The Annals of Probability
1982, Vol. 10, No. 2, 423–434

CRAMÉR TYPE LARGE DEVIATIONS FOR LINEAR COMBINATIONS OF ORDER STATISTICS

M. VANDEMAELE AND N. VERAVERBEKE

Limburgs Universitair Centrum, Belgium

Let $T_n = n^{-1} \sum_{i=1}^n c_{in} X_{in}$ be a linear combination of order statistics and put $T_n^* = (T_n - E(T_n))/\sqrt{\text{Var}(T_n)}$. Sufficient conditions on the c_{in} and on the moments of the underlying distribution are established under which the ratio $P(T_n^* > x)/(1 - \Phi(x))$ tends to 1, either uniformly in the range $-A \le x \le c\sqrt{\ln n}$ ($A \ge 0$, $c > 0$) (moderate deviation theorem) or uniformly in the range $-A \le x \le o(n^\alpha)(A \ge 0)$ (Cramér type large deviation theorem). The proof relies on Helmers' approximation method and on the corresponding results for U-statistics.

1. Introduction. Linear combinations of order statistics (or L-statistics) are statistics of the form

$$T_n = n^{-1} \sum_{i=1}^n c_{in} X_{in} \tag{1}$$

where the weights c_{in} ($i = 1, \cdots, n$; $n = 1, 2, \cdots$) are real numbers and where $X_{1n} \le X_{2n} \le \cdots \le X_{nn}$ are the order statistics of a sequence $X_1, \cdots, X_n$ of independent random variables with the same distribution function F.

L-statistics have been studied by many authors under different sets of conditions on the weights and on the underlying random variable X with distribution function F. For instance, Stigler (1974) considered L-statistics of the form

$$T'_n = n^{-1} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{in} \tag{2}$$

where J is some real-valued, bounded and continuous function on (0, 1). He proved that $(T'_n - E(T'_n))/\sqrt{\text{Var}(T'_n)}$ is asymptotically standard normal if $E(X^2) < \infty$ and if

$$\sigma^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(x))J(F(y))[F(\min(x, y)) - F(x)F(y)]\, dx\, dy \tag{3}$$

is strictly positive. His assumptions also imply that

$$\lim_{n\to\infty} n \text{ Var}(T'_n) = \sigma^2.$$

Under the somewhat stronger assumption that the function J in (2) satisfies a Lipschitz condition of order 1 on (0, 1), Helmers (1980) proved that

$$\sup_x \left| P\left(\frac{T'_n - E(T'_n)}{\sqrt{\text{Var}(T'_n)}} \le x\right) - \Phi(x) \right| = O(n^{-1/2})$$

provided $E|X|^3 < \infty$ and $\sigma^2 > 0$; Φ denotes the standard normal distribution function. In the Ph.D. thesis of Helmers (1978), the same result was also obtained for general L-statistics of the form (1) with weights c_{in} satisfying the following condition

Received July 1980; revised February 1981.

AMS 1980 *subject classifications.* Primary 60F10; secondary 62G30.

Key words and phrases. Large deviations, linear combinations of order statistics, moderate deviations, U-statistics.

$$\max_{1\le i\le n}\left|c_{in} - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\,ds\right| = O(n^{-1})$$

where J is a real-valued function on (0, 1) which is Lipschitz of order 1 on (0, 1). In the proof of this Berry-Esseen result, Helmers constructs U-statistics related to the given L-statistic and carries over the Berry-Esseen result for U-statistics due to Callaert and Janssen (1978). Throughout this paper we will assume the following (weaker) condition:

CONDITION (*): There exists a real-valued function J on (0, 1), which is Lipschitz of order 1 on (0, 1) and which is such that for $n \to \infty$:

$$\sum_{i=1}^{n}\left[c_{in} - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\,ds\right]^2 = O(n^{-1}).$$

This condition on the weights still guarantees asymptotic normality. Indeed,

$$T_n = T_n' + n^{-1}\sum_{i=1}^{n}\left[c_{in} - n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\,ds\right]X_{in}$$
$$+ n^{-1}\sum_{i=1}^{n}\left[n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\,ds - J\left(\frac{i}{n+1}\right)\right]X_{in}.$$

It follows from Stigler's result and condition (*) that $\sigma^{-1}n^{1/2}(T_n - E(T_n))$ and also $(T_n - E(T_n))/\sqrt{\mathrm{Var}(T_n)}$ are asymptotically standard normal if $E(X^2) < \infty$ and if σ^2, given by (3), is strictly positive. Also $\lim_{n\to\infty} n\,\mathrm{Var}(T_n) = \sigma^2$.

The purpose of this paper is to study the behaviour of the ratios

$$\frac{P(\sigma^{-1}n^{1/2}(T_n - E(T_n)) > x)}{1 - \Phi(x)}$$

and

$$\frac{P\left(\dfrac{T_n - E(T_n)}{\sqrt{\mathrm{Var}(T_n)}} > x\right)}{1 - \Phi(x)}$$

when x depends on n and tends to $+\infty$ as $n \to +\infty$. In particular, we establish sufficient conditions under which these ratios tend to 1, either uniformly in the range $-A \le x \le c\sqrt{\ln n}$ ($A \ge 0$, $c > 0$) (moderate deviation theorem) or uniformly in the range $-A \le x \le o(n^{\alpha})$($A \ge 0$, $\alpha > 0$) (Cramér-type large deviation theorem). In the situation of properly standardized sums of independent random variables, these kind of theorems were initiated by Cramér (1938) and refined by Petrov (1975) (in the large deviation case) and by Rubin and Sethuraman (1965) and Amosova (1972) (in the moderate deviation case).

