

CRAMÉR TYPE MODERATE DEVIATIONS FOR STUDENTIZED U-STATISTICS ^{*, **, ***}

TZE LENG LAI¹, QI-MAN SHAO² AND QIYING WANG³

Abstract. Let T_n be a Studentized U-statistic. It is proved that a Cramér type moderate deviation $P(T_n \geq x)/(1 - \Phi(x)) \rightarrow 1$ holds uniformly in $x \in [0, o(n^{1/6})]$ when the kernel satisfies some regular conditions.

Mathematics Subject Classification. 60F10, 60F05.

Received September 20, 2008. Revised June 2, 2009.

1. INTRODUCTION AND MAIN RESULTS

Let X, X_1, X_2, \dots, X_n be a sequence of independent and identically distributed (i.i.d.) random variables, and let $h(x_1, x_2)$ be a real-valued symmetric Borel measurable function. Assume that $\theta = Eh(X_1, X_2)$. An unbiased estimator of θ is the Hoeffding [7] U-statistic

$$U_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} h(X_i, X_j). \quad (1.1)$$

The U-statistic elegantly and usefully generalizes the notion of a sample mean. Typical examples include

- (i) sample mean: $h(x_1, x_2) = \frac{1}{2}(x_1 + x_2)$;
- (ii) sample variance: $h(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$;
- (iii) Gini's mean difference: $h(x_1, x_2) = |x_1 - x_2|$;
- (iv) one-sample Wilcoxon's statistic: $h(x_1, x_2) = 1(x_1 + x_2 \leq 0)$.

The non-degenerate U-statistic shares many limiting properties with the sample mean. For example, if $Eh^2(X_1, X_2) < \infty$ and $\sigma_1^2 = \text{Var}(g(X_1)) > 0$, where

$$g(x) = Eh(x, X), \quad (1.2)$$

Keywords and phrases. Moderate deviation, u-statistic, studentized.

* *Research partially supported by NSF-MCS-0305749.*

** *Research partially supported by Hong Kong RGC - CERG 602206 and 602608.*

*** *Research partially supported by an Australian Research Council (ARC) discovery project.*

¹ Department of Statistics, Stanford University Stanford, CA CA 94305-4065, USA; lait@stanford.edu

² Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon Hong Kong, P.R. China; maqmshao@ust.hk

³ School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia; qiying@maths.usyd.edu.au

then the central limit theorem holds, *i.e.*,

$$\sup_x |P\left(\frac{\sqrt{n}}{2\sigma_1}(U_n - \theta) \leq x\right) - \Phi(x)| \rightarrow 0, \tag{1.3}$$

where $\Phi(x)$ is the standard normal distribution function. A systematic presentation of the theory of U-statistics was given in [10]. We refer the study on uniform Berry-Esseen bound for U-statistics to Alberink and Bentkus [1, 2], Wang and Weber [17] and the references there. One can also refer to Borovskich and Weber [4,5] for large deviations. However, since σ_1 is typically unknown, it is necessary to estimate σ_1 first and then substitute it in (1.3). Therefore, what used in practice is actually the following studentized U-statistic (see, *e.g.*, Arvesen [3])

$$T_n = \sqrt{n}(U_n - \theta)/R_n, \tag{1.4}$$

where

$$R_n^2 = \frac{4(n-1)}{(n-2)^2} \sum_{i=1}^n (q_i - U_n)^2 \quad \text{with} \quad q_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n h(X_i, X_j). \tag{1.5}$$

One can refer to Wang, Jing and Zhao [16] on uniform Berry-Esseen bound for studentized U-statistics. Also see Callaert and Veraverbeke [6] and Zhao [18]. We also refer to Vandemaële and Veraverbeke [14] and Wang [15] for the Cramér type moderate deviation.

