# LARGE DEVIATIONS FOR A LINEAR COMBINATION OF U-STATISTICS

Hajime YAMATO

Received April 5, 2002; revised May 20, 2002

Abstract. As an estimator of an estimable parameter, we consider a linear combination of U-statistics introduced by Toda and Yamato (2001). As a special case, this statistic includes the V-statistic and LB-statistic. In case that the kernel is not degenerate, we show some large deviations for this linear combination of U-statistics.

**1 Introduction** Let $\theta(F)$ be an estimable parameter of an unknown distribution $F$ which has a symmetric kernel $g(x_1,...,x_k)$ of degree $k(\geq 2)$ and $X_1,\dots,X_n$ be a random sample of size $n$ from the distribution $F$. We assume that the kernel $g$ is not degenerate.

As an estimator of $\theta(F)$, Toda and Yamato (2001) introduces a linear combination $Y_n$ of U-statistics as follows: Let $w(r_1,\dots,r_j;k)$ be a nonnegative and symmetric function of positive integers $r_1,\dots,r_j$ such that $j=1,\dots,k$ and $r_1+\cdots+r_j=k$, where $k$ is the degree of the kernel $g$ and fixed. We assume that at least one of $w(r_1,\dots,r_j;k)$'s is positive. For $j=1,\dots,k$, let $g_{(j)}(x_1,...,x_j)$ be the kernel given by

(1.1)
$$g_{(j)}(x_1,\dots,x_j)=\frac{1}{d(k,j)}\sum\nolimits^{+}_{r_1+\cdots+r_j=k} w(r_1,\dots,r_j;k)g(\underbrace{x_1,\dots,x_1}_{r_1},\dots,\underbrace{x_j,\dots,x_j}_{r_j}),$$

where the summation $\sum^{+}_{r_1+\cdots+r_j=k}$ is taken over all positive integers $r_1,...,r_j$ satisfying $r_1+\cdots+r_j=k$ with $j$ and $k$ fixed and $d(k,j)=\sum^{+}_{r_1+\cdots+r_j=k}w(r_1,\dots,r_j;k)$ for $j=1,2,...,k$. Let $U_n^{(j)}$ be the U-statistic associated with this kernel $g_{(j)}(x_1,\dots,x_j;k)$ for $j=1,\dots,k$. The kernel $g_{(j)}(x_1,\dots,x_j;k)$ is symmetric because of the symmetry of $w(r_1,\dots,r_j;k)$. If $d(k,j)$ is equal to zero for some $j$, then the associated $w(r_1,\dots,r_j;k)$'s are equal to zero. In this case, we let the corresponding statistic $U_n^{(j)}$ be zero. The statistics $Y_n$ is given by

$$Y_n=\frac{1}{D(n,k)}\sum_{j=1}^{k} d(k,j)\binom{n}{j}U_n^{(j)}, \tag{1.2}$$

where $D(n,k)=\sum_{j=1}^{k}d(k,j)\binom{n}{j}$. Since $w$'s are nonnegative and at least one of them is positive, $D(n,k)$ is positive. Note that $U_n^{(k)}=U_n$ for $w(1,\dots,1;k)>0$, because of $g_{(k)}=g$.

For example, let $w$ be the function given by $w(1,1,\dots,1;k)=1$ and $w(r_1,\dots,r_j;k)=0$ for positive integers $r_1,\dots,r_j$ such that $j=1,\dots,k-1$ and $r_1+\cdots+r_j=k$. Then the

2000 *Mathematics Subject Classification.* Primary 62G20; Secondary 60F10.
*Key words and phrases.* Estimable parameter, Large deviation, linear combination of U-statistics, V-statistics.

corresponding statistic $Y_n$ is equal to U-statistic $U_n$, which is given by

$$U_n = \binom{n}{k}^{-1} \sum_{1 \le j_1 < \cdots < j_k \le n} g(X_{j_1}, \ldots, X_{j_k}), \tag{1.3}$$

where $\sum_{1 \le j_1 < \cdots < j_k \le n}$ denotes the summation over all integers $j_1, \ldots, j_k$ satisfying $1 \le j_1 < \cdots < j_k \le n$.
Let $w$ be the function given by $w(r_1, \ldots, r_j; k) = 1$ for positive integers $r_1, \ldots, r_j$ such that $j = 1, \ldots, k$ and $r_1 + \cdots + r_j = k$. Then the corresponding statistic $Y_n$ is equal to the LB-statistic $B_n$ given by

$$B_n = \binom{n+k-1}{k}^{-1} \sum_{r_1 + \cdots + r_n = k} g(\underbrace{X_1, \ldots, X_1}_{r_1}, \ldots, \underbrace{X_n, \ldots, X_n}_{r_n}), \tag{1.4}$$

where $\sum_{r_1 + \cdots + r_n = k}$ denotes the summation over all non-negative integers $r_1, ..., r_n$ satisfying $r_1 + \cdots + r_n = k$.

