The Annals of Probability
1996, Vol. 24, No. 1, 491–503

THE BERRY–ESSEEN BOUND FOR STUDENT'S STATISTIC[1]

BY V. BENTKUS AND F. GÖTZE

University of Bielefeld

We prove the Berry–Esseen bound for the Student t-statistic. Under the assumption of a third moment this bound coincides (up to an absolute constant) with the classical Berry–Esseen bound for the mean. In general the distribution of the Student statistic converges to the standard normal distribution function at least as fast as the distribution of the mean, and sometimes faster. For example, rates of convergence can be proved if the underlying distribution is in the domain of attraction of the normal law.

1. Introduction and results. Let $X_1, \dots, X_N$ denote i.i.d. mean-zero random variables. Write

$$\sigma^2 = \mathbf{E}X_1^2, \qquad \beta_s = \mathbf{E}|X_1|^s$$

and

$$\overline{X} = N^{-1}(X_1 + \cdots + X_N), \qquad \hat{\sigma}^2 = N^{-1} \sum_{i=1}^{N} \left(X_i - \overline{X}\right)^2.$$

Assume that $\sigma > 0$, and define Student's statistic by

$$t = \overline{X}/\hat{\sigma} \quad \text{for } N \geq 2,$$

where $\overline{X}$ denotes the sample mean and $\hat{\sigma}$ the sample variance (if $\hat{\sigma} = 0$, then, for instance, set $t = 0$). The following Berry–Esseen bound is a consequence of our general estimates of the convergence rate of the distribution of $\sqrt{N}t$ to the standard normal distribution function. Write

$$\delta_N = \sup_x |\mathbf{P}\{\sqrt{N}t < x\} - \Phi(x)|.$$

THEOREM 1.1. *There exists an absolute constant $c > 0$ such that*

$$\delta_N \leq \frac{c\beta_3}{\sigma^3\sqrt{N}}.$$

Theorem 1.1 is an easy corollary of the following theorem.

THEOREM 1.2. *There exists an absolute constant c such that*

$$\delta_N \leq c\sigma^{-2}\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > \sigma^2 N\} + cN^{-1/2}\sigma^{-3}\mathbf{E}|X_1|^3\mathbf{I}\{X_1^2 \leq \sigma^2 N\}.$$

In particular, $\delta_N = o(N^{-s/2})$ provided $\beta_{2+s} < \infty$ and $0 \leq s < 1$.

Received August 1994; revised March 1995.
[1]Research supported by the SFB 343 at Bielefeld.
AMS 1991 *subject classifications*. Primary 62E20; secondary 60F05.
Key words and phrases. Student's statistic, Berry–Esseen bound, convergence rate, central limit theorem, self-normalized sums.

In order to formulate our main result, we introduce a special normalization and truncated moments. For a given random variable X_1 and a natural number N, define the number $a^2 = a_N^2(X_1)$ by the truncated second moment equation

$$a^2 = \sup\{b : \mathbf{E}X_1^2\mathbf{I}\{X_1^2 \le bN\} \ge b\}, \qquad a \ge 0. \tag{1.1}$$

We shall need the following well-known properties of a.

LEMMA 1.3. *The number a in* (1.1) *exists for any random variable X_1 and any N, and a is the largest solution of the equation*

$$a^2 = \mathbf{E}X_1^2\mathbf{I}\{X_1^2 \le a^2N\}. \tag{1.2}$$

If $\sigma < \infty$, then $a \le \sigma$. If $\sigma = \infty$, then $a_N \to \infty$ as $N \to \infty$. If $\mathbf{P}\{X_1 = 0\} < 1$, then a is positive for sufficiently large N. If $\mathbf{E}X_1^2 = 1$ and $a^2 < \frac{1}{2}$, then $\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > N/2\} \ge \frac{1}{2}$. Furthermore, $a^2(tX_1) = t^2a^2(X_1)$, for all $t \in \mathbf{R}$.

Write

$$Y_i = a^{-1}N^{-1/2}X_i\mathbf{I}\{X_i^2 \le a^2N\} \quad \text{for } 1 \le i \le N.$$

Note that Y_i does not change if we replace X_i by τX_i, $\tau > 0$. Furthermore, $|Y_i| \le 1$ and $\mathbf{E}Y_i^2 = 1/N$.

Our main result is the following estimate, which holds without assuming the existence of moments.

THEOREM 1.4. *There exists an absolute constant $c > 0$ such that*

$$\delta_N \le 2N\mathbf{P}\{X_1^2 > a^2N\} + cN|\mathbf{E}Y_1| + cN\mathbf{E}|Y_1|^3,$$

whenever the largest solution a of (1.1) *is positive.*

COROLLARY 1.5. *Fix a sequence $X_1, X_2, \dots$ of i.i.d. random variables. If*

$$\lim_{\tau\to\infty} \frac{\tau^2\mathbf{P}\{X_1^2 > \tau^2\}}{\mathbf{E}X_1^2\mathbf{I}\{X_1^2 \le \tau^2\}} = 0, \tag{1.3}$$

then $\delta_N \to 0$ as $N \to \infty$.