A special case of the studentized U-statistics is the Student t-statistic with $h(x_1, x_2) = (x_1 + x_2)/2$. Although the t-statistic has a close relationship with the classical standardized partial sum, it has been found that the t-statistic enjoys much better limiting properties. For example, Shao [11] proves that the large deviation always holds for t-statistic without any moment assumption and Shao [12] further shows that a Cramér type moderate deviation is valid under only a finite third moment. Jing, *et al.* [8] proved a Cramér type moderate deviation result (for independent random variables) under a Lindeberg type condition. Jing, *et al.* [9] obtained the saddlepoint approximation without any moment condition. Thus, it is natural to ask whether similar results hold for the studentized U-statistics. The main objective of this paper is to show that the studentized U-statistics share similar properties like the student t-statistic does when the kernel satisfies

$$h^2(x_1, x_2) - \theta \leq c_0[\sigma_1^2 + g^2(x_1) - \theta + g^2(x_2) - \theta] \tag{1.6}$$

for some $c_0 > 0$. This condition is satisfied by the typical examples of U-statistics listed at the beginning of this section.

Theorem 1.1. *Assume $0 < \sigma_1^2 < \infty$ and that (1.6) holds for some $c_0 > 0$. Then, for any x_n with $x_n \rightarrow \infty$ and $x_n = o(n^{1/2})$,*

$$\ln P(T_n \geq x_n) \sim -x_n^2/2. \tag{1.7}$$

If in addition $E|g(X_1)|^3 < \infty$, then

$$P(T_n \geq x) = (1 - \Phi(x))[1 + o(1)] \tag{1.8}$$

holds uniformly in $x \in [0, o(n^{1/6})]$.

Assume $\theta = 0$. Write $S_n = \sum_{j=1}^n g(X_j)$ and $V_n^2 = \sum_{j=1}^n g^2(X_j)$. It is known (see Shao [11]) that

$$\ln P(S_n/V_n \geq x_n) \sim -x_n^2/2 \tag{1.9}$$

for any x_n with $x_n \rightarrow \infty$ and $x_n = o(n^{1/2})$. It is also known (see Jing *et al.* [8]) that if $E|g(X_1)|^3 < \infty$, then

$$P(S_n/V_n \geq x) = (1 - \Phi(x))\left[1 + O(1)(1 + x^3)n^{-1/2}\right] \tag{1.10}$$

for $0 \leq x \leq O(n^{1/6})$. The following theorem shows that the studentized U -statistic T_n can be approximated by the self-normalized sum S_n/V_n under the condition (1.6). As a result, (1.7) and (1.8) follow from (1.9) and (1.10), together with (1.11) below, respectively.

Theorem 1.2. *Assume that $\theta = 0$, $0 < \sigma_1^2 = Eg^2(X_1) < \infty$ and the kernel $h(x_1, x_2)$ satisfies the condition (1.6). Then there exists a constant $\eta > 0$ depending only on σ_1^2 and c_0 such that, for all $4/(n-1) \leq \epsilon_n < 1$, $0 \leq x \leq \sqrt{n}/3$ and n sufficiently large,*

$$\begin{aligned} P[S_n/V_n \geq (1 + \epsilon_n)x] - 5\sqrt{2}(n+2)e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}} &\leq P(T_n \geq x) \\ &\leq P[S_n/V_n \geq (1 - \epsilon_n)x] + 5\sqrt{2}(n+2)e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}}. \end{aligned} \quad (1.11)$$

This paper is organized as follows. In the next section we will prove the main theorems. A technical proposition will be postponed to Section 3.

2. PROOFS OF THEOREMS

We start with some preliminaries. Write

$$T_n^* = \sqrt{n}U_n/R_n^*, \quad (2.1)$$

where $R_n^{*2} = \frac{4(n-1)}{(n-2)^2} \sum_{i=1}^n q_i^2$. Observe that

$$\sum_{i=1}^n (q_i - U_n)^2 = \sum_{i=1}^n q_i^2 - 2U_n \sum_{i=1}^n q_i + nU_n^2 = \sum_{i=1}^n q_i^2 - nU_n^2.$$

We have

$$T_n = \frac{T_n^*}{\left(1 - \frac{4(n-1)}{(n-2)^2} T_n^{*2}\right)^{1/2}} \quad (2.2)$$

and

$$\{T_n \geq x\} = \left\{T_n^* \geq \frac{x}{\left[1 + 4x^2(n-1)/(n-2)^2\right]^{1/2}}\right\}. \quad (2.3)$$