Let $w$ be the function given by $w(r_1, \ldots, r_j; k) = k!/(r_1! \cdots r_j!)$ for positive integers $r_1, \ldots, r_j$ such that $j = 1, \ldots, k$ and $r_1 + \cdots + r_j = k$. Then the corresponding statistic $Y_n$ is equal to the V-statistic $V_n$ given by

$$V_n = \frac{1}{n^k} \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} g(X_{j_1}, \ldots, X_{j_k}). \tag{1.5}$$

(See Toda and Yamato (2001)).
Let $w$ be the function given by $w(r_1, \ldots, r_j; k) = k!/(r_1 \cdots r_j)$ for positive integers $r_1, \ldots, r_j$ such that $j = 1, \ldots, k$ and $r_1 + \cdots + r_j = k$. Then, for example, the corresponding statistic $Y_n$ for the third central moment of the distribution $F$ is given by

$$S_n = \frac{n}{n^2+1} \sum_{i=1}^{n} (X_i - \bar{X})^3,$$

where $\bar{X}$ is the sample mean of $X_1, \ldots, X_n$ (see Nomachi et al. (2002)).

In Section 2 for the U-statistic, we quote probability inequalities and tail probability, which are also known as the large deviations, from Serfling (1980), Christofides (1991), Vandemaele and Veraverbeke (1982) and Borovskikh (1996).
Our purpose is to show large deviations for the statistic $Y_n$ given by (1.3), using the results for the U-statistic stated in Section 2. These are shown in Section 3.

**2 Large Deviations for U-statistics** We shall quote some large deviations for U-statistics. Put $\sigma^2 = Var\big(g(X_1, \ldots, X_k)\big)$ and assume that $\sigma^2 > 0$. We denote $[n/k]$ by $m$, where $[x]$ is the greatest integer not greater than $x$.

**Lemma 2.1** (Serfling (1980, p.201)) *Assume that $a \le g(x_1, \ldots, x_k) \le b$, where $a$ and $b$ are constants. Then, for $t > 0$ and $n \ge k$,*

$$P(U_n - \theta \ge t) \le \exp\Big(-\frac{2mt^2}{(b-a)^2}\Big) \tag{2.1}$$

*and*

$$P(U_n - \theta \geq t) \leq \exp\left( - \frac{mt^2}{2\left(\sigma^2 + \frac{1}{3}(b-a)t\right)}\right). \tag{2.2}$$

We note that by using Markov's inequality to $P((U_n - \theta \geq t) = P(e^{s(U_n - \theta - t)} \geq 1)$, $s > 0$ we get

$$P(U_n - \theta \geq t) \leq E[e^{s(U_n - \theta - t)}] \quad \text{for} \quad s > 0. \tag{2.3}$$

The inequality (2.1) is derived as follows : The right-hand side of (2.3) is less than or equal to $e^{Q(s)}$ for $s > 0$ where $Q(s) = (b-a)^2 s^2/(8m) - st$, and the minimum value over $s > 0$ of $Q(s)$ is given by the right-hand side of (2.1) (see, for example, Serfling (1980, p. 201)). In this sense, the inequality (2.1) is equivalent to

$$\inf_{s>0} E[e^{s(U_n - \theta - t)}] \leq \left( - \frac{2mt^2}{(b-a)^2}\right). \tag{2.4}$$

Similarly, the inequality (2.2) is equivalent to

$$\inf_{s>0} E[e^{s(U_n - \theta - t)}] \leq \exp\left( - \frac{mt^2}{2\left(\sigma^2 + \frac{1}{3}(b-a)t\right)}\right). \tag{2.5}$$

