BULL. AUSTRAL. MATH. SOC. 60F10
VOL. 63 (2001) [151-165]

LARGE DEVIATION RESULTS FOR A U-STATISTICAL SUM WITH PRODUCT KERNEL

Y.V. BOROVSKIKH AND N.C. WEBER

Large deviation theorems are proved for non-degenerate U-statistical sums of degree m with kernel $h(x_1,\dots,x_m) = x_1 \cdots x_m$ under the Cramér condition and under the Linnik condition. The method of proof uses truncation and the contraction technique.

1. INTRODUCTION AND THEOREMS

Let $X_1,\cdots,X_n$ be independent and identically distributed random variables with distribution function F. Assume $EX_1 = \mu \neq 0$ and $0 < \sigma^2 = E(X_1-\mu)^2 < \infty$. Consider the U-statistic of the form

$$U_n = \binom{n}{m}^{-1} \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} X_{i_1}\cdots X_{i_m}.$$

According to the Hoeffding decomposition

$$U_n - \mu^m = \sum_{c=1}^{m} \frac{m(m-1)\cdots(m-c+1)}{n(n-1)\cdots(n-c+1)}\mu^{m-c} \sum_{1\leqslant i_1<\cdots<i_c\leqslant n} (X_{i_1}-\mu)\cdots(X_{i_c}-\mu). \tag{1}$$

For $m=1$, $U_n - \mu = n^{-1}\sum_{i=1}^{n}(X_i-\mu)$ is the usual sum of independent and identically distributed random variables. Large deviation results with Cramér series for such sums have been developed in many papers including [**1, 4, 6, 7, 8**], both under the Cramér and Linnik conditions and under violation of these conditions.

Let $F_n(x) = P\left(\sqrt{n}(U_n-\mu^m)/(m\sigma|\mu|^{m-1}) \leqslant x\right)$, $x \in R$. By the central limit theorem for non-degenerate U-statistics, $F_n(x) \to \Phi(x)$ uniformly in x as $n \to \infty$, where $\Phi(x)$ denotes the standard normal distribution function. Hence for any fixed m and x

$$\bigl(1-F_n(x)\bigr)/\bigl(1-\Phi(x)\bigr) = \bigl(1+o(1)\bigr), \tag{2}$$

as $n \to \infty$. We are interested in (2) when x tends to infinity together with n and $x = o(\sqrt{n})$. All existing large deviation results for U-statistical sums with x in this range require the kernels to be bounded and so they do not apply here. For our special case we have the following result where the coefficients λ_{km} are defined via equation (18).

Received 15th June, 2000

Theorem 1. *Let $x = o(\sqrt{n})$ and suppose that the Cramér condition*

$$E\exp(a|X_1|) < \infty \tag{3}$$

is satisfied for some $a > 0$. Then for any fixed $m \geqslant 1$

$$(1 - F_n(x))/(1 - \Phi(x)) = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right)$$

with the Cramér series

$$\lambda_m(u) = \sum_{k=0}^{\infty}\lambda_{km}u^k$$

which converges for $0 \leqslant u \leqslant \varepsilon$ for some $\varepsilon > 0$.

Further, for $0 < \alpha < 1/2$, let $s = [4\alpha/(1-2\alpha)]$ denote the integer part of $4\alpha/(1-2\alpha)$. Introduce the truncated Cramér series

$$\lambda_m^{[s]}(u) = \lambda_{0m} + \lambda_{1m}u + \cdots + \lambda_{sm}u^s.$$

Theorem 2. *Let $x = o(n^\alpha)$ and suppose that the Linnik condition*

$$E\exp\left\{a|X_1|^{4\alpha/(2\alpha+1)}\right\} < \infty \tag{4}$$

is satisfied for some $a > 0$. Then for any fixed $m \geqslant 1$

$$(1 - F_n(x))/(1 - \Phi(x)) = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right).$$

2. Proofs

In the following we can suppose that $x \geqslant 1$. In fact, by [2, Theorem 6.3.2]

$$\sup_x|F_n(x) - \Phi(x)| \leqslant 36\left(\sigma^{-3}E|X_1 - \mu|^3 + m\exp\{2(\sigma/\mu)^2\}\right)\frac{1}{\sqrt{n}}$$

for all $1 \leqslant m \leqslant \sqrt{n}$. Further $1 - \Phi(x) \geqslant \pi^{-2}$ for $0 \leqslant x \leqslant 1$. Hence for any fixed $m \geqslant 1$ and for $0 \leqslant x \leqslant 1$

$$(1 - F_n(x))/(1 - \Phi(x)) = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right).$$

Proof of Theorem 1: The contraction technique for U-statistics of degree 2, proposed in [5] , was extended in [3] to U-statistics with bounded kernels of any degree. Following their approach we define the operator $T_t^{m-1}\colon g \to T_t^{m-1}g$, by

$$T_t^{m-1}g = \frac{E\left\{\prod_{s=1}^{m-1}e^{tg(X_s)}h(X_1,\ldots,X_{m-1},\cdot)\right\}}{\left(Ee^{tg(X_1)}\right)^{m-1}}, \quad t \in R$$

for any fixed integer $m = 2, 3 \ldots$, where $g(X_1)$ satisfies the Cramér condition, that is, $E\exp\big(a|g(X_1)|\big) < \infty$ for some $a > 0$.

The following lemma shows that under the Cramér condition (3), for sufficiently small t, the equation $g = T_t^{m-1} g$ with kernel

$$h(x_1, \ldots, x_m) = x_1 \cdots x_m. \tag{5}$$

has a unique solution g_t, that is,

$$g_t = T_t^{m-1} g_t. \tag{6}$$

□

LEMMA 1. *If the Cramér condition (3) is satisfied then the function*

$$g_t(x) = \mu_m^{m-1}(t) \cdot x, \qquad x \in R$$

is the solution of (6) for kernel (5), where $\mu_m(t)$ is an analytical function in the region $0 \leqslant t \leqslant b$ for some sufficiently small $b > 0$.