We now establish a relationship between T_n^* and S_n/V_n . To do this, further let $\psi(x_1, x_2) = h(x_1, x_2) - g(x_1) - g(x_2)$,

$$\Delta_n = \frac{1}{n-1} \sum_{1 \leq i < j \leq n} \psi(X_i, X_j), \quad W_n^{(i)} = \sum_{\substack{j=1 \\ j \neq i}}^n \psi(X_i, X_j), \quad \Lambda_n^2 = \sum_{i=1}^n (W_n^{(i)})^2.$$

It is easy to see that

$$nU_n/2 = S_n + \Delta_n. \quad (2.4)$$

Also observe that $\sum_{\substack{j=1 \\ j \neq i}}^n h(X_i, X_j) = (n-2)g(X_i) + S_n + W_n^{(i)}$ and

$$\begin{aligned} \frac{(n-1)(n-2)^2}{4} R_n^{*2} &= \sum_{i=1}^n \left(\sum_{\substack{j=1 \\ j \neq i}}^n h(X_i, X_j) \right)^2 \\ &= (n-2)^2 V_n^2 + \Lambda_n^2 + (3n-4)S_n^2 \\ &\quad + 2(n-2) \sum_{i=1}^n g(X_i)W_n^{(i)} + 2S_n \sum_{i=1}^n W_n^{(i)}. \end{aligned}$$

Therefore, using $|\sum_{i=1}^n g(X_i)W_n^{(i)}| \leq V_n \Lambda_n$,

$$|S_n \sum_{i=1}^n W_n^{(i)}| \leq |S_n| \sqrt{n} \Lambda_n \quad \text{and} \quad \Lambda_n^2 \leq n \max_{1 \leq i \leq n} |W_n^{(i)}|^2$$

by the Hölder’s inequality, we have

$$R_n^{*2} = \frac{4}{n-1} V_n^2 (1 + \delta_n), \tag{2.5}$$

where

$$\begin{aligned} |\delta_n| &\leq \frac{1}{(n-2)^2} \left[\frac{\Lambda_n^2}{V_n^2} + \frac{3n S_n^2}{V_n^2} + \frac{2n \Lambda_n}{V_n} + 2\sqrt{n} \frac{|S_n| \Lambda_n}{V_n^2} \right] \\ &\leq \frac{1}{(n-2)^2} \left(\frac{\Lambda_n^2}{V_n^2} + \frac{4n \Lambda_n}{V_n} + \frac{3n S_n^2}{V_n^2} \right). \end{aligned} \tag{2.6}$$

By (2.4)–(2.5) and (2.1)

$$T_n^* = \frac{S_n + \Delta_n}{d_n V_n (1 + \delta_n)^{1/2}}, \tag{2.7}$$

where $d_n = \sqrt{n/(n-1)}$.

Next proposition shows that Δ_n and δ_n are negligible.

Proposition 2.1. *There exist constants $\delta_0 > 0$ and $\delta_1 > 0$, depending only on σ_1^2 and c_0 , such that for all $y > 0$*

$$P(|\delta_n| \geq y) \leq 4\sqrt{2}(n+2) \exp(-\delta_0 \min\{1, y\} n) \tag{2.8}$$

and

$$P(|\Delta_n| \geq y V_n) \leq \sqrt{2}(n+2) \exp(-\delta_1 \min\{n, y \sqrt{n}\}). \tag{2.9}$$

The proof of Proposition 2.1 is postponed to Section 3. We mention that the proof is based on exponential inequalities for self-normalized sums of martingale difference sequence (Lems. 3.1 and 3.4) and for self-normalized sums of independent random variables (Lems. 3.2 and 3.3). These inequalities are interesting in their own rights.

We are now ready to prove our main results.