**Lemma 2.2** (Christofides (1991)) (a) *Assume that there exists $M > 0$ such that $E(g(X_1, \dots, X_k) - \theta)^r \leq r!\sigma^2 M^{r-2}/2$ for $r = 2, 3, \dots$. Then for $t > 0$,*

$$P(U_n - \theta \geq t) \leq \exp\left( - \frac{m}{2M^2}\left(\sqrt{2tM + \sigma^2} - \sigma\right)^2\right). \tag{2.6}$$

(b) *Assume that $a \leq g(x_1, \dots, x_k) \leq b$, where $a$ and $b$ are constants. Then for $t > 0$,*

$$P(U_n - \theta \geq t) \leq \exp\left( - \frac{9m}{2(b-a)^2}\left(\sqrt{\frac{2}{3}t(b-a) + \sigma^2} - \sigma\right)^2\right). \tag{2.7}$$

For the kernel $g(x_1, \dots, x_k)$, we put

$$\psi_l(x_1, \dots, x_l) = E\left(g(X_1, \dots, X_k) \mid X_1 = x_1, \dots, X_l = x_l\right), \quad l = 1, ..., k.$$

For $l = 2, 3, ..., k$, we put

$$g^{(1)}(x_1) = \psi_1(x_1) - \theta,$$

$$g^{(l)}(x_1, \dots, x_l) = \psi_l(x_1, \dots, x_l) - \sum_{i=1}^{l-1} \sum_{1 \leq j_1 < \cdots < j_i \leq l} g^{(i)}(x_{j_1}, \dots, x_{j_i}) - \theta.$$

We suppose $\sigma_1^2 = Var(\psi_1(X_1)) > 0$. Let $\Phi(x)$ be the standard normal distribution function. It satisfies the following relation.

$$1 - \Phi(x \pm (\ln n)^{-2}) = \left(1 - \Phi(x)\right)\left(1 + o\left(\frac{1}{\ln n}\right)\right) \tag{2.8}$$

uniformly in the range $-A \leq x \leq c\sqrt{\ln n}$, where $A \geq 0$ and $c > 0$ (see, Vandemaele and Veraverberke (1982)).

**Lemma 2.3** (Vandemaele and Veraverberke (1982), Lemma 1) (a) *If* $E \mid g(X_1, \ldots, X_k) \mid^p < \infty$ *for some* $p > 2 + c^2$ $(c > 0)$, *then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\Big(1 + o\Big(\frac{1}{\ln n}\Big)\Big) \tag{2.9}$$

*uniformly in the range* $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$.

(b) *If for all* $p = 1, 2, \cdots$ $E \mid g(X_1, \ldots, X_k) \mid^p < K^p p^{\gamma p}$ (*where* $K$ *and* $\gamma \ge 0$ *are constants not depending on* $p$), *then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\big(1 + o(1)\big) \tag{2.10}$$

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

The proposition (a) can be strengthened as follows:

**Lemma 2.4** (Borovskikh(1996)) *Suppose that*

$$E \mid g^{(1)}(X_1) \mid^p < \infty, \quad p > 2 + c^2$$

*and*

$$E \mid g^{(l)}(X_1, \ldots, X_l) \mid^{c_l + c^2} < \infty, \quad l = 2, \ldots, k,$$

*where* $c_l = 2l/(2l - 1)$ *and some constant* $c > 0$. *Then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\Big(1 + o\Big(\frac{1}{\ln n}\Big)\Big)$$

*uniformly in the range* $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$.

As a corollary of this Lemma, the following is obtained.

**Lemma 2.5** (Borovskikh(1996)) *Under the same conditions as Lemma* 2.4,

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > c\sqrt{\ln n}\Big) = \frac{1}{\sqrt{2\pi c^2 \ln n}} n^{-\frac{c^2}{2}}\Big(1 + O\Big(\frac{1}{\ln n}\Big)\Big).$$

## 3 Large Deviations for Y-statistics

We put for $j = 1, \ldots, k$

$$\theta_j = E\big(g_{(j)}(X_1, \ldots, X_j)\big)$$

and

$$\sigma_{(j)}^2 = Var\big(g_{(j)}(X_1, \ldots, X_j)\big).$$

We note that $\theta_k = \theta$ and $\sigma_{(k)}^2 = \sigma^2$. We put $\tau^2 = \max\{\sigma_{(1)}^2, \ldots, \sigma_{(k)}^2\}$. We show some probability inequalities for a linear combinations of U-statistics. The first two probability inequalities correspond to Lemma 1.