PROOF: If some function $g_t(x)$ is the solution of (6) then we can write

$$g_t(x) = \frac{\Big(E\big(e^{tg_t(X_1)} X_1\big)\Big)^{m-1}}{\big(Ee^{tg_t(X_1)}\big)^{m-1}} \cdot x.$$

Let

$$\mu_m(t) = \frac{E\big(e^{tg_t(X_1)} X_1\big)}{Ee^{tg_t(X_1)}}.$$

Hence, for $\mu_m(t)$ we obtain the equation

$$\mu_m(t) = \frac{E\big(e^{t\mu_m^{m-1}(t)X_1} X_1\big)}{Ee^{t\mu_m^{m-1}(t)X_1}}. \tag{7}$$

Futher we shall prove that this equation has a unique analytical solution for small t. To show this let

$$\xi = t\mu_m^{m-1}(t). \tag{8}$$

Then (7) can be written as

$$t = \xi \cdot \left(\frac{Ee^{\xi X_1}}{E\big(X_1 e^{\xi X_1}\big)}\right)^{m-1}. \tag{9}$$

Under condition (3) the functions $E\exp(\xi X_1)$, $E\big(X_1\exp(\xi X_1)\big)$ and $E\exp(\xi X_1)/E\big(X_1\exp(\xi X_1)\big)$ are analytic for $|\xi| \leqslant a/2$ and therefore in (9)

$$t = \xi \cdot \sum_{k=0}^{\infty} c_{km}\xi^k, \tag{10}$$

where the power series converges for small $|\xi|$. For any sufficiently small $|t|$ the equation (10) has a unique real solution $\xi = \xi(t)$, where

$$\xi(t) = t \cdot \sum_{k=0}^{\infty} d_{km} t^k \tag{11}$$

with convergent series for small $|t|$. From (8) and (11) we obtain

$$\mu_m^{m-1}(t) = \sum_{k=0}^{\infty} d_{km} t^k. \tag{12}$$

From (12) we see that for any $m \geqslant 2$

$$\mu_m(t) = \sum_{k=0}^{\infty} \mu_{km} t^k \tag{13}$$

where the power series converges for sufficiently small t. This proves Lemma 1. ▯

Let $\psi(t) = \ln E \exp\{t g_t(X_1)\}$, where $g_t(x)$ is given in Lemma 1. Define F_t by

$$dF_t = e^{tg_t - \psi(t)} dF.$$

Let P_t denote the probability measure under which $X_1, X_2, \ldots$ are independent and identically distributed from F_t and let E_t denote the expectation under P_t. Hence, in (7)

$$\mu_m(t) = E_t X_1. \tag{14}$$

Let $t = t(x/\sqrt{n})$, where $t(x/\sqrt{n})$ is the solution of the equation

$$\mu_m^m(t) - \mu^m = \frac{x}{\sqrt{n}} m\sigma |\mu|^{m-1} \tag{15}$$

with $x\,n^{-1/2} \to 0$. The existence of this solution is guaranteed by (13). In fact,

$$\mu_m^m(t) = \mu^m + \sum_{k=1}^{\infty} \beta_{km} t^k,$$

where the power series converges for sufficiently small t. From (15) we obtain

$$\sum_{k=1}^{\infty} \beta_{km} t^k = \frac{x}{\sqrt{n}} m\sigma |\mu|^{m-1}.$$

By the inversion theorem for analytic functions we get

$$t\left(\frac{x}{\sqrt{n}}\right) = \sum_{k=1}^{\infty} t_{km} \left(\frac{x}{\sqrt{n}}\right)^k \tag{16}$$

with convergent series for small $x n^{-1/2}$.

Further, let $U_n(t)$ denote the U-statistical sum U_n, where $X_1, \ldots, X_n$ are independent and identically distributed with distribution function F_t. By analogy with (1) and by using Lemma 1 we have the Hoeffding decomposition under P_t

$$U_n(t) = \mu_m^m(t) + \frac{m}{n}\sum_{i=1}^{n}\left(g_t(X_i) - \mu_m^m(t)\right)$$
$$+ \sum_{c=2}^{m} \frac{m(m-1)\cdots(m-c+1)}{n(n-1)\cdots(n-c+1)}\mu_m^{m-c}(t) \sum_{1\leqslant i_1<\cdots<i_c\leqslant n} (X_{i_1} - \mu_m(t))\ldots(X_{i_c} - \mu_m(t)).$$

Let

$$\sigma_1 = \frac{1}{n}\sum_{j=1}^{n}(g_t(X_j) - \mu_m^m(t)) \quad \text{and} \quad v_n = \sum_{c=2}^{m}\lambda_c(t)\pi_c,$$

where

$$\lambda_c(t) = \frac{(m-1)\cdots(m-c+1)}{(n-1)\cdots(n-c+1)}\cdot n^{c-1}\mu_m^{-m(c-1)}(t),$$

and

$$\pi_c = n^{-c}\sum_{1\leqslant i_1<\cdots<i_c\leqslant n}(g_t(X_{i_1}) - \mu_m^m(t))\cdots(g_t(X_{i_c}) - \mu_m^m(t)).$$

Then

$$U_n(t) = \mu_m^m(t) + m\sigma_1 + mv_n.$$

Hence using (15)

$$U_n(t) - \frac{x}{\sqrt{n}}m\sigma|\mu|^{m-1} = m\sigma_1 + mv_n + \mu^m$$

and

$$\begin{aligned}
(17)\quad & P\left(U_n - \mu^m > \frac{x}{\sqrt{n}}m\sigma|\mu|^{m-1}\right)\\
&= \int I\left(U_n(y_1, \ldots, y_n) - \mu^m > \frac{x}{\sqrt{n}}m\sigma|\mu|^{m-1}\right)\prod_{i=1}^{n}dF(y_i)\\
&= \int \exp\left\{\sum_{i=1}^{n}(\psi(t) - tg_t(y_i))\right\}I\left(U_n - \mu^m > \frac{x}{\sqrt{n}}m\sigma|\mu|^{m-1}\right)\prod_{i=1}^{n}dF_t(y_i)\\
&= E_t\left[e^{n\psi(t)-nt\mu_m^m(t)-tn\sigma_1}I\left(U_n(t) - \frac{x}{\sqrt{n}}m\sigma|\mu|^{m-1} - \mu^m > 0\right)\right]\\
&= e^{n\psi(t)-nt\mu_m^m(t)}\cdot J_n,
\end{aligned}$$

where $J_n = E_t e^{-tn\sigma_1}I(\sigma_1 + v_n > 0)$.