Theorem 1.2 easily follows from Theorem 1.4. Indeed, without loss of generality we may assume that $\sigma^2 = 1$. In the case $a^2 < \frac{1}{2}$ there is nothing to prove because $\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > N/2\} \ge \frac{1}{2}$, by Lemma 1.3, and $\delta_N \le 1 \le 2\mathbf{E}X_1^2\mathbf{I}\{X_1^2 > N/2\}$. In the case $a^2 \ge \frac{1}{2}$, Lemma 1.3 implies $\frac{1}{2} \le a^2 \le 1 = \sigma^2$, and the result follows from Theorem 1.4 by an application of the Chebyshev inequality.

REMARK. All results remain true if we replace $\sqrt{N}t$ by the so-called self-normalized sums

$$(X_1 + \cdots + X_N)/\sqrt{X_1^2 + \cdots + X_N^2}.$$

In this case proofs are the same; the only exception is the truncation Lemma 2.2, which becomes simpler. Self-normalized sums are extensively studied [see, e.g., Efron (1969), Logan, Mallows, Rice and Shepp (1973), LePage, Woodroofe and Zinn (1981), Griffin and Kuelbs (1989, 1991) Hahn, Kuelbs and Weiner (1990) and Griffin and Mason (1991)].

Condition (1.3) is the weakest known sufficient condition for the central limit theorem for Student's t-statistic [Maller (1981) and Csörgő and Mason (1987)]. It is necessary in the symmetric case [see Griffin and Mason (1991)], and it characterizes the domain of attraction of the normal law. Whether this condition is also necessary in general remains open [see Logan, Mallows, Rice and Shepp (1973) and Griffin and Mason (1991)].

Convergence rates and Edgeworth expansions for Student's and related statistics were considered by Chibisov (1980, 1984) and Slavova (1985), Helmers and van Zwet (1982), van Zwet (1984), Helmers (1985), Bhattacharya and Ghosh (1978), Hall (1987, 1988), Friedrich (1989), Bhattacharya and Denker (1990) and Bentkus, Götze and van Zwet (1994), among others.

Chibisov (1980, 1984) and Slavova (1985) proved that there exists a finite function f such that $\delta_N \le f(\beta_3/\sigma^3)/\sqrt{N}$. Theorem 1.1 improves this result. For a fixed sequence $X_1, X_2, \dots,$ Hall (1988) obtained an asymptotic result (without the explicit estimate) which is comparable in this case with Theorem 1.2 but assuming that, for some $y \in \mathbf{R}$,

$$(*) \qquad \mathbf{E}p(y)(1-p(y)) > 0, \quad \text{where } p(y) = \mathbf{P}\{X - y > 0 \,|\, |X-y|\}.$$

The paper of Hall (1988) contains a number of interesting corollaries which one can derive from results like Theorems 1.2 and 1.4. In the symmetric case, Hall (1988) proves, without condition $(*)$, an asymptotic result comparable with Theorem 1.4. It is interesting to notice that in this case there is a very fast and short reduction to sums of independent random variables by symmetrization and conditioning arguments (see the proof of Theorem 1.4 in the symmetric case in Section 2).

Other estimates are obtained as special cases of results for general symmetric asymptotically normal statistics. The result of Helmers and van Zwet (1982) implies that $\delta_N = O(N^{-1/2})$ provided $\beta_{9/2} < \infty$. The general result of van Zwet (1984) yields that $\delta_N \le c\beta_4/(\sigma^4\sqrt{N})$. Friedrich (1989) improved this result to $\Delta_N \le c\beta_{10/3}/(\sigma^{10/3}\sqrt{N})$. Bentkus, Götze and Zitikis (1994) obtained a lower estimate showing that the result of Friedrich (1989) is final, and therefore the latter bound for δ_N is the best possible result that can be derived from general results. Our proofs are related to the approach developed by Götze and van Zwet (1992) and Bentkus, Götze and van Zwet (1994).

2. Proofs. The proofs are rather straightforward, with the exception of the proof of Theorem 1.4.

PROOF OF LEMMA 1.3. The set of b satisfying the inequality in (1.1) is nonempty since $b = 0$ satisfies it.

Let us show that a is finite. For any given random variable X_1 there exists a monotone function $h(t) \geq 1$, $t \geq 0$, such that $\mathbf{E}h(|X_1|) < \infty$, $\lim_{t \to \infty} h(t) = \infty$ and such that the function $t^2/h(t)$ is increasing. Therefore, for b satisfying the inequality in (1.1),

$$b \leq \mathbf{E}X_1^2 \mathbf{I}\{X_1^2 \leq bN\} \leq bN\mathbf{E}h(|X_1|)/h(\sqrt{bN}),$$

which contradicts $a = \infty$.

It follows from the theorem of dominated convergence that a is a solution of (1.2).