**Theorem 3.1** *Assume that $a \le g(x_1, \dots, x_k) \le b$, where $a$ and $b$ are constants. Then, for $t > 0$ and $n \ge k$,*

$$P(Y_n - EY_n \ge t) \le \exp\Big(-\frac{2mt^2}{(b-a)^2}\Big) \tag{3.1}$$

*and*

$$P(Y_n - EY_n \ge t) \le \exp\bigg(-\frac{mt^2}{2\big(\tau^2 + \frac{1}{3}(b-a)t\big)}\bigg). \tag{3.2}$$

**Proof** By the similar method to (2.3), for $s > 0$ we have

$$P(Y_n - EY_n \ge t) \le E(e^{s(Y_n - EY_n - t)}) = E\exp\Big(\sum_{j=1}^{k} \frac{d(k,j)}{D(n,k)} \binom{n}{j} s(U_n^{(j)} - \theta_j - t)\Big).$$

Because we may regard $d(k,j)/D(n,k))\binom{n}{j}$ as a probability function on $j = 1, \cdots, k$, we can use Jensen's inequality to the exponent of the right-hand side and get the inequality

$$P(Y_n - EY_n \ge t) \le \sum_{j=1}^{k} \frac{d(k,j)}{D(n,k)} \binom{n}{j} E(e^{s(U_n^{(j)} - \theta_j - t)}), \quad s > 0. \tag{3.3}$$

Since $a \le g_{(j)} \le b$ and $E(U_n^{(j)}) = \theta_j$ $(j = 1, \dots, k)$, applying (2.4) to the expectation of the right-hand side and using $m = [n/k] \le [n/j]$ $(j = 1, \dots, k)$, we get

$$\inf_{s>0} E(e^{s(U_n^{(j)} - \theta_j - t)}) \le \exp\Big(-\frac{2\big[\frac{n}{j}\big]t^2}{(b-a)^2}\Big) \le \exp\Big(-\frac{2mt^2}{(b-a)^2}\Big).$$

Using this result to the right-hand side of (3.3), we get (3.1) by the relation $D(n,k) = \sum_{j=1}^{n} d(k,j)\binom{n}{j}$.

Applying (2.5) to the expectation of the right-hand side of (3.3) and using $m \le [n/j]$ and $\tau^2 \ge \sigma_j^2$ $(j = 1, \dots, k)$, we get

$$\inf_{s>0} E(e^{s(U_n^{(j)} - \theta_j - t)}) \le \exp\bigg(-\frac{\big[\frac{n}{j}\big]t^2}{2\big(\sigma_{(j)}^2 + \frac{1}{3}(b-a)t\big)}\bigg) \le \exp\Big(-\frac{mt^2}{2\big(\tau^2 + \frac{1}{3}(b-a)t\big)}\Big).$$

Using this result to the right-hand side of (3.3), we get (3.2). □

**Theorem 3.2** (a) *Assume that there exists $M > 0$ such that $E(g_{(j)}(X_1, \dots, X_j) - \theta_j)^r \le r!\sigma^2 M^{r-2}/2$ for $j = 1, \dots, k$ and $r = 2, 3, \dots$. Then for $t > 0$,*

$$P(Y_n - EY_n \ge t) \le \exp\Big(-\frac{m}{2M^2}\big(\sqrt{2tM + \tau^2} - \tau\big)^2\Big). \tag{3.4}$$

(b) *Assume that $a \le g(x_1, \dots, x_k) \le b$, where $a$ and $b$ are constants. Then for $t > 0$,*

$$P(Y_n - EY_n \ge t) \le \exp\bigg(-\frac{9m}{2(b-a)^2}\Big(\sqrt{\frac{2}{3}t(b-a) + \tau^2} - \tau\Big)^2\bigg). \tag{3.5}$$

**Proof** A U-statistic can be represented as an average of averages of i.i.d. random variables (see, for example, Serfling (1980, p.180) and Borovskikh (1996, p.14)). For the random variable having mean zero, variance $\sigma^2 > 0$ and satisfying $E(X^r) \le r!\sigma^2 M^{r-2}/2$ for $r = 2, 3, \ldots$, its moment generating function satisfies $E(e^{sX}) \le \exp\{\sigma^2 s^2/[2(1 - sM)]\}$ for $0 < s < 1/M$. Using these two facts, it can be shown that for $j = 1, \ldots, k$,

$$Ee^{s(U_n^{(j)} - \theta_j - t)} \le \exp\Big( - st + \frac{\sigma_j^2 s^2}{2([\frac{n}{j}] - sM)}\Big), \quad 0 < s < \frac{m}{M}\Big( \le \frac{[\frac{n}{j}]}{M}\Big),$$