Using (16) and the definitions of $\psi(t)$ and $\mu_m(t)$ we have

$$\begin{aligned}
(18)\qquad n\psi(t) - nt\mu_m^m(t) &= -\frac{x^2}{2} + \lambda_{0m}\frac{x^3}{\sqrt{n}} + \lambda_{1m}\frac{x^4}{n} + \ldots\\
&= -\frac{x^2}{2} + \frac{x^3}{\sqrt{n}}\lambda_m\left(\frac{x}{\sqrt{n}}\right).
\end{aligned}$$

Next we shall estimate J_n. Let $\xi_j = g_t(X_j) - \mu_m^m(t)$, $j = 1, \ldots, n$, and introduce the truncated random variables

$$\overline{\xi}_j = \xi_j I(|\xi_j| \leqslant x\sqrt{n}), \quad j = 1, \ldots, n.$$

Then $J_n = J_n(\xi_1, \ldots, \xi_n)$ and $\overline{J}_n = J_n(\overline{\xi}_1, \ldots, \overline{\xi}_n) = E_t e^{-tn\overline{\sigma}_1} I(\overline{\sigma}_1 + \overline{v}_n > 0)$.

Lemma 2. *If $x \geqslant 1$ and $xn^{-1/2} \to 0$ as $n \to \infty$, then*

$$J_n = \overline{J}_n + e^{x^2/2}(1 - \Phi(x)) \cdot O\left(\frac{x}{\sqrt{n}}\right).$$

Proof: Denote

$$f(\xi_1, \ldots, \xi_n) = e^{-tn\sigma_1} I(\sigma_1 + v_n > 0).$$

Note

$$\begin{aligned} J_n - \overline{J}_n &= E_t\big[f(\xi_1, \ldots, \xi_n) - f(\overline{\xi}_1, \ldots, \overline{\xi}_n)\big] \\ &= \sum_{j=1}^n E_t\big[f(\overline{\xi}_1, \ldots, \overline{\xi}_{j-1}, \xi_j, \ldots, \xi_n) - f(\overline{\xi}_1, \ldots, \overline{\xi}_j, \xi_{j+1}, \ldots, \xi_n)\big] \\ &= \sum_{j=1}^n E_t I(|\xi_j| > x\sqrt{n}) \big[f(\overline{\xi}_1, \ldots, \overline{\xi}_{j-1}, \xi_j, \xi_{j+1} \ldots, \xi_n) \\ &\qquad\qquad - f(\overline{\xi}_1, \ldots, \overline{\xi}_{j-1}, 0, \xi_{j+1}, \ldots, \xi_n)\big]. \end{aligned}$$

Thus

$$|J_n - \overline{J}_n| \leqslant \sum_{j=1}^n E_t\Big[I(|\xi_1| > x\sqrt{n})(e^{t|\xi_1|} + 1)\Big] \left(E_t e^{t\xi_1}\right)^{n-j} \left(E_t e^{t\overline{\xi}_1}\right)^{j-1}. \tag{19}$$

Since

$$e^y = 1 + y + \theta y^2 e^{|y|}, \quad |\theta| \leqslant 1, \quad y \in R \tag{20}$$

under condition (3), $E_t e^{t\xi_1} \leqslant \exp\{c_1 x^2/n\}$, $E_t e^{t\overline{\xi}_1} \leqslant \exp\{c_1 x^2/n\}$, and

$$\begin{aligned} E_t I(|\xi_1| > x\sqrt{n}) e^{t|\xi_1|} &\leqslant \left(E_t e^{2t|\xi_1|}\right)^{1/2} \Big(P_t(|\xi_1| > x\sqrt{n})\Big)^{1/2} \\ &\leqslant \exp\{c_1 x^2/n\} \Big(P_t(|\xi_1| > x\sqrt{n})\Big)^{1/2} \leqslant c_2 e^{-c_3 x\sqrt{n}} \end{aligned}$$

for some positive constants $c_i > 0$, which do not depend on x and n. Hence in (19)

$$\begin{aligned} |J_n - \overline{J}_n| &\leqslant 2c_2 n \exp\{c_1 x^2 - c_3 x\sqrt{n}\} \\ &= e^{x^2/2}(1 - \Phi(x)) O\left(\frac{x}{\sqrt{n}}\right), \end{aligned}$$

since

$$e^{x^2/2}\left(1-\Phi(x)\right)=O\left(\frac{1}{x}\right),\ x\to\infty.$$

This proves Lemma 2. ∎

Consider $\overline{J}_n$. The second Waring formula gives the following representation for $\overline{\pi}_c$:

$$\overline{\pi}_c=\sum_{\{i_s\geqslant 0: i_1+2i_2+\cdots+ci_c=c\}}\prod_{s=1}^{c}(-1)^{i_s-1}(s^{i_s}i_s!)^{-1}\,\overline{\sigma}_s{}^{i_s},$$

where

$$\overline{\sigma}_s=n^{-s}\sum_{k=1}^{n}\overline{\xi}_k^s,\qquad s=1,\ldots,c.$$

Next we introduce the truncated random variables

$$\widetilde{\sigma}_j=\overline{\sigma}_jI\left(|\overline{\sigma}_j|\leqslant\delta\right),\qquad j=1,\ldots,m$$

for some sufficiently small $\delta>0$. In addition, $\widetilde{J}_n$, $\widetilde{v}_n$, $\widetilde{\pi}_c$ are obtained from $\overline{J}_n$, $\overline{v}_n$, $\overline{\pi}_c$, respectively, by substituting $\widetilde{\sigma}_j$ for $\overline{\sigma}_j$.