More recently, both types of theorems were obtained for U-statistics by Malevich and Abdalimov (1979). Since then their Cramér-type large deviation result has been sharpened by Vandemaele (1980).

The paper is organized in the following way: in Section 2 we state the two theorems; the proofs are given in Section 5. The two main tools in the proofs, namely the corresponding results for U-statistics and Helmers' approximation method, are given in Sections 3 and 4 respectively. Technical details which are used in the proofs several times are contained in the Lemmas A1 through A3 of the appendix (Section 6).

2. The results.

THEOREM 1. *Let* $T_n = n^{-1} \sum_{i=1}^n c_{in} X_{in}$ *be an L-statistic with weights satisfying condition* (*). *Assume that* $\sigma^2 > 0$. *If* $E|X|^p < \infty$ *for some* $p > 2 + c^2$ $(c > 0)$ *then, uniformly in the range* $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$

(i) $P(\sigma^{-1} n^{1/2}(T_n - E(T_n)) > x) = [1 - \Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right]$

(ii) $P\left(\frac{T_n - E(T_n)}{\sqrt{\operatorname{Var}(T_n)}} > x\right) = [1 - \Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right]$.

THEOREM 2. *Let* $T_n = n^{-1} \sum_{i=1}^n c_{in} X_{in}$ *be an L-statistic with weights satisfying condition* (*). *Assume that* $\sigma^2 > 0$. *If for all* $p = 1, 2, \cdots : E|X|^p \le K^p p^{\gamma p}$ (*where* K *and* $\gamma \ge 0$ *are constants not depending on* p) *then, uniformly in the range* $-A \le x \le o(n^\alpha)$ *with* $A \ge 0$ *and* $\alpha = \frac{1}{2(3 + 2\gamma)}$

(i) $P(\sigma^{-1} n^{1/2}(T_n - E(T_n)) > x) = [1 - \Phi(x)][1 + o(1)]$

(ii) $P\left(\frac{T_n - E(T_n)}{\sqrt{\operatorname{Var}(T_n)}} > x\right) = [1 - \Phi(x)][1 + o(1)]$.

3. The corresponding results for U-statistics. Basic facts in the proofs of Theorems 1 and 2 are the corresponding results for U-statistics, i.e. statistics of the form

$$U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_m \le n} h(X_{i_1}, \cdots, X_{i_m}),$$

where $X_1, \cdots, X_n$ are independent and identically distributed random variables and $h(x_1, \cdots, x_m)$ is a function symmetric in its $m \le n$ arguments such that $E[h(X_1, \cdots, X_m)] = 0$ and such that $g(X_1) = E[h(X_1, \cdots, X_m) \mid X_1]$ has a strictly positive variance σ_g^2. These results are stated in the following lemma.

LEMMA 1. *Let* $U_n = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1}, \cdots, X_{i_m})$ *be a U-statistic with* $E[h(X_1, \cdots, X_m)] = 0$ *and* $\sigma_g^2 > 0$.
(a) *if* $E|h(X_1, \cdots, X_m)|^p > \infty$ *for some* $p > 2 + c^2 (c > 0)$ *then*

$$P(m^{-1}\sigma_g^{-1} n^{1/2} U_n > x) = [1 - \Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right] \tag{4}$$

uniformly in the range $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$;
(b) *if for all* $p = 1, 2, \cdots : E|h(X_1, \cdots, X_m)|^p \le K^p p^{\gamma p}$ (*where* K *and* $\gamma \ge 0$ *are constants not depending on* p) *then*

$$P(m^{-1}\sigma_g^{-1} n^{1/2} U_n > x) = [1 - \Phi(x)][1 + o(1)] \tag{5}$$

uniformly in the range $-A \le x \le o(n^\alpha)(A \ge 0)$ *with* $\alpha = \frac{1}{2(3 + 2\gamma)}$.

Both results in Lemma 1 were originally obtained by Malevich and Abdalimov (1979). The (b) part, however, is already a sharpening due to Vandemaele (1980).

It should be noticed that in these papers, (4) and (5) are proved to hold uniformly in the ranges $0 \le x \le c\sqrt{\ln n}$ and $0 \le x \le o(n^\alpha)$ respectively. However, for (5), it follows immediately from the Berry-Esseen theorem of Callaert and Janssen (1978) that the range can be extended to $-A \le x \le o(n^\alpha)$.

That (4) holds in the wider range $-A \le x \le c\sqrt{\ln n}$ can be seen from the following argument.

