Elementary Symmetric Polynomials of Increasing Order

A.J. van Es[1] and R. Helmers[2]

[1] Mathematical Institute, University of Amsterdam, Roetersstraat 15, 1018 BW Amsterdam, The Netherlands

[2] Centre for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Summary. The asymptotic behaviour of elementary symmetric polynomials $S_n^{(k)}$ of order k, based on n independent and identically distributed random variables $X_1, \ldots, X_n$, is investigated for the case that both k and n get large. If $k = o(n^{\frac{1}{2}})$, then the distribution function of a suitably normalised $S_n^{(k)}$ is shown to converge to a standard normal limit. The speed of this convergence to normality is of order $\mathcal{O}(kn^{-\frac{1}{2}})$, provided $k = \mathcal{O}(\log^{-1} n \log_2^{-1} n n^{\frac{1}{2}})$ and certain natural moment assumptions are imposed. This order bound is sharp, and cannot be inferred from one of the existing Berry-Esseen bounds for U-statistics. If $k \to \infty$ at the rate $n^{\frac{1}{2}}$ then a non-normal weak limit appears, provided the X_j's are positive and $S_n^{(k)}$ is standardised appropriately. On the other hand, if $k \to \infty$ at a rate faster than $n^{\frac{1}{2}}$ then it is shown that for positive X_j's there exists no linear norming which causes $S_n^{(k)}$ to converge weakly to a nondegenerate weak limit.

1. Introduction and Results

Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with common distribution function F and

$$\mu = EX_1 \neq 0, \ 0 < \sigma^2 = \sigma^2(X_1) < \infty. \tag{1.1}$$

Let for any $1 \leq k \leq n$,

$$S_n^{(k)} = \binom{n}{k}^{-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \cdots \sum X_{i_1} \ldots X_{i_k} \tag{1.2}$$

and let

$$F_n^{(k)}(x) = P\left(\left\{\frac{n^{\frac{1}{2}}(S_n^{(k)} - \mu^k)}{(k\mu^{k-1}\sigma)} \leq x\right\}\right) \tag{1.3}$$

for real x. Let Φ denote the standard normal distribution function (df).

Clearly $S_n^{(k)}$ is a U-statistic of degree k with kernel h given by $h(x_1, ..., x_k)$ $=x_1 ... x_k$. If k is fixed, then Hoeffding's CLT for U-statistics (Hoeffding (1948)) yields that $F_n^{(k)}$ converges in distribution to Φ, as $n \to \infty$, provided assumption (1.1) is satisfied. We shall show that the result remains valid if $k=k(n) \to \infty$, whenever $k=\mathcal{o}(n^{\frac{1}{2}})$.

Theorem 1. *If* $\mu=EX_1 \neq 0, 0<\sigma^2=\sigma^2(X_1)<\infty$ *and* $k=\mathcal{o}(n^{\frac{1}{2}})$, *then*

$$\lim_{n \to \infty} \sup_x |F_n^{(k)}(x)-\Phi(x)|=0. \qquad (1.4)$$

The speed of convergence to normality in (1.4) is determined in

Theorem 2. *If* $\mu=EX_1 \neq 0,$ $0<\sigma^2=\sigma^2(X_1),$ $E|X_1|^3<\infty$ *and* k $=\mathcal{O}(\log^{-1} n \log_2^{-1} n n^{\frac{1}{2}})$, *then, as* $n \to \infty$,

$$\sup_x |F_n^{(k)}(x)-\Phi(x)|=\mathcal{O}(kn^{-\frac{1}{2}}). \qquad (1.5)$$

If k is fixed, then (1.5) yields the classical rate $n^{-\frac{1}{2}}$ for the accuracy of the normal approximation. For this very special case Theorem 2 is a simple consequence of the Berry-Esseen theorem for U-statistics of degree k given by van Zwet (1984). On the other hand, if $k=k(n) \to \infty$ but $k=\mathcal{O}(\log^{-1} n \log_2^{-1} n^{\frac{1}{2}})$, then – perhaps somewhat surprisingly – the order bound (1.5) cannot be inferred from one of the existing Berry-Esseen bounds for U-statistics. Application of Corollary 4.1 of van Zwet (1984) yields the bound $\mathcal{O}(k^2 \left(1+\dfrac{\sigma^2}{\mu^2}\right)^k n^{-\frac{1}{2}})$, for the l.h.s. of (1.5), which is of course much worse then the bound given in (1.5). A related Berry-Esseen bound for U-statistics of Friedrich (1985) also does not give us (1.5). It appears that the dependence on k in these bounds is not optimal. In contrast, the bound $\mathcal{O}(kn^{-\frac{1}{2}})$ established in Theorem 2 is sharp. To see this we note that, if $k=k(n) \to \infty$ at a rate slower than $\log^{-1} n \log_2^{-1} n n^{\frac{1}{2}}$ and the moment assumptions of Theorem 2 are satisfied, we obtain a valid one-term Edgeworth expansion $\tilde{F}_n^{(k)}(x)=\Phi(x)+\frac{1}{6}n^{-\frac{1}{2}}\phi(x)(1-x^2)\{\sigma^{-3} E(X_1-\mu)^3+3(k$ $-1)\sigma\mu^{-1}\}$ for $F_n^{(k)}$ with uniform remainder $\mathcal{o}(kn^{-\frac{1}{2}})$. A proof of this assertion may be found in the appendix. There is no need for the usual requirement that F is non-lattice. However, if k is fixed, then we must add the assumption that F is non-lattice, in order to guarantee that our expansion is valid uniformly. This latter statement can be inferred from a recent paper of Maesono (1987), where an Edgeworth expansion for U-statistics of fixed degree k is established. For the case $k=2$ we also refer to Theorem 1.2 of Bickel, Götze and van Zwet (1986), whereas the classical case $k=1$ is of course well-known and treated, e.g., in Feller (1971).