LEMMA 3. *If $x\geqslant 1$ and $xn^{-1/2}\to 0$, then*

$$\overline{J}_n=\widetilde{J}_n+e^{x^2/2}\left(1-\Phi(x)\right)O\left(\frac{x}{\sqrt{n}}\right).$$

PROOF: Denote $\varphi(\overline{\sigma}_1,\ldots,\overline{\sigma}_m)=e^{-tn\overline{\sigma}_1}I(\overline{\sigma}_1+\overline{v}_n>0)$.

Arguing as in the proof of Lemma 2

$$\begin{aligned}\overline{J}_n-\widetilde{J}_n&=E_t\left[\varphi(\overline{\sigma}_1,\ldots,\overline{\sigma}_m)-\varphi(\widetilde{\sigma}_1,\ldots,\widetilde{\sigma}_m)\right]\\&=\sum_{k=1}^{m}E_tI\left(|\overline{\sigma}_k|>\delta\right)\left[\varphi(\widetilde{\sigma}_1,\ldots,\widetilde{\sigma}_{k-1},\overline{\sigma}_k,\overline{\sigma}_{k+1},\ldots,\overline{\sigma}_m)\right.\\&\qquad\left.-\varphi(\widetilde{\sigma}_1,\ldots,\widetilde{\sigma}_{k-1},0,\overline{\sigma}_{k+1},\ldots,\overline{\sigma}_m)\right]\end{aligned}$$

from which we obtain

$$\begin{aligned}|\overline{J}_n-\widetilde{J}_n|&\leqslant 2\sum_{k=1}^{m}E_tI\left(|\overline{\sigma}_k|>\delta\right)e^{tn|\overline{\sigma}_1|}\\&\leqslant 2\sum_{k=1}^{m}\left(E_te^{2tn|\overline{\sigma}_1|}\right)^{1/2}\cdot\left(P_t\left(|\overline{\sigma}_k|>\delta\right)\right)^{1/2}.\end{aligned}$$

Here

$$\begin{aligned}E_te^{2tn|\overline{\sigma}_1|}&\leqslant E_te^{2tn\overline{\sigma}_1}+E_te^{-2tn\overline{\sigma}_1}\\&=\left(E_te^{2t\overline{\xi}_1}\right)^n+\left(E_te^{-2t|\overline{\xi}_1|}\right)^n\leqslant c_4e^{c_5x^2}\end{aligned}$$

and

$$(21)\qquad P_t(|\bar\sigma_k| > \delta) \leqslant e^{-\delta x\sqrt{n}} E_t \exp\{x\sqrt{n}|\bar\sigma_k|\}$$
$$\leqslant e^{-\delta x\sqrt{n}} E_t \exp\{x\sqrt{n}\bar\sigma_k\} + e^{-\delta x\sqrt{n}} E_t \exp\{-x\sqrt{n}\bar\sigma_k\}.$$

Now

$$E_t \exp\{x\sqrt{n}\bar\sigma_k\} = \left(E_t \exp\left\{\frac{x\sqrt{n}}{n^k}\bar\xi_1^k\right\}\right)^n,$$

and using (20)

$$(22)\qquad E_t \exp\left\{\frac{x\sqrt{n}}{n^k}\bar\xi_1^k\right\} \leqslant 1 + \frac{x\sqrt{n}}{n^k}E_t\bar\xi_1^k + \frac{x^2}{n^{2k-1}}E_t\left(\bar\xi_1^{2k}\exp\left\{\frac{x\sqrt{n}}{n^k}|\bar\xi_1^k|\right\}\right).$$

For $k = 1$, $|E_t\bar\xi_1| \leqslant c_6(x\sqrt{n})^{-1}$, and

$$E_t\bar\xi_1^2 \exp\left\{\frac{x}{\sqrt{n}}|\bar\xi_1|\right\} \leqslant (E_t\bar\xi_1^4)^{1/2}\left(E_t\exp\left\{\frac{2x}{\sqrt{n}}|\bar\xi_1|\right\}\right)^{1/2} \leqslant c_7.$$

Hence

$$(23)\qquad E_t \exp\{x\sqrt{n}\bar\sigma_1\} \leqslant \exp\{c_6 + c_7x^2\}.$$

Now consider (22) when $k \geqslant 2$. In this case $|E_t\bar\xi_1^k| \leqslant c_8$, and

$$E_t\left(\bar\xi_1^{2k}\exp\left\{\frac{x\sqrt{n}}{n^k}|\bar\xi_1^k|\right\}\right) \leqslant \left(E_t\bar\xi_1^{4k}\right)^{1/2}\left(E_t\exp\left\{\frac{2x\sqrt{n}}{n^k}|\bar\xi_1^k|\right\}\right)^{1/2}$$
$$\leqslant \left(E_t\bar\xi_1^{4k}\right)^{1/2}\left(E_t\exp\left\{2\left(\frac{x}{\sqrt{n}}\right)^k|\bar\xi_1^k|\right\}\right)^{1/2} \leqslant c_9.$$

Hence for $k \geqslant 2$,

$$(24)\qquad E_t \exp\{x\sqrt{n}\bar\sigma_k\} \leqslant \exp\left\{c_8\frac{x}{\sqrt{n}} + c_9\left(\frac{x}{n^{k-1}}\right)^2\right\}$$

for some positive constants c_i. These estimates combined with (21) prove Lemma 3. ☐

Next we shall estimate $\widetilde{J}_n$.