In the proof of Theorem 1 in Malevich and Abdalimov (1979), $m^{-1}\sigma_g^{-1}n^{1/2}U_n$ is written as $\sigma_g^{-1}n^{-1/2}\sum_{i=1}^n g(X_i) + R_n$ where R_n satisfies: $E|R_n|^p \leq Cn^{-p/2}$ for some constant C. Application of the classical Slutsky argument yields:

$$\begin{aligned} &P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x + (\ln n)^{-2}) - P(|R_n| > (\ln n)^{-2}) \\ &\leq P(n^{-1}\sigma_g^{-1}n^{1/2}U_n > x) \\ &\leq P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x - (\ln n)^{-2}) + P(|R_n| > (\ln n)^{-2}). \end{aligned}$$

Using the general fact that $(1-\Phi(x))^{-1} = O(c_n e^{c_n^2/2})$ uniformly in $-A \leq x \leq c_n$ ($A > 0$, $c_n > 0$) and Lemma A1, we have

$$P(|R_n| > (\ln n)^{-2}) = [1-\Phi(x)]o\left(\frac{1}{\ln n}\right)$$

and

$$1-\Phi(x \pm (\ln n)^{-2}) = [1-\Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right],$$

both uniformly in the range $-A \leq x \leq c\sqrt{\ln n}$. Hence it is sufficient to show that

$$P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x) = [1-\Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right]$$

uniformly in the range $-A \leq x \leq c\sqrt{\ln n}$.

Since $E|g(X_i)|^p < \infty$ for $p > 2 + c^2 (c > 0)$, it follows from Theorem 3.4.1 of Ibragimov and Linnik (1971) that

$$\sup_x|P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x) - (1-\Phi(x))| = O(n^{-\delta/2})$$

with $0 < \delta < 1$. Hence uniformly in the range $-A \leq x \leq c'\sqrt{\ln n}$, with $0 < c' < \sqrt{\delta}$:

$$P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x) = [1-\Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right].$$

From the proof of Theorem 1 of Amosova (1972), it follows that uniformly in $M \leq x \leq c\sqrt{\ln n}$ ($M > 0$)

$$P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x) = [1-\Phi(x)]\left[1 + o\left(\frac{1}{x^2}\right)\right].$$

Hence also in $c'\sqrt{\ln n} \leq x \leq c\sqrt{\ln n}$ we have

$$P(\sigma_g^{-1}n^{-1/2}\textstyle\sum_{i=1}^n g(X_i) > x) = [1-\Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right].$$

4. Helmers' construction. In this section we briefly sketch Helmers' method to approximate L-statistics by U-statistics. We also state and prove some lemmas which will be used later on.

The first step is to approximate T_n by the statistic V_n given by

$$V_n = \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\,ds\,X_{in}.$$

For this first approximation we will need the following lemma.

LEMMA 2. *If $E|X|^p < \infty$ for some $p \geq 2$, then for some constant C (independent of p and n)*

$$E|T_n - V_n|^p \leq C^p E|X|^p n^{-p}.$$

PROOF. Since $T_n - V_n = n^{-1} \sum_{i=1}^n a_{in} X_{in}$ with $a_{in} = c_{in} - n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\, ds$, we have for $p \geq 2$:

$$|T_n - V_n|^p \leq n^{-p} (\sum_{i=1}^n a_i^2)^{p/2} (\sum_{i=1}^n X_i^2)^{p/2} \leq n^{-p} (\sum_{i=1}^n a_i^2)^{p/2} n^{p/2-1} \sum_{i=1}^n |X_i|^p.$$

From condition (*):

$$(\sum_{i=1}^n a_i^2)^{p/2} \leq C^p n^{-p/2},$$

where C is some constant. Hence the result follows.

The second step is to approximate V_n from above and below by statistics W_{n-} and W_{n+}, namely for all $n \geq 1$

$$W_{n-} \leq V_n \leq W_{n+}.$$

The precise construction of W_{n+} and W_{n-} can be found in Helmers (1978, 1980). They are of the following form

$$W_{n+} = n^{-1} \sum_{i=1}^n h_1(U_i) + Kn^{-2} \sum_{i=1}^n \sum_{j=1}^n h(U_i, U_j) + C$$

$$W_{n-} = n^{-1} \sum_{i=1}^n h_1(U_i) - Kn^{-2} \sum_{i=1}^n \sum_{j=1}^n h(U_i, U_j) - C.$$

Here C and $K > 0$ are constants; $U_1, \cdots, U_n$ are independent random variables which are uniformly distributed on (0, 1) and

$$h_1(u) = -\int_0^1 J(s)[\chi_{(0,s]}(u) - s]\, dF^{-1}(s) \qquad (0 < u < 1)$$

$$h(u, v) = \int_0^1 [\chi_{(0,s]}(u) - s][\chi_{(0,s]}(v) - s]\, dF^{-1}(s) \qquad (0 < u, v < 1)$$

(χ_E denotes the indicator of set E).

