The Annals of Statistics
1998, Vol. 26, No. 4, 1540–1569

EMPIRICAL EDGEWORTH EXPANSIONS FOR SYMMETRIC STATISTICS[1]

BY HEIN PUTTER AND WILLEM R. VAN ZWET

University of Leiden and University of North Carolina, Chapel Hill

In this paper the validity of a one-term Edgeworth expansion for Studentized symmetric statistics is proved. We propose jackknife estimates for the unknown constants appearing in the expansion and prove their consistency. As a result we obtain the second-order correctness of the empirical Edgeworth expansion for a very general class of statistics, including U-statistics, L-statistics and smooth functions of the sample mean. We illustrate the application of the bootstrap in the case of a U-statistic of degree two.

1. Introduction. Let $(\mathscr{X}, \mathscr{A})$ be a measurable space and P a probability measure on $(\mathscr{X}, \mathscr{A})$. Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables, taking values in $\mathscr{X}$ with unknown common distribution P. Let $t_N: \mathscr{X}^N \times \mathscr{P} \to \mathbb{R}$ be symmetric as a function on $\mathscr{X}^N$, that is, for every $x_1, \dots, x_N \in \mathscr{X}$ and every permutation $\{\alpha_1, \dots, \alpha_N\}$ of $\{1, \dots, N\}$, we have

$$t_N(x_1, \dots, x_N; P) = t_N(x_{\alpha_1}, \dots, x_{\alpha_N}; P).$$

It will be assumed throughout this paper that

$$T_N = t_N(X_1, \dots, X_N; P) \tag{1.1}$$

is a random variable with expectation

$$ET_N = 0 \quad \text{for all } N \tag{1.2}$$

and variance $\sigma^2(T_N) = \sigma_N^2$, satisfying

$$0 < c \le \sigma_N^2 \le C < \infty \quad \text{for all } N \tag{1.3}$$

for finite positive constants c and C. Suppose that T_N/σ_N converges in distribution to a standard normal distribution. Typically, the accuracy of the normal approximation is of the order $N^{-1/2}$ as N tends to infinity. In this paper we shall focus on second-order approximations, that is, on approximations with error $o(N^{-1/2})$ to the distribution functions of T_N/σ_N and also of the Studentized version T_N/S_N, where S_N^2 is an estimator of σ_N^2. We shall accomplish this by first proving the validity of a (one-term) Edgeworth

---

Received July 1995; revised May 1997.

[1]Supported by the Netherlands Organization for Scientific Research (NWO) and the Sonderforschungsbereich 343 "Diskrete Strukturen in der Mathematik" at the University of Bielefeld, Germany.

AMS 1991 *subject classifications*. Primary 62E20, 62G09.

Key words and phrases. Edgeworth expansion, Hoeffding's decomposition, jackknife, bootstrap.

expansion with remainder $o(N^{-1/2})$ and then estimating the unknown constants in the expansion. The procedure is therefore called an *empirical Edgeworth expansion*.

As an estimate S_N^2 of σ_N^2, we shall use the jackknife estimator of variance, introduced by Quenouille (1949, 1956) and Tukey (1958). Let us suppose that we have one additional observation X_{N+1} at our disposal, and define for $i = 1, \dots, N$,

$$T_N^{(i)} = t_N(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_{N+1}; P), \qquad T_N^{(N+1)} = T_N \tag{1.4}$$

and

$$\overline{T}_N = \frac{1}{N+1} \sum_{i=1}^{N+1} T_N^{(i)}. \tag{1.5}$$

The jackknife estimator of variance S_N^2 is then defined as

$$S_N^2 = \sum_{i=1}^{N+1} \left(\overline{T}_N - T_N^{(i)}\right)^2. \tag{1.6}$$

We shall make extensive use of the properties of Hoeffding's decomposition [Hoeffding (1948)]. For $k = 1, 2, \dots,$ let Ω_k denote the set of integers from 1 up to k and define for a set $A \subset \Omega_N$,

$$E(T_N|A) = E(T_N|X_i, i \in A). \tag{1.7}$$

Next, for $D \subset \Omega_N$, define

$$T_{N,D} = \sum_{A \subset D} (-1)^{|D|-|A|} E(T_N|A). \tag{1.8}$$

Here $|A|$ denotes the cardinality of a set A and the summation in (1.8) is over all subsets A of D including the empty set. In this way we obtain for instance

$$\begin{aligned} T_{N,\varnothing} &= ET_N, \\ T_{N,i} = T_{N,\{i\}} &= E(T_N|X_i) - ET_N, \\ T_{N,\{i,j\}} &= E(T_N|X_i, X_j) - E(T_N|X_i) - E(T_N|X_j) + ET_N. \end{aligned} \tag{1.9}$$

The Hoeffding decomposition of T_N is given by

$$T_N = \sum_{D \subset \Omega_N} T_{N,D} = \sum_{i=1}^{N} T_{N,\{i\}} + \sum\sum_{1 \le i < j \le N} T_{N,\{i,j\}} + \sum\sum\sum_{1 \le i < j < k \le N} T_{N,\{i,j,k\}} + \cdots.$$

For notational convenience we shall write T_{Ni} instead of $T_{N,\{i\}}$ and T_{Nij} instead of $T_{N,\{i,j\}}$. Define two real numbers λ_1 and λ_2 as

$$\lambda_1 = N^{3/2}\sigma_N^{-3} ET_{N1}^3, \qquad \lambda_2 = N^{5/2}\sigma_N^{-3} ET_{N1}T_{N2}T_{N12}. \tag{1.10}$$

Then the Edgeworth expansion for the distribution function of T_N/σ_N is given by

$$G_N(x) = \Phi(x) - \frac{\lambda_1 + 3\lambda_2}{6\sqrt{N}}(x^2 - 1)\phi(x). \tag{1.11}$$

Note that $(\lambda_1 + 3\lambda_2)N^{-1/2}$ serves as an approximation to the third cumulant of T_N/σ_N. The Edgeworth expansion to the distribution function of T_N/S_N is given by

$$H_N(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\big((2x^2+1)\lambda_1 + 3(x^2+1)\lambda_2\big). \tag{1.12}$$

We shall prove the following results:

THEOREM 1.1. *Suppose that there exist real numbers* $c > 0$, $C > 0$, $p > 3$, $r > 2$, *sequences* $\{\delta_N\}_{N=1}^{\infty}$ *with* $\delta_N \searrow 0, \{\tau_N\}$ *with* $\tau_N \to \infty$ *and a positive continuous function* χ *on* $(0,\infty)$, *such that* (1.2) *and* (1.3) *are satisfied and*

$$E|N^{1/2}T_{N1}|^p \le C, \tag{1.13}$$

$$E|N^{3/2}T_{N12}|^r \le C, \tag{1.14}$$

$$\sum_{k=3}^{N} \binom{N}{k} ET^2_{N\Omega_k} \le \delta_N N^{-3/2} \tag{1.15}$$

and

$$|Ee^{itN^{1/2}T_{N1}}| \le 1 - \chi(t) < 1 \qquad \forall\, t \in (0,\tau_N) \text{ for } N = 1,2,\dots. \tag{1.16}$$

Then there exists a sequence $\varepsilon_N \searrow 0$, *depending only on* $c, C, p, r, \{\delta_N\}, \{\tau_N\}$ *and* χ, *such that for* $N = 2,3,\dots,$

$$\sup_{x\in\mathbb{R}} |P(T_N/\sigma_N \le x) - G_N(x)| \le \varepsilon_N N^{-1/2}. \tag{1.17}$$

THEOREM 1.2. *Suppose that there exist real numbers* $c > 0$, $C > 0$, $p > 3$, $r > 2$, *a sequence* $\{\tau_N\}$ *with* $\tau_N \to \infty$ *and a positive continuous function* χ *on* $(0,\infty)$ *such that* (1.2), (1.3), (1.13), (1.14) *and* (1.16) *are satisfied and*

$$\begin{aligned}&\sum_{k=3}^{N} \binom{N-1}{k-1} ET^2_{N\Omega_k} \le CN^{-3},\\ &\sum_{k=3}^{N} \binom{N-2}{k-2} ET^2_{N\Omega_k} \le CN^{-7/2}.\end{aligned} \tag{1.18}$$

Then there exists a sequence $\varepsilon_N \searrow 0$, *depending only on* $c, C, p, r, \{\tau_N\}$ *and* χ, *such that for* $N = 2,3,\dots,$

$$\sup_{x\in\mathbb{R}} |P(T_N/S_N \le x) - H_N(x)| \le \varepsilon_N N^{-1/2}. \tag{1.19}$$

The empirical Edgeworth expansions are obtained by replacing the constants λ_1 and λ_2 in (1.11) and (1.12) by estimates. The estimation of λ_1 is straightforward and very similar to the estimation of $\sigma^2(T_N)$. Recall that, with one additional observation X_{N+1} from P, $T_N^{(i)}$ and $\overline{T}_N$ are defined as in (1.4) and (1.5).

To estimate λ_2 we assume that we have two additional observations from P: X_{N+1} and X_{N+2}. Let T_N be as in (1.1). Define, for $1 \le i < j \le N+2$,

$$T_N^{(i,j)} = T_N^{(j,i)} = t_N\big(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_{N+2}; P\big), \tag{1.20}$$

with X_i and X_j replaced by X_{N+1} and X_{N+2}. Furthermore, let

$$\overline{T}_N^{(i)} = \frac{1}{N+1} \sum_{\substack{j=1 \\ j \ne i}}^{N+2} T_N^{(i,j)}, \tag{1.21}$$

$$\overline{\overline{T}}_N = \frac{1}{N+2} \sum_{i=1}^{N+2} \overline{T}_N^{(i)} = \binom{N+2}{2}^{-1} \sum\sum_{1 \le i < j \le N+2} T_N^{(i,j)}. \tag{1.22}$$

We propose the following jackknife estimates for λ_1 and λ_2:

$$\hat{\lambda}_1 = \sqrt{N} \sum_{i=1}^{N+1} \big(\overline{T}_N - T_N^{(i)}\big)^3 \Big/ S_N^3 \tag{1.23}$$

and

$$\begin{aligned} \hat{\lambda}_2 = 2\sqrt{N} \sum\sum_{1 \le i < j \le N+2} & \Big(\overline{\overline{T}}_N - \overline{T}_N^{(i)} - \overline{T}_N^{(j)} + T_N^{(i,j)}\Big) \\ & \times \big(\overline{T}_N - T_N^{(i)}\big)\big(\overline{T}_N - T_N^{(j)}\big) \Big/ S_N^3. \end{aligned} \tag{1.24}$$

By substituting the estimators $\hat{\lambda}_1$ and $\hat{\lambda}_2$ for λ_1 and λ_2 in the Edgeworth expansions G_N and H_N, defined in (1.11) and (1.12), we obtain the empirical Edgeworth expansions

$$\hat{G}_N(x) = \Phi(x) - \frac{\hat{\lambda}_1 + 3\hat{\lambda}_2}{6\sqrt{N}}(x^2 - 1)\phi(x) \tag{1.25}$$

and

$$\hat{H}_N(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\big((2x^2+1)\hat{\lambda}_1 + 3(x^2+1)\hat{\lambda}_2\big). \tag{1.26}$$

The following result asserts the validity of the empirical Edgeworth expansion.