Lemma 4. *If $x \geqslant 1$ and $x = o(\sqrt{n})$, $n \to \infty$, then*

$$\widetilde{J}_n = e^{x^2/2}(1 - \Phi(x))\left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right).$$

Proof: Write

$$\widetilde{\pi}_c = \widetilde{\sigma}_1 \cdot \alpha_c + \beta_c, \qquad c \geqslant 2$$

https://doi.org/10.1017/S0004972700019213 Published online by Cambridge University Press

where

$$\alpha_c = \sum_{\substack{i_1+2i_2+\cdots+ci_c=c \\ i_1\geqslant 1,\, i_s\geqslant 0,\, s=2,\ldots,c}} \left(\prod_{s=1}^{c}(-1)^{i_s-1}(s^{i_s}i_s!)^{-1}\right)\widetilde{\sigma}_1^{i_1-1}\widetilde{\sigma}_2^{i_2}\cdots\widetilde{\sigma}_c^{i_c},$$

$$\beta_c = \sum_{2i_2+\cdots+ci_c=c} (-1)^{c+i_2+\cdots+i_c}\left(\prod_{s=2}^{c}(s^{i_s}i_s!)^{-1}\right)\widetilde{\sigma}_2^{i_2}\cdots\widetilde{\sigma}_c^{i_c}.$$

Then

$$\widetilde{v}_n = \widetilde{\sigma}_1\alpha + \beta, \tag{25}$$

where

$$\alpha = \sum_{c=2}^{m}\lambda_c(t)\alpha_c \quad \text{and} \quad \beta = \sum_{c=2}^{m}\lambda_c(t)\beta_c.$$

Note that $|\alpha| \leqslant a_1\cdot\delta \leqslant 1/2$ for some constant $a_1 > 0$ and sufficiently small δ. Further, taking account of (25) we have

$$\widetilde{J}_n = E_t e^{-tn\widetilde{\sigma}_1}I(\widetilde{\sigma}_1 + \widetilde{v}_n > 0) = E_t e^{-tn\widetilde{\sigma}_1}I\left(\widetilde{\sigma}_1 > -\frac{\beta}{1+\alpha}\right) = \sum_{i=1}^{5}\Delta_i,$$

where

$$\Delta_1 = E_t e^{-tn\sigma_1}I(\sigma_1 > 0),$$
$$\Delta_2 = E_t e^{-tn\overline{\sigma}_1}I(\overline{\sigma}_1 > 0) - E_t e^{-tn\sigma_1}I(\sigma_1 > 0),$$
$$\Delta_3 = E_t e^{-tn\widetilde{\sigma}_1}I(\widetilde{\sigma}_1 > 0) - E_t e^{-tn\overline{\sigma}_1}I(\overline{\sigma}_1 > 0),$$
$$\Delta_4 = E_t e^{-tn\widetilde{\sigma}_1}I\left(\widetilde{\sigma}_1 > -\frac{\beta}{1+\alpha}\right)I\left(\frac{|\beta|}{1+\alpha}\leqslant\frac{1}{n}\right) - E_t e^{-tn\widetilde{\sigma}_1}I(\widetilde{\sigma}_1 > 0),$$
$$\Delta_5 = E_t e^{-tn\widetilde{\sigma}_1}I\left(\widetilde{\sigma}_1 > -\frac{\beta}{1+\alpha}\right)I\left(\frac{|\beta|}{1+\alpha} > \frac{1}{n}\right).$$

ESTIMATE Δ_1. Let $\sigma_t^2 = E_t\big(g_t(X_1) - \mu_m^m(t)\big)^2$, $x_t = \sqrt{n}t\sigma_t$,

$$S = \frac{1}{\sigma_t\sqrt{n}}\sum_{i=1}^{n}\big(g_t(X_1) - \mu_m^m(t)\big), \quad \text{and} \quad \Phi_n(y) = P_t(S\leqslant y), \quad y\in R.$$

Note that by Lemma 1, (15) and (16)

$$\sigma_t = \sigma|\mu|^{m-1} + O\left(\frac{x}{\sqrt{n}}\right), \qquad t = \frac{x}{\sigma|\mu|^{m-1}\sqrt{n}}\left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right), \quad \text{and}$$

$$x_t = x + \frac{x^2}{\sqrt{n}}O(1). \tag{26}$$

https://doi.org/10.1017/S0004972700019213 Published online by Cambridge University Press

By definition

$$(27)\qquad \Delta_1 = \int e^{-x_t y} I(y>0)\,d\Phi_n(y) = \int_0^\infty e^{-x_t y}\,d\Phi(y) + \int_0^\infty e^{-x_t y}\,d\big(\Phi_n(y)-\Phi(y)\big).$$

Recall $e^z = 1 + z\int_0^1 e^{uz}du$, $z\in R$, so using (26) we have

$$\int_0^\infty e^{-x_t y}\,d\Phi(y) = \int_0^\infty \exp\Big\{-xy\Big(1+O\Big(\frac{x}{\sqrt{n}}\Big)\Big)\Big\}d\Phi(y) = \int_0^\infty e^{-xy}\,d\Phi(y) - \frac{x^2}{\sqrt{n}}\int_0^1\int_0^\infty O(1)y\exp\Big\{-xy-\frac{x^2}{\sqrt{n}}yuO(1)\Big\}d\Phi(y)du.$$

Here

$$\int_0^\infty e^{-xy}\,d\Phi(y) = e^{x^2/2}\big(1-\Phi(x)\big)$$

and

$$\frac{x^2}{\sqrt{n}}\int_0^\infty y\exp\Big\{-xy-\frac{x^2}{\sqrt{n}}yuO(1)\Big\}d\Phi(y) \leqslant \frac{x^2}{\sqrt{n}}\int_0^\infty y\exp\Big\{-xy-\frac{x^2}{\sqrt{n}}yuO(1)\Big\}dy = \frac{x^2}{\sqrt{n}}\Big(x+\frac{x^2}{\sqrt{n}}uO(1)\Big)^{-2} = e^{x^2/2}\big(1-\Phi(x)\big)O\Big(\frac{x}{\sqrt{n}}\Big).$$

Consequently,

$$\int_0^\infty e^{-x_t y}\,d\Phi(y) = e^{x^2/2}\big(1-\Phi(x)\big)\Big(1+O\Big(\frac{x}{\sqrt{n}}\Big)\Big).$$

Further, in (27)

$$\int_0^\infty e^{-x_t y}\,d\big(\Phi_n(y)-\Phi(y)\big) = -\big(\Phi_n(0)-\Phi(0)\big) + x_t\int_0^\infty \big(\Phi_n(y)-\Phi(y)\big)e^{-x_t y}dy.$$

By the Berry-Esseen theorem (see, for example, [7])

$$\sup_y\big|\Phi_n(y)-\Phi(y)\big| \leqslant \sigma_t^{-3}E\big|g_t(X_1)-\mu_m^m(t)\big|^3\frac{1}{\sqrt{n}}.$$