(Christofides (1991, p.258–259)). Because of $m \le [n/j]$ and $\tau^2 \ge \sigma_j$ $(j = 1, \ldots, k)$ , we have

$$Ee^{s(U_n^{(j)} - \theta_j - t)} \le \exp\Big( - st + \frac{\tau^2 s^2}{2(m - sM)}\Big), \quad 0 < s < \frac{m}{M}.$$

Thus by (3.3) we have

$$P(Y_n - EY_n \ge t) \le \exp\Big( - st + \frac{\tau^2 s^2}{2(m - sM)}\Big), \quad 0 < s < \frac{m}{M}.$$

Putting $y = m - sM(> 0)$, the exponent of the right-hand side is equal to

$$\frac{1}{2M^2}\Big(\frac{\tau m}{\sqrt{y}} - \sqrt{(2tM + \tau^2)y}\Big)^2 + \frac{m}{M^2}\Big(\sqrt{2tM + \tau^2} \cdot \tau - (tM + \tau^2)\Big).$$

The minimum value over $y > 0$ of this function is given by the second term which is equal to the exponent of the right-hand side of (3.4).

Under the condition of (b), the condition of (a) is satisfied with $M = (b - a)/3$ (Christofides (1991, p.259)). Therefore the inequality (3.5) is obtained from (3.4) by replacing $M$ with $(b - a)/3$. □

**Theorem 3.3** (a) *If* $E \mid g(X_1, \ldots, X_k) \mid^p < \infty$ *and* $E \mid g(X_{j_1}, \ldots, X_{j_k}) \mid^{p-2} < \infty$, $1 \le j_1 \le \cdots \le j_k \le k$, *for some* $p > 2 + c^2$ $(c > 0)$, *then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(Y_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\Big(1 + o\Big(\frac{1}{\ln n}\Big)\Big) \tag{3.6}$$

*uniformly in the range* $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$.

(b) *If for all* $p = 1, 2, \cdots$ $E \mid g(X_{j_1}, \ldots, X_{j_k}) \mid^p < K^p p^{\gamma p}$, $1 \le j_1 \le \cdots \le j_k \le k$, *(where* $K$ *and* $\gamma \ge 0$ *are constants not depending on* $p$*), then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(Y_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\big(1 + o(1)\big) \tag{3.7}$$

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

If $d(k, k) = w(1, \ldots, 1; k) > 0$, then there exists a constant $\beta(\ge 0)$ such that

$$\frac{d(k,k)}{D(n,k)}\binom{n}{k} = 1 - \frac{\beta}{n} + O(\frac{1}{n^2}) \tag{3.8}$$

and

$$\sum_{j=1}^{k-1} \frac{d(k,j)}{D(n,k)} \binom{n}{j} = \frac{\beta}{n} + O\Big(\frac{1}{n^2}\Big). \tag{3.9}$$

For the U-statistic $U_n$, $\beta = 0$. In the following proof of Theorem 3.3 (b), we assume that

$$\beta > 0,$$

because the corresponding large deviations for the U-statistic are given in Section 2. For the V-statistic $V_n$ and the S-statistic $S_n$, $\beta = k(k-1)/2$. For the LB-statistic $B_n$, $\beta = k(k-1)$.

As stated in Toda and Yamato (2001, p.229), we can write

$$Y_n = U_n + R_n$$

and $R_n$ satisfies the following:If $E \mid g(X_{j_1}, \ldots, X_{j_k}) \mid^r < \infty$ for $r > 0$ and any integers $j_1, \ldots, j_k$ $(1 \leq j_1 \leq \cdots \leq j_k \leq k)$, then

$$E \mid R_n \mid^r \leq \frac{C_1}{n^r}, \tag{3.10}$$

where $C_1$ is a generic constant (this relation holds even if $r$ is not integer by the same reason as its proof of Toda and Yamato (2001)).