The third step is to relate W_{n+} and W_{n-} to appropriate U-statistics U_{n+} and U_{n-}, defined by

$$U_{n+} = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h_+(U_i, U_j)$$

$$U_{n-} = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h_-(U_i, U_j)$$

where for $0 < u, v < 1$:

$$h_\pm (u, v) = h_1(u) + h_1(v) \pm 2Kh(u, v).$$

We have from Helmers (1980) that

$$E[h_\pm (U_i, U_j)] = 0, \qquad E[h_\pm (U_i, U_j) \mid U_i] = h_1(U_i), \qquad \text{Var}\, h_1(U_i) = \sigma^2.$$

LEMMA 3. *If $E|X|^p < \infty$ for some $p \geq 1$, then for some constant C independent of p*

$$E|h_\pm (U_1, U_2)|^p \leq C^p E|X|^p.$$

PROOF. Since

$$|h_\pm (U_1, U_2)| \leq |h_1(U_1)| + |h_1(U_2)| + 2K|h(U_1, U_2)|$$

we have that for $p \geq 1$

$$E|h_\pm(U_1, U_2)|^p \leq 3^{p-1}[2E|h_1(U_1)|^p + (2K)^p E|h(U_1, U_2)|^p].$$

From Helmers (1976):

$$E|h_1(U_1)|^p \leq C_1^p E|X|^p, \qquad E|h(U_1, U_2)|^p \leq C_2^p E|X|^p$$

where C_1 and C_2 are constants. Hence the result follows.

Define

$$\bar{U}_{n+} = \frac{n-1}{2n} U_{n+}, \qquad \bar{U}_{n-} = \frac{n-1}{2n} U_{n-}.$$

Then

$$\bar{U}_{n\pm} = \frac{n-1}{2n}\left[\frac{2}{n}\sum_{i=1}^n h_1(U_i) \pm 2K\binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} h(U_i, U_j)\right] \tag{6}$$

$$= \frac{n-1}{n^2}\sum_{i=1}^n h_1(U_i) \pm \frac{2K}{n^2}\sum_{1\leq i<j\leq n} h(U_i, U_j).$$

On the other hand

$$W_{n\pm} - E(W_{n\pm}) = n^{-1}\sum_{i=1}^n h_1(U_i) \pm \frac{2K}{n^2}\sum_{1\leq i<j\leq n} h(U_i, U_j) \tag{7}$$

$$\pm \frac{K}{n^2}\sum_{i=1}^n \{h(U_i, U_i) - E[h(U_i, U_i)]\}.$$

Subtraction of (6) and (7) gives

$$W_{n\pm} - E(W_{n\pm}) - U_{n\pm} = n^{-2}\sum_{i=1}^n \{h_1(U_i) \pm K[h(U_i, U_i) - E[h(U_i, U_i)]]\}. \tag{8}$$

The precise relationship between

$$P(\sigma^{-1}n^{1/2}(T_n - E(T_n)) > x)$$

and the introduced statistics V_n, W_{n+}, W_{n-}, $\bar{U}_{n+}$ and $\bar{U}_{n-}$ is now given in the following lemma.

LEMMA 4. *Let $\varepsilon_n > 0$, $\delta_n > 0$ and put for all real x*

$$\underset{\sim}{x}(n) = x + \sigma^{-1}n^{1/2}E(T_n - W_{n-}) + \sigma^{-1}n^{1/2}(\varepsilon_n + \delta_n)$$

$$\bar{x}(n) = x + \sigma^{-1}n^{1/2}E(T_n - W_{n+}) - \sigma^{-1}n^{1/2}(\varepsilon_n + \delta_n).$$

Then

$$P(\sigma^{-1}n^{1/2}\bar{U}_{n-} > \underset{\sim}{x}(n)) - P(|W_{n-} - E(W_{n-}) - \bar{U}_{n-}| > \delta_n) - P(|T_n - V_n| > \varepsilon_n)$$

$$\leq P(\sigma^{-1}n^{1/2}(T_n - E(T_n)) > x).$$

$$\leq P(\sigma^{-1}n^{1/2}\bar{U}_{n+} > \bar{x}(n)) + P(|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| > \delta_n) + P(|T_n - V_n| > \varepsilon_n).$$

PROOF.

$$P(\sigma^{-1}n^{1/2}(T_n - E(T_n)) > x) = P(T_n > x\sigma n^{-1/2} + E(T_n)).$$

Applying the classical Slutsky argument to $T_n = V_n + (T_n - V_n)$ gives, for $\varepsilon_n > 0$, that $P(T_n > x\sigma n^{-1/2} + E(T_n))$ is bounded above and below by respectively

$$P(V_n > x\sigma n^{-1/2} + E(T_n) - \varepsilon_n) + P(|T_n - V_n| > \varepsilon_n) \tag{9}$$

and

$$P(V_n > x\sigma n^{-1/2} + E(T_n) + \varepsilon_n) - P(|T_n - V_n| > \varepsilon_n). \tag{10}$$

Since $W_{n-} \leq V_n \leq W_{n+}$, the first term in (9) may be replaced by the larger quantity

$$P(W_{n+} - E(W_{n+}) > x\sigma n^{-1/2} + E(T_n - W_{n+}) - \varepsilon_n) \tag{11}$$

and the first term in (10) by the smaller quantity

$$P(W_{n-} - E(W_{n-}) > x\sigma n^{-1/2} + E(T_n - W_{n-}) + \varepsilon_n). \tag{12}$$

Applying the Slutsky argument again to

$$W_{n\pm} - E(W_{n\pm}) = \bar{U}_{n\pm} + [W_{n\pm} - E(W_{n\pm}) - \bar{U}_{n\pm}]$$

gives that, for $\delta_n > 0$, (11) is not larger than

$$P(\sigma^{-1}n^{1/2}\bar{U}_{n+} > \bar{x}(n)) + P(|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| > \delta_n)$$

and that (12) is not smaller than

$$P(\sigma^{-1}n^{1/2}\bar{U}_{n-} > \underline{x}(n)) - P(|W_{n-} - E(W_{n-}) - \bar{U}_{n-}| > \delta_n).$$

Combining the above inequalities gives the result of the lemma.