Hence,

$$\Big|\int_0^\infty e^{-x_t y}\,d\big(\Phi_n(y)-\Phi(y)\big)\Big| \leqslant 2\sigma_t^{-3}E\big|g_t(X_1)-\mu_m^m(t)\big|^3\frac{1}{\sqrt{n}} = e^{x^2/2}\big(1-\Phi(x)\big)O\Big(\frac{x}{\sqrt{n}}\Big),$$

and so we have

$$\Delta_1 = e^{x^2/2}\big(1-\Phi(x)\big)\Big(1+O\Big(\frac{x}{\sqrt{n}}\Big)\Big).$$

https://doi.org/10.1017/S0004972700019213 Published online by Cambridge University Press

ESTIMATE Δ_2. Let $\Psi(\xi_1,\dots,\xi_n)=e^{-tn\sigma_1}I(\sigma_1>0)$. Then

$$\begin{aligned}\Delta_2 &= E_t\big[\Psi(\overline{\xi}_1,\dots,\overline{\xi}_n)-\Psi(\xi_1,\dots,\xi_n)\big]\\ &=\sum_{j=1}^{n}E_tI\big(|\xi_j|>x\sqrt{n}\big)\big[\Psi(\overline{\xi}_1,\dots,\overline{\xi}_{j-1},0,\xi_{j+1},\dots,\xi_n)\\ &\qquad -\Psi(\overline{\xi}_1,\dots,\overline{\xi}_{j-1},\xi_j,\xi_{j+1},\dots,\xi_n)\big]\end{aligned}$$

and therefore

$$|\Delta_2|\leqslant 2\sum_{j=1}^{n}E_tI\big(|\xi_j|>x\sqrt{n}\big)=2nP_t\big(|\xi_1|>x\sqrt{n}\big)\leqslant 2E_t\xi_1^4n^{-1}.$$

Hence,

$$\Delta_2=e^{x^2/2}(1-\Phi(x))O\Big(\frac{x}{\sqrt{n}}\Big).$$

ESTIMATE Δ_3. By definition, $\Delta_3=E_tI\big(|\overline{\sigma}_1|>\delta\big)\big(1-e^{-tn\overline{\sigma}_1}I(\overline{\sigma}_1>0)\big)$ and so from (21) and (23) we have

$$\Delta_3=e^{x^2/2}(1-\Phi(x))O\Big(\frac{x}{\sqrt{n}}\Big).$$

ESTIMATE Δ_4. Clearly

$$-E_te^{-tn\widetilde{\sigma}_1}I\Big(0\leqslant\widetilde{\sigma}_1\leqslant\frac{1}{n}\Big)\leqslant\Delta_4\leqslant E_te^{-tn\widetilde{\sigma}_1}I\Big(-\frac{1}{n}\leqslant\widetilde{\sigma}_1\leqslant 0\Big).$$

That is,

$$\begin{aligned}(28)\qquad |\Delta_4| &\leqslant e^tE_tI\big(|\widetilde{\sigma}_1|\leqslant n^{-1}\big)\\ &= e^t\Big[E_tI\big(|\overline{\sigma}_1|\leqslant\delta\big)I\big(|\widetilde{\sigma}_1|\leqslant n^{-1}\big)+E_tI\big(|\overline{\sigma}_1|>\delta\big)I\big(|\widetilde{\sigma}_1|\leqslant n^{-1}\big)\Big]\\ &\leqslant e^t\Big[P_t\big(|\overline{\sigma}_1|\leqslant n^{-1}\big)+P_t\big(|\overline{\sigma}_1|>\delta\big)\Big].\end{aligned}$$

But

$$P_t\big(|\overline{\sigma}_1|>\delta\big)\leqslant e^{-\delta x\sqrt{n}}E_te^{x\sqrt{n}|\overline{\sigma}_1|}$$

and so from (21) and (23)

$$(29)\qquad P_t\big(|\overline{\sigma}_1|>\delta\big)=e^{x^2/2}(1-\Phi(x))O\Big(\frac{x}{\sqrt{n}}\Big).$$

Further in (28)

$$P_t\Big(|\overline{\sigma}_1|\leqslant\frac{1}{n}\Big)=P_t(x_n\leqslant\overline{s}\leqslant y_n),$$

https://doi.org/10.1017/S0004972700019213 Published online by Cambridge University Press

where

$$x_n = -(1 + nE_t\bar{\xi}_1)/(\bar{\sigma}_t\sqrt{n}), \qquad y_n = (1 - nE_t\bar{\xi}_1)/(\bar{\sigma}_t\sqrt{n}),$$
$$\bar{s} = \sum_{i=1}^{n}(\bar{\xi}_i - E_t\bar{\xi}_1)/(\bar{\sigma}_t\sqrt{n}), \quad \text{and} \quad \bar{\sigma}_t^2 = E_t(\bar{\xi}_1 - E_t\bar{\xi}_1)^2.$$

Let $\overline{\Phi}_n(y) = P(\bar{s} \leqslant y)$. Then

$$(30) \qquad P_t\left(|\bar{\sigma}_1| \leqslant \frac{1}{n}\right) \leqslant 2\sup_y\left|\overline{\Phi}_n(y) - \Phi(y)\right| + \Phi(y_n) - \Phi(x_n)$$
$$\leqslant 2E_t|\bar{\xi}_1 - E_t\bar{\xi}_1|^3/(\bar{\sigma}_t^3\sqrt{n}) + 2(\bar{\sigma}_t\sqrt{n})^{-1}$$
$$= e^{x^2/2}(1 - \Phi(x))O\left(\frac{x}{\sqrt{n}}\right).$$

Hence, from (28)–(30) we find

$$\Delta_4 = e^{x^2/2}(1 - \Phi(x))O\left(\frac{x}{\sqrt{n}}\right).$$

Estimate Δ_5. Since $|\alpha| \leqslant 1/2$, then

$$\Delta_5 \leqslant E_t e^{2tn|\beta|} I\big(2n|\beta| \geqslant 1\big).$$

By Hölder's inequality

$$(31) \qquad \Delta_5 \leqslant \big(E_t e^{4tn|\beta|}\big)^{1/2}\Big(P_t\big(2n|\beta| \geqslant 1\big)\Big)^{1/2}.$$