**Proof of Theorem 3.3** Since $Y_n - \theta = U_n - \theta + R_n$, for any $\varepsilon > 0$

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > x + \varepsilon\Big) - P\Big(\frac{\sqrt{n}}{k\sigma_1} \mid R_n \mid > \varepsilon\Big)$$

$$\leq P\Big(\frac{\sqrt{n}}{k\sigma_1}(Y_n - \theta) > x\Big) \tag{3.11}$$

$$\leq P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n - \theta) > x - \varepsilon\Big) + P\Big(\frac{\sqrt{n}}{k\sigma_1} \mid R_n \mid > \varepsilon\Big).$$

At first we shall show (3.6). Using Markov's inequality and (3.10), for $\varepsilon = (\ln n)^{-2}$ we have

$$P\Big(\frac{\sqrt{n}}{k\sigma_1} \mid R_n \mid > \varepsilon\Big) \leq C_2 \frac{(\ln n)^{2(p-2)}}{n^{(p-2)/2}}, \tag{3.12}$$

where $C_2(> 0)$ is a generic constant. For a large $x > 0$, $1 - \Phi(x) \approx (\sqrt{2\pi}x)^{-1} e^{-x^2/2}$ (see, for example, Johnson et al. (1994)). Hence for $-A \leq x \leq c\sqrt{\ln n}$, we have $1/\big(1 - \Phi(x)\big) \leq O((\ln n)^{1/2} n^{c^2/2})$. By this relation, (3.12) and $p - c^2 > 2$, we have

$$\frac{P\big(\frac{\sqrt{n}}{k\sigma_1} \mid R_n \mid > \varepsilon\big)}{1 - \Phi(x)} = O\Big(\frac{(\ln n)^{2(p-2)+1/2}}{n^{(p-2-c^2)/2}}\Big).$$

Thus we have

$$\frac{P\big(\frac{\sqrt{n}}{k\sigma_1} \mid R_n \mid > \varepsilon\big)}{1 - \Phi(x)} = o\big((\ln n)^{-1}\big). \tag{3.13}$$

By (2.9), we have

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n-\theta)>x\pm\varepsilon\Big)=\big(1-\Phi(x\pm\varepsilon)\big)\Big(1+o\big((\ln n)^{-1}\big)\Big).$$

Taking $\varepsilon=(\ln n)^{-2}$ and using (2.8), we get

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n-\theta)>x\pm\varepsilon\Big)=\big(1-\Phi(x)\big)\Big(1+o\big((\ln n)^{-1}\big)\Big). \tag{3.14}$$

Applying (3.13) and (3.14) to (3.11), we get (3.6).

Now we shall prove (3.7). By the condition on moments of $g$, we have

$$\big(E\mid U_n^{(j)}\mid^p\big)^{1/p}\leq Kp^{\gamma},\quad j=1,\dots,k.$$

Therefore by (3.8), (3.9) and Minkowski's inequality, we have

$$\{E\mid R_n\mid^p\}^{1/p}\leq Kp^{\gamma}\Big(\frac{2\beta}{n}+O\big(\frac{1}{n^2}\big)\Big).$$

Therefore, for $p=1,2,\dots$ we have

$$E\mid\sqrt{n}R_n\mid^p\leq(2\beta K)^p n^{-p/2}p^{p\gamma+1}\Big(1+O\big(\frac{1}{n}\big)\Big).$$

By Markov's inequality, for $\varepsilon=n^{-\alpha}$ and $p=cn^{(1-2\alpha)/(2+2\gamma)}$ $(\alpha=1/\{2(3+2\gamma)\})$, we have

$$P(\mid\sqrt{n}R_n\mid\geq\varepsilon)\leq O\Big(\big(2\beta Kn^{\alpha-\frac{1}{2}}p^{\gamma}\big)^p\cdot p\Big)=O\Big(e^{p[\ln(2\beta K)-\ln p]+\ln p}\Big) \tag{3.15}$$

Let $p_n$ be a positive sequence such that $p_n\to 0$ and $p_nn^{\alpha}\to\infty$. Then by the same reason stated in the first part, $1/\big(1-\Phi(p_nn^{\alpha})\big)\approx\sqrt{2\pi}p_nn^{\alpha}e^{p_n^2n^{2\alpha}/2}$. Since $\varepsilon=n^{-\alpha}$, $p=cn^{(1-2\alpha)/(2+2\gamma)}$ and $\alpha=1/\{2(3+2\gamma)\}$, we have

$$\frac{P\big(\frac{\sqrt{n}}{k\sigma_1}\mid R_n\mid>\varepsilon\big)}{1-\Phi(p_nn^{\alpha})}=O\Bigg(p_n\exp\Big(\ln c+n^{2\alpha}\big(\frac{p_n^2}{2}+C_3-(2\alpha c-\frac{3\alpha}{n^{2\alpha}})\ln n\big)\Big)\Bigg),$$

whose exponent diverges to $-\infty$ as $n\to\infty$ because of $\alpha c>0$, where $C_3$ is a generic constant depending on $c$, $\beta$ and $K$. Hence the left-hand side converges to 0 as $n\to\infty$. Thus we get