5. Proofs.

PROOF OF THEOREM 1(i). Applying Lemma 4 with the choice $\varepsilon_n = \delta_n = n^{-1/2}(\ln n)^{-2}$, it suffices to show that, uniformly in the range $-A \leq x \leq c\sqrt{\ln n}$:

$$P(\sigma^{-1}n^{1/2}\bar{U}_{n+} > \bar{x}(n)) = [1 - \Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right] \tag{13}$$

$$P(\sigma^{-1}n^{1/2}\bar{U}_{n-} > \underline{x}(n)) = [1 - \Phi(x)]\left[1 + o\left(\frac{1}{\ln n}\right)\right] \tag{14}$$

$$\frac{P(|W_{n\pm} - E(W_{n\pm}) - U_{n\pm}| > n^{-1/2}(\ln n)^{-2})}{1 - \Phi(x)} = o\left(\frac{1}{\ln n}\right) \tag{15}$$

$$\frac{P(|T_n - V_n| > n^{-1/2}(\ln n)^{-2})}{1 - \Phi(x)} = o\left(\frac{1}{\ln n}\right). \tag{16}$$

To prove (13), we first remark that

$$\bar{x}(n) = x + \sigma^{-1}n^{1/2}E(T_n - W_{n+}) - 2\sigma^{-1}(\ln n)^{-2} = x + O((\ln n)^{-2}).$$

Indeed,

$$|E(T_n - W_{n+})| \leq [E(T_n - V_n)^2]^{1/2} + E|V_n - W_{n+}| = O(n^{-1})$$

using Lemma 2 and the fact that $E|V_n - W_{n+}| = O(n^{-1})$ (Helmers (1980)). Now

$$\frac{P(\sigma^{-1}n^{1/2}\bar{U}_{n+} > \bar{x}(n))}{1 - \Phi(x)} = \frac{P\left(2^{-1}\sigma^{-1}n^{1/2}U_{n+} > \frac{n}{n-1}\bar{x}(n)\right)}{1 - \Phi\left(\frac{n}{n-1}\bar{x}(n)\right)} \cdot \frac{1 - \Phi\left(\frac{n}{n-1}\bar{x}(n)\right)}{1 - \Phi(x)}. \tag{17}$$

Since

$$\frac{n}{n-1}\bar{x}(n) = x + O((\ln n)^{-2}) \tag{18}$$

we have that, for $-A \leq x \leq c\sqrt{\ln n}$,

$$-A' \leq \frac{n}{n-1}\bar{x}(n) \leq c'\sqrt{\ln n}$$

for some $A' \geq 0$ and $c' > 0$, such that $2 + c'^2 < p$.

Since $2^{-1}\sigma^{-1}n^{1/2}U_{n+}$ is a properly normalized U-statistic, we can use Lemma 1 and Lemma 3 to obtain that the first factor in (17) is $1 + o\left(\frac{1}{\ln n}\right)$.

Also (18) and Lemma A1 give that the second factor in (17) is $1 + o\left(\frac{1}{\ln n}\right)$. So (13) is proved. We omit the proof of (14) since it is completely similar.

The proof of (15) will be given only for the + signs. From (8) we know that

$$W_{n+} - E(W_{n+}) - \bar{U}_{n+} = n^{-2} \sum_{i=1}^{n} Y_i$$

with $Y_i = h_1(U_i) + K[h(U_i, U_i) - E[h(U_i, U_i)]]$.

Remark that the Y_i are independent and identically distributed with zero mean. The assumption $E|X|^p < \infty$ entails that $E|Y_i|^p < \infty$; indeed, from Helmers (1976) we know

(19) $$E|h_1(U_i)|^p \leq C_1^p E|X|^p$$

(20) $$E|h(U_i, U_i)|^p \leq C_2^p E|X|^p$$

for some constants C_1 and C_2. Hence $E|W_{n+} - E(W_{n+}) - U_{n+}|^p = n^{-2p}E|\sum_{i=1}^{n} Y_i|^p$ and since $p \geq 2$ we can apply the theorem of Dharmadhikari, Fabian and Jogdeo (1968) to get $E|\sum_{i=1}^{n} Y_i|^p \leq C_3 n^{p/2}$ for some constant C_3. Hence

$$\frac{P(|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| \geq n^{-1/2}(\ln n)^{-2})}{1 - \Phi(x)} \leq C_3 \frac{n^{-p}(\ln n)^{2p}}{1 - \Phi(x)}$$

and this is clearly $o\left(\frac{1}{\ln n}\right)$ for $-A \leq x \leq c\sqrt{\ln n}$. For the proof of (16) we use Lemma 2 to obtain that for $p > 2 + c^2$

$$P(|T_n - V_n| > n^{-1/2}(\ln n)^{-2}) \leq C^p E|X|^p n^{-p/2}(\ln n)^{2p}.$$

Hence (16) follows.