THEOREM 1.3. *Let $X_1, \dots, X_N$ be an i.i.d. sample from P, let T_N be a symmetric random variable, defined as in* (1.1). *Let the jackknife estimator of variance S_N^2 be defined as in* (1.6), *$\hat{\lambda}_1$ and $\hat{\lambda}_2$ as in* (1.23) *and* (1.24), *and $\hat{G}_N$ and $\hat{H}_N$ as in* (1.25) *and* (1.26). *Suppose that there exist real numbers $c > 0$, $C > 0$, $p > 3$, $r > 2$, a positive continuous function χ on $(0, \infty)$ and a sequence $\tau_N \to \infty$, such that* (1.2), (1.3), (1.13), (1.14) *and* (1.16) *are satisfied and*

$$\sum_{k=3}^{N} \binom{N-2}{k-2} E T_{N\Omega_k}^2 \le C N^{-4}. \tag{1.27}$$

Then there exist sequences $\delta_N \searrow 0$ *and* $\varepsilon_N \searrow 0$, *which depend only on* c, C, p, r, χ *and the sequence* $\{\tau_N\}$, *such that for* $N = 2, 3, \dots,$

$$P\Big(\sup_{x\in\mathbb{R}}\big|P(T_N/\sigma_N \le x) - \hat{G}_N(x)\big| \ge \varepsilon_N N^{-1/2}\Big) \le \delta_N, \tag{1.28}$$

$$P\Big(\sup_{x\in\mathbb{R}}\big|P(T_N/S_N \le x) - \hat{H}_N(x)\big| \ge \varepsilon_N N^{-1/2}\Big) \le \delta_N. \tag{1.29}$$

The typical situation to which these empirical Edgeworth expansions may be applied is the following: let $\theta = \theta(P)$ be a parameter of interest and suppose that $U_N = u_N(X_1, \dots, X_N)$ is an unbiased estimator of θ. As $X_1, \dots, X_N$ are i.i.d., we may safely restrict attention to symmetric functions u_N. Let

$$T_N = \sqrt{N}(U_N - \theta) \tag{1.30}$$

and suppose that with $\sigma_N^2 = \sigma^2(T_N)$, $T_N/\sigma_N \to^{\mathscr{D}} \mathscr{N}(0,1)$. We wish to obtain a second-order correct confidence interval for θ. For the special case of the sample mean it is well known that it is important to base inference on a pivotal random variable, that is, on a random variable whose limiting distribution does not depend on any unknown quantities. If σ_N^2 is known, we might take T_N/σ_N; if σ_N^2 is unknown, we have to use T_N/S_N, with S_N^2 an appropriate estimator of σ_N^2. The confidence interval for θ may then be based on the quantiles of the empirical Edgeworth expansion of the distribution function of T_N/σ_N or T_N/S_N.

Results similar to Theorem 1.2 were obtained earlier in a paper by Helmers (1991) in the special case of Studentized U-statistics of degree two. The results of this paper may be used to prove Helmers' result under weaker moment conditions, but more importantly, the class of statistics for which the Edgeworth expansions are established is considerably larger and includes, for instance, L-statistics, smooth functions of the sample mean and smooth functionals of the empirical distribution function.

All the results in this section are formulated as inequalities for fixed, but arbitrary, N. Since the constants in the conclusions are not specified, however, they should be viewed as purely asymptotic results. The reason for phrasing the assumptions and conclusions in these results in such a laborious way is that we want to define uniformity classes. The constants and sequences appearing in the conclusions of the results depend only on the constants and sequences appearing in the assumptions and in particular *not on N*. This allows us to consider a sequence of problems, indexed by N, where for every N the random variables X_i may be different, as well as their distributions P, the functions $T_N = t_N(X_1, \dots, X_N; P)$ and so on, as long as the conditions continue to be satisfied for the same fixed constants and sequences for every N. The conclusions of the theorems are then also true for every N and the asymptotic assertion follows, uniformly in P and $\{T_N\}$.

The jackknife estimator of variance as we have defined it in (1.4)–(1.6) may seem somewhat awkward, since with $N + 1$ observations from P, for practical purposes we would then wish to approximate the distribution of T_{N+1} instead of T_N, and the variance of T_{N+1} would be the object of interest. In our notation, the familiar delete-one jackknife would coincide with S^2_{N-1}. A comparison of (4.13) with the same expression for $N-1$ yields that the delete-one jackknife S^2_{N-1} can also be used in Theorem 1.2 if for sequences $\delta_N \searrow 0$ and $\varepsilon_N \searrow 0$,

$$P\left(\left|\frac{T_N}{\sigma_N} - \frac{T_N}{\sigma_{N-1}}\right| \geq \varepsilon_N N^{-1/2}\right) \leq \delta_N N^{-1/2}.$$

In view of (1.3) this is easily seen to be true, provided that $|\sigma_N^2 - \sigma_{N-1}^2| \leq \tilde{\varepsilon}_N N^{-3/4}$ for a sequence $\tilde{\varepsilon}_N \searrow 0$.

In (1.2) it is assumed that $ET_N = 0$ for all N. This condition excludes interesting standardized statistics, such as many L-statistics and smooth functionals of the empirical, for which typically $ET_N = \mathscr{O}(N^{-1/2})$. Suppose that $ET_N = \beta_N$ and write $\tilde{T}_N = T_N - \beta_N$. Then, apart from the first (constant) term, the Hoeffding decompositions of T_N and $\tilde{T}_N$ coincide. An inspection of the proofs of Theorems 1.1 and 1.2 shows that if $|\beta_N| \leq \varepsilon_N N^{-1/4}$ for a sequence $\varepsilon_N \searrow 0$, then the Edgeworth expansions of T_N/σ_N and T_N/S_N require an additional term of $-(\beta_N/\sigma_N)\phi(x)$. Thus, for instance, the Edgeworth expansion of T_N/S_N becomes

$$\Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\left((2x^2+1)\lambda_1 + 3(x^2+1)\lambda_2 - 6\sqrt{N}\,\frac{\beta_N}{\sigma_N}\right).$$

Of course, to obtain an empirical Edgeworth expansion, one would proceed to estimate β_N.

The class of jackknife-type estimators that we consider in this paper has the desirable property that every evaluation needed to compute it, such as $T_N^{(i)}$ and $T_N^{(i,j)}$, is based on exactly N observations. This avoids the problem of relating the Hoeffding decompositions of T_N and T_{N-1}. Unfortunately, Quenouille's (1956) jackknife estimator of bias is essentially based on this difference between the Hoeffding decomposition of T_N and T_{N-1}. It is not surprising, therefore, that the type of estimators that we consider in this paper are not suited to estimate bias. We shall therefore not address bias estimation here and insist that $ET_N = 0$.

The remainder of this paper is organized as follows. In Section 2 we discuss some important examples. In Section 3 we show how Theorem 1.2 can be applied to prove second-order correctness for bootstrapping Studentized U-statistics of degree two. Section 4 contains the proof of Theorem 1.2. In Section 5 we prove the consistency of the jackknife estimators of the quantities appearing in the Edgeworth expansions. Finally, Section 6 contains a technical lemma.

2. Applications. We shall consider some important applications of Theorem 1.3: U-statistics, L-statistics, smooth functions of the sample mean and smooth functionals of the empirical distribution function. Berry–Esseen bounds for U- and L-statistics were established in van Zwet (1984) and in the second example we shall make use of the results in that paper. Recall that the jackknife estimator of variance S_N^2 is defined as in (1.6), $\hat{\lambda}_1$ and $\hat{\lambda}_2$ as in (1.23) and (1.24) and $\hat{G}_N$ and $\hat{H}_N$ as in (1.25) and (1.26).

APPLICATION 1 (U-statistics). Let $X_1, \dots, X_N$ be i.i.d. random variables assuming values in a measurable space $(\mathscr{X}, \mathscr{A})$ with common distribution P, and let $h\colon \mathscr{X}^m \to \mathbb{R}$ be a measurable function which is symmetric in its arguments, with

$$Eh(X_1, \dots, X_m) = \theta \quad \text{and} \quad Eh^2(X_1, \dots, X_m) < \infty.$$

Let

$$U_N = \binom{N}{m}^{-1} \sum_{1 \le i_1 < \cdots < i_m \le N} \cdots \sum h(X_{i_1}, \dots, X_{i_m})$$

be a U-statistic of degree m and put

$$T_N = \sqrt{N}(U_N - \theta).$$

Define

$$g(x) = E(h(X_1, \dots, X_m) | X_1 = x) - \theta, \tag{2.1}$$

$$\psi(x, y) = E(h(X_1, \dots, X_m) | X_1 = x, X_2 = y) - g(x) - g(y) - \theta. \tag{2.2}$$

Theorem 1.3 implies the following corollary.

COROLLARY 2.1. *Suppose that the distribution of $g(X_1)$ is nonlattice and suppose that there exist $p > 3$ and $r > 2$ such that $E|g(X_1)|^p < \infty$ and $E|\psi(X_1, X_2)|^r < \infty$. Then*

$$\sqrt{N} \sup_{x \in \mathbb{R}} \left| P(T_N/\sigma_N \le x) - \hat{G}_N(x) \right| \to_P 0, \tag{2.3}$$

$$\sqrt{N} \sup_{x \in \mathbb{R}} \left| P(T_N/S_N \le x) - \hat{H}_N(x) \right| \to_P 0. \tag{2.4}$$

The proof is quite straightforward, following the proof of Corollary 4.1 of van Zwet (1984), and is therefore omitted.

APPLICATION 2 (L-statistics). Let $X_1, \dots, X_N$ be i.i.d. random variables with common distribution function F. Let $c_1, \dots, c_N$ be a sequence of real numbers, let $X_{1:N}, \dots, X_{N:N}$ denote the order statistics of $X_1, \dots, X_N$, and define the L-statistic

$$L_N = N^{-1} \sum_{i=1}^{N} c_i X_{i:N}.$$

Suppose that L_N is an unbiased estimate of θ and define

$$T_N = \sqrt{N}(L_N - \theta).$$

Suppose that there exist real numbers a, b and c such that

$$\max_{1\le i\le N}|c_i|\le a,\qquad N\max_{2\le i\le N}|c_i-c_{i-1}|\le b,$$
$$N^2\max_{3\le i\le N}|c_i-2c_{i-1}+c_{i-2}|\le c. \tag{2.5}$$

This corresponds to the case of smooth weights. Assumption (2.5) is fulfilled if there exists a function J: $(0,1)\to\mathbb{R}$ with bounded second derivative such that $c_i=J(i/(N+1))$. Theorem 1.3 implies the following corollary.

COROLLARY 2.2. *Suppose that* (2.5) *is satisfied*, $E|X_1|^p<\infty$ *for some* $p>3$, $\sigma^2(T_N)\ge c'$ *for some* $c'>0$ *and all* N *and* $E(N^{1/2}T_N|X_1)$ *satisfies* (1.16). *Then* (2.3) *and* (2.4) *hold.*

To prove Corollary 2.2 we start by deriving a representation for the terms in the Hoeffding decomposition of T_N which is of interest in its own right. Define i.i.d. uniform random variables $U_1,\dots,U_N$ and take $X_i=F^{-1}(U_i)$, where $F^{-1}(t)=\inf\{x\colon F(x)\ge t\}$ denotes the left-continuous version of the inverse of F. We let $U_{1:N},\dots,U_{N:N}$ denote the order statistics of $U_1,\dots,U_N$. Then we have the relation

$$X_{j+1:N}-X_{j:N}=\sum_{\substack{A\subset\Omega_N\\|A|=j}}\int_0^1\prod_{i\in A}\big(\mathbf{1}_{[U_i,1)}(t)\big)\prod_{i\in A^c}\big(\mathbf{1}_{(0,U_i)}(t)\big)\,dF^{-1}(t). \tag{2.6}$$

To see why this relation holds, note that the integrand is zero for a fixed t unless t is between the largest of the U_i's with i in A and the smallest of the U_i's with i not in A, and this can only occur if the U_i's with i in A happen to be the j smallest among $U_1,\dots,U_N$. For the only A for which the integrand is not identically equal to zero, the integral yields $\int_{[U_{j:N},U_{j+1:N})}dF^{-1}(t)=X_{j+1:N}-X_{j:N}$. Now we obtain

$$\begin{aligned}
&\big[X_{j+1:N}-X_{j:N}\big]_D\\
&\quad=\sum_{\substack{A\subset\Omega_N\\|A|=j}}\int_0^1\prod_{i\in A\cap D}\big(\mathbf{1}_{[U_i,1)}(t)-t\big)t^{|A\cap D^c|}\\
&\qquad\qquad\times\prod_{i\in A^c\cap D}\big(\mathbf{1}_{(0,U_i)}(t)-(1-t)\big)(1-t)^{|A^c\cap D^c|}\,dF^{-1}(t)\\
&\quad=\sum_{\substack{A\subset\Omega_N\\|A|=j}}\int_0^1\prod_{i\in D}\big(\mathbf{1}_{[U_i,1)}(t)-t\big)(-1)^{|A^c\cap D|}t^{|A\cap D^c|}(1-t)^{|A^c\cap D^c|}\,dF^{-1}(t).
\end{aligned} \tag{2.7}$$