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}\mid R_n\mid>n^{-\alpha}\Big)=(1-\Phi(x))\cdot o(1) \tag{3.16}$$

uniformly in $-A\leq x\leq o(n^{\alpha})$. By (2.10), for a sufficiently small $\varepsilon>0$ we have

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(U_n-\theta)>x\pm\varepsilon\Big)=(1-\Phi(x\pm\varepsilon))(1+o(1)).$$

Using the approximation $1-\Phi(x)\approx(\sqrt{2\pi}x)^{-1}e^{-x^2/2}$ for a large $x>0$, we have

$$1-\Phi(x\pm n^{-\alpha})=(1-\Phi(x))(1+o(1)) \tag{3.17}$$

uniformly in $-A\leq x\leq o(n^{\alpha})$, which is shown by the method similar to (3.16). Therefore, applying (3.16) and (3.17) to (3.11), we can get (3.7). □

Noting that the conditions on $R_n$ in the following Corollary are as same as (a) of Theorem 3.3, the proposition (a) can be strengthened by Lemma 2.4.

**Corollary 3.4** *Suppose that*

$$E \mid g(X_{j_1}, \dots , X_{j_k}) \mid^{p-2} < \infty, \quad (1 \le j_1 \le \cdots \le j_k \le k), \quad p > 2 + c^2,$$

$$E \mid g^{(1)}(X_1) \mid^{p} < \infty, \quad p > 2 + c^2,$$

*and*

$$E \mid g^{(l)}(X_1, \dots , X_l) \mid^{c_l + c^2} < \infty, \quad l = 2, \dots , k,$$

*where* $c_l = 2l/(2l-1)$ *and some constant* $c > 0$. *Then*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(Y_n - \theta) > x\Big) = \big(1 - \Phi(x)\big)\Big(1 + o\big(\frac{1}{\ln n}\big)\Big)$$

*uniformly in the range* $-A \le x \le c\sqrt{\ln n}$ $(A \ge 0)$.

As a corollary of this result, the following is obtained because of $1 - \Phi(c\sqrt{\ln n}) = (2\pi c^2 \ln n)^{-1/2} n^{-c^2/2}(1 + O((\ln n)^{-1}))$.

**Corollary 3.5** *Under the same conditions as Corollary* 3.4*,*

$$P\Big(\frac{\sqrt{n}}{k\sigma_1}(Y_n - \theta) > c\sqrt{\ln n}\Big) = \frac{1}{\sqrt{2\pi c^2 \ln n}} n^{-\frac{c^2}{2}}\Big(1 + O\big(\frac{1}{\ln n}\big)\Big).$$

**Acknowledgments**
The author would like to express his hearty thanks to the referee for his careful reading of the manuscript and the useful comments.

## References

[1] Borovskikh, Yu.V. (1996), *U-statistics in Banach spaces*, VSP, Utrecht.

[2] Christofides, T. S. (1991), Probability inequalities with exponential bounds for U-statistics, *Statist. Prob. Letters*, **12**, 257–261.

[3] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994), *Continuous univariate distributions*, Vol. 1, 2nd ed., John Wiley, New York.

[4] Nomachi, T. , Kondo, M. and Yamato, H. (2002), Higher order efficiency of linear combinations of U-statistics as estimators of estimable parameters, *Scientiae Mathematicae Japonicae*, (online, **6**, 95–106.)

[5] Serfling, R. J. (1980), *Approximation theorems of mathematical statistics*, John Wiley and Sons, New York.

[6] Toda, K. and Yamato, H. (2001). Berry-Esseen bounds for some statistics including LB-statistic and V-statistic. *J. Japan Statist. Soc.*, **31**, No. 2, 225–237.

[7] Vandemaele, M. and Veraverbeke, N. (1982), Cramér type large deviations for linear combinations of order statistics, *Ann. Prob.*, **10**, 423–434.

*Author*:
Department of Mathematics and Computer Science, Kagoshima University, Kagoshima 890–0065, Japan