PROOF OF THEOREM 2(i). We again apply Lemma 4. For the choice $\varepsilon_n = \delta_n = n^{-\alpha-1/2}$, we have to show that, uniformly in the range $-A \leq x \leq o(n^{\alpha})$:

(21) $$P(\sigma^{-1}n^{1/2}\bar{U}_{n+} > \bar{x}(n)) = [1 - \Phi(x)][1 + o(1)]$$

(22) $$P(\sigma^{-1}n^{1/2}\bar{U}_{n-} > \underline{x}(n)) = [1 - \Phi(x)][1 + o(1)]$$

(23) $$P(|W_{n\pm} - E(W_{n\pm}) - \bar{U}_{n\pm}| > n^{-\alpha-1/2}) = o(1 - \Phi(x))$$

(24) $$P(|T_n - V_n| > n^{-\alpha-1/2}) = o(1 - \Phi(x)).$$

We first prove (21). (The proof of (22) is similar.) Remark that

$$\bar{x}(n) = x + \sigma^{-1}n^{1/2}E(T_n - W_{n+}) - 2\sigma^{-1}n^{-\alpha} = x + O(n^{-\alpha})$$

(indeed: $E(T_n - W_{n+}) = O(n^{-1})$ as shown in the proof of Theorem 1(i)). Also

(25) $$\frac{n}{n-1}\bar{x}(n) = x + O(n^{-\alpha})$$

so that, for x in the prescribed range:

$$-A' \leq \frac{n}{n-1}\bar{x}(n) \leq o(n^{\alpha})$$

for some $A' \geq 0$.

Looking now at the equality (17), we easily see that the first factor tends to 1. Indeed, apply Lemma 1 to the properly normalized U-statistic $2^{-1}\sigma^{-1}n^{1/2}U_{n+}$. The required condition on $h_+(U_1, U_2)$ is satisfied because of Lemma 3 and the assumption of the theorem. Also, (25) and Lemma A1 give that the second factor in (17) tends to 1. To prove

(23) (for the + signs only), we have as in the proof of Theorem 1(i) that

$$E|W_{n+} - E(W_{n+}) - \bar{U}_{n+}|^p = n^{-2p}E|\textstyle\sum_{i=1}^n Y_i|^p.$$

Since

$$E|Y_i|^p \le 3^{p-1}[E|h_1(U_i)|^p + K^pE|h(U_i, U_i)|^p + C^p]$$

(where C is a constant) and because of (19) and (20), we have that the assumption of the theorem entails that

$$E|Y_i|^p \le K_1^p p^{\gamma p}$$

where K_1 is a constant.

Hence by Lemma A2, for all $n = 1, 2, \cdots$ and for all real $p \in [2, 2n]$:

$$E|\textstyle\sum_{i=1}^n Y_i|^p \le K_2^p p^{(\gamma+1/2)p} n^{p/2}.$$

This gives

$$E|n^{3/2}(W_{n+} - E(W_{n+}) - U_{n+})|^p \le K_2^p p^{(\gamma+1/2)p}$$

and therefore

$$\begin{aligned} P(|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| > n^{-\alpha-1/2}) &= P(n^{3/2}|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| > n^{1-\alpha}) \\ &\le P(n^{3/2}|W_{n+} - E(W_{n+}) - \bar{U}_{n+}| > n^{2\alpha(\gamma+1/2)}) \\ &\quad \left(\text{since } 1 - \alpha > 2\alpha(\gamma+1/2) \text{ for } \alpha = \frac{1}{2(3+2\gamma)}\right) \\ &= o(1 - \Phi(x)) \end{aligned}$$

uniformly in the range $-A \le x \le o(n^\alpha)$ (Lemma A3).

To prove (24), we use the assumption of the theorem together with Lemma 2 to find

$$E|n(T_n - V_n)|^p \le C^p p^{\gamma p}.$$

Therefore

$$\begin{aligned} P(|T_n - V_n| > n^{-\alpha-1/2}) &= P(n|T_n - V_n| > n^{-\alpha+1/2}) \\ &\le P(n|T_n - V_n| > n^{2\alpha\gamma}) \\ &\quad \left(\text{since } \frac{1}{2} - \alpha > 2\alpha\gamma \text{ for } \alpha = \frac{1}{2(3+2\gamma)}\right) \\ &= o(1 - \Phi(x)) \end{aligned}$$

uniformly in the range $-A \le x \le o(n^\alpha)$ (Lemma A3).

PROOF OF THEOREM 1(ii) AND 2(ii). We write

$$\frac{P\left(\dfrac{T_n - E(T_n)}{\sqrt{\mathrm{Var}(T_n)}} > x\right)}{1 - \Phi(x)} = \frac{P(\sigma^{-1}n^{1/2}(T_n - E(T_n)) > \lambda_n x)}{1 - \Phi(\lambda_n x)} \cdot \frac{1 - \Phi(\lambda_n x)}{1 - \Phi(x)} \tag{26}$$

where $\lambda_n = \sigma^{-1}n^{1/2}\sqrt{\mathrm{Var}(T_n)}$.

It will be shown that

$$\lambda_n = 1 + O(n^{-1/2}). \tag{27}$$

This enables us to apply the appropriate (i) part of the theorem to the first factor in (26) and to apply Lemma A1 to the second factor of (26), giving the desired result.