After the present paper was submitted for publication Dr. K.O. Friedrich informed us that a different proof of our Theorem 2 can be obtained by applying the very general Theorem 2.1 – a Berry-Esseen theorem for arbitrary statistics – of his PH.D. thesis (cf. Friedrich (1985)). In fact, Friedrich's approach leads to a slightly improved Theorem 2: the restriction $k=\mathcal{O}(\log^{-1} n \log_2^{-1} nn^{\frac{1}{2}})$ can

be replaced by the weaker requirement $k = O(\log^{-1} n n^{\frac{1}{2}})$. On the other hand, in view of the remarks concerning the one-term Edgeworth expansion for elementary symmetric polynomials of increasing order, we can strengthen the order bound (1.5) to obtain a more precise assertion:

$$\sup_x |F_n^{(k)}(x) - \Phi(x)| \sim c \, \frac{\sigma}{\mu} \, \frac{k}{\sqrt{n}} \tag{1.6}$$

provided $k \to \infty$ at a rate $o(n^{\frac{1}{2}} \log^{-1} n \log_2^{-1} n)$. The constant c is equal to $(2\pi e^3)^{-\frac{1}{2}} = 0.089 \dots$ Thus, to first order, the deviation from normality depends only on the coefficient of variation $\frac{\sigma}{\mu}$ of the underlying distribution of the X_i's.

In Sect. 2 we supplement the foregoing discussion, by an example establishing the sharpness of the bound (1.5) for zero-one random variables $X_1, X_2, \dots$ in a more direct way.

If $k = k(n) \to \infty$ at the rate $n^{\frac{1}{2}}$ or faster, then the asymptotic behaviour of $S_n^{(k)}$ becomes completely different from the one described in the Theorems 1 and 2. From a weak limit theorem for the k-th root of $S_n^{(k)}$ due to Székely (1982), which considers only strictly positive X_j's, we derive for positive X_j's the following two results.

Theorem 3. *If* $P(X_1 > 0) = 1$, $0 < \sigma^2 = \sigma^2(X_1) < \infty$, *and* $k \sim c n^{\frac{1}{2}}$ *for some constant* $c > 0$ *and* $n \to \infty$, *then there exists a sequence of positive numbers* $\{s_n\}$, *such that* $s_n \to \mu$ *and*

$$s_n^{-k} S_n^{(k)} \xrightarrow[d]{} \exp\left(\frac{c\sigma}{\mu} N\right) \tag{1.7}$$

where $\xrightarrow[d]{}$ *denotes convergence in distribution and* N *stands for a standard normal random variable.*

Theorem 4. *If* $P(X_1 > 0) = 1$, $0 < \sigma^2 = \sigma^2(X_1) < \infty$, *and* $k n^{-\frac{1}{2}} \to \infty$ *as* $n \to \infty$, *then there exists no linear norming which causes* $S_n^{(k)}$ *to converge in distribution to a nondegenerate limit distribution.*

Theorem 4 tells us that for positive X_j there exists no linear norming of $S_n^{(k)}$ that converges in distribution to a nondegenerate limit if $k n^{-\frac{1}{2}} \to \infty$. In view of Székely's (1982) result (see Lemma 5 in the next section) it turns out that for positive X_j if $k n^{-\frac{1}{2}} \to \infty$ taking the k-th root is essentially required to obtain nondegenerate weak limits, while if $k = o(n^{\frac{1}{2}})$ there exist linear normings of both $S_n^{(k)}$ and its k-th root that converge to a nondegenerate limit. The border case $k \sim c n^{\frac{1}{2}} (c > 0)$ is treated in Theorem 3. Again Székely's (1982) result can be used to find that for positive X_j the random variable $S_n^{(k)}$ can be appropriately scaled to yield a nondegenerate weak limit.

The limit behaviour of $S_n^{(k)}$ in the case $k \sim c n (0 \le c \le 1)$ was investigated in detail by van Es (1986), Halász and Székely (1976), Székely (1974, 1982) and

Móri and Székely (1982). In the latter paper it is shown that in the degenerate case ($\mu = 0$ and $P(|X_1| = \text{constant}) = 1$) nondegenerate limit distributions exist of linear normings of $S_n^{(k)}$ even if $k \sim cn (0 \leq c \leq 1)$.

Proofs

Proof of Theorem 1. The Hoeffding decomposition for elementary symmetric polynomials yields

$$S_n^{(k)} - \mu^k = \sum_{r=1}^{k} H_r(X_1, \ldots, X_n) \tag{2.1}$$

where

$$H_r(X_1, \ldots, X_n) = \binom{n}{k}^{-1} \binom{n-r}{k-r} \mu^{k-r} \sum_{1 \leq j_1 < \ldots < j_r \leq n} \prod_{i=1}^{r} (X_{j_i} - \mu) \tag{2.2}$$

as given by Karlin and Rinott (1982), page 496. Define

$$q_r = \frac{\sigma^2(H_{r+1})}{\sigma^2(H_r)}, r = 1, 2, \ldots, k-1. \tag{2.3}$$

A simple computation using (2.2) yields

$$q_r = \frac{\sigma^2}{\mu^2} (k-r)^2 / ((r+1)(n-r)), \quad r = 1, 2, \ldots, k-1. \tag{2.4}$$

Since the summands of the r.h.s. of (2.1) are uncorrelated we find

$$\sigma^2(S_n^{(k)}) = \sum_{r=1}^{k} \sigma^2(H_r(X_1, \ldots, X_n))$$

$$= \sigma^2(H_1(X_1, \ldots, X_n))(1 + q_1 + q_1 q_2 + \cdots + q_1 q_2 \cdots q_{k-1}).$$