Next, write

$$\begin{aligned}
\sum_{i=1}^N c_iX_{i:N}&=\bar c\sum_{i=1}^N X_i+\sum_{i=1}^N(c_i-\bar c)X_{i:N}\\
&=\sum_{j=1}^{N-1}a_j\big(X_{j+1:N}-X_{j:N}\big)+\bar c\sum_{i=1}^N X_i,
\end{aligned} \tag{2.8}$$

where $a_j = -\sum_{i=1}^{j}(c_i - \bar{c})$, for $j = 1, \dots, N-1$, and 0 otherwise. Write $|A \cap D^c| = l$ and $|A \cap D| = m$, with $j = l + m$. For the first term on the right in (2.8), (2.7) yields

$$\begin{aligned}
&\left[\sum_{j=1}^{N-1} a_j (X_{j+1:N} - X_{j:N})\right]_D \\
&\quad = \int_0^1 \prod_{i \in D} \left(\mathbf{1}_{[U_i,1)}(t) - t\right) \sum_{l=0}^{N-|D|} \binom{N-|D|}{l} t^l (1-t)^{N-|D|-l} \\
&\qquad\qquad \times \sum_{m=0}^{|D|} \binom{|D|}{m} (-1)^{|D|-m} a_{l+m}\, dF^{-1}(t) \qquad (2.9)\\
&\quad = \int_0^1 \prod_{i \in D} \left(\mathbf{1}_{[U_i,1)}(t) - t\right) \sum_{l=0}^{N-|D|} \mathscr{B}(l; N-|D|, t)\, \Delta^{|D|}(a_l)\, dF^{-1}(t),
\end{aligned}$$

where $\mathscr{B}(l; N-|D|, t)$ denotes the probability that a binomial random variable with parameters $N - |D|$ and t equals l and $\Delta^{|D|}(a_l)$ is the $|D|$th difference of a_l, defined recursively by

$$\Delta(a_l) = a_{l+1} - a_l, \qquad \Delta^{\nu}(a_l) = \Delta\left(\Delta^{\nu-1}(a_l)\right).$$

Taking $D = \{i\}$ and $D = \{i, j\}$ and using (2.8) we find that

$$\begin{aligned}
T_{Ni} = -N^{-1/2} \sum_{l=1}^{N} c_l \int_0^1 &\left(\mathbf{1}_{[U_i,1)}(t) - t\right) \\
&\times \binom{N-1}{l-1} t^{l-1} (1-t)^{N-l}\, dF^{-1}(t),
\end{aligned} \qquad (2.10)$$

$$\begin{aligned}
T_{Nij} = -N^{-1/2} \sum_{l=2}^{N} (c_l - c_{l-1}) \int_0^1 &\left(\mathbf{1}_{[U_i,1)}(t) - t\right)\left(\mathbf{1}_{[U_j,1)}(t) - t\right) \\
&\times \binom{N-2}{l-2} t^{l-2} (1-t)^{N-l}\, dF^{-1}(t).
\end{aligned} \qquad (2.11)$$

Arguing as in van Zwet (1984) we find

$$E|T_{N1}|^p \le a^p 2^{p-1} E|X_1|^p N^{-(1/2)p}, \qquad (2.12)$$

$$E|T_{N12}|^r \le b^r 2^{2r} E|X_1|^r N^{-(3/2)r}. \qquad (2.13)$$

In Putter (1994) it is shown that

$$\sum_{k=3}^{N} \binom{N-2}{k-2} E T_{N\Omega_k}^2 \le 45 c^2 N^{-4} E X_1^2. \qquad (2.14)$$

Since by (2.12), (2.13) and (2.14), $\sigma^2(T_N)$ is bounded, application of Theorem 1.3 completes the proof of the corollary. □

APPLICATION 3 (Smooth functions of the sample mean). Let $X_1, \dots, X_N$ be i.i.d. mean zero random variables taking values in a real separable Banach space $\mathbf{B}$. Let $H\colon \mathbf{B} \to \mathbb{R}$ and define

$$T_N = \sqrt{N}\left(H\left(\overline{X}_N\right) - EH\left(\overline{X}_N\right)\right).$$

Let $H^{(s)}(x)$ denote the sth Fréchet derivative of H at the point $x \in \mathbf{B}$, where $H^{(s)}(x)h_1 \cdots h_s$ is the s-linear continuous symmetric form with arguments $h_1, \dots, h_s \in \mathbf{B}$. Define $\|H^{(s)}(x)\|$ to be the supremum of $H^{(s)}(x)h_1 \cdots h_s$ over all $h_1, \dots, h_s \in \mathbf{B}$ with $\|h_i\| = 1$ and let

$$\|H^{(s)}\|_\infty = \sup_{x \in \mathbf{B}} \left\| H^{(s)}(x) \right\|. \tag{2.15}$$

Since $\sum_{k=3}^N \binom{N-3}{k-3} ET_{N\Omega_k}^2 \le CN^{-5}$ implies $\sum_{k=3}^N \binom{N-2}{k-2} ET_{N\Omega_k}^2 \le CN^{-4}$, the following is a consequence of the results in Bentkus, Götze and van Zwet (1997) and Theorem 1.3.

COROLLARY 2.3. *Suppose that $H'(0)X_1$ satisfies Cramér's condition*

$$\limsup_{|t| \to \infty} \left| E\exp\{itH'(0)X_1\}\right| < 1 \tag{2.16}$$

and suppose that $E\|X_1\|^p < \infty$ for some $p > 3$. Suppose furthermore that H is three times Fréchet differentiable with $\sum_{i=1}^3 \|H^{(i)}\|_\infty$ finite. Then (2.3) *and* (2.4) *hold.*

APPLICATION 4 (Smooth functionals of the empirical distribution function). Let $X_1, \dots, X_N$ be real-valued i.i.d. random variables with common distribution function F. Let $F_N(x) = N^{-1}\sum_{i=1}^N \mathbf{I}_{\{X_i \le x\}}$ denote the empirical distribution function of $X_1, \dots, X_N$. Let $\mathbf{B}$ be the space of cadlag functions from $\mathbb{R}$ to $\mathbb{R}$, let H be a map from $\mathbf{B}$ to $\mathbb{R}$ and define

$$T_N = \sqrt{N}\left(H(F_N) - EH(F_N)\right).$$

Let $H^{(s)}(F)$ denote the sth Fréchet derivative of H at the point $F \in \mathbf{B}$ with s-linear continuous symmetric form $H^{(s)}(F)h_1 \cdots h_s$ for $h_1, \dots, h_s \in \mathbf{B}$. Define $\|H^{(s)}(F)\|$ to be the supremum of $H^{(s)}(F)h_1 \cdots h_s$ over all $h_1, \dots, h_s \in \mathbf{B}$ with $\|h_i\| = 1$ and define $\|H^{(s)}\|_\infty = \sup_{F \in \mathbf{B}} \|H^{(s)}(F)\|$ as in (2.15). The results in Bentkus, Götze and van Zwet (1994) and Theorem 1.3 imply the following corollary.

COROLLARY 2.4. *Suppose that $H'(F)(\mathbf{I}_{\{X_i \le x\}} - F(x))$ satisfies Cramér's condition as in* (2.16). *Suppose furthermore that H is three times Fréchet differentiable with $\sum_{i=1}^3 \|H^{(i)}\|_\infty$ finite. Then* (2.3) *and* (2.4) *hold.*

3. The bootstrap. The results in Section 1 have been formulated in such a way that the conclusions hold uniformly for all T_N and P satisfying the assumptions of the theorems for fixed constants and sequences (cf. the discussion following Theorem 1.3). This allows an application to the bootstrap. For every N, we take as our underlying distribution the empirical distribution P_N based on the observed sample $X_1, \dots, X_N$. Then we need to check the moment assumptions of the various terms in the Hoeffding decomposition under P_N. When the structure of T_N is not too complicated, it is possible to relate these moment assumptions under P_N to moment assumptions of the corresponding terms in the Hoeffding decomposition of T_N under P. We shall illustrate this for a U-statistic of degree two. This case has been studied earlier by Helmers (1991). For more complicated statistics, verification of the nonlattice condition (3.2) for the linear part in the Hoeffding decomposition of the bootstrap statistic may pose considerable problems.

Let $X_1, \dots, X_N$ be a sequence of i.i.d. random variables with common distribution P, let h: $\mathbb{R}^2 \to \mathbb{R}$ be a symmetric kernel with $Eh(X_1, X_2) = \theta$ and $Eh^2(X_1, X_2) < \infty$ and define the U-statistic

$$U_N = \binom{N}{2}^{-1} \sum\sum_{1 \le i < j \le N} h(X_i, X_j).$$

Let $T_N = \sqrt{N}(U_N - \theta)$ and define

$$\begin{aligned} g(x) &= E(h(X_1, X_2) | X_1 = x) - \theta, \\ \psi(x, y) &= h(x, y) - g(x) - g(y) - \theta. \end{aligned} \tag{3.1}$$

Hoeffding's decomposition of T_N is then given by

$$T_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} g(X_i) + \frac{2}{\sqrt{N}(N-1)} \sum\sum_{1 \le i < j \le N} \psi(X_i, X_j).$$

Let S_N^2 be the jackknife estimator of variance of T_N and define $F_N(x) = P(T_N / S_N \le x)$. To define the bootstrap approximation to F_N, let $X_1^*, \dots, X_N^*$ be an i.i.d. sample from the empirical distribution P_N and define

$$U_N^* = \binom{N}{2}^{-1} \sum\sum_{1 \le i < j \le N} h(X_i^*, X_j^*),$$

$$\theta_N^* = E(U_N^* | X_1, \dots, X_N) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} h(X_i, X_j),$$

$$T_N^* = \sqrt{N}(U_N^* - \theta_N^*), \qquad F_N^*(x) = P(T_N^* / S_N^* \le x | X_1, \dots, X_N),$$

where S_N^{*2} is given by

$$U_N^{(i)^*} = \binom{N}{2}^{-1} \sum\sum_{\substack{1 \le j < k \le N+1 \\ j, k \ne i}} h(X_j^*, X_k^*), \qquad T_N^{(i)^*} = \sqrt{N}\left(U_N^{(i)^*} - \theta_N^*\right),$$

$$\overline{T}_N^* = \frac{1}{N+1} \sum_{i=1}^{N+1} T_N^{(i)^*}, \qquad S_N^{*2} = \sum_{i=1}^{N+1} \left(\overline{T}_N^* - T_N^{(i)^*}\right)^2.$$

The Hoeffding decomposition of T_N^* can be expressed as

$$T_N^* = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_N(X_i^*) + \frac{2}{\sqrt{N}(N-1)} \sum\sum_{1 \le i < j \le N} \psi_N(X_i^*, X_j^*),$$

where the functions g_N and ψ_N are defined by

$$g_N(X_1^*) = E^*(h(X_1^*, X_2^*)|X_1^*) - \theta_N^*,$$
$$\psi_N(X_1^*, X_2^*) = h(X_1^*, X_2^*) - g_N(X_1^*) - g_N(X_2^*) - \theta_N^*.$$

Here E^* denotes expectation under P_N, conditionally given $X_1, \dots, X_N$. It is easily seen that the functions g_N and ψ_N may be expressed in terms of the functions g and ψ as follows:

$$g_N(X_1^*) = g(X_1^*) - \bar{g} + \bar{\psi}(X_1^*) - \bar{\bar{\psi}},$$
$$\psi_N(X_1^*, X_2^*) = \psi(X_1^*, X_2^*) - \bar{\psi}(X_1^*) - \bar{\psi}(X_2^*) + \bar{\bar{\psi}},$$

where

$$\bar{g} = N^{-1} \sum_{i=1}^{N} g(X_i), \qquad \bar{\psi}(x) = N^{-1} \sum_{i=1}^{N} \psi(x, X_i), \qquad \bar{\bar{\psi}} = N^{-1} \sum_{j=1}^{N} \bar{\psi}(X_j).$$