The other statements of the lemma are obvious. □

PROOF OF THEOREM 1.1. The result is an immediate consequence of Theorem 1.2. □

PROOFS OF THEOREM 1.2 AND COROLLARY 1.5. The results follow from Theorem 1.4. □

Thus it remains to prove Theorem 1.4. Here we shall assume that, for a sufficiently small absolute constant $c_0 > 0$,

$$N|\mathbf{E}Y_1| \leq c_0, \qquad N\mathbf{E}|Y_1|^3 \leq c_0, \tag{2.1}$$

since otherwise the result follows from the obvious estimate

$$\sup_x |\mathbf{P}\{\sqrt{N}t < x\} - \Phi(x)| \leq 1.$$

Write

$$Y = Y_1 + \cdots + Y_N \quad \text{and} \quad \eta = \eta_1 + \cdots + \eta_N, \quad \text{where } \eta_i = Y_i^2 - N^{-1}.$$

LEMMA 2.1. *Let $1 \leq m \leq N$ and assume that* (2.1) *holds. There exists an absolute constant c and a constant $c(p)$ depending only on p such that*

$$\mathbf{E}|Y_1 + \cdots + Y_m|^p \leq c(p) \quad \text{for } p \geq 0,$$

$$\mathbf{E}(\eta_1 + \cdots + \eta_m)^2 \leq \frac{m}{N} n\mathbf{E}|Y_1|^3,$$

$$\mathbf{E}U^2 \leq c\left(\frac{m}{N}\right)^2 N\mathbf{E}|Y_1|^3, \quad \text{where } U = \sum_{i \neq j, 1 \leq i, j \leq m} Y_i \eta_j.$$

PROOF. It is well known that [see Petrov (1987)]

$$\mathbf{E}|\theta_1 + \cdots + \theta_m|^p \leq c(p) m\mathbf{E}|\theta_1|^p + c(p)\left(m\mathbf{E}\theta_1^2\right)^{p/2} \quad \text{for } p \geq 2,$$

for any i.i.d. mean-zero random variables $\theta_1, \ldots, \theta_m$. This inequality implies the first estimate of the lemma via elementary calculations using $|Y_1| \leq 1$ and $N\mathbf{E}Y_1^2 = 1$. The proofs of the second and the third estimates of the lemma are elementary as well (note that $\mathbf{E}\eta_i = 0$). □

Let $g\colon \mathbf{R} \to \mathbf{R}$ denote a function which is infinitely differentiable with bounded derivatives such that

$$\frac{1}{8} \le g(x) \le 2 \quad \text{for all } x \in \mathbf{R} \quad \text{and} \quad g(x) = \frac{1}{\sqrt{|x|}} \quad \text{for } \frac{1}{4} \le |x| \le \frac{7}{4}.$$

Define the statistic

$$S = Yg(1+\eta).$$

LEMMA 2.2 (Truncation). *Assume that* (2.1) *holds. There exists an absolute constant c such that*

$$\delta_N \le \sup_x |\mathbf{P}\{S < x\} - \Phi(x)| + 2N\mathbf{P}\{X_1^2 > a^2 N\} + cN|\mathbf{E}Y_1| + cN\mathbf{E}|Y_1|^3$$

whenever $a = a_N(X_1) > 0$.

PROOF. If $a > 0$, we may write

$$\sqrt{N}t = \frac{\sqrt{N}\,\overline{X}/a}{\sqrt{p^2/a^2 - (\overline{X}/a)^2}}, \quad \text{where } p^2 = (X_1^2 + \cdots + X_N^2)/N.$$

The complement of the event

$$\{X_1^2 \le a^2 N, \dots, X_N^2 \le a^2 N\}$$

has probability less than

$$N\mathbf{P}\{X_1^2 > a^2 N\}.$$

Therefore we may replace $\mathbf{P}\{\sqrt{N}t < x\}$ by $\mathbf{P}\{S_0 < x\}$, where the statistic S_0 is defined by

$$S_0 = Y/\sqrt{s^2 - N^{-1}Y^2}, \quad \text{where } s^2 = Y_1^2 + \cdots + Y_N^2.$$

Lemma 2.1 implies that

$$\mathbf{E}Y^2 \le c, \qquad \mathbf{P}\{|s^2 - 1| \ge \tfrac{1}{2}\} \le 4\mathbf{E}\eta^2 \le 4N\mathbf{E}|Y_1|^3.$$

Thus the events

$$\{s^2 \le \tfrac{1}{2}\}, \qquad \{s^2 \ge \tfrac{3}{2}\} \quad \text{and} \quad \{N^{-1}Y^2 \ge \tfrac{1}{4}\}$$

occur with probabilities less than $cN\mathbf{E}|Y_1|^3$ (note that $N^{-1/2} \le N\mathbf{E}|Y_1|^3$). Therefore we can replace $\mathbf{P}\{S_0 < x\}$ by $\mathbf{P}\{S_1 < x\}$, where $S_1 = Yg(s^2 - N^{-1}Y^2)$. The function g has bounded derivatives. Expanding in powers of $N^{-1}Y^2$, we obtain $S_1 = S + R$, where $|R| \le cN^{-1}|Y|^3$. Chebyshev's inequality and Lemma 2.1 imply $\mathbf{P}\{|R| > N^{-1/2}\} \le cN^{-1/2}$; but again $N^{-1/2} \le N\mathbf{E}|Y_1|^3$, and the result of the lemma follows. □

By ξ we shall denote a standard normal variable. We can write $\xi = \xi_1 + \cdots + \xi_N$, where $\xi_1, \dots, \xi_N$ are i.i.d. centered normal random variables such that $\mathbf{E}\xi_1^2 = 1/N$. We shall assume that ξ and $\xi_1, \dots, \xi_N$ are independent of all other random variables.