Because of (2.4) and the assumption $k = o(n^{\frac{1}{2}})$ we see that for fixed k and n, and n sufficiently large,

$$q_r < c \frac{k^2}{rn}, \quad r = 1, 2, \ldots, k-1, \tag{2.6}$$

for some constant $c > 0$. This implies

$$0 \leq \sum_{r=2}^{k} q_1 \cdots q_{r-1} \leq \sum_{r=2}^{\infty} \frac{1}{(r-1)!} \left(\frac{ck^2}{n}\right)^{r-1} = e^{c\frac{k^2}{n}} - 1 = o(1) \tag{2.7}$$

as $n \to \infty$. In view of (2.5) this yields that

$$\sigma^2(S_n^{(k)}) \sim \sigma^2(H_1(X_1, \ldots, X_n)). \tag{2.8}$$

In other words:

$$H_1(X_1, \ldots, X_n) = k n^{-1} \mu^{k-1} \sum_{i=1}^{n} (X_i - \mu) \qquad (2.9)$$

is the dominant term in the decomposition (2.1). Application of the CLT of Lindeberg-Levy to $H_1(X_1, \ldots, X_n)$ shows that (1.4) is valid if $S_n^{(k)} - \mu^k$ in (1.3) is replaced by $H_1(X_1, \ldots, X_n)$. This completes the proof.

Proof of Theorem 2. To establish (1.5) we rewrite (2.1) as

$$S_n^{(k)} - \mu^k = U_n + R_n \qquad (2.10)$$

where

$$U_n = H_1(X_1, \ldots, X_n) + H_2(X_1, \ldots, X_n) \qquad (2.11)$$

with H_1 as in (2.9)

$$H_2(X_1, \ldots, X_n) = \frac{k(k-1)}{n(n-1)} \mu^{k-2} \sum\sum_{1 \leq i < j \leq n} (X_i - \mu)(X_j - \mu) \qquad (2.12)$$

and

$$R_n = \sum_{r=3}^{k} H_r(X_1, \ldots, X_n) \qquad (2.13)$$

is a remainder term of lower order. Clearly U_n is a U-statistic of degree 2 with kernel

$$\tfrac{1}{2} k \mu^{k-1}(x + y - 2\mu) + \tfrac{1}{2} k(k-1) \mu^{k-2}(x - \mu)(y - \mu). \qquad (2.14)$$

It follows directly from (2.10) that to prove (1.5) it suffices to show that both

$$\sup_x \left| P\left(\left\{ \frac{n^{\frac{1}{2}} U_n}{(k \mu^{k-1} \sigma)} \leq x \right\} \right) - \Phi(x) \right| = \mathcal{O}(k n^{-\frac{1}{2}}) \qquad (2.15)$$

and

$$P\left(\left\{ \frac{|R_n|}{(k \mu^{k-1} \sigma)} \geq k n^{-1} \right\} \right) = \mathcal{O}(k n^{-\frac{1}{2}}) \qquad (2.16)$$

are satisfied.

To prove (2.15) we shall need characteristic functions (ch.f) arguments. We shall only treat the case that $k \to \infty$. If k is fixed then Theorem 2 is a simple special case of Corollary 4.1 of van Zwet (1984). Let $\rho_n^{(k)}(t)$ denote the ch.f of $n^{\frac{1}{2}} U_n/(k \mu^{k-1} \sigma)$, i.e.

$$\rho_n^{(k)}(t) = E \exp(i t n^{\frac{1}{2}} U_n/(k \mu^{k-1} \sigma)). \qquad (2.17)$$

The usual argument based on Esseen's smoothing lemma implies that

$$\int_{|t| \leq k^{-\frac{1}{2}} n^{\frac{1}{2}}} |t|^{-1} |\rho_n^{(k)}(t) - e^{-\frac{1}{2} t^2}| \, dt = \mathcal{O}(k n^{-\frac{1}{2}}) \qquad (2.18)$$

and

$$\int_{k^{-\frac{1}{4}}n^{\frac{1}{4}} \leq |t| \leq k^{-1}n^{\frac{1}{4}}} |t|^{-1} |\rho_n^{(k)}(t) - e^{-\frac{1}{2}t^2}| \, dt = \mathcal{O}(kn^{-\frac{1}{4}}) \qquad (2.19)$$

together will yield (2.15). The statements (2.18) and (2.19) may be proved by modifying the proof of Theorem 2.1 of Helmers and van Zwet (1982). Our proof of (2.19) will require the assumption $k = \mathcal{O}(\log^{-1} n \, \log_2^{-1} n n^{\frac{1}{4}})$. For details of the proof the interested reader is referred to the appendix.

It remains to establish (2.16). An application of Chebyshev's inequality yields:

$$P\left(\left\{\frac{|R_n|}{(k\mu^{k-1}\sigma)} \geq kn^{-1}\right\}\right) \leq n^2 k^{-4} \mu^{-2k+2} \sigma^{-2} \sigma^2(R_n). \qquad (2.20)$$

Using (2.13) and an argument like (2.5)–(2.7) yields

$$\sigma^2(R_n) = \sum_{r=3}^{k} \sigma^2(H_r(X_1, \ldots, X_n))$$

$$= \sigma^2(H_1(X_1, \ldots, X_n)) \sum_{r=3}^{k} q_1 \cdots q_{r-1}$$

$$= k^2 n^{-1} \mu^{2k-2} \sigma^2 \sum_{r=3}^{k} q_1 \cdots q_{r-1}$$

$$\leq k^2 n^{-1} \mu^{2k-2} \sigma^2 \sum_{r=3}^{k} \frac{1}{(r-1)!} \left(\frac{ck^2}{n}\right)^{r-1}$$

$$\leq c^2 k^6 n^{-3} \mu^{2k-2} \sigma^2 \sum_{r=3}^{k} \frac{1}{(r-1)!} \left(\frac{ck^2}{n}\right)^{r-3}$$

$$\leq c^2 k^6 n^{-3} \mu^{2k-2} \sigma^2 2^{-1} \sum_{r=1}^{\infty} \frac{1}{(r-1)!} \left(\frac{ck^2}{n}\right)^{r-1}$$

$$\leq c^2 k^6 n^{-3} \mu^{2k-2} \sigma^2 2^{-1} e^{c\frac{k^2}{n}}$$

$$= \mathcal{O}(k^6 n^{-3} \mu^{2k-2}). \qquad (2.21)$$

Together with (2.20) this gives

$$P\left(\left\{\frac{|R_n|}{(k\mu^{k-1}\sigma)} \geq kn^{-1}\right\}\right) = \mathcal{O}(k^2 n^{-1}) \qquad (2.22)$$

which implies (2.16). This completes the proof.