COROLLARY 3.1. *Suppose that the distribution of $g(X_1)$ is nonlattice and suppose that there exist $p > 3$ and $r > 2$ such that $E|g(X_1)|^p < \infty$, $E|\psi(X_1, X_2)|^r < \infty$ and $E|\psi(X_1, X_1)|^{r/2} < \infty$. Then*

$$\sqrt{N} \sup_{x \in \mathbb{R}} |F_N^*(x) - F_N(x)| \to_P 0.$$

PROOF. It is clear from the proof of Theorem 3 of Helmers (1991) that the nonlattice condition on the distribution of $g(X_1)$ implies almost surely a nonlattice condition on the distribution of $g_N(X_1^*)$, which is uniform for large N; that is, for every $0 < a < A < \infty$, there exists $\varepsilon > 0$ such that

$$\limsup_N \sup_{a \le |t| \le A} |E^* \exp(itg_N(X_1^*))| \le 1 - \varepsilon \quad \text{a.s.} \tag{3.2}$$

We proceed to show that $E^*|g_N(X_1^*)|^p$ is bounded in probability. Note that

$$E^*|g_N(X_1^*)|^p \le c_p\Big(E^*|g(X_1^*)|^p + |\bar{g}|^p + E^*|\bar{\psi}(X_1^*)|^p + |\bar{\bar{\psi}}|^p\Big).$$

Since $E|g(X_1)|^p < \infty$, we have by the law of large numbers,

$$E^*|g(X_1^*)|^p \to_P E|g(X_1)|^p, \qquad |\bar{g}|^p \to_P 0.$$

Next,

$$\begin{aligned} E^*|\bar{\psi}(X_1^*)|^p + |\bar{\bar{\psi}}|^p &= N^{-1} \sum_{i=1}^{N} \left| N^{-1} \sum_{j=1}^{N} \psi(X_i, X_j) \right|^p \\ &\quad + \left| N^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} \psi(X_i, X_j) \right|^p \\ &\le 2N^{-1} \sum_{i=1}^{N} \left| N^{-1} \sum_{j=1}^{N} \psi(X_i, X_j) \right|^p. \end{aligned} \tag{3.3}$$

It suffices therefore to show that $N^{-1}\sum_{i=1}^N |N^{-1}\sum_{j=1}^N \psi(X_i, X_j)|^p$ goes to zero in probability. Write

$$N^{-1}\sum_{i=1}^N \left|N^{-1}\sum_{j=1}^N \psi(X_i, X_j)\right|^p$$
$$= 2^{p-1}N^{-p-1}\sum_{i=1}^N |\psi(X_i, X_i)|^p + 2^{p-1}N^{-1}\sum_{i=1}^N \left|N^{-1}\sum_{j\neq i}\psi(X_i, X_j)\right|^p.$$

The first of these terms tends to zero almost surely, and a fortiori in probability, by the Marcinkiewicz strong law of large numbers (if $E|\psi(X_1, X_1)|^{p/(p+1)} < \infty$). Applying Lemma A.1 we see that there exists $\delta > 0$ and random variables $\tilde{\psi}(X_i, X_j)$, for $1 \le i < j \le N$ such that $|\tilde{\psi}(X_i, X_j)| \le N^{1-\delta}$ and

$$P\Big(\tilde{\psi}(X_i, X_j) = \psi(X_i, X_j), 1 \le i < j \le N\Big) = 1 - o(1).$$

It follows that

$$P\left(N^{-1}\sum_{i=1}^N \left|N^{-1}\sum_{j\neq i}\psi(X_i, X_j)\right|^p \ge \varepsilon\right)$$
$$= P\left(N^{-1}\sum_{i=1}^N \left|N^{-1}\sum_{j\neq i}\tilde{\psi}(X_i, X_j)\right|^p \ge \varepsilon\right) + o(1)$$
$$\le P\left(N^{(1-\delta)(p-2)-3}\sum_{i=1}^N \left(\sum_{j\neq i}\tilde{\psi}(X_i, X_j)\right)^2 \ge \varepsilon\right) + o(1)$$
$$= P\left(N^{(1-\delta)(p-2)-3}\sum_{i=1}^N \left(\sum_{j\neq i}\psi(X_i, X_j)\right)^2 \ge \varepsilon\right) + o(1)$$
$$\le \varepsilon^{-1}N^{(1-\delta)(p-2)-3}E\sum_{i=1}^N \left(\sum_{j\neq i}\psi(X_i, X_j)\right)^2 + o(1)$$
$$\le \varepsilon^{-1}N^{(1-\delta)(p-2)-1}E\psi^2(X_1, X_2) + o(1) \to 0$$

if $3 < p < 2 + (1-\delta)^{-1}$. It follows that there exists $p > 3$ such that

$$E^*\left|\bar{\psi}(X_1^*)\right|^p + \left|\bar{\bar{\psi}}\right|^p \to_P 0. \tag{3.4}$$

Next, we show that $E^*|\psi_N(X_1^*, X_2^*)|^r$ is bounded in probability. Note that

$$E^*\left|\psi_N(X_1^*, X_2^*)\right|^r \le c_r\Big(E^*\left|\psi(X_1^*, X_2^*)\right|^r + 2E^*\left|\bar{\psi}(X_1^*)\right|^r + |\bar{\bar{\psi}}|^r\Big).$$

First of all,

$$E^*\left|\psi(X_1^*, X_2^*)\right|^r = N^{-2}\sum_{i=1}^N\sum_{j=1}^N \left|\psi(X_i, X_j)\right|^r$$
$$\le N^{-2}\sum_{i=1}^N |\psi(X_i, X_i)|^r + \binom{N}{2}^{-1}\sum\sum_{1\le i<j\le N}\left|\psi(X_i, X_j)\right|^r.$$

The first of these terms tends to zero almost surely by the Marcinkiewicz strong law of large numbers, the second to $E|\psi(X_1, X_2)|^r < \infty$ by the strong law of large numbers for U-statistics. Finally, for $r \le p$, we have by (3.4),

$$2E^*\left|\overline{\psi}(X_1^*)\right|^r + |\overline{\overline{\psi}}|^r \to_P 0.$$

Application of Theorem 1.2 shows that uniformly in x,

$$F_N^*(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\left[(2x^2+1)\lambda_1^* + 3(x^2+1)\lambda_2^*\right] + \sigma(N^{-1/2}),$$

where λ_1^* and λ_2^* are given by

$$\lambda_1^* = E^* g_N^3(X_1^*), \qquad \lambda_2^* = E^* g_N(X_1^*) g_N(X_2^*) \psi_N(X_1^*, X_2^*).$$

Arguments, similar to the ones employed for the terms $E^*|g_N(X_1^*)|^p$ and $E^*|\psi_N(X_1^*, X_2^*)|^r$ show that $\lambda_i^* \to_P \lambda_i$ for $i = 1, 2$. Since the functions $(2x^2+1)\phi(x)$ and $(x^2+1)\phi(x)$ are bounded in x, the corollary is proved. □

4. Edgeworth expansions. The proofs of this section rely heavily on a result on Edgeworth expansions for U-statistics, obtained by Bickel, Götze and van Zwet (1986). Let $X_1, \dots, X_N$ be i.i.d. random variables assuming values in a measurable space $(\mathscr{X}, \mathscr{A})$ with common distribution P, and let h: $\mathscr{X} \times \mathscr{X} \to \mathbb{R}$ be a measurable function which is symmetric in its two arguments, with $Eh(X_1, X_2) = \theta$ and $Eh^2(X_1, X_2) < \infty$. Let

$$U_N = \binom{N}{2}^{-1} \sum\sum_{1 \le i < j \le N} h(X_i, X_j)$$

be a U-statistic of degree two. Define the functions g and ψ as in (3.1) and

$$\begin{aligned} &\lambda_1 = \sigma_g^{-3} E g^3(X_1), \\ &\lambda_2 = \sigma_g^{-3} E g(X_1) g(X_2) \psi(X_1, X_2), \qquad \kappa_3 = \lambda_1 + 3\lambda_2. \end{aligned} \tag{4.1}$$

Then $\kappa_3 N^{-1/2}$ serves as an approximation to the third cumulant of $U_N/\sigma(U_N)$. The (one-term) Edgeworth expansion for the distribution function of $U_N/\sigma(U_N)$ is given by

$$G_N(x) = \Phi(x) - \frac{\kappa_3}{6\sqrt{N}}(x^2 - 1)\phi(x). \tag{4.2}$$

The validity of this expansion was first proved by Janssen (1978) and Callaert, Janssen and Veraverbeke (1980). The best result to date has been obtained by Bickel, Götze and van Zwet (1986), who proved the following theorem.

THEOREM 4.1 [Bickel, Götze and van Zwet (1986)]. *Suppose that there exist real numbers* $C > 0$, $p > 3$, $r > 2$, *and a positive continuous function* χ *on* $(0, \infty)$, *such that*

$$E\left|g(X_1)\right|^p \le C; \tag{4.3}$$

$$E\left|\psi(X_1, X_2)\right|^r \le C; \tag{4.4}$$

and

$$|E\exp(itg(X_1))| \le 1 - \chi(t) < 1 \qquad \forall t > 0. \tag{4.5}$$

Then there exists a sequence $\varepsilon_N \searrow 0$, *depending only on* p, C, r *and* χ, *such that for* $N = 2, 3, \dots,$

$$\sup_{x\in\mathbb{R}} |P(U_N/\sigma(U_N) \le x) - G_N(x)| \le \varepsilon_N N^{-1/2}. \tag{4.6}$$

Bickel, Götze and van Zwet (1986) prove this theorem under a slightly weaker moment assumption on g, namely $E|g(X_1)|^3\mathbf{I}_{\{|g(X_1)|\ge t\}} \to 0$ as $t \to \infty$. This is implied by (4.3), since for $t > 0$ and $p > 3$,

$$\begin{aligned} E|g(X_1)|^3\mathbf{I}_{\{|g(X_1)|\ge t\}} &\le \{E|g(X_1)|^p\}^{3/p}\{P(|g(X_1)| \ge t)\}^{(p-3)/p} \\ &\le \{E|g(X_1)|^p\}^{3/p}\left\{\frac{E|g(X_1)|^p}{t^p}\right\}^{(p-3)/p} = \frac{E|g(X_1)|^p}{t^{p-3}}. \end{aligned}$$

Note that, like the results of Section 1, the present theorem is formulated in such a way that the conclusion is valid uniformly for any class (h, P) for which the assumptions are satisfied for fixed C, p, r and χ. The uniformity is important when we want to apply this result to symmetric statistics where the functions g and ψ in the Hoeffding decomposition of the statistic depend on N. It allows us to consider sequences $\{g_N\}$, $\{\psi_N\}$ and $\{P_N\}$ as long as (4.3)–(4.5) are satisfied for fixed constants C, p and r, and for a fixed function χ for every N.

Assumption (4.5) ensures that the distribution of $g(X_1)$ is nonlattice. However, it is clear from the proof of Bickel, Götze and van Zwet (1986) that the behavior of the characteristic function of $g = g_N$ is irrelevant for $t > \tau_N$, if $\tau_N \to \infty$. Thus, assumption (4.5) can actually be relaxed. Sufficient is the existence of a sequence $\tau_N \to \infty$ and a positive continuous function χ on $(0, \infty)$, such that

$$|E\exp(itg_N(X_1))| \le 1 - \chi(t) \qquad \forall\, t \in (0, \tau_N) \text{ for all } N. \tag{4.7}$$

Of course the sequence $\{\varepsilon_N\}$ in Theorem 4.1 will then also depend on the sequence $\{\tau_N\}$. In particular, in Theorem 1.2 we require the less restrictive assumption (4.7). This is because the proof employs truncation of the linear term $g_N(X_1)$ (corresponding to $N^{1/2}T_{N1}$), which will destroy (4.5), while (4.7) is still fulfilled for the truncated $g_N(X_1)$ if it is fulfilled for the original $g_N(X_1)$ (cf. the discussion after Corollary 4.2).

Generalization of Theorem 4.1 to symmetric statistics is now straightforward.