Proof of Theorem 1.2. Since $x^2 \leq n/9$ and $0 \leq \epsilon_n < 1$, it is easy to show that, for $0 \leq x \leq \sqrt{n}/3$,

$$\tau_n \equiv \left(1 - \frac{\epsilon_n}{4}\right)^{1/2} \tau'_n \geq 1 - \epsilon_n/2,$$

whenever n is sufficiently large, where $\tau'_n = \sqrt{\frac{n}{n-1}} \left[1 + \frac{4x^2(n-1)}{(n-2)^2}\right]^{-1/2}$. Hence it follows from (2.3), (2.7) and Proposition 2.1 that

$$\begin{aligned} P(T_n \geq x) &\leq P[S_n/V_n \geq (1 - \epsilon_n)x] + P\left\{|\Delta_n|/V_n \geq x(\epsilon_n - 1) + x \tau'_n (1 + \delta_n)^{1/2}\right\} \\ &\leq P[S_n/V_n \geq (1 - \epsilon_n)x] + P\left\{|\Delta_n|/V_n \geq x(\epsilon_n - 1) + x \tau_n\right\} + P\left\{|\delta_n| \geq \epsilon_n/4\right\} \\ &\leq P[S_n/V_n \geq (1 - \epsilon_n)x] + P\left\{|\Delta_n|/V_n \geq x \epsilon_n/2\right\} + P\left\{|\delta_n| \geq \epsilon_n/4\right\} \\ &\leq P[S_n/V_n \geq (1 - \epsilon_n)x] + 5\sqrt{2}(n+2) e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}}, \end{aligned}$$

where $\eta > 0$ is a constant depending only on σ_1^2 and c_0 . This proves the upper bound of (1.11). Similarly, for the lower bound of (1.11)

$$\begin{aligned}
P(S_n/V_n \geq (1 + \epsilon_n)x) &\leq P(T_n \geq x) + P\left\{|\Delta_n|/V_n \geq x(1 + \epsilon_n) - x\tau'_n(1 + \delta_n)^{1/2}\right\} \\
&\leq P(T_n \geq x) + P\left\{|\Delta_n|/V_n \geq x(1 + \epsilon_n) - x\tau'_n(1 + \epsilon_n/4)^{1/2}\right\} + P(|\delta_n| \geq \epsilon_n/4) \\
&\leq P(T_n \geq x) + P\left\{|\Delta_n|/V_n \geq x(1 + \epsilon_n) - x(1 + 1/(n-1))(1 + \epsilon_n/8)\right\} \\
&+ P(|\delta_n| \geq \epsilon_n/4) \leq P(T_n \geq x) + P(|\Delta_n|/V_n \geq x\epsilon_n/2) + P(|\delta_n| \geq \epsilon_n/4) \\
&\leq P(T_n \geq x) + 5\sqrt{2}(n+2)e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}}.
\end{aligned}$$

The proof of Theorem 1.2 is now complete.

Proof of Theorem 1.1. Theorem 1.1 follows from (1.9)–(1.10) and Theorem 1.2 by a suitable choice of ϵ_n , together with some routine calculations. Indeed, by the central limit theorem for T_n (see, e.g., Thm. 3.1, [16]), the result (1.8) is obvious when $0 \leq x \leq 1$. In order to prove (1.8) for $x \in [1, o(n^{1/6})]$, we choose $\epsilon_n = \max\{\epsilon'_n x/\sqrt{n}, n^{-1/8}\}$, where ϵ'_n is a sequence of constants satisfying $\epsilon'_n \rightarrow \infty$ and $\epsilon'_n x^3/\sqrt{n} \rightarrow 0$ for $x \in [1, o(n^{1/6})]$. It is readily to see that

$$\min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\} \geq \sqrt{n}\epsilon_n x \geq \max\{\epsilon'_n x^2, n^{3/8} x\},$$

and hence uniformly in $x \in [1, o(n^{1/6})]$,

$$n e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}} = o[1 - \Phi(x)], \quad (2.10)$$

when n is sufficiently large, where we have used a well-known fact: for $x > 0$,

$$\frac{1}{\sqrt{2\pi}} \left(\frac{1}{x} - \frac{1}{x^3} \right) e^{-x^2/2} \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-x^2/2}.$$

The meaning of (2.10) is that for all sequence u_n with $u_n = o(n^{1/6})$

$$\lim_{n \rightarrow \infty} \sup_{x \in