Proof of Theorem 3. The following result of Székely (1982) will be used.

Lemma 5. *If* $P(X_1 > 0) = 1$, $0 < \sigma^2 = \sigma^2(X_1) < \infty$ *and* $k = o(n)$, *then, as,* $n \to \infty$

$$n^{\frac{1}{2}}((S_n^{(k)})^{1/k} - s_n) \xrightarrow[d]{} N(0, \sigma^2) \qquad (2.23)$$

where $\{s_n\}$ denotes a sequence of positive numbers, such that $s_n \to \mu$, and $N(0, \sigma^2)$ stands for a normal random variable with expectation zero and variance σ^2.

Define

$$V_n = n^{\frac{1}{2}} ((S_n^{(k)})^{1/k} - s_n) \tag{2.24}$$

and note that because of (2.23) V_n is asymptotically $N(0, \sigma^2)$. In addition we have that $s_n + n^{-\frac{1}{2}} V_n \geq 0$. A Taylor expansion argument yields

$$
\begin{aligned}
S_n^{(k)} &= \exp[k \log(s_n + n^{-\frac{1}{2}} V_n)] \\
&= \exp[k(\log s_n + n^{-\frac{1}{2}} W_n)] \\
&= s_n^k \exp(kn^{-\frac{1}{2}} W_n)
\end{aligned}
\tag{2.25}
$$

where W_n is asymptotically $N(0, \sigma^2 \mu^{-2})$ distributed. This directly yields (1.7) and the proof of the theorem is complete.

Proof of Theorem 4. In view of (2.25) it suffices to prove the statement of the theorem for the random variables

$$Z_n = \exp(kn^{-\frac{1}{2}} W_n). \tag{2.26}$$

Let us suppose that there exist sequences $\{a_n\}$ and $\{b_n\}$ with $a_n > 0$ such that $a_n(Z_n - b_n)$ converges in distribution to a nondegenerate limit distribution F. Then there exist three points $-\infty < u_1 < u_2 < u_3 < \infty$, continuity points of F, with $0 \leq F(u_1) < F(u_2) < F(u_3) \leq 1$.

It follows from this and (2.26) that

$$
\begin{aligned}
&\lim_{n \to \infty} P(\{a_n(Z_n - b_n) \leq u_i\}) \\
&= \lim_{n \to \infty} P\left(W_n \leq n^{\frac{1}{2}} k^{-1} \log\left(\frac{u_i + a_n b_n}{a_n}\right)\right) = F(u_i)
\end{aligned}
\tag{2.27}
$$

for $i = 1, 2, 3$. This implies

$$
\begin{aligned}
&\lim_{n \to \infty} n^{\frac{1}{2}} k^{-1} \left(\log\left(\frac{u_2 + a_n b_n}{a_n}\right) - \log\left(\frac{u_1 + a_n b_n}{a_n}\right)\right) \\
&= \frac{\sigma}{\mu} \{\Phi^{-1}(F(u_2)) - \Phi^{-1}(F(u_1))\} > 0
\end{aligned}
\tag{2.28}
$$

and hence

$$\lim_{n \to \infty} n^{\frac{1}{2}} k^{-1} \log\left(\frac{u_2 + a_n b_n}{u_1 + a_n b_n}\right) > 0. \tag{2.29}$$

Since $n^{\frac{1}{2}} k^{-1} \to 0$, evidently $a_n b_n \to -u_1$. But the same argument also holds for u_2, and u_3, implying $a_n b_n \to -u_2$, which is in contradiction with $a_n b_n \to -u_1$. This completes the proof of the Theorem.

An example establishing the sharpness of the bound (1.5).

Consider i.i.d. zero-one random variables $X_1, X_2, \ldots$ with $P(X_1 = 0) = P(X_1 = 1) = \frac{1}{2}$. The elementary symmetric polynomial $S_n^{(k)}$ then reduces to $S_n^{(k)} = 0$ if $E_n < k$, $S_n^{(k)} = \binom{E_n}{k} / \binom{n}{k}$ if $E_n \geq k$ where E_n denotes the number of ones in $X_1, \ldots, X_n$. Of course E_n is binomially distributed with parameters n and $\frac{1}{2}$. This example is also treated in van Es (1986) for the case $k \sim cn$ $(0 < c < 1)$. We prove the following lemma.

Lemma 6. *If $P(X_1 = 0) = P(X_1 = 1) = \frac{1}{2}$ and $k = o(n^{\frac{1}{2}})$ then there exists a sequence x_n, n odd, satisfying*

$$x_n = \frac{1}{2}(k+1) n^{-\frac{1}{2}}(1 + o(1)) \tag{2.30}$$

and

$$F_n^{(k)}(x_n) - \Phi(x_n) = (2\pi)^{-\frac{1}{2}} \frac{1}{2}(k+1) n^{-\frac{1}{2}}(1 + o(1)) \tag{2.31}$$

as $n \to \infty$.