In order to prove (27), we write

$$\lambda_n^2 = \sigma^{-2} n \frac{\mathrm{Var}(T_n)}{\mathrm{Var}(V_n)} \cdot \frac{\mathrm{Var}(V_n)}{\mathrm{Var}(W_{n+})} \cdot \frac{\mathrm{Var}(W_{n+})}{\mathrm{Var}(\bar{U}_{n+})} \cdot \mathrm{Var}(\bar{U}_{n+}) \tag{28}$$

where V_n, W_{n+}, and $\bar{U}_{n+}$ are defined before.
We know from Stigler (1974) that (see (2))

$$\lim_{n\to\infty} n \,\mathrm{Var}(T'_n) = \sigma^2. \tag{29}$$

Since J is Lipschitz of order 1 on (0, 1):

$$\sum_{i=1}^{n} \left[J\left(\frac{i}{n+1}\right) - n \int_{\frac{i-1}{n}}^{\frac{i}{n}} J(s)\, ds \right]^2 = O(n^{-1})$$

and hence, as in the proof of Lemma 2,

$$\mathrm{Var}(V_n - T'_n) \le E(V_n - T'_n)^2 = O(n^{-2}). \tag{30}$$

It follows from (29) and (30) that:

$$\lim_{n\to\infty} n \,\mathrm{Var}(V_n) = \sigma^2.$$

By Lemma 2:

$$\mathrm{Var}(V_n - T_n) = O(n^{-2})$$

and hence:

$$\mathrm{Var}(T_n) = \mathrm{Var}(V_n)[1 + O(n^{-1/2})]. \tag{31}$$

The behavior of the factors $\mathrm{Var}(V_n)/\mathrm{Var}(W_{n+})$ and $\mathrm{Var}(W_{n+})/\mathrm{Var}(\bar{U}_{n+})$ can be taken over from Helmers (1980):

$$\mathrm{Var}(V_n) = \mathrm{Var}(W_{n+})[1 + O(n^{-1/2})] \tag{32}$$

$$\mathrm{Var}(W_{n+}) = \mathrm{Var}(U_{n+})[1 + O(n^{-1/2})]. \tag{33}$$

As to $\mathrm{Var}(\bar{U}_{n+})$ we have

$$\begin{aligned}\mathrm{Var}(\bar{U}_{n+}) &= \frac{(n-1)^2}{4n^2} \mathrm{Var}(U_{n+}) \\ &= \frac{(n-1)^2}{4n^2}\left[\frac{4(n-2)}{n(n-1)}\sigma^2 + \frac{2}{n(n-1)} E[h_+^2(U_1, U_2)]\right] \qquad (34)\end{aligned}$$

(since U_{n+} is a U-statistic)

$$= \sigma^2 n^{-1}[1 + O(n^{-1})].$$

Combining (31) through (34) gives that $\lambda_n^2 = 1 + O(n^{-1/2})$ or

$$\lambda_n - 1 = \frac{\lambda_n^2 - 1}{\lambda_n + 1} = O(n^{-1/2}).$$

6. APPENDIX.

LEMMA A1. *Let* λ_n *and* μ_n *be such that* $\lim_{n\to\infty} \lambda_n = 1$ *and* $\lim_{n\to\infty} \mu_n = 0$. *Put* $b_n = \max\{(|\lambda_n - 1|)^{1/2}, |\mu_n|\}$. *Let* $a_n \ge \varepsilon > 0$ *be such that* $\lim_{n\to\infty} a_n \cdot b_n = 0$. *Then*

$$[1 - \Phi(\lambda_n x + \mu_n)] = [1 - \Phi(x)][1 + O(a_n \cdot b_n)] \text{ uniformly in the range}$$

$$-A \le x \le a_n, \quad (A \ge 0).$$

PROOF. Using Taylor's expansion, we can write that

$$1 - \Phi(\lambda_n x + \mu_n) = 1 - \Phi(x) - \Phi'(\xi)[(\lambda_n - 1)x + \mu_n]$$

where $\xi = x + \vartheta[(\lambda_n - 1)x + \mu_n]$ and $|\vartheta| < 1$. This gives

$$\left| \frac{1 - \Phi(\lambda_n x + \mu_n)}{1 - \Phi(x)} - 1 \right| \le \frac{|\lambda_n - 1|\,|x| + \mu_n}{1 - \Phi(x)} \cdot \frac{e^{-\xi^2/2}}{\sqrt{2\pi}}. \tag{35}$$

If $x \le 1$, the right hand side of (35) is majorized by a constant times $|\lambda_n - 1| + |\mu_n|$ which is $O(a_n \cdot b_n)$. If $x \ge 1$, we use the fact that $e^{-x^2/2}/x\sqrt{2\pi}[1 - \Phi(x)]$ is bounded, to obtain that the right hand side of (35) is majorized by a constant times

$$a_n \cdot [|\lambda_n - 1| a_n + |\mu_n|] \cdot e^{|\lambda_n - 1| a_n^2 + |\mu_n| a_n}$$

which is easily seen to be $O(a_n \cdot b_n)$.