PROOF OF THEOREM 1.1. By (1.15) and Chebyshev's inequality,

$$P\left(\left|\sum_{|D|\ge 3} T_{ND}\right| \ge \delta_N^{1/3}N^{-1/2}\right) \le \frac{\delta_N N^{-3/2}}{\delta_N^{2/3}N^{-1}} \le \delta_N^{1/3}N^{-1/2},$$

so that this part can be neglected. What remains is a U-statistic of degree two and we can apply Theorem 4.1 using the correspondence

$$N^{1/2}T_{Ni} = g(X_i), \qquad N^{3/2}T_{Nij} = \psi(X_i, X_j).$$

This proves Theorem 1.1. □

Before we set out to prove Theorem 1.2, some preliminary remarks are in order. To prove Edgeworth expansions under weak moment conditions, truncation is a well-established technique. Lemma A.1, stated and proved in the Appendix, is the truncation lemma we shall find useful. First we apply it to the random variables T_{Ni}. Note that for our purposes events with probability $o(N^{-1/2})$ may be neglected. We have $E|T_{Ni}|^p \le CN^{-p/2}$ for some $p > 3$, by (1.13). In Lemma A.1 we therefore choose $s = p > 3$ and $\eta = N^{-(p-3)/4p}$ to obtain the following corollary.

COROLLARY 4.2. *Suppose that $E|N^{1/2}T_{N1}|^p \le C$ for some $p > 3$. Then there exist i.i.d. $T'_{N1}, \dots, T'_{N,N+1}$ with $T'_{Ni} = \varrho_N(T_{Ni})$ satisfying*

$$|T'_{Ni}| \le N^{-(p-3)/4p} = o(1); \tag{4.8}$$

$$\begin{aligned} P(T'_{Ni} = T_{Ni}, i = 1, \dots, N+1) &\ge 1 - 2C\frac{N+1}{N}N^{(p-1)/4} \\ &\ge 1 - 4CN^{-(p-1)/4} = 1 - o(N^{-1/2}); \end{aligned} \tag{4.9}$$

$$ET'_{Ni} = 0; \tag{4.10}$$

$$\begin{aligned} N^{t/2}E|T'_{Ni} - T_{Ni}|^t &\le 2^{2t+1}CN^{-((p+3)(p-t)/4p)} \\ &= \begin{cases} o(N^{-(3(p-t)/2p)}), & 0 < t < p, \\ \mathcal{O}(1), & t = p. \end{cases} \end{aligned} \tag{4.11}$$

Throughout the proofs in the sequel we shall assume that the T_{Ni} have been replaced by their truncated versions T'_{Ni}. For simplicity we delete the prime in T'_{Ni}, thus in effect assuming that (4.8), (4.10) and (4.11) are satisfied for the original T_{Ni}.

In the formulation of the theorems, however, both the assumptions and the conclusions are stated in terms of the original T_{Ni}. We should therefore make sure in the first place that the assumptions for the original T_{Ni} imply the same assumptions for the truncated T_{Ni}. Secondly, having conducted the proof with the truncated random variables, we obtain a conclusion for the truncated random variables, and hence we also have to check that the conclusion of the theorem is still true when we replace the truncated T_{Ni} by the original T_{Ni}.

To see that all this is justified in Theorem 1.2, we note that the probability that the substitution affects the values of any of the T_{Ni}, and hence of T_N or S_N^2, is $o(N^{-1/2})$ uniformly in P and T_N satisfying (1.13). Exceptional events with probabilities of this order of magnitude are allowed in all our results.

It remains to check that the substitution does not affect the assumptions or conclusions of our results in any other way. Clearly, (4.11) guarantees that (1.13) implies that $E|N^{1/2}T'_{N1}|^p \le 2^{p-1}C(2^{2p+1}+1)$, so that assumption (1.13) is satisfied for T'_{N1}. In Theorem 1.2 we shall also encounter the assumption that there exist a sequence $\tau_N \to \infty$ and a positive function χ on $(0,\infty)$, such that $|E\exp\{itN^{1/2}T_{N1}\}| \le 1-\chi(t)$ for all $t \in (0,\tau_N)$ for all N. Now (4.9) and (4.11) imply

$$\begin{aligned}\left|E\exp\left\{itN^{1/2}T'_{N1}\right\} - E\exp\left\{itN^{1/2}T_{N1}\right\}\right| &\le 8CN^{-(p-1)/4}\\ &\le 8CN^{-1/2} \quad \text{for all } t,\end{aligned}$$

and

$$\left|E\exp\left\{itN^{1/2}T'_{N1}\right\}\right| \le 1 - 1/3NET'^2_{N1}t^2 \qquad \text{if } |t| \le \frac{E\left(N^{1/2}T'_{N1}\right)^2}{E|N^{1/2}T'_{N1}|^3}.$$

Since the remaining assumptions ensure that $NET^2_{N1} \sim \sigma^2_N \ge c > 0$, and (4.11) yields

$$N|ET'^2_{N1} - ET^2_{N1}| \le 32CN^{-((p+3)(p-2)/4p)}, \tag{4.12}$$

we find that $E(N^{1/2}T'_{N1})^2$ and $E|N^{1/2}T'_{N1}|^3$ are bounded away from zero and infinity. Hence, for some $t_0 > 0$ and integer N_0, $(E(N^{1/2}T'_{N1})^2/E|N^{1/2}T'_{N1}|^3) \ge t_0$ for all $N \ge N_0$. Now, for $0 < t < \tau_N$,

$$\left|E\exp\left\{itN^{1/2}T'_{N1}\right\}\right| \le 1-\chi(t) + 8CN^{-1/2} \le 1-\chi(t)/2,$$

if t is such that $\chi(t) \ge 16CN^{-1/2}$. Since χ is a positive continuous function, it has a positive minimum on every closed interval $K = [t_0, \tau]$, and hence we may choose a $\tau = \tau'_N < \tau_N$ for sufficiently large N, so that the minimum over K of $\chi(t)$ is still larger than $16CN^{-1/2}$. The fact that $16CN^{-1/2} \to 0$ as $N \to \infty$, allows us to choose τ'_N tending to infinity, as N tends to infinity. Defining a function

$$\tilde{\chi}(t) = \min\left(\frac{\chi(\mathrm{t})}{2}, \frac{\mathrm{c}}{4}\mathrm{t}^2\right),$$

it follows that there exists a sequence $\tau'_N \to \infty$, such that for sufficiently large $N \ge N_0$, $|E\exp\{itN^{1/2}T'_{N1}\}| \le 1-\tilde{\chi}(t)$ for all $t \in (0,\tau'_N)$. It is easy to see that the function $\tilde{\chi}$, the sequence $\{\tau'_N\}$ and N_0 depend only on P and $\{T_N\}$ through the various constants and sequences in our assumptions (including $\{\tau_N\}$) and this is therefore all we need.

Finally, in the conclusion of Theorem 1.2, the quantities σ^2_N, $\lambda_1 = N^{3/2}\sigma_N^{-3}ET^3_{N1}$ and $\lambda_2 = N^{5/2}\sigma_N^{-3}ET_{N1}T_{N2}T_{N12}$ occur. Application of (4.11) shows that changing T'_{Ni} into T_{Ni} affects σ^2_N, λ_1 and λ_2 only to order $o(N^{-1/2})$, $o(1)$ and $o(1)$, respectively, which does not alter our conclusions.

Summarizing, we conclude that in the proof of Theorem 1.2 *we may assume without loss of generality that* (4.8), (4.10) *and* (4.11) *hold for the original* T_{Ni}.

We noted earlier that all results (though they should be viewed as asymptotic results) are formulated as inequalities for fixed, but arbitrary N. In the proofs, we should then formally also be working with inequalities that are true for every N. Phrases like $o(N^{-1})$ are, strictly speaking, not allowed, because there is no N tending to infinity. To work with inequalities throughout the proofs would, however, become extremely tedious. To avoid these laborious formulations from occurring throughout the proofs and the more informal passages in the text, we shall use o and $\mathcal{O}$ symbols all the same, agreeing that they are uniform in everything that satisfies the assumptions for fixed constants and sequences. Thus, the statement $a_N = o(N^{-1})$ for example, should be read as $Na_N \le \varepsilon_N$, for some sequence ε_N tending to zero, depending only on the constants and sequences in the assumptions. Similarly, $a_N = \mathcal{O}(N^{-1})$ will stand for $Na_N \le C$, for some constant C, depending only on the constants and sequences in the assumptions.

In order to prove Theorem 1.2 we shall need the following lemma.

LEMMA 4.3. *Suppose that there exist numbers* $p > 3$, $r > 2$, $c > 0$ *and* $C > 0$, *such that for* $N = 2, 3, \dots,$ (1.2), (1.3), (1.13), (1.14) *and* (1.18) *are satisfied. Then there exist sequences* $\delta_N \searrow 0$ *and* $\varepsilon_N \searrow 0$, *depending on* T_N *and* P *only through* p, r, c *and* C, *such that for* $N = 2, 3, \dots,$

$$P\left(\left|S_N^2 - \sigma_N^2 - \sum_{i=1}^{N+1} (K_{Ni} + L_{Ni}) - \sum\sum_{1 \le i < j \le N+1} M_{Nij}\right| \ge \varepsilon_N N^{-1/2}\right) \le \delta_N N^{-1/2}, \tag{4.13}$$

where

$$K_{Ni} = f_{N1}(X_i) = 2NE(T_{Nij}T_{Nj}|X_i), \qquad j \ne i, \tag{4.14}$$

$$L_{Ni} = f_{N2}(X_i) = T_{Ni}^2 - ET_{N1}^2, \tag{4.15}$$

and for $i \ne j$,

$$\begin{aligned} M_{Nij} &= f_{N3}(X_i, X_j) \\ &= 2T_{Nij}(T_{Ni} + T_{Nj}) - 2E(T_{Nij}T_{Nj}|X_i) - 2E(T_{Nij}T_{Ni}|X_j). \end{aligned} \tag{4.16}$$

Lemma 4.3 should be compared for instance with identity (A8) in Callaert and Veraverbeke (1981). It asserts essentially that the difference between S_N^2 and σ_N^2 can be expressed as a U-statistic plus a remainder term which is of negligible order for our purposes. For a proof we refer to Putter (1994).

We are now prepared to prove Theorem 1.2.

PROOF OF THEOREM 1.2. In view of Corollary 4.2, we assume that $|T_{Ni}| \le a_N = N^{-(p-3)/4p}$. Let V_N denote the U-statistic

$$V_N = \sum_{i=1}^{N+1} (K_{Ni} + L_{Ni}) + \sum\sum_{1 \le i < j \le N+1} M_{Nij}.$$

Under the assumptions of the present theorem, Lemma 4.3 asserts that

$$P(|S_N^2 - \sigma_N^2 - V_N| \ge \varepsilon_N N^{-1/2}) \le \delta_N N^{-1/2} \tag{4.17}$$

for sequences $\varepsilon_N \searrow 0$ and $\delta_N \searrow 0$.

We begin by bounding the moments of the terms in V_N. We have

$$|K_{Ni}| \le 2N\Big[E\Big(|T_{Nij}|^{p/(p-1)}|X_i\Big)\Big]^{(p-1)/p}\Big[E|T_{Nj}|^p\Big]^{1/p}$$
$$\le 2C^{1/p}N^{1/2}\Big[E\Big(|T_{Nij}|^{p/(p-1)}|X_i\Big)\Big]^{(p-1)/p} \quad \text{for } i \ne j.$$

Since $r(p-1)/p > 1$ for $r > 2$ and $p > 3$, this implies that

$$E|K_{N1}|^r \le 2^r C^{r/p} N^{r/2} E|T_{N12}|^r \le 2^r C^{1+r/p} N^{-r}, \tag{4.18}$$

and in view of an inequality of Dharmadhikari, Fabian and Jogdeo (1968),

$$E\left|\sum_{i=1}^{N+1} K_{Ni}\right|^r \le C' N^{-r/2}, \tag{4.19}$$

for an appropriate constant C'. Similarly,

$$E|L_{N1}|^{p/2} \le 2^{p/2} E|T_{N1}|^p \le 2^{p/2} C N^{-p/2}, \tag{4.20}$$

$$E|L_{N1}|^t \le 2^t E|T_{N1}|^{2t} \le 2^t C a_N^{2t-p} N^{-p/2} \quad \text{for } t > p/2, \tag{4.21}$$

$$E\left(\sum_{i=1}^{N+1} L_{Ni}\right)^2 \le C' a_N^{4-p} N^{-(p-2)/2} \quad \text{if } p < 4, \tag{4.22}$$

$$E|M_{N12}|^r \le 2^{3r} E|T_{N1}T_{N12}|^r \le 2^{3r} C a_N^r N^{-3r/2}, \tag{4.23}$$

$$E\left|\sum\sum_{1\le i<j\le N+1} M_{Nij}\right|^r \le C' a_N^r N^{-r/2}. \tag{4.24}$$

