*. Pitman, London.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, Berlin.

WALLACE, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29** 635–654.

DEPARTMENT OF STATISTICS
FACULTY OF ECONOMICS AND COMMERCE
AUSTRALIAN NATIONAL UNIVERSITY
G.P.O. BOX 4
CANBERRA, A.C.T. 2601
AUSTRALIA