LEMMA A2. *Let $Y_j (j = 1, 2, \cdots)$ be a sequence of martingale summands (i.e. $E(Y_1) = 0$; $E(Y_j \mid Y_1, \cdots, Y_{j-1}) = 0, j \ge 2$). If for all $j = 1, 2, \cdots$ and all $p = 1, 2, \cdots$*

$$E|Y_j|^p \le K^p p^{\gamma p}$$

(where K and $\gamma \ge 0$ are constants, not depending on j or p) then for all $n = 1, 2, \cdots$ and all real $k \in [2, 2n]$

$$E|\textstyle\sum_{j=1}^n Y_j|^k \ge K'^k k^{(\gamma + 1/2)k} n^{k/2}$$

(where K' is a constant, not depending on n or k).

PROOF. For $k \in [2, 2n]$, define m such that $2(m - 1) < k \le 2m$ and $m \in \{1, \cdots, n\}$. Now

$$[E|\textstyle\sum_{j=1}^n Y_j|^k]^{\frac{2m}{k}} \le E|\sum_{j=1}^n Y_j|^{2m}.$$

Since $m \in \{1, \cdots, n\}$, we can apply a lemma of Bickel (1974) to obtain that

$$E|\textstyle\sum_{j=1}^n Y_j|^{2m} \le (4em)^m n^m \max_{1 \le j \le n} E|Y_j|^{2m} \le (4em)^m n^m K^{2m} (2m)^{\gamma 2m} \quad \text{(by assumption)}.$$

Hence

$$\begin{aligned} E|\textstyle\sum_{j=1}^n Y_j|^k &\le (4em)^{k/2} K^k (2m)^{\gamma k} n^{k/2} \\ &\le (4ek)^{k/2} K^k (2k)^{\gamma k} n^{k/2} \text{ (since } m \le k) \\ &= K'^k k^{(\gamma + 1/2)k} n^{k/2} \end{aligned}$$

with $K' = 2^{\gamma+1} e^{1/2} K$.

LEMMA A3. *Let $Z_n (n = 1, 2, \cdots)$ be a sequence of random variables. If for all $n = 1, 2, \cdots$ and all real $k \in [2, 2n]$*

$$E|Z_n|^k \le K^k k^{\gamma k}$$

(where K and $\gamma \ge 0$ are constants, not depending on n or k) then for $\gamma \ge c > 0$

$$P(|Z_n| > n^c) = o(1 - \Phi(x))$$

uniformly in the range $-A \le x \le o(n^{c/2\gamma}) (A \ge 0)$.

PROOF. For all $n = 1, 2, \cdots$ and all real $k \in [2, 2n]$

$$P(|Z_n| > n^c) \le n^{-kc} E|Z_n|^k \le (n^{-c} K k^{\gamma})^k.$$

We specify the choice of k as follows:

$$k = \rho n^{c/\gamma} \text{ with } \rho = \min\left(\left(\frac{e^{-1}}{K}\right)^{\frac{1}{\gamma}}, 2\right).$$

Since $c \leq \gamma$ we have that $k \leq 2n$ and since $c > 0$ we have that $k \geq 2$ (for sufficiently large n). With this choice of k we have that

$$n^{-c}Kk^{\gamma} \leq e^{-1}$$

so that

$$P(|Z_n| > n^c) \leq e^{-k} = e^{-\rho n^{c/\gamma}}.$$

Hence we obtain that for $-A \leq x \leq o(n^{c/2\gamma})$:

$$\frac{P(|Z_n| > n^c)}{1 - \Phi(x)} \leq o(n^{c/2\gamma})e^{-\rho n^{c/\gamma} + \frac{1}{2}o(n^{c/\gamma})}$$

which tends to 0 as $n \to \infty$.

REFERENCES

[1] Amosova, N. N. (1972). On limit theorems for the probabilities of moderate deviations. *Vestnik Leningrad. Univ.* **13** 5–14 (in Russian).

[2] Bickel, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

[3] Callaert, H. and Janssen, P. (1978). The Berry-Esseen theorem for U-statistics. *Ann. Statist.* **6** 417–421.

[4] Cramer, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Sci. Ind. n°* **736** 5–23.

[5] Dharmadhikari, S. W., Fabian, V. and Jogdeo, K. (1968). Bounds on the moments of martingales. *Ann. Math. Statist.* **39** 1719–1723.

[6] Helmers, R. (1976). Edgeworth expansions for linear combinations of order statistics with smooth weight functions. Preprint, SW 44/76. Mathematisch Centrum, Amsterdam.

[7] Helmers, R. (1978). Edgeworth expansions for linear combinations of order statistics. Ph.D. Thesis, Mathematisch Centrum, Amsterdam.

[8] Helmers, R. (1981). A Berry-Esseen theorem for linear combinations of order statistics. *Ann. Probability* **9** 342–347.

[9] Ibragimov, I. A. and Linnik, Yu. V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

[10] Malevich, T. L. and Abdalimov, B. (1979). Large deviation probabilities for U-statistics. *Theor. Probability Appl.* **24** 215–219.

[11] Petrov, V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

[12] Rubin, H. and Sethuraman, J. (1965). Probabilities of moderate deviations. *Sankhya Series A*, **27** 325–346.

[13] Stigler, S. M. (1974). Linear functions of order statistics with smooth weight functions. *Ann. Statist.* **2** 676–693.

[14] Vandemaele, M. (1980). On probabilities of large deviations for U-statistics. (to appear).

Department of Mathematics
Limburgs Universitair Centrum
B-3610 Diepenbeek (Belgium)