Now, first, (4.19), (4.22) and (4.24) imply that $EV_N^2 = o(N^{-1/2})$, so by (1.2) and (4.17) we find that for every $\varepsilon > 0$,

$$P\left(\left|\frac{S_N^2 - \sigma_N^2}{\sigma_N^2}\right| \ge \varepsilon\right) = o(N^{-1/2}).$$

Since also

$$P\left(\left|\sum_{|D|\ge 3} T_{ND}\right| \ge \varepsilon_N N^{-1/2}\right) \le \frac{\sum_{k=3}^N \binom{N}{k} E T_{N\Omega_k}^2}{\varepsilon_N^2 N^{-1}} = \mathscr{O}\left(\frac{1}{\varepsilon_N^2 N}\right),$$

we find that there exist $\varepsilon_N \searrow 0$, such that

$$P\left(\left|\frac{T_N}{S_N} - \frac{\sum_{|D|\le 2} T_{ND}}{S_N}\right| \ge \varepsilon_N N^{-1/2}\right) = o(N^{-1/2}),$$

and

$$F_N(x) = P\left(\frac{T_N}{S_N} \le x\right) \le P\left(\frac{\sum_{|D|\le 2} T_{ND}}{S_N} \le x + o(N^{-1/2})\right) + o(N^{-1/2}),$$

$$F_N(x) = P\left(\frac{T_N}{S_N} \le x\right) \ge P\left(\frac{\sum_{|D|\le 2} T_{ND}}{S_N} \le x - o(N^{-1/2})\right) + o(N^{-1/2}).$$

We have to show that $\sup_x |F_N(x) - H_N(x)| = o(N^{-1/2})$, and as $H_N(x)$ has bounded derivative, this is equivalent to showing the same thing with F_N replaced by the distribution function of $\sum_{|D| \le 2} T_{ND}/S_N$. Hence T_N may be replaced by the U-statistic

$$U_N = \sum_{i=1}^{N} T_{Ni} + \sum\sum_{1 \le i < j \le N} T_{Nij},$$

and $F_N(x)$ by $P(U_N/S_N \le x)$.

Next we write

$$P\left(\frac{U_N}{S_N} \le x\right) = P\left(\frac{U_N}{\sigma_N} \le x\frac{S_N}{\sigma_N}\right).$$

Because $xH_N'(x)$ is bounded, $H_N(x(1 + o(N^{-1/2}))) = H_N(x) + o(N^{-1/2})$, and in view of (4.17) we may replace S_N^2 in $P\ (U_N/\sigma_N \le x(S_N/\sigma_N))$ by $\sigma_N^2 + V_N$. Therefore, we may replace $F_N(x)$ by

$$P\left(\frac{U_N}{\sigma_N} \le x\frac{\left(\sigma_N^2 + V_N\right)^{1/2}}{\sigma_N}\right) = P\left(\frac{U_N}{\sigma_N} - x\left\{\left(1 + \frac{V_N}{\sigma_N^2}\right)^{1/2} - 1\right\} \le x\right).$$

For $|z| \le 4/5$ we have $1 + (z/2) - (z^2/4) \le (1 + z)^{1/2} \le 1 + z/2$ and hence we have, with probability $1 - o(N^{-1/2})$,

$$\frac{V_N}{2\sigma_N^2} - \frac{V_N^2}{4\sigma_N^4} \le \left(1 + \frac{V_N}{\sigma_N^2}\right)^{1/2} - 1 \le \frac{V_N}{2\sigma_N^2}.$$

As

$$V_N^2 \le 3\left\{\left(\sum_{i=1}^{N+1} K_{Ni}\right)^2 + \left(\sum_{i=1}^{N+1} L_{Ni}\right)^2 + \left(\sum\sum_{1 \le i < j \le N+1} M_{Nij}\right)^2\right\}$$

and

$$P\left(\left(\sum_{i=1}^{N+1} K_{Ni}\right)^2 + \left(\sum\sum_{1 \le i < j \le N+1} M_{Nij}\right)^2 \ge \varepsilon_N N^{-1/2}\right) \le \delta_N N^{-1/2}$$

for some $\varepsilon_N \searrow 0$ and $\delta_N \searrow 0$ by (4.19) and (4.24), we see that

$$\frac{V_N}{2\sigma_N^2} - 3\frac{\left(\sum_{i=1}^{N+1} L_{Ni}\right)^2}{4\sigma_N^4} + o(N^{-1/2}) \le \left(1 + \frac{V_N}{\sigma_N^2}\right)^{1/2} - 1 \le \frac{V_N}{2\sigma_N^2}$$

with probability $1 - o(N^{-1/2})$. By the same argument as before, it follows that we have to show that

$$\sup_x \left| P\left(\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2} \le x\right) - H_N(x)\right| = o(N^{-1/2}) \tag{4.25}$$

and

$$\sup_x \left| P\left(\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2} + \frac{3x\left(\sum_{i=1}^{N+1} L_{Ni}\right)^2}{4\sigma_N^4} \le x\right) - H_N(x)\right| = o(N^{-1/2}). \tag{4.26}$$

The first of these is a consequence of Theorem 4.1. Obviously $(U_N/\sigma_N) - (xV_N/2\sigma_N^2)$ is a U-statistic of degree two. Moreover, because $a_N = N^{-(p-3)/4p}$ and $p > 3$, (4.23) implies

$$E|N^{3/2} x M_{N12}|^r \le C' a_N^r (\log N)^r = o(1) \quad \text{if } |x| \le \log N.$$

Because of (4.21), we find that for $|x| \le \log N$,

$$E|N^{1/2} x L_{N1}|^p \le 2^p C a_N^p (\log N)^p = o(1).$$

At this point we apply Lemma A.1 again to $K_{Ni} = 2NE(T_{Nij}T_{Nj}|X_i)$. By (4.18), $E|K_{N1}|^r \le C' N^{-r}$. In Lemma A.1 we now choose $s = r > 2$ and $\eta = N^{-1/4-((r-2)/2r)}$. It follows that there exist i.i.d. $K'_{N1}, \dots, K'_{N,N+1}$ with $K'_{Ni} = \varrho_N(K_{Ni})$ satisfying

$$|K'_{Ni}| \le N^{-1/4-((r-2)/2r)}; \tag{4.27}$$

$$\begin{aligned} P(K'_{Ni} = K_{Ni},\ i = 1, \dots, N+1) &\\ \ge 1 - 2C' \frac{N+1}{N} N^{-r/4} & \\ \ge 1 - 4C' N^{-r/4} = 1 - o(N^{-1/2}); & \end{aligned} \tag{4.28}$$

$$EK'_{Ni} = 0; \tag{4.29}$$

$$\begin{aligned} N^t E|K'_{Ni} - K_{Ni}|^t &\le 2^{2t+1} C N^{-((r+4)(r-t)/4r)} \\ &= \begin{cases} o(N^{-(3(r-t))/2r}), & 0 < t < r, \\ \mathcal{O}(1), & t = r. \end{cases} \end{aligned} \tag{4.30}$$

We shall treat the K_{Ni} as we did the T_{Ni}, that is, we shall delete the prime and assume throughout the proof of this theorem that the original K_{Ni} satisfy (4.27), (4.29) and (4.30). The probability that this makes any difference is $o(N^{-1/2})$. The only place where the K_{Ni} occur in the conclusion of the proof is in $\lambda_2 = N^{5/2} ET_{N1}T_{N2}T_{N12} = N^{3/2} ET_{N1}K_{N1}$, and we have to check that

$$N^{3/2} |ET_{N1}K_{N1} - ET_{N1}K'_{N1}| \to 0.$$

This is a consequence of (1.13) and (4.30), since

$$\begin{aligned} N^{3/2} |ET_{N1}K_{N1} - ET_{N1}K'_{N1}| &\le N^{3/2} \big|ET_{N1}(K_{N1} - K'_{N1})\big| \\ &\le N^{3/2} \big(ET_{N1}^2\big)^{1/2} \Big(E(K_{N1} - K'_{N1})^2\Big)^{1/2} \\ &= o(N^{-(3(r-2))/4r}). \end{aligned}$$

Now we find that $E|N^{1/2} x K_{Ni}|^{p'} = o(1)$, for $|x| \le \log N$ and any $p' \le 3r$. It follows that for $|x| \le \log N$, $(U_N/\sigma_N) - (xV_N/2\sigma_N^2)$ satisfies assumptions (4.3) and (4.4) of Theorem 4.1.

To check assumption (4.7) for $(U_N - (x/2\sigma_N)V_N)$, we have to consider the characteristic function $\psi(t) = E\exp\{itN^{1/2}\tilde{T}_{N1}\}$, where $\tilde{T}_{N1} = T_{N1} -$

$(x/2\sigma_N)(K_{N1}+L_{N1})$. Of course

$$|\psi(t)| \le 1 - 1/3E\left(N^{1/2}\tilde{T}_{N1}\right)^2 t^2 \quad \text{if } |t| \le \frac{E\left(N^{1/2}\tilde{T}_{N1}\right)^2}{E|N^{1/2}\tilde{T}_{N1}|^3}.$$

We have shown above that $E|N^{1/2}x(K_{N1}+L_{N1})|^p \to 0$ uniformly in $|x| \le \log N$ for some $p > 3$. Also $E(N^{1/2}T_{N1})^2$ and $E|N^{1/2}T_{N1}|^3$ are bounded away from zero and infinity by (1.2), (1.13) and (1.18). Hence there exist $t_0 > 0$ and N_0 such that

$$|\psi(t)| \le 1 - c/4t^2 \quad \text{for } 0 < |t| \le t_0 \text{ and } N \ge N_0.$$

Next we note that

$$|\psi(t)| \le 1 - \chi(t) + \frac{|t|}{2\sigma_N} E\left|N^{1/2}x(K_{N1}+L_{N1})\right|.$$

The expectation on the right tends to zero uniformly for $x \le \log N$. Since $\chi(t)/t$ is positive and continuous on $(0,\infty)$, we find that there exist $\tau_N' \to \infty$, $\tau_N' < \tau_N$, such that

$$|\psi(t)| \le 1 - 1/2\chi(t) \quad \text{for } t_0 \le |t| \le \tau_N' \text{ and } N \ge N_1.$$

It follows that assumption (4.7) is satisfied uniformly in $|x| \le \log N$. Hence, by Theorem 4.1,

$$\sup_{|x|\le \log N}\left|P\left(\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2} \le x\right) - \tilde{G}_N(x)\right| = o(N^{-1/2}),$$

uniformly in all T_N and P satisfying our assumptions. Here

$$\tilde{G}_N(x) = \Phi\left(\frac{x}{\sigma_x}\right) - \frac{\kappa_3}{6}N^{-1/2}(x^2-1)\phi(x),$$

$$\begin{aligned} \sigma_x^2 &= \sigma^2\left(\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2}\right) \\ &= 1 - \frac{Nx}{\sigma_N^3}\left(ET_{N1}^3 + 2NET_{N12}T_{N1}T_{N2}\right) + o(N^{-1/2}), \end{aligned} \tag{4.31}$$

$$\kappa_3 = \sigma_N^{-3}\left(N^{3/2}ET_{N1}^3 + 3N^{5/2}ET_{N12}T_{N1}T_{N2}\right),$$

and it follows that $\tilde{G}_N(x)$ may be replaced by

$$H_N(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{N}}\left[\lambda_1(2x^2+1) + 3\lambda_2(x^2+1)\right],$$

$$\lambda_1 = N^{3/2}\sigma_N^{-3}ET_{N1}^3, \qquad \lambda_2 = N^{5/2}\sigma_N^{-3}ET_{N12}T_{N1}T_{N2}.$$

For $x \le -\log N$, $H_N(x) = \mathscr{O}(N^{-c})$, and for $x \ge \log N$, we have that $1 - H_N(x) = \mathscr{O}(N^{-c})$ for every $c > 0$, so monotonicity of a distribution function implies (4.25)

$$\sup_x \left| P\left(\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2} \le x \right) - H_N(x) \right| = o(N^{-1/2}).$$