LEMMA 2.3. *Assume that* (2.1) *holds. Let* $H\colon \mathbf{R} \to \mathbf{C}$ *denote an infinitely many times differentiable function with bounded derivatives. Then, for* $1 \le k \le N$,

$$\begin{aligned}(2.2)\qquad &|\mathbf{E}H((Y_1 + \cdots + Y_k)g(1 + \eta_1 + \cdots + \eta_k)) - \mathbf{E}H(\xi_1 + \cdots + \xi_k)| \\ &\quad \le cc_H N\big(|\mathbf{E}Y_1| + \mathbf{E}|Y_1|^3\big),\end{aligned}$$

where $c_H = \|H'\|_\infty + \|H''\|_\infty + \|H'''\|_\infty$ *and* $\|H\|_\infty = \sup_x |H(x)|$.

PROOF. We shall prove the lemma for $k = N$ only. We have to show that $|\mathbf{E}H(S) - \mathbf{E}H(\xi)|$ is bounded from above as in (2.2). It is sufficient to prove that

$$(2.3)\qquad |\mathbf{E}H(S) - \mathbf{E}H(Y)| \le cc_H N\big(|\mathbf{E}Y_1| + \mathbf{E}|Y_1|^3\big),$$

$$(2.4)\qquad |\mathbf{E}H(Y) - \mathbf{E}H(\xi)| \le cc_H N\big(|\mathbf{E}Y_1| + \mathbf{E}|Y_1|^3\big).$$

The proof of (2.4) is easy [see Bentkus, Götze, Paulauskas and Račkauskas (1991)]. Let us prove (2.3). Expanding in powers of η, we get

$$\mathbf{E}H(S) = \mathbf{E}H(Yg(1 + \eta)) = \mathbf{E}H(Y + \mathbf{E}_\theta Y\eta g'(1 + \theta_1\eta)),$$

where $\theta_1, \theta_2, \dots,$ denote an i.i.d. sequence of random variables uniformly distributed on $[0, 1]$, independent of all other random variables. Here $\mathbf{E}_\theta$ stands for the conditional expectation given all r.v. but $\theta_1, \theta_2, \dots$. Let us split the sum $Y\eta$ into the sum of its diagonal and remaining part,

$$Y\eta = D + U, \quad \text{where } D = \sum_{i=1}^{N} Y_i\eta_i \text{ and } U = \sum_{i \ne j,\, 1 \le i, j \le N} Y_i\eta_j.$$

Expanding in powers of $\mathbf{E}_\theta Dg'(1 + \theta_1\eta)$, we may replace $\mathbf{E}H(Yg(1 + \eta))$ by $\mathbf{E}H(Y + \mathbf{E}_\theta Ug'(1 + \theta_1\eta))$. The error of such a replacement is bounded from above by $cc_H\mathbf{E}|D|$, which does not exceed the right-hand side of (2.3). Similarly, expanding in powers of $\mathbf{E}_\theta\theta_1\eta$ and $\mathbf{E}_\theta U\theta_1\eta g''(1 + \theta_1\theta_2\eta)$, we may replace $\mathbf{E}H(Y + \mathbf{E}_\theta Ug'(1 + \theta_1\eta))$ by $\mathbf{E}H(Y - U/2)$ since $g'(1) = -\frac{1}{2}$. The error does not exceed

$$cc_H\mathbf{E}|U|\,|\eta| \le cc_H\big(\mathbf{E}U^2\mathbf{E}\eta^2\big)^{1/2} \le cc_H N\mathbf{E}|Y_1|^3$$

(see Lemma 2.1 for the estimates of $\mathbf{E}U^2$ and $\mathbf{E}\eta^2$). Expanding in powers of $U/2$, we may replace $\mathbf{E}H(Y - U/2)$ by $\mathbf{E}H(Y) - \mathbf{E}UH'(Y)/2$, and it remains to estimate $\mathbf{E}UH'(Y)/2$ only. Due to the symmetry and the i.i.d. assumption, we have

$$\mathbf{E}UH'(Y) = (N^2 - N)\mathbf{E}Y_1\eta_2 H'(Y).$$

In order to estimate $N^2|\mathbf{E}Y_1\eta_2 H'(Y)|$, we first expand in powers of Y_1 and subsequently in powers of η_2. Thus we get

$$N^2|\mathbf{E}Y_1\eta_2 H'(Y)| \le cc_H N^2\big(|\mathbf{E}Y_1|\mathbf{E}\eta_2^2 + \mathbf{E}Y_1^2\mathbf{E}\eta_2^2\big) \le cc_H N\mathbf{E}|Y_1|^3,$$

which concludes the proof of (2.3) and of the lemma. □

PROOF OF THEOREM 1.4. Without loss of generality we shall assume that (2.1) is fulfilled.

An application of the truncation Lemma 2.2 reduces our proof to the verification of the following inequality

$$\sup_x|\mathbf{P}\{S < x\} - \mathbf{P}\{\xi < x\}| \le cN|\mathbf{E}Y_1| + cN\mathbf{E}|Y_1|^3 \tag{2.5}$$

whenever $a = a_N(X_1) > 0$.