Using the inequality

$$|y_1 \cdots y_m| \leqslant q_1^{-1}|y_1|^{q_1} + \cdots + q_m^{-1}|y_m|^{q_m}, \quad q_1^{-1} + \cdots + q_m^{-1} = 1, \quad y_i \in R$$

we obtain

$$(32) \qquad E_t e^{4tn|\beta|} \leqslant a_2 \sum_{c=2}^{m} \sum_{2i_2+\cdots+ci_c=c} E_t \exp\left\{a_3 x\sqrt{n}|\tilde{\sigma}_2^{i_2} \cdots \tilde{\sigma}_c^{i_c}|\right\}$$

for some positive constants a_i. Without loss of generality we can suppose that in (32) all $i_s \geqslant 1$, $s = 2, \ldots, m$. Otherwise, if $i_k = 0$ for some $k \geqslant 2$ then $\tilde{\sigma}_k^{i_k} = 1$. Hence, in (32), if $\delta < 1$,

$$E_t \exp\left\{a_3 x\sqrt{n}|\tilde{\sigma}_2^{i_2} \cdots \tilde{\sigma}_c^{i_c}|\right\} \leqslant E_t \exp\left\{a_3 x\sqrt{n}\delta^{i_2+\cdots+i_c-1}|\tilde{\sigma}_j|\right\}$$
$$\leqslant E_t \exp\left\{a_3 x\sqrt{n}|\tilde{\sigma}_j|\right\}, \quad j = 2, \ldots, c,$$

and so using (24) we get

$$E_t e^{4tn|\beta|} \leqslant c_{10}.$$

Further in (31) we have for any integer $l \geqslant 8$

$$(33)\qquad P_t\big(2n|\beta| \geqslant 1\big) \leqslant (2n)^l E_t|\beta|^l \leqslant (2m)^l \sum_{c=2}^{m} |\lambda_c(t)|^l \{n^l E_t|\beta_c|^l\}.$$

Let $r_j \geqslant 1$, $j = 1, \dots, p$, be the non-zero terms in the solution of the equation $2i_2 + \dots + ci_c = c$, $c \geqslant 2$. Hence $i_{k_1} = r_1, \dots, i_{k_p} = r_p$ for some sequence $k_1, \dots, k_p$ with $k_j \geqslant 2$. Then in (33)

$$E_t|\beta_c|^l \leqslant c_{11} \sum_{(r_1,\dots,r_p)} E_t\big(|\widetilde{\sigma}_{k_1}|^{lr_1} \cdots |\widetilde{\sigma}_{k_p}|^{lr_p}\big) \leqslant c_{11} \sum_{(r_1,\dots,r_p)} \sum_{s=1}^{p} E_t|\overline{\sigma}_{k_s}|^{plr_s}.$$

It is easy to see that

$$n^l E_t|\overline{\sigma}_{k_s}|^{plr_s} \leqslant c_{12} n^{-l/2}, \quad s = 1, \dots, p.$$

Hence in (33), $P_t\big(2n|\beta| \geqslant 1\big) \leqslant c_{13} n^{-1}$. Substituting into (31) gives

$$\Delta_5 = e^{x^2/2}\big(1 - \Phi(x)\big) O\Big(\frac{x}{\sqrt{n}}\Big).$$

Combining the bounds for the Δ_i completes the proof of Lemma 4. ∎

Theorem 1 follows from (17), (18) and Lemmas 2,3,4.

PROOF OF THEOREM 2: At first we introduce the truncated random variables

$$\overline{X}_j = X_j I\big(|X_j| \leqslant n^{(1/2)+\alpha}\big), \quad j = 1, \dots, n$$

and define the U-statistic $\overline{U}_n = U_n(\overline{X}_1, \dots, \overline{X}_n)$. The proof of Theorem 2 follows from Lemmas 5 and 6 below. ∎

LEMMA 5. *Assume condition (4) holds. Then for* $1 \leqslant x \leqslant o(n^\alpha)$

$$\Big|P\Big(U_n - \mu^m > \frac{x}{\sqrt{n}} m\sigma|\mu|^{m-1}\Big) - P\Big(\overline{U}_n - \overline{\mu}^m > \frac{x}{\sqrt{n}} m\overline{\sigma}|\overline{\mu}|^{m-1}\Big)\Big| = \exp\Big\{\frac{x^3}{\sqrt{n}} \lambda_m^{[s]}\Big(\frac{x}{\sqrt{n}}\Big)\Big\}\big(1 - \Phi(x)\big) O\Big(\frac{x}{\sqrt{n}}\Big),$$

where $\overline{\mu} = \mu(\overline{X}_1) = E\overline{X}_1$, $\overline{\sigma}^2 = \sigma^2(\overline{X}_1) = E(\overline{X}_1 - \overline{\mu})^2$.

PROOF: By analogy with the proof of Lemma 2 we can write the inequality

$$\Big|P\Big(U_n - \mu^m > \frac{x}{\sqrt{n}} m\sigma|\mu|^{m-1}\Big) - P\Big(\overline{U}_n - \overline{\mu}^m > \frac{x}{\sqrt{n}} m\overline{\sigma}|\overline{\mu}|^{m-1}\Big)\Big| \leqslant 2nP\big(|X_1| > n^{(1/2)+\alpha}\big).$$

Here

$$P\big(|X_1| > n^{(1/2)+\alpha}\big) \leqslant e^{-an^{2\alpha}} E\exp\{a|X_1|^{4\alpha/(2\alpha+1)}\}.$$

https://doi.org/10.1017/S0004972700019213 Published online by Cambridge University Press