To prove (4.26) we have to show that the presence of the term $3x(\sum_{i=1}^{N+1} L_{Ni})^2/(4\sigma_N^4)$ in (4.26) does not influence the expansion. As before, it is sufficient to consider $|x| \le \log N$. Without loss of generality we assume that $3 < p < 4$. Consider Hoeffding's decomposition of $x(\sum_{i=1}^{N+1} L_{Ni})^2$:

$$\begin{aligned} x\left(\sum_{i=1}^{N+1} L_{Ni} \right)^2 &= x(N+1)EL_{N1}^2 + x\sum_{i=1}^{N+1} \left(L_{Ni}^2 - EL_{N1}^2 \right) \\ &\quad + 2x \sum\sum_{1 \le i < j \le N+1} L_{Ni}L_{Nj}. \end{aligned} \tag{4.32}$$

By (4.22) the constant term satisfies

$$|x|(N+1)EL_{N1}^2 = \mathscr{O}\left(a_N^{4-p} N^{-(p-2)/2} \log N \right) = o(N^{-1/2}),$$

so it does not influence the expansion. By (4.21),

$$\begin{aligned} E\left| N^{1/2} x\left(L_{N1}^2 - EL_{N1}^2 \right) \right|^p &\le \left(2N^{1/2} \log N \right)^p E|L_{N1}|^{2p} \\ &\le 2^{3p} C a_N^{3p} (\log N)^p = o(1) \end{aligned} \tag{4.33}$$

and for $r' = 2p/3 > 2$, (4.21) ensures that

$$\begin{aligned} E|N^{3/2} 2x L_{N1} L_{N2}|^{r'} &= \left(2N^{3/2} \log N \right)^{r'} \left(E|L_{N1}|^{r'} \right)^2 \\ &= \mathscr{O}\left(a_N^{2(2r'-p)} (\log N)^{r'} \right) = o(1). \end{aligned} \tag{4.34}$$

It follows that the terms $x(L_{Ni}^2 - EL_{N1}^2)$ and $2xL_{Ni}L_{Nj}$ in the Hoeffding decomposition (4.32) satisfy the same assumptions (1.13) and (1.14) as T_{Ni} and T_{Nij}. Also the presence of a term $(L_{N1}^2 - EL_{N1}^2)$ cannot affect the bound (1.16) on the characteristic function for the same reason that the presence of K_{N1} and L_{N1} cannot. Hence,

$$\frac{U_N}{\sigma_N} - \frac{xV_N}{2\sigma_N^2} + \frac{3x\left(\sum_{i=1}^{N+1} L_{Ni} \right)^2}{4\sigma_N^4} \tag{4.35}$$

has an Edgeworth expansion with remainder $o(N^{-1/2})$ uniformly for $|x| \le \log N$. This Edgeworth expansion is of the same form as the expansion $H_N(x)$ in (4.25), and in view of (4.33) and (4.34), the presence of the term $3x(\sum_{i=1}^{N+1} L_{Ni})^2/(4\sigma_N^4)$ can only influence the expansion through the variance of the random variable (4.35) [cf. (4.31)]. However, the term of largest order that can contribute to this variance is of order

$$\begin{aligned} |x|N\left(\left| \operatorname{cov}\left(T_{N1}, L_{N1}^2 \right) \right| \right) &= \mathscr{O}\left(N(\log N) E|T_{N1}|^5 \right) \\ &= \mathscr{O}\left(a_N^{5-p} (\log N) N^{-(p-2)/2} \right) = o(N^{-1/2}), \end{aligned}$$

which does not change the expansion $H_N(x)$. This proves (4.26) and the theorem. □

5. Empirical Edgeworth expansions. In this section we prove the consistency of the jackknife estimates $\hat{\lambda}_1$ and $\hat{\lambda}_2$ [as in (1.23) and (1.24)] of the quantities λ_1 and λ_2 appearing in the Edgeworth expansions $G_N(x)$ and $H_N(x)$ of the distribution functions of T_N/σ_N and T_N/S_N, respectively, thus validating the empirical Edgeworth expansion.

For the proofs we again assume that $T_N = t_N(X_1, \dots, X_N; P)$ is a symmetric random variable with $ET_N = 0$, $ET_N^2 < \infty$ and $0 < c \le \sigma^2(T_N) \le C < \infty$ for some positive numbers c and C.

LEMMA 5.1. *Suppose that there exist constants $p > 3$, $c > 0$ and $C > 0$ such that* (1.2) *and* (1.3) *are satisfied*, $E|N^{1/2}T_{N1}|^p \le C$, *and* $\sum_{k=2}^{N}\binom{N-1}{k-1}ET_{N\Omega_k}^2 \le CN^{-2}$. *Then there exist sequences $\delta_N \searrow 0$ and $\varepsilon_N \searrow 0$, which depend only on p, c and C, such that for $N = 2, 3, \dots$,*

$$P\left(N^{1/2}\left|\sum_{i=1}^{N+1}\left(\overline{T}_N - T_N^{(i)}\right)^3 - NET_{N1}^3\right| \ge \varepsilon_N\right) \le \delta_N.$$

PROOF. Obviously all we have to prove is that for every $\varepsilon > 0$,

$$P\left(N^{1/2}\left|\sum_{i=1}^{N+1}\left(\overline{T}_N - T_N^{(i)}\right)^3 - NET_{N1}^3\right| \ge \varepsilon\right) \to 0$$

uniformly for fixed p, c and C. Write $\overline{T}_N - T_N^{(i)} = V_{Ni} + W_{Ni}$, with

$$V_{Ni} = \sum_{\substack{D\subset\Omega_{N+1}\\ |D|=1}}\left(\mathbf{1}_D(i) - \frac{|D|}{N+1}\right)T_{ND}$$

and

$$W_{Ni} = \sum_{\substack{D\subset\Omega_{N+1}\\ |D|\ge 2}}\left(\mathbf{1}_D(i) - \frac{|D|}{N+1}\right)T_{ND}.$$

Then V_{Ni} can be expressed as $T_{Ni} - (1/(N+1))\sum_{i=1}^{N+1}T_{Ni} = T_{Ni} - \Delta_N$, and an inequality of Dharmadhikari, Fabian and Jogdeo (1968) ensures that $E|\Delta_N|^3 = \mathscr{O}(N^{-3})$. Expansion of $\sum_{i=1}^{N+1}(\overline{T}_N - T_N^{(i)})^3$ yields

$$\sum_{i=1}^{N+1}\left(\overline{T}_N - T_N^{(i)}\right)^3 = \sum_{i=1}^{N+1}V_{Ni}^3 + 3\sum_{i=1}^{N+1}V_{Ni}^2W_{Ni} + 3\sum_{i=1}^{N+1}V_{Ni}W_{Ni}^2 + \sum_{i=1}^{N+1}W_{Ni}^3.$$

First we show that

$$N^{1/2}\left(\sum_{i=1}^{N+1}V_{Ni}^3 - \sum_{i=1}^{N+1}T_{Ni}^3\right) \to_P 0.$$

This follows easily from the fact that for $i = 1, \dots, N+1$,

$$\begin{aligned} E|V_{Ni}^3 - T_{Ni}^3| &= E|-3T_{Ni}^2\Delta_N + 3T_{Ni}\Delta_N^2 - \Delta_N^3| \\ &\le 3\left(E|T_{N1}|^3\right)^{2/3}\left(E|\Delta_N|^3\right)^{1/3} + 3\left(E|T_{N1}|^3\right)^{1/3}\left(E|\Delta_N|^3\right)^{2/3} + E|\Delta_N|^3 \\ &= \mathscr{O}(N^{-2}). \end{aligned}$$

Since $E|N^{1/2}T_{N1}|^3 \le C^{3/p}$, it also follows that $E|N^{1/2}V_{N1}|^3 \le C_1$ for some $C_1 > 0$ and all N.

Next, define

$$R_{N1} = N^{1/2} \sum_{i=1}^{N+1} W_{Ni}^3, \qquad R_{N2} = N^{1/2} \sum_{i=1}^{N+1} V_{Ni} W_{Ni}^2,$$
$$R_{N3} = N^{1/2} \sum_{i=1}^{N+1} V_{Ni}^2 W_{Ni}.$$

Since $E|N^{1/2}V_{N1}|^3 \le C_1$ and $E(NW_{N1})^2 \le N^2 \sum_{k=2}^{N} \binom{N-1}{k-1} ET_{N\Omega_k}^2 \le C$, two applications of Lemma A.1 yield that for every $\eta_N \to \infty$ there exist V'_{Ni} and W'_{Ni} for $i = 1, \dots, N+1$, such that

$$|V'_{Ni}| \le \eta_N N^{-1/6}, \qquad |W'_{Ni}| \le \eta_N N^{-1/2}$$

and

$$P(V'_{Ni} = V_{Ni},\, i = 1, \dots, N+1) = 1 - o(1),$$
$$P(W'_{Ni} = W_{Ni},\, i = 1, \dots, N+1) = 1 - o(1).$$

Hence, we may replace V_{Ni} by V'_{Ni} and W_{Ni} by W'_{Ni} in R_{N1}, R_{N2} and R_{N3}. For these remainder terms to go to zero in probability, it suffices to show that $E|R'_{Nj}| \to 0$, for $j = 1, 2, 3$ with

$$R'_{N1} = N^{1/2} \sum_{i=1}^{N+1} W'_{Ni} W_{Ni}^2, \qquad R'_{N2} = N^{1/2} \sum_{i=1}^{N+1} V'_{Ni} W_{Ni}^2,$$
$$R'_{N3} = N^{1/2} \sum_{i=1}^{N+1} V'_{Ni} V_{Ni} W_{Ni}.$$

This is easy, since for $\eta_N = o(N^{1/6})$,

$$E|R'_{N1}| \le \eta_N N E W_{N1}^2 \to 0,$$
$$E|R'_{N2}| \le \eta_N N^{4/3} E W_{N1}^2 \to 0,$$
$$E|R'_{N3}| \le \eta_N N^{4/3} \left(EV_{N1}^2\right)^{1/2} \left(EW_{N1}^2\right)^{1/2} \to 0.$$

Finally, using an inequality of von Bahr and Esseen (1965), we have for every $\varepsilon > 0$, as N tends to infinity,

$$P\left(\left| N^{1/2} \sum_{i=1}^{N+2} \left(T_{Ni}^3 - ET_{N1}^3\right) \right| > \varepsilon \right) \le \frac{C'(N+1)E|T_{N1}|^p}{\left(\varepsilon N^{-1/2}\right)^{p/3}} \to 0.$$

Since the uniformity of the convergence of the various terms is easily checked, the lemma is proved. □

LEMMA 5.2. *Suppose that there exist constants* $c > 0$ *and* $C > 0$, *such that* (1.2) *and* (1.3) *are satisfied*, $E|N^{1/2}T_{N1}|^3 \le C$, $E(N^{3/2}T_{N12})^2 \le C$, *and* $\sum_{k=3}^{N} \binom{N-2}{k-2} ET_{N\Omega_k}^2 \le CN^{-4}$. *Then there exist sequences* $\delta_N \searrow 0$ *and* $\varepsilon_N \searrow 0$, *which depend only on c and C, such that for $N = 2, 3, \ldots,$*

$$P\left(N^{1/2}\left|\sum\sum_{1\le i<j\le N+2}\left\{\left(\overline{\overline{T}}_N - \overline{T}_N^{(i)} - \overline{T}_N^{(j)} + T_N^{(i,j)}\right)\right.\right.\right.$$
$$\left.\left.\left.\left(\overline{T}_N - T_N^{(i)}\right)\left(\overline{T}_N - T_N^{(j)}\right) - ET_{N1}T_{N2}T_{N12}\right\}\right| \ge \varepsilon_N\right) \le \delta_N.$$

PROOF. Write

$$\overline{\overline{T}}_N - \overline{T}_N^{(i)} - \overline{T}_N^{(j)} + T_N^{(i,j)} = T_{Nij} - \frac{1}{N+1}\left(T_{Ni} + T_{Nj}\right) + \Delta_{Nij},$$
$$\overline{T}_N - T_N^{(i)} = T_{Ni} + Z_{Ni}.$$

Then

$$\left(\overline{\overline{T}}_N - \overline{T}_N^{(i)} - \overline{T}_N^{(j)} + T_N^{(i,j)}\right)\left(\overline{T}_N - T_N^{(i)}\right)\left(\overline{T}_N - T_N^{(j)}\right)$$
$$= \left(T_{Nij} - \frac{1}{N+1}\left(T_{Ni} + T_{Nj}\right)\right)\left(T_{Ni} + Z_{Ni}\right)\left(T_{Nj} + Z_{Nj}\right)$$
$$+ \Delta_{Nij}\left(T_{Ni} + Z_{Ni}\right)\left(T_{Nj} + Z_{Nj}\right).$$

Straightforward calculations show [cf. Putter (1994)] that under the conditions of the present lemma, $E\Delta_{Nij}^2 = \mathscr{O}(N^{-4})$. It also follows from the proof of Lemma 5.1 that $EZ_{Ni}^2 = \mathscr{O}(N^{-2})$.