In order to prove (2.5), let us apply the Berry–Esseen inequality for characteristic functions and Lemma 2.3. Write

$$f(\tau) = \mathbf{E}\exp\{i\tau S\} = \mathbf{E}\exp\{i\tau Y g(1+\eta)\},$$
$$\phi(\tau) = \mathbf{E}\exp\{i\tau\xi\} = \exp\{-\tau^2/2\}.$$

Thus (2.5) is a consequence of

$$\int_{C_1\le|\tau|\le T}|f(\tau) - \phi(\tau)|\,d\tau/|\tau| \le cN|\mathbf{E}Y_1| + cN\mathbf{E}|Y_1|^3, \tag{2.6}$$

where $T = c_1/(N\mathbf{E}|Y_1|^3)$ and where we may choose the absolute constant C_1 sufficiently large and the absolute constant $c_1 > 0$ sufficiently small. We may assume that the interval (C_1, T) is nonempty since otherwise the Theorem holds like (2.1).

Define the natural number

$$m = m(\tau) \sim \frac{C_2 N\ln|\tau|}{\tau^2} \quad \text{for } C_1 \le |\tau| \le T,$$

where C_2 is a sufficiently large absolute constant. Such a number with $2 \le m \le N/2$ exists due to our choice of constants.

Throughout the proof we shall write $A \simeq B$ if

$$\int_{C_1\le|\tau|\le T}|A - B|\,d\tau/|\tau| \le cN\big(|\mathbf{E}Y_1| + \mathbf{E}|Y_1|^3\big).$$

Thus (2.6) means that we have to prove that $f \simeq \phi$. We shall prove this relation in several steps. In the first step we shall replace f by an expectation containing as a factor a product of m conditionally independent random characteristic functions [see (2.10)–(2.13)]. This product will ensure the convergence of the integral [see (2.16)]. In the next step we replace this product by a nonrandom product [see (2.18)]. An application of Lemma 2.3 will then conclude the proof. Let us split

$$Y = X + Z, \qquad \eta = \gamma + \rho, \quad \text{where } X = Y_1 + \cdots + Y_m \text{ and } \gamma = \eta_1 + \cdots + \eta_m.$$

Then

$$f(\tau) = \mathbf{E}\exp\{i\tau Xg(1+\gamma+\rho) + i\tau Zg(1+\gamma+\rho)\}. \tag{2.7}$$

Let us show that

$$\begin{aligned} f(t) \simeq f_1(t) = \mathbf{E}\exp\{&itXg(1+\rho) + itUg'(1+\rho) \\ &+ itZg(1+\rho) + itZ\gamma g'(1+\rho)\}, \end{aligned} \tag{2.8}$$

where

$$U = \sum_{j \neq k,\, 1 \le j,\, k \le m} Y_j \eta_k .$$

Expanding in powers of γ, we have

$$Xg(1+\gamma+\rho) = Xg(1+\rho) + X\gamma g'(1+\theta\gamma+\rho),$$

$$Zg(1+\gamma+\rho) = Zg(1+\rho) + Z\gamma g'(1+\rho) + Z\gamma^2 g''(1+\theta\gamma+\rho)/2,$$

where $|\theta| \le 1$. The random variables Z and γ are independent. Applying Lemma 2.1 to bound moments and using the boundedness of the derivatives of g, we have

$$\mathbf{E}|\tau Z\gamma^2 g''(1+\theta\gamma+\rho)| \le c|\tau|\mathbf{E}|Z|\mathbf{E}\gamma^2 \le c\frac{|\tau|m}{N}N\mathbf{E}|Y_1|^3, \tag{2.9}$$

The factor $|\tau|m/N$ in (2.9) allows us to remove the term $i\tau Z\gamma^2 g''(1+\theta\gamma+\rho)/2$ in the exponent in (2.7) since $|\tau|m/N = O(|\tau|^{-1}\ln|\tau|)$ and thus the integral with respect to the measure $d\tau/|\tau|$ is convergent as $|\tau| \to \infty$.

Let us split the sum $X\gamma$ into its diagonal and nondiagonal parts,

$$X\gamma = D + U, \qquad D = \sum_{j=1}^{m} Y_j \eta_j .$$

We have

$$\mathbf{E}|\tau Dg'(1+\theta\gamma+\rho)| \le 2|\tau|m\mathbf{E}|Y_1|^3,$$

which allows us to remove $i\tau Dg'(1+\theta\gamma+\rho)$ in the exponent in (2.7). We have

$$Ug'(1+\theta\gamma+\rho) = Ug'(1+\rho) + R,$$

where

$$|R| \le c\mathbf{E}|U\gamma| \le c\big(\mathbf{E}U^2\mathbf{E}\gamma^2\big)^{1/2} \le cm\big(|\mathbf{E}Y_1| + \mathbf{E}|Y_1|^3\big),$$

and we arrive at (2.8).