To prove Lemma 6 we first note that (1.3) can be rewritten as

$$F_n^{(k)}(x) = P(g_{k,n}(E_n) - n^{\frac{1}{2}} k^{-1} \leq x) \tag{2.32}$$

where the function $g_{k,n} \colon \mathbb{N} \to \mathbb{R}$ is defined by

$$g_{k,n}(m) = 0 \quad \text{if } m < k, \qquad g_{k,n}(m) = n^{\frac{1}{2}} k^{-1} 2^k \binom{n}{k}^{-1} \binom{m}{k} \quad \text{if } m \geq k. \tag{2.33}$$

Since $g_{k,n}$ is increasing on the set $\{k, k+1, \ldots\}$ it follows that

$$F_n^{(k)}(g_{k,n}(m) - n^{\frac{1}{2}} k^{-1}) = P(E_n \leq m), \ m \geq k. \tag{2.34}$$

But this probability is exactly equal to $\frac{1}{2}$ if we take $m = \frac{1}{2}(n-1)$ (for n odd). Thus defining

$$x_n = g_{k,n}(\tfrac{1}{2}(n-1)) - n^{\frac{1}{2}} k^{-1} \tag{2.35}$$

we obtain

$$F_n^{(k)}(x_n) - \Phi(x_n) = \frac{1}{2} - \Phi(x_n). \tag{2.36}$$

In the appendix it is shown that (x_n) satisfies (2.30). A simple Taylor expansion argument then directly yields, in view of (2.36), relation (2.31). This completes the proof of Lemma 6. Note that (2.31) can also be deduced from the Edgeworth expansion type result, mentioned in Section 1, for the case that $k \to \infty$ but at a rate slower than $n^{\frac{1}{2}} \log^{-1} n \log_2^{-1} n$. For k is fixed this is not true, because F is lattice.

Appendix

Proof of (2.18) and (2.19). We will indicate how (2.18) and (2.19) can be obtained by modifying the proof of Theorem 2.1 of Helmers and van Zwet (1982).

Define functions g and ψ by

$$g(x) = \tfrac{1}{2} k \mu^{k-1} (x - \mu) \tag{A.1}$$

and

$$\psi(x, y) = \tfrac{1}{2} k(k-1) \mu^{k-2} (x - \mu)(y - \mu) \tag{A.2}$$

and note that the U-statistic U_n (cf. (2.11) and (2.14)) can be written as

$$U_n = 2n^{-1} \sum_{i=1}^{n} g(X_i) + 2(n(n-1))^{-1} \sum\sum_{1 \leq i < j \leq n} \psi(X_i, X_j). \tag{A.3}$$

We first prove (2.18).

Let χ denote the ch.f. of $\sigma^{-1}(X_1 - \mu)$, thus

$$\chi(t) = E \exp\{it\sigma^{-1}(X_1 - \mu)\} \tag{A.4}$$

and let

$$a = \frac{E|g(X_1)|^3}{(Eg^2(X_1))^{3/2}} = \frac{E|X_1 - \mu|^3}{\sigma^3}. \tag{A.5}$$

To proceed we follow the argument given in Helmers and van Zwet (1982), with some minor changes. Using (2.17) and the elementary inequality $|e^{ix} - 1 - ix| \leq 2x^2$ for all real x, we find

$$\begin{aligned}
\rho_n^{(k)}(t) &= E \exp(itn^{\frac{1}{2}} U_n/(k\mu^{k-1}\sigma)) \\
&= E \exp\{itn^{\frac{1}{2}}(U_n - \Delta_n(m))/(k\mu^{k-1}\sigma)\} \\
&\quad \cdot (1 + itn^{\frac{1}{2}}\Delta_n(m)/(k\mu^{k-1}\sigma)) + \tilde{R}_n
\end{aligned} \tag{A.6}$$

where

$$\Delta_n(m) = 2(n(n-1))^{-1} \sum_{i=1}^{m} \sum_{j=i+1}^{n} \psi(X_i, X_j)$$

and

$$|\tilde{R}_n| = O\left(mn^{-2} t^2 \frac{E\psi^2(X_1, X_2)}{Eg^2(X_1)} \right) \tag{A.7}$$

uniformly for any integer $m = 1, 2, \ldots, n-1$ and for all real t.

Taking $m = (n-1)$ first and using the fact that

$$E(\psi(X_1, X_2)|X_1) = E(\psi(X_1, X_2)|X_2) = 0 \text{ a.s.,}$$

and the independence of X_1, X_2, ..., we obtain after a little calculation

$$|En^{\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}\Delta_n(n-1)\exp\{2itn^{-\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}\sum_{i=1}^{n}g(X_i)\}|$$

$$\leqq n^{\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}|E\psi(X_1,X_2)\prod_{i=1}^{2}[\exp(2itn^{-\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}g(X_i))-1]|$$

$$\cdot|E\exp(itn^{-\frac{1}{2}}\sigma^{-1}(X_1-\mu))|^{n-2}$$

$$\leqq n^{-\frac{1}{2}}(k\mu^{k-1}\sigma)^{-3}t^2 E|g(X_1)g(X_2)\psi(X_1,X_2)|\cdot|E\exp(itn^{\frac{1}{2}}\sigma^{-1}(X_1-\mu))|^{n-2}.$$
$$(A.8)$$

Using (A.1) and the well-known inequality

$$|E\exp(itn^{-\frac{1}{2}}\sigma^{-1}(X_1-\mu))|\leqq\exp\left(-\frac{t^2}{3n}\right)$$

for all $|t|\leqq a^{-1}n^{\frac{1}{2}}$ we easily conclude that, uniformly for all $|t|\leqq a^{-1}n^{\frac{1}{2}}$, the quantity in (A.8) is of order

$$\mathcal{O}\left(t^2 n^{-\frac{1}{2}}\frac{E|g(X_1)g(X_2)\psi(X_1,X_2)|}{(Eg^2(X_1))^{\frac{3}{2}}}\exp\left(-\frac{(n-2)}{3n}t^2\right)\right).$$

Combining this with (A.6) and (A.7) directly yields

$$|\rho_n^{(k)}(t)-\chi^n(n^{-\frac{1}{2}}t)|=\mathcal{O}\left(\frac{E\psi^2(X_1,X_2)}{Eg^2(X_1)}n^{-1}t^2+\frac{E|g(X_1)g(X_2)\psi(X_1,X_2)|}{(Eg^2(X_1))^{3/2}}\right.$$

$$\left.\cdot n^{-\frac{1}{2}}|t|^3\exp\left(-\frac{(n-2)}{3n}t^2\right)\right)$$
$$(A.9)$$

as $n\to\infty$. In view of (A.1) and (A.2) it is easily checked that

$$\frac{E\psi^2(X_1,X_2)}{Eg^2(X_1)}=\mathcal{O}(k^2),\qquad\frac{E|g(X_1)g(X_2)\psi(X_1,X_2)|}{(Eg^2(X_1))^{3/2}}=\mathcal{O}(k)\qquad(A.10)$$

as $k\to\infty$.