Since $x \geqslant 1$ and $xn^{-\alpha} \to 0$ as $n \to \infty$ then

$$\begin{aligned} e^{-an^{2\alpha}} &= e^{-(a/2)n^{2\alpha}} \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right) - \frac{x^2}{2}\right\} \\ &\qquad \exp\left\{-n^{2\alpha}\left[\frac{a}{2} - \left(\frac{x}{n^\alpha}\right)^2\left(\frac{1}{2} - \frac{x}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right)\right)\right]\right\} \\ &= e^{-(a/2)n^{2\alpha}} \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right)\right\}(1-\Phi(x))O(1), \end{aligned}$$

completing the proof of Lemma 5. □

Lemma 6. *If $1 \leqslant x \leqslant o(n^\alpha)$ and condition (4) is satisfied then*

$$P\left(\overline{U}_n - \overline{\mu}^m > \frac{x}{\sqrt{n}}m\overline{\sigma}|\overline{\mu}|^{m-1}\right) = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right)\right\}(1-\Phi(x))\left(1+O\left(\frac{x}{\sqrt{n}}\right)\right).$$

Proof: We shall apply Theorem 1. In the following $\overline{\mu}$, $\overline{\sigma}$, $\overline{t}$, $\overline{\psi}, \ldots$ denote μ, σ, t, $\psi, \ldots$ respectively, where, instead of X_1, we substitute $\overline{X}_1$. For example, $\overline{\mu} = E\overline{X}_1$ and $\overline{\sigma}^2 = E(\overline{X}_1 - \overline{\mu})^2$. At first we note that for $|z| \leqslant (1/2)an^{-(1/2)+\alpha}$,

$$\begin{gathered} E\exp(z\overline{X}_1) \leqslant E\exp\{a|X_1|^{4\alpha/(2\alpha+1)}\}, \\ \left|E\exp(z\overline{X}_1)\overline{X}_1\right|^2 \leqslant EX_1^2 E\exp\{a|X_1|^{4\alpha/(2\alpha+1)}\}. \end{gathered}$$

Hence, given condition (4), the function $E\exp(z\overline{X}_1)/E\exp(z\overline{X}_1)\overline{X}_1$ in (9) is majorised by an analytic function uniformly in n for $|z| \leqslant (1/2)an^{-(1/2)+\alpha}$. Further, by analogy with (18)

$$n\overline{\psi}(\overline{t}) - n\overline{t}\overline{\mu}_m^m(\overline{t}) = -\frac{x^2}{2} + \frac{x^3}{\sqrt{n}}\overline{\lambda}_m\left(\frac{x}{\sqrt{n}}\right).$$

Therefore, under condition (4), applying Theorem 1 for $1 \leqslant x \leqslant o(n^\alpha)$

$$P\left(\overline{U}_n - \overline{\mu}^m > \frac{x}{\sqrt{n}}m\overline{\sigma}|\overline{\mu}|^{m-1}\right) = (1-\Phi(x))\exp\left\{\frac{x^3}{\sqrt{n}}\overline{\lambda}_m\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1+O\left(\frac{x}{\sqrt{n}}\right)\right),$$

with Cramér series

$$\overline{\lambda}_m(u) = \sum_{k=0}^{\infty}\overline{\lambda}_{km}u^k \tag{34}$$

which converges for $0 \leqslant u \leqslant \varepsilon n^{-(1/2)+\alpha}$ and sufficiently small $\varepsilon > 0$. The coefficients $\overline{\lambda}_{km} = \lambda_{km}(\overline{X}_1)$ depend on the moments of the truncated random variable $\overline{X}_1$.

Let

$$\overline{\lambda}_m(u) = \sum_{k=0}^{s}\overline{\lambda}_{km}u^k + \rho(u),$$

where $\rho(u) = \sum_{k=s+1}^{\infty} \overline{\lambda}_{km} u^k$. Note $s+1 \geqslant 4\alpha/(1-2\alpha)$ and in (34) the series is convergent and uniformly bounded. Hence for $x = o(n^\alpha)$

$$\frac{x^3}{\sqrt{n}}\rho\left(\frac{x}{\sqrt{n}}\right) = O(1)\frac{x^3}{\sqrt{n}}\left(\frac{x}{\sqrt{n}}\right)^{s+1} = O\left(\frac{x}{\sqrt{n}}\right).$$

Moreover

$$\sum_{k=0}^{s} \overline{\lambda}_{km}\left(\frac{x}{\sqrt{n}}\right)^k = \lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right) + O\left(\frac{1}{n}\right).$$

Thus

$$\exp\left\{\frac{x^3}{\sqrt{n}}\overline{\lambda}_m\left(\frac{x}{\sqrt{n}}\right)\right\} = \exp\left\{\frac{x^3}{\sqrt{n}}\lambda_m^{[s]}\left(\frac{x}{\sqrt{n}}\right)\right\}\left(1+O\left(\frac{x}{\sqrt{n}}\right)\right).$$

This proves Lemma 6. □

References

[1] H. Cramér, 'Sur un nouveaux théorème limit de la théorie des probabilités', *Actualités Sci. Indust.* **736** (1938), 5–23.

[2] Y.V. Borovskikh, *U-statistics in Banach Spaces* (VSP, Utrecht, The Netherlands, 1996).

[3] Y.V. Borovskikh and N.C. Weber, *On large deviations of U-statistics*, Technical Report 00-4 (2000) (School of Mathematics and Statistics, University of Sydney, Australia, 2000).

[4] I.A. Ibragimov and Y.V. Linnik, *Independent and stationary sequences of random variables* (Wolters-Noordhoff Publishing, Groningen, The Netherlands, 1971).

[5] R.W. Keener, J. Robinson and N.C. Weber, 'Tail probability approximations for U-statistics', *Statist. Probab. Lett.* **37** (1998), 59–65.

[6] Y.V. Linnik, *The collected works. Probability theory* (Nauka Publishers, Leningrad, 1981).

[7] V.V. Petrov, *Sums of independent random variables* (Springer-Verlag, Berlin, Heidelberg, New York, 1975).

[8] L. Saulis and V. Statulevicius, *Limit theorems on large deviations* (Mokslas Publishers, Vilnius, 1989).

Department of Applied Mathematics
Transport University
190031, St. Petersburg
Russia
e-mail: borovski@spb.cityline.ru

School of Mathematics and Statistics, F07
University of Sydney
New South Wales 2006
Australia
e-mail: neville@maths.usyd.edu.au