We will show that (2.8) implies

$$f \simeq f_1 \simeq f_2 + f_3 \quad \text{for some } f_3 \simeq 0, \tag{2.10}$$

where

$$f_2(\tau) = \mathbf{E}\exp\{i\tau Xg(1+\rho) + i\tau Zg(1+\rho) + i\tau Z\gamma g'(1+\rho)\}$$

Expanding in powers of $i\tau Ug'(1+\rho)$ and estimating $\mathbf{E}U^2$ by Lemma 2.1, we obtain (2.10) with

$$f_3(\tau) = i\tau\mathbf{E}Ug'(1+\rho)\exp\{i\tau Xg(1+\rho) + i\tau Zg(1+\rho) + i\tau Z\gamma g'(1+\rho)\}.$$

Expanding in powers of $i\tau Z\gamma g'(1+\rho)$ and estimating the remainder $c\tau^2\mathbf{E}|Z|\mathbf{E}|\gamma U|$ by Lemma 2.1, we have

$$f_3(\tau) \simeq f_4(\tau) = i\tau\mathbf{E}Ug'(1+\rho)\exp\{i\tau Xg(\rho) + i\tau Zg(1+\rho)\}.$$

Due to the symmetry and the i.i.d. assumption, we have

$$f_4(\tau) = i\tau(m^2-m)\mathbf{E}Y_1\eta_2 g'(1+\rho)\exp\{i\tau Xg(1+\rho) + i\tau Zg(1+\rho)\}.$$

Conditioning, we have

$$|f_4(\tau)| \le c|\tau| m^2\mathbf{E}|G_1(\tau)|\,|G_2(\tau)|,$$

where

$$G_1(\tau) = \mathbf{E}_1 Y_1 \exp\{i\tau Y_1 g(1+\rho)\}, \qquad G_2(\tau) = \mathbf{E}_2\eta_2\exp\{i\tau Y_2 g(1+\rho)\},$$

and where $\mathbf{E}_1$ (resp., $\mathbf{E}_2$) denotes the conditional expectation given all random variables independent of Y_1 (resp., Y_2). Expanding in powers of $i\tau Y_1 g(1+\rho)$, we get

$$|G_1(\tau)| \le |\mathbf{E}Y_1| + c|\tau|N^{-1} \le c|\tau|N^{-1},$$

if we note that $|\tau| \ge C_1 > 1$ and use (2.1). Similarly,

$$|G_2(\tau)| \le c|\tau|\mathbf{E}|Y_1\eta_1| \le c|\tau|N^{-1}N\mathbf{E}_1|Y_1|^3.$$

Therefore

$$|f_4(\tau)| \le c|\tau|^3 m^3 N^{-2} N\mathbf{E}_1|Y_1|^3,$$

and $f_3 \simeq f_4 \simeq 0$, which concludes the proof of (2.10).

Let us derive from (2.10) that

$$f \simeq f_5 + f_6, \tag{2.11}$$

where

$$f_5(\tau) = \mathbf{E}\exp\{i\tau Xg(1+\rho) + i\tau Zg(1+\rho)\},$$
$$f_6(\tau) = i\tau m\mathbf{E}Z\eta_1 g'(1+\rho)\exp\{i\tau Xg(1+\rho) + i\tau Zg(1+\rho)\}.$$

Using the inequality $|\exp\{iA\} - 1 - iA| \le |A|^{3/2}$, with $A = \tau Z\gamma g'(1+\rho)$, and the bound $\mathbf{E}|Z|^{3/2} \le c$, as well as

$$\mathbf{E}|\gamma|^{3/2} \le cm\mathbf{E}|\eta_1|^{3/2} \le cm\mathbf{E}|Y_1|^3,$$

we obtain (2.11) with

$$f_6(\tau) = i\tau\mathbf{E}Z\gamma g'(1+\rho)\exp\{\tau Xg(1+\rho) + i\tau Zg(1+\rho)\}.$$

Now the symmetry and the i.i.d. assumption together imply (2.11).

Write

$$F(\tau) = \mathbf{E}_1\exp\{i\tau Y_1\}, \qquad G(\tau) = \mathbf{E}_1\exp\{i\tau Y_1 g(1+\rho)\}$$

and

$$G_1(\tau) = \mathbf{E}_1\eta_1\exp\{i\tau Y_1 g(1+\rho)\}.$$

Conditioning, we can rewrite f_5 and f_6 in (2.11) as

$$f_5(\tau) = \mathbf{E}\exp\{i\tau Zg(1+\rho)\}G^m(\tau), \tag{2.12}$$

$$f_6(\tau) = i\tau m\mathbf{E}Zg'(1+\rho)\exp\{i\tau Zg(1+\rho)\}G^{m-1}(\tau)G_1(\tau). \tag{2.13}$$

Expanding, we have

(2.14) $$|G_1(\tau)| \le c|\tau|\mathbf{E}|Y_1|^3.$$

The function $|G(\tau)|^2$ is the characteristic function of a symmetric random variable with variance b_2 such that

$$N^{-1}g^2(1+\rho) \le b_2 \le 2N^{-1}g^2(1+\rho),$$

and third absolute moment bounded from above by $4g^3(1+\rho)\mathbf{E}|Y_1|^3$. The function g satisfies $\frac{1}{8} \le g \le 2$. Therefore, as is well known (and easy to show),

(2.15) $$|G(\tau)| \le \exp\{-c_2\tau^2/N\} \quad \text{for } |\tau| \le c_3/\big(N\mathbf{E}|Y_1|^3\big)$$

with some absolute constants $c_2 > 0$ and $c_3 > 0$. Without loss of generality we may assume that $m/4$ is an integer. Obviously $|F(\tau)|$ is bounded from above by the right-hand side of (2.15) on the same interval, and with some absolute constant $c_4 > 0$ we have

(2.16) $$\max\{|F(\tau)|^{m/4}; |G(\tau)|^{m/4}\} \le \exp\{-c_4 m\tau^2/N\} = |\tau|^{-c_4C_2} \le |\tau|^{-10},$$

for $C_1 \le |\tau| \le T$, provided we choose the constant C_2 sufficiently large.