The proof of the classical Berry-Esseen theorem for sums of i.i.d. random variables ensures that

$$\int_{|t|\leqq a^{-1}n^{\frac{1}{2}}}|t|^{-1}|\chi^n(n^{-\frac{1}{2}}t)-e^{-\frac{1}{2}t^2}|dt=\mathcal{O}(n^{-\frac{1}{2}})\qquad(A.11)$$

as $n\to\infty$.

Together the relations (A.9)–(A.11) directly imply (2.18). It remains to prove (2.19). Instead of (A.9) we now employ the following bound: Uniformly for any integer $m = 1, \ldots, n-2$ and for all $|t| \leq a^{-1} n^{\frac{1}{2}}$

$$|\rho_n^{(k)}(t)| = \mathcal{O}\left\{\frac{E\psi^2(X_1, X_2)}{Eg^2(X_1)}\, mn^{-2}t^2\right.$$

$$\left. + \exp\left(-\frac{mt^2}{3n}\right) + \frac{E|\psi(X_1, X_2)|}{(Eg^2(X_1))^{\frac{1}{2}}}\, mn^{-\frac{1}{2}}|t|\exp\left(-\frac{(m-2)t^2}{3n}\right)\right\} \quad \text{(A.12)}$$

as $n \to \infty$.

The bound (A.12) follows by an argument similar to the one leading to (A.10). For any fixed $m \leq n-2$ we employ, instead of (A.8), the inequality

$$|E(k\mu^{k-1}\sigma)^{-1}\Delta_n(m)\exp\{itn^{\frac{1}{2}}(U_n - \Delta_n(m))/(k\mu^{k-1}\sigma)\}|$$

$$\leq \left|2n^{-\frac{1}{2}}(n-1)^{-1}(k\mu^{k-1}\sigma)\sum_{i=1}^{m}\sum_{j=i+1}^{n} E\psi(X_i, X_j)\right.$$

$$\left. \cdot \exp\left\{2itn^{-\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}\left(\sum_{k=1}^{n} g(X_k) + (n-1)^{-1}\sum\sum_{m<k<l\leq n}\psi(X_k, X_l)\right)\right\}\right|$$

$$\leq \frac{m}{n^{\frac{1}{2}}}\frac{E|\psi(X_1, X_2)|}{(Eg^2(X_1))^{-\frac{1}{2}}}|E\exp(itn^{-\frac{1}{2}}\sigma^{-1}(X_1 - \mu))|^{m-2} \quad \text{(A.13)}$$

from which (A.12) is easily verified. Using (A.1) and (A.2) once more we also have that

$$\frac{E|\psi(X_1, X_2)|}{(Eg^2(X_1))^{\frac{1}{2}}} = \mathcal{O}(k) \quad \text{(A.14)}$$

as $k \to \infty$.

It follows now directly from (A.10), (A.12) and (A.14) that

$$\int_{k^{-\frac{1}{2}}n^{1/4}\leq|t|\leq k^{-1}n^{\frac{1}{2}}} |t|^{-1}|\rho_n^{(k)}(t) - e^{-\frac{1}{2}t^2}|\,dt$$

$$= \mathcal{O}\left\{\frac{k^2}{n}\log n \log\left(\frac{\sqrt{n}}{k}\right) + n^{-c/3}\log\left(\frac{\sqrt{n}}{k}\right) + k^{3/2}n^{1/4-c/3}\log n\right\}$$

$$= \mathcal{O}\left(\frac{k}{\sqrt{n}}\right) \quad \text{(A.15)}$$

provided we take $m = \left[\dfrac{cn\log n}{t^2}\right]$ in (A.12) with c sufficiently large and by requiring $k = \mathcal{O}(\log^{-1} n \log_2^{-1} n n^{\frac{1}{2}})$. This completes the proof of relation (2.19).

Proof of the validity of the Edgeworth expansion $\tilde{F}_n^{(k)}$

We derive the one-term Edgeworth expansion for $F_n^{(k)}$, referred to in the discussion after Theorem 2, for the case $k \to \infty$. Suppose that the moment assumptions of Theorem 2 are satisfied and, in addition that $k = o(\log^{-1} n \log_2^{-1} n n^{\frac{1}{2}})$.

To begin with we note that the argument leading to (2.22) can easily be modified to find that

$$P\left(\frac{|R_n|}{(k\mu^{k-1}\sigma)}\geq(\log n)^{-\frac{1}{2}}kn^{-1}\right)=o(kn^{-\frac{1}{2}}).\qquad(A.16)$$

Thus, R_n is of negligible order of magnitude for our present purposes as well. It remains to show that $\tilde{F}_n^{(k)}$ is an Edgeworth expansion for $n^{\frac{1}{2}}U_n/(k\mu^{k-1}\sigma)$, where U_n is as in (2.11), with uniform remainder $o(kn^{-\frac{1}{2}})$. To prove this we again employ Esseen's smoothing lemma and find that it suffices to show that (2.18) and (2.19) remain true, if we replace the region of integration $|t|\leq k^{-\frac{1}{2}}n^{1/4}$ in (2.18) by $|t|\leq(\log n)^{-\frac{1}{2}}k^{-\frac{1}{2}}n^{1/4}$, and, similarly in (2.19) the interval $k^{-\frac{1}{2}}n^{1/4}$ $\leq|t|\leq k^{-1}n^{\frac{1}{2}}$ by $(\log n)^{-\frac{1}{2}}k^{-\frac{1}{2}}n^{1/4}\leq|t|\leq Mk^{-1}n^{\frac{1}{2}}$, for any fixed constant $M>0$, and the two big $\mathcal{O}$'s on the r.h.s. of (2.18) and (2.19) by little o's. Finally we replace the term $e^{-\frac{1}{2}t^2}$ appearing in the common integrand of (2.18) and (2.19) by

$$\tilde{\rho}_n^{(k)}(t)=e^{-\frac{1}{2}t^2}\left\{1+\frac{(it)^3}{6n^{\frac{1}{2}}}(\sigma^{-3}E(X_1-\mu)^3+3(k-1)\sigma\mu^{-1})\right\}.\qquad(A.17)$$

Note that $\tilde{\rho}_n^{(k)}$ is the Fourier-Stieltjes transform of $\tilde{F}_n^{(k)}$.