First we shall show that

$$N^{1/2}\sum\sum_{i<j}\Delta_{Nij}\left(T_{Ni} + Z_{Ni}\right)\left(T_{Nj} + Z_{Nj}\right)$$
$$= N^{1/2}\sum\sum_{i<j}\Delta_{Nij}\left(T_{Ni}T_{Nj} + T_{Ni}Z_{Nj} + T_{Nj}Z_{Ni} + Z_{Ni}Z_{Nj}\right)$$

tends to zero in probability. Applying Lemma A.1 we see that there exist T'_{Ni} such that $|T'_{Ni}| \le N^{-1/8}$ and

$$P(T'_{Ni} = T_{Ni} \text{ for } i = 1, \ldots, N+2) = 1 - o(1),$$

so it suffices to show that

$$E\left|N^{1/2}\sum\sum_{i<j}\Delta_{Nij}\left(T'_{Ni}T_{Nj} + T'_{Ni}Z_{Nj} + T'_{Nj}Z_{Ni} + Z_{Ni}Z_{Nj}\right)\right| \to 0.$$

This is true, as the above expression is less than or equal to

$$\tfrac{1}{2}N^{5/2}N^{-1/8}E\left|\Delta_{N12}\left(T_{N1} + Z_{N1} + Z_{N2}\right)\right| + \tfrac{1}{2}N^{5/2}E|\Delta_{N12}Z_{N1}Z_{N2}|$$
$$\le N^{19/8}\left(E\Delta_{N12}^2\right)^{1/2}\left\{\left(ET_{N1}^2\right)^{1/2} + \left(EZ_{N1}^2\right)^{1/2}\right\}$$
$$+ \tfrac{1}{2}N^{5/2}\left(E\Delta_{N12}^2\right)^{1/2}EZ_{N1}^2$$
$$= \mathscr{O}\left(N^{-1/8} + N^{-3/2}\right).$$

It remains to consider

$$N^{1/2} \sum\sum_{i<j} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right) (T_{Ni} + Z_{Ni})(T_{Nj} + Z_{Nj})$$
$$= N^{1/2} \sum\sum_{i<j} T_{Ni} T_{Nj} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right)$$
$$+ N^{1/2} \sum\sum_{i<j} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right) (T_{Ni} Z_{Nj} + T_{Nj} Z_{Ni})$$
$$+ N^{1/2} \sum\sum_{i<j} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right) Z_{Ni} Z_{Nj}.$$

We can deal with the second and third of these terms in a similar way and see that we have to show that

$$E \left| N^{1/2} \sum\sum_{i<j} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right) (T'_{Ni} Z_{Nj} + T'_{Nj} Z_{Ni}) \right| \to 0$$

and

$$E \left| N^{1/2} \sum\sum_{i<j} \left(T_{Nij} - \frac{1}{N+1} (T_{Ni} + T_{Nj}) \right) Z'_{Ni} Z_{Nj} \right| \to 0,$$

with $|T'_{Ni}| \le N^{-1/8}$ and $|Z'_{Ni}| \le N^{-1/3}$. The above expectations can be bounded, respectively, by

$$N^{5/2} N^{-1/8} E \left| \left(T_{N12} - \frac{1}{N+1} (T_{N1} + T_{N2}) \right) Z_{N1} \right|$$
$$\le N^{19/8} \left(E \left(T_{N12} - \frac{1}{N+1} (T_{N1} + T_{N2}) \right)^2 \right)^{1/2} \left(E Z_{N1}^2 \right)^{1/2}$$
$$= \mathscr{O}(N^{-1/8})$$

and

$$\tfrac{1}{2} N^{5/2} N^{-1/3} E \left| \left(T_{N12} - \frac{1}{N+1} (T_{N1} + T_{N2}) \right) Z_{N1} \right|$$
$$\le \tfrac{1}{2} N^{13/6} \left\{ E \left(T_{N12} - \frac{1}{N+1} (T_{N1} + T_{N2}) \right)^2 \right\}^{1/2} \left(E Z_{N1}^2 \right)^{1/2}$$
$$= \mathscr{O}(N^{-1/6}).$$

Finally, note that under the moment assumptions of the present theorem

$$E |T_{N1} T_{N2} T_{N12}|^{6/5} \le \left(E |T_{N1} T_{N2}|^3 \right)^{2/5} \left(E T_{N12}^2 \right)^{3/5} \le C^{7/5} N^{-3},$$
$$E \left| \frac{1}{N+1} T_{N1}^2 T_{N2} \right|^{6/5} \le N^{-6/5} \left(E |T_{N1}|^3 \right)^{6/5} \le C^{6/5} N^{-3}.$$

Now define

$$L_{Ni} = E\big(T_{Ni}T_{Nj}T_{Nij}|X_i\big) - \frac{1}{N+1}E\big(T_{Ni}T_{Nj}(T_{Ni}+T_{Nj})|X_i\big) - ET_{N1}T_{N2}T_{N12},$$

$$M_{Nij} = T_{Ni}T_{Nj}T_{Nij} - \frac{1}{N+1}T_{Ni}T_{Nj}(T_{Ni}+T_{Nj}) - L_{Ni} - L_{Nj} - ET_{N1}T_{N2}T_{N12}.$$

Obviously we have $E|L_{N1}|^{6/5} \le C'N^{-3}$ and $E|M_{N12}|^{6/5} \le C'N^{-3}$ for an appropriate C'. Now

$$N^{1/2}\sum\sum_{1\le i<j\le N+2}\left(T_{Ni}T_{Nj}\left(T_{Nij} - \frac{1}{N+1}(T_{Ni}+T_{Nj})\right) - ET_{N1}T_{N2}T_{N12}\right)$$

$$= N^{1/2}(N+1)\sum_{i=1}^{N+2}L_{Ni} + N^{1/2}\sum\sum_{1\le i<j\le N+2}M_{Nij}$$

and

$$P\left(\left|N^{1/2}(N+1)\sum_{i=1}^{N+2}L_{Ni}\right| > \varepsilon\right) \le \frac{C''(N+2)E|L_{N1}|^{6/5}}{\big(\varepsilon N^{-1/2}(N+1)^{-1}\big)^{6/5}} \to 0,$$

$$P\left(\left|N^{1/2}\sum\sum_{1\le i<j\le N+2}M_{Nij}\right| > \varepsilon\right) \le \frac{C''\binom{N+2}{2}E|M_{N12}|^{6/5}}{(\varepsilon N^{-1/2})^{6/5}} \to 0,$$

by the inequality of von Bahr and Esseen (1965). One easily checks the uniformity of the various convergences and the lemma is proved. □

Proof of Theorem 1.3. The conditions of Theorem 1.3 allow application of Lemmas 5.1 and 5.2, which together with Lemma 4.3 and Slutsky's lemma imply consistency of the estimators $\hat{\lambda}_1$ and $\hat{\lambda}_2$. Combining this with Theorems 1.1 and 1.2 proves the theorem. □

APPENDIX

In this Appendix we prove a truncation lemma that is used several times in the proofs of Theorem 1.2, Corollary 3.1 and Lemmas 5.1 and 5.2.

Lemma A.1. *Let Y be a random variable with $EY = 0$ and $E|Y|^s = \nu_s < \infty$ for some $s \ge 1$. Then, for every $\eta > 0$, there exists $Y' = \varrho(Y)$ such that*

$$|Y'| \le \eta \quad a.s.;$$

$$P(Y' \ne Y) \le \frac{2\nu_s}{\eta^s};$$

$$EY' = 0;$$
$$E|Y' - Y|^t \le \frac{2^{2t+1}\,\nu_s}{\eta^{s-t}} \quad \textit{for every } 0 < t \le s.$$

PROOF. Choose $\eta > 0$ and define

$$Y'' = \begin{cases} -\eta, & \text{if } Y < -\eta, \\ Y, & \text{if } -\eta \le Y \le \eta, \\ \eta, & \text{if } Y > \eta. \end{cases}$$

We have $|Y''| \le \eta$ a.s.,

$$P(Y'' \ne Y) = P(|Y| > \eta) \le \frac{\nu_s}{\eta^s}$$

and

$$E|Y'' - Y|^t = E(|Y| - \eta)^t \mathbf{I}_{\{|Y| \ge \eta\}}$$
$$\le (E|Y|^s)^{t/s} P(|Y| \ge \eta)^{1-t/s} \quad \text{for } 0 < t \le s.$$

Next we change Y'' slightly to make its expectation vanish. By the above we have $|EY''| \le \nu_s/\eta^{s-1}$ because $EY = 0$. Assume without loss of generality that $EY'' < 0$ and change the value of Y'' on a set of probability $\le \nu_s/\eta^s$ where $-\eta \le Y'' < 0$ to the value $+\eta$ until the expectation equals zero. This can actually be done, since otherwise this process would end with a nonnegative random variable with negative expectation. Call the resulting variable $Y' = \varrho(Y)$.

Obviously, $|Y'| \le \eta$ and $EY' = 0$. Also, $P(Y'' \ne Y') \le \nu_s/\eta^s$ so $P(Y' \ne Y) \le 2\nu_s/\eta^s$. Since we have obtained Y' from Y'' by changing the value of Y'' by at most 2η on a set of probability at most ν_s/η^s, we have $E|Y' - Y''|^t \le (\nu_s/\eta^s)(2\eta)^t$. Hence,

$$E|Y' - Y|^t \le (2^{t-1} \vee 1)\{E|Y'' - Y|^t + E|Y' - Y|^t\}$$
$$\le \frac{\nu_s}{\eta^{s-t}}(1 + 2^t)(2^{t-1} \vee 1) \le 2^{(2t+1)} \frac{\nu_s}{\eta^{s-t}}.$$

This proves the lemma. □

Acknowledgments. The authors are indebted to Dimitri Chibisov for a number of illuminating discussions and to the Editor, Associate Editor and two referees for their thorough and helpful reviews.

REFERENCES

BENTKUS, V., GÖTZE, F. and VAN ZWET, W. R. (1997). An Edgeworth expansion for symmetric statistics. *Ann. Statist.* **25** 851–896.

BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The Edgeworth expansion for *U*-statistics of degree 2. *Ann. Statist.* **14** 1463–1484.

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

DHARMADHIKARI, S. W., FABIAN, V. and JOGDEO, K. (1968). Bounds on the moments of martingales. *Ann. Math. Statist.* **39** 1719–1723.

HELMERS, R. (1991). On the Edgeworth expansion and the bootstrap approximation for a Studentized U-statistic. *Ann. Statist.* **19** 470–484.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

JANSSEN, P. (1978). The Berry–Esseen theorem and an asymptotic expansion for U-statistics, Ph.D. thesis, Limburg Univ. Centre (in Dutch).

PUTTER, H. (1994). Consistency of resampling methods. Ph.D. thesis, Univ. Leiden.

QUENOUILLE, M. H. (1949). Approximate tests of correlation in time series. *J. Roy. Statist. Soc. Ser. B* **11** 68–84.

QUENOUILLE, M. H. (1956). Notes on bias in estimation. *Biometrika* **43** 353–360.

TUKEY, J. W. (1958). Bias and confidence in not-quite large samples. *Ann. Math. Statist.* **29** 614 (abstract).

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

VON BAHR, B. and ESSEEN C. G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36** 299–303.

NATEC
AMSTERDAM MEDICAL CENTER
UNIVERSITY OF AMSTERDAM
MEIBERGDREEF 9
1105 AZ AMSTERDAM
E-MAIL: h.putter@amc.uva.nl

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LEIDEN
P.O. BOX 9512
8300 RA LEIDEN
THE NETHERLANDS
E-MAIL: vanzwet@wi.leidenuniv.nl