Relations (2.13), (2.14) and (2.16) together imply that $f_6 \simeq 0$, and thus we can rewrite (2.11) as

(2.17) $$f \simeq f_5$$

with f_5 defined by (2.12). Due to (2.16), we may ignore powers of τ as factors in the following estimates. Indeed, the integral will converge as $\tau \to \infty$ provided we have factors $G^{m/4}$ or $F^{m/4}$ balancing these powers.

Let us derive from (2.17) that

(2.18) $$f \simeq f_7 + f_8,$$

where

(2.19) $$\begin{aligned} f_7(\tau) &= F^m(\tau)\mathbf{E}\exp\{i\tau Zg(1+\rho)\}, \\ |f_8(\tau)| &\le c|\tau|^3 mN^{-1}|F^{m-1}(\tau)\mathbf{E}\rho\exp\{i\tau Zg(1+\rho)\}|. \end{aligned}$$

We may assume that $k = m/2$ is an integer, and we may write

$$G^m(\tau) = G^k(\tau)G^k(\tau) \quad \text{with } G^k(\tau) = \mathbf{E}_X\exp\{i\tau Xg(1+\rho)\},$$

where now we write $X = Y_1 + \cdots + Y_k$ and where $\mathbf{E}_X$ denotes the conditional expectation given all random variables but X. Expanding first in powers of ρ and subsequently expanding the exponential, we have

$$G^k(\tau) = \mathbf{E}_X\exp\{i\tau X\} - \frac{i\tau\rho}{2}\mathbf{E}_X X\exp\{i\tau X\} + R,$$

with $|R| \le c\mathbf{E}_X|\tau X|\rho^2 + c\mathbf{E}_x\tau^2X^2\rho^2 \le c\tau^2\rho^2$. However, $\mathbf{E}\rho^2 \le cN\mathbf{E}|Y_1|^3$, and we may replace $G^k(\tau)$ in (2.12) by

$$\mathbf{E}_X\exp\{i\tau X\} - \frac{i\tau\rho}{2}\mathbf{E}_X X\exp\{i\tau X\} = F^k(\tau) - \frac{im\tau\rho}{4}F^{k-1}(\tau)G_2(\tau),$$

where $G_2(\tau) = \mathbf{E}_1 Y_1 \exp\{i\tau Y_1\}$ satisfies

$$|G_2(\tau)| \le |\tau|\,|\mathbf{E}Y_1| + \tau^2 N^{-1} \le c\tau^2 N^{-1}. \tag{2.20}$$

Now the balancing factor F^{k-1} is present, and arguing similarly we get (2.18) with

$$f_8(\tau) = -\frac{i\tau m}{2} G_2(\tau) F^{m-1}(\tau) \mathbf{E}\rho \exp\{i\tau Z g(1+\rho)\}.$$

Applying (2.20) we obtain (2.19), which concludes the proof of (2.18).

The symmetry and the i.i.d. assumption together imply

$$\begin{aligned} &|\mathbf{E}\rho \exp\{i\tau Zg(1+\rho)\}| \\ &\quad \le (N-m)|\mathbf{E}\eta_N \exp\{i\tau Zg(1+\rho)\}| \le cN\mathbf{E}|Y_1|^3, \end{aligned} \tag{2.21}$$

by Taylor expansions in powers of Y_N and η_N. Due to (2.21), $f_8 \simeq 0$, and it follows from (2.18) that $f \simeq f_7$. Let us apply Lemma 2.3 to $\mathbf{E}\exp\{i\tau Zg(1+\rho)\}$. We obtain

$$f(\tau) \simeq f_7(\tau) \simeq F^m(\tau)\exp\{-(1 - mN^{-1})\tau^2/2\}.$$

According to our choice $m \le N/2$ and therefore $\exp\{-(1 - mN^{-1})\tau^2/2\} \le \exp\{-\tau^2/4\}$. Thus using Lemma 2.3 we may replace $F^m(\tau)$ by $\exp\{-m\tau^2/(2N)\}$. We obtain $f \simeq \phi$, which concludes the proof of the theorem. □

PROOF OF THEOREM 1.4 IN THE SYMMETRIC CASE [cf. Griffin and Mason (1991)]. Define the statistic

$$T = Y/s, \quad \text{where } s^2 = Y_1^2 + \cdots + Y_N^2.$$

A small modification of the truncation Lemma 2.2 shows that it is sufficient to demonstrate that

$$\delta_N' = \sup_x |\mathbf{P}\{T < x\} - \Phi(x)| \le cN\mathbf{E}|Y_1|^3. \tag{2.22}$$

Let $\varepsilon_1, \dots, \varepsilon_N$ be i.i.d. symmetric Bernoulli random variables independent of all other random variables and such that $\mathbf{P}\{|\varepsilon_1| = 1\} = 1$. Then, due to symmetry,

$$\delta_N' \le 2\mathbf{P}\{s^2 < \tfrac{1}{2}\} + \mathbf{E}\mathbf{I}\{s^2 \ge \tfrac{1}{2}\} \sup_x \left|\mathbf{P}_\varepsilon\{\sqrt{N} Y_\varepsilon/s < x\} - \Phi(x)\right|.$$