Let (2.18)* and (2.19)* denote the above-mentioned modifications of (2.18) and (2.19). We first prove (2.18)*. Taylor expanding $\rho_n^{(k)}(t)$ we find that uniformly for all $|t|\leq(\log n)^{-\frac{1}{2}}k^{-\frac{1}{2}}n^{1/4}$

$$|\rho_n^{(k)}(t)-\tilde{\rho}_n^{(k)}(t)|$$
$$=\mathcal{O}\left(\frac{E\psi^2(X_1,X_2)}{Eg^2(X_1)}n^{-1}t^2\right)+o(kn^{-\frac{1}{2}}|t|P(|t|)\exp(-\frac{1}{4}t^2))\qquad(A.18)$$

as $n\to\infty$, where P is a fixed polynomial.

To check this we follow the pattern of the proof of Theorem 1.2 of Bickel, Götze and van Zwet (1986) (see also Callaert, Janssen and Veraverbeke (1980)) with some slight modifications. The details of the proof of (A.18) are as follows: Because of (A.3) we know that

$$\rho_n^{(k)}(t)=E\exp(itn^{\frac{1}{2}}U_n/(k\mu^{k-1}\sigma))$$

$$=E\exp\left(it\left\{2n^{-\frac{1}{2}}\sum_{i=1}^n g(X_i)+2n^{-\frac{3}{2}}(n-1)^{-1}\sum\sum_{1\leq i<j\leq n}\psi(X_i,X_j)\right\}/(k\mu^{k-1}\sigma)\right)$$

with g and ψ as in (A.1) and (A.2). Taylor expanding $\rho_n^{(k)}(t)$ now yields

$$\rho_n^{(k)}(t)=E\exp\left(it2n^{-\frac{1}{2}}(k\mu^{k-1}\sigma)^{-1}\sum_{i=1}^n g(X_i)\right)$$

$$\cdot\left[1+\frac{it2n^{-\frac{3}{2}}(n-1)^{-1}}{(k\mu^{k-1}\sigma)}\sum\sum_{1\leq i<j\leq n}\psi(X_i,X_j)\right]$$

$$+\mathcal{O}\left(\frac{t^2n^{-3}}{(k\mu^{k-1}\sigma)^2}E(\sum\sum_{1\leq i<j\leq n}\psi(X_i,X_j))^2\right).$$

After some easy calculations, exploiting the independence present, this reduces to

$$
\rho_n^{(k)}(t) = \chi^n(tn^{-\frac{1}{2}}) + it(k-1)\mu^{-1}\sigma^{-1}n^{\frac{1}{2}}2^{-1}\chi^{n-2}(tn^{-\frac{1}{2}})
$$
$$
\cdot E\exp\{itn^{-\frac{1}{2}}\sigma^{-1}((X_1-\mu)+(X_2-\mu))\}\cdot(X_1-\mu)(X_2-\mu)
$$
$$
+ \mathcal{O}\left(\frac{E\psi^2(X_1,X_2)}{Eg^2(X_1)}n^{-1}t^2\right).
$$

By a simple Taylor expansion argument

$$
E\exp\{itn^{-\frac{1}{2}}\sigma^{-1}((X_1-\mu)+(X_2-\mu))\}\cdot(X_1-\mu)(X_2-\mu)
$$
$$
= -\frac{t^2}{n}\sigma^2(1+\delta(1)) \quad \text{for } |t| = \delta(n^{\frac{1}{2}}).
$$

We also need the well-known fact that (cf. Feller (1971), p. 539) for all $|t| \leq \varepsilon n^{\frac{1}{2}}$ and some $\varepsilon > 0$,

$$
\chi^n(tn^{-\frac{1}{2}}) = e^{-\frac{1}{2}t^2}\left\{1 + \frac{(it)^3}{6n^{\frac{1}{2}}}\frac{E(X_1-\mu)^3}{\sigma^3}\right\}
$$
$$
+ \delta(n^{-\frac{1}{2}}|t|P(|t|)e^{-\frac{1}{4}t^2})
$$

where P is a fixed polynomial. Also note that $\chi^{-2}(tn^{-\frac{1}{2}}) = 1 + \mathcal{O}(t^2n^{-1})$ for $|t| = \delta(\sqrt{n})$. Combining all these results we obtain

$$
\rho_n^{(k)}(t) = \exp(-\frac{1}{2}t^2)\left\{1 + \frac{(it)^3}{6n^{\frac{1}{2}}}\frac{E(X_1-\mu)^3}{\sigma^3}\right\}
$$
$$
+ it(k-1)\mu^{-1}\sigma^{-1}n^{\frac{1}{2}}2^{-1}
$$
$$
\frac{(it)^2}{n}\sigma^2(1+\delta(1))\exp(-\frac{1}{2}t^2)
$$
$$
+ \delta(n^{-\frac{1}{2}}|t|P(|t|)\exp(-\frac{1}{4}t^2))
$$
$$
+ \mathcal{O}\left(\frac{E\psi^2(X_1,X_2)}{Eg^2(X_1)}t^2n^{-1}\right)
$$

from which (A.18) directly follows. Combining (A.18) with the first part of (A.10) directly yields (2.18)*. To prove (2.19)* we simply follow the calculations leading to (A.15) and the argument following it, to find that with minor changes the same proof also yields (2.19)*, provided we require $k = \mathcal{O}(\log^{-1}n \log_2^{-1}nn^{\frac{1}{2}})$ instead of $k = \mathcal{O}(\log^{-1}n \log_2^{-1}nn^{\frac{1}{2}})$. Since $Mk^{-1}n^{\frac{1}{2}} = \mathcal{O}(n^{\frac{1}{2}})$, as $k \to \infty$, we do not need a non-lattice condition here. This completes the proof of (2.19)* and the validity of the expansion $\bar{F}_n^{(k)}$ is established.