Here $Y_\varepsilon = \varepsilon_1 Y_1 + \cdots \varepsilon_N Y_N$, and $\mathbf{P}_\varepsilon$ denotes the conditional probability given $Y_1, \dots, Y_N$. Applying conditionally the well-known result [see Petrov (1975)]

$$\sup_x |\mathbf{P}\{Z_1 + \cdots + Z_N < x\} - \Phi(x)| \le c\mathbf{E}|Z_1|^3 + \cdots + c\mathbf{E}|Z_N|^3, \tag{2.23}$$

valid for independent mean-zero random variables $Z_1, \dots, Z_N$ such that $\mathbf{E}Z_1^2 + \cdots + \mathbf{E}Z_N^2 = 1$, we obtain (2.22).

Let us note that we need (2.23) only for Bernoulli random variables, that is, for random variables which assume at most two values. □

REFERENCES

BENTKUS, V., GÖTZE, F., PAULAUSKAS, V. and RAČKAUSKAS, A. (1991). The accuracy of Gaussian approximation in Banach spaces. *Encyclopaedia of Mathematical Sciences* **81**. Springer, Berlin.

BENTKUS, V., GÖTZE, F. and VAN ZWET, W. (1994). An Edgeworth expansion for symmetric statistics. Preprint 94-020, SFB 343, Univ. Bielefeld.

BENTKUS, V., GÖTZE, F. and ZITIKIS, R. (1994). Lower estimates of the convergence rate for U-statistics. *Ann. Probab.* **22** 1707–1714.

BHATTACHARYA, R. N. and DENKER, M. (1990). *Asymptotic Statistics*. Birkhäuser, Boston.

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

CHIBISOV, D. M. (1980). Asymptotic expansion for the distribution of a statistic admitting a stochastic expansion I. *Theory Probab. Appl.* **25** 732–744.

CHIBISOV, D. M. (1984). Asymptotic expansions and deficiences of tests. In *Proc. Internat. Congress Math.* (Z. Cieselski and C. Olech, eds.) **2** 1063–1079. North-Holland, Amsterdam.

CSÖRGŐ and MASON, D. M. (1987). Approximations of weighted empirical processes with applications to extreme, trimmed and self-normalized sums. In *Proceedings of the First World Congress of the Bernoulli Society* **2** 811–819. VNU Sci, Utrecht.

EFRON, B. (1969). Student's t-test under symmetry conditions. *J. Amer. Statist. Assoc.* 1278–1302.

FRIEDRICH, K. O. (1989). A Berry–Esseen bound for functions of independent random variables. *Ann. Statist.* **17** 170–183.

GÖTZE, F. and VAN ZWET, W. (1992). Edgeworth expansions for asymptotically linear statistics. SFB 343, Preprint 91-034, Univ. Bielefeld. (Revised version.)

GRIFFIN, P. S. and KUELBS, J. D. (1989). Self-normalized laws of the iterated logarithm. *Ann. Probab.* **17** 1571–1601.

GRIFFIN, P. S. and KUELBS, J. D. (1991). Some extensions of the LIL via self-normalizations. *Ann. Probab.* **19** 380–395.

GRIFFIN, P. S. and MASON, D. M. (1991). On the asymptotic normality of self-normalized sums. *Math. Proc. Cambridge Philos. Soc.* **109** 597–610.

HAHN, M. G., KUELBS, J. and WEINER, D. C. (1990). The asymptotic joint distribution of self-normalized censored sums and sums of squares. *Ann. Probab.* **18** 1284–1341.

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

HALL, P. (1988). On the effect of random norming on the rate of convergence in the central limit theorem. *Ann. Probab.* **16** 1265–1280.

HELMERS, R. (1985). The Berry–Esseen bound for studentized U-statistics. *Canad. J. Statist.* **13** 79–82.

HELMERS, R. and VAN ZWET, W. (1982). The Berry–Esseen bound for U-statistics. In *Statistical Decision Theory and Related Topics III.* (S. S. Gupta and J. O. Berger, eds.) **1** 497–512. Academic Press, New York.

LEPAGE, R., WOODROOFE, M. and ZINN, J. (1981). Convergence to a stable distribution via order statistics. *Ann. Probab.* **9** 624–632.

LOGAN, B., MALLOWS, C., RICE, S. and SHEPP, L. (1973). Limit distributions of self-normalized sums. *Ann. Probab.* **1** 788–809.

MALLER, A. (1981). A theorem on products of random variables, with application to regression. *Austral. J. Statist.* **23** 177–185.

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

PETROV, V. V. (1987). *Limit Theorems for Sums of Independent Random Variables*. Nauka, Moscow. (In Russian.)

SLAVOVA, V. V. (1985). On the Berry–Esseen bound for Student's statistics. *Stability Problems for Stochastic Models. Lecture Notes in Math.* **1155** 335–390. Springer, Berlin.

VAN ZWET, W. R. (1984). A Berry–Esseen bound for symmetric statistics. *Z. Wahrsch Verw Gebiete* **66** 425–440.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 100131
33501 BIELEFELD 1
GERMANY
E-mail: bentkus@mathematik.uni-bielefeld.de
goetze@mathematik.uni-bielefeld.de