Proof of (2.30). We first treat the case $k \to \infty$. Using the refinement of Stirling's formula given on page 54 of Feller (1968), we find, for odd n,

$$g_{k,n}(\tfrac{1}{2}(n-1)) = n^{\frac{1}{2}} k^{-1} 2^k \binom{n}{k}^{-1} \binom{\frac{1}{2}(n-1)}{k}$$

$$= n^{\frac{1}{2}} k^{-1} 2^k n^{-n-\frac{1}{2}}(n-k)^{n-k+\frac{1}{2}}(\tfrac{1}{2}(n-1))^{\frac{1}{2}(n-1)+\frac{1}{2}}$$

$$\cdot (\tfrac{1}{2}(n-1)-k)^{-\frac{1}{2}(n-1)+k-\frac{1}{2}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right)$$

$$= n^{\frac{1}{2}} k^{-1} \exp\left\{(n-k+\tfrac{1}{2})\log\left(1-\frac{k}{n}\right)+\tfrac{1}{2}n\log\left(1-\frac{1}{n}\right)\right.$$

$$\left. -(\tfrac{1}{2}n-k)\log\left(1-\frac{2k+1}{n}\right)\right\}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)$$

$$= n^{\frac{1}{2}} k^{-1} \exp\left\{-\tfrac{1}{2}\frac{k^2}{n}(1+o(1))\right\}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)$$

$$= n^{\frac{1}{2}} k^{-1}\left(1-\tfrac{1}{2}\frac{k^2}{n}(1+o(1))\right)\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)$$

$$= n^{\frac{1}{2}} k^{-1} - \tfrac{1}{2}kn^{-\frac{1}{2}}(1+o(1)) + \mathcal{O}(n^{-\frac{1}{2}}k^{-1}) \tag{A.19}$$

and therefore

$$x_n = g_{k,n}(\tfrac{1}{2}(n-1)) - n^{\frac{1}{2}}k^{-1} = -\tfrac{1}{2}(k+1)n^{-\frac{1}{2}}(1+o(1)) \tag{A.20}$$

and (2.30) is proved. If k is fixed the calculation simplifies to

$$g_{k,n}(\tfrac{1}{2}(n-1)) = \frac{n^{\frac{1}{2}} k^{-1} 2^k (\tfrac{1}{2}(n-1))(\tfrac{1}{2}(n-1)-1)\dots(\tfrac{1}{2}(n-1)-k+1)}{n(n-1)\dots(n-k+1)}$$

$$= n^{\frac{1}{2}} k^{-1}\left(1-\tfrac{1}{2}\frac{k(k+1)}{n}+\mathcal{O}\left(\frac{1}{n^2}\right)\right)$$

$$= n^{\frac{1}{2}} k^{-1} - \tfrac{1}{2}(k+1)n^{-\frac{1}{2}} + \mathcal{O}(n^{-3/2}) \tag{A.21}$$

which also yields (2.30) in this case.

References

1. Bickel, P.J., Götze, F., Zwet, W.R. van: The Edgeworth expansion for U-statistics of degree two. Ann. Stat. **14**, 1463–1484 (1986)
2. Callaert, H., Janssen, P., Veraverbeke, N.: An Edgeworth expansion for U-statistics. Ann. Stat. **8**, 299–312 (1980)
3. Es, A.J. van: On the weak limits of elementary symmetric polynomials. Ann. Probab. **14**, 677–695 (1986)
4. Feller, W.: An Introduction to probability theory and its applications, vol. 1. New York: Wiley 1968
5. Feller, W.: An Introduction to probability theory and its applications, vol. 2. New York: Wiley 1971

6. Friedrich, K.O.: Berry-Esseen Abschätzungen für Nichtparametrische Statistiken, Ph.D. thesis, Albert-Ludwigs Universität, Freiburg im Breisgau (1985)

7. Halász, G., Székely, G.J.: On the elementary symmetric polynomials of independent random variables. Acta Math. Acad. Sci. Hungary, 28, 397–400 (1976)

8. Helmers, R., Zwet, W.R. van: The Berry-Esseen bound for U-statistics. In: Gupta, S.S., Berger, J.O. (eds.) Statistical decision theory and related topics III, vol. I, pp. 492–512. New York: Academic 1982

9. Hoeffding, W.: A class of statistics with asymptotically normal distribution. Ann. Math. Stat., 19, 293–325 (1948)

10. Karlin, S., Rinott, Y.: Applications of ANOVA type decompositions for conditional variance statistics including jackknife estimates. Ann. Stat., 10, 485–501 (1982)

11. Maesono, Y.: Edgeworth expansion for one sample U-statistics. Bulletin of Informatics and Cybernetics, 22, 189–197 (1987)

12. Móri, T.F., Székely, G.J.: Asymptotic behaviour of symmetric polynomial statistics. Ann. Probab., 10, 124–131 (1982)

13. Székely, G.J.: On the polynomials of independent random variables. Coll. Math. Soc. J. Bolyai, Limit theorems of probability theory (ed. P. Révész), Keszthely, Hungary, pp. 365–371 (1974)

14. Székely, G.J.: A limit theorem for elementary symmetric polynomials of independent random variables. Z. Wahrscheinlichkeitstheor. Ver. Geb. 59, 355–359 (1982)

15. Zwet, W.R. van: A Berry-Esseen bound for symmetric statistics. Z. Wahrscheinlichkeitstheor. Verw. Geb. 66, 425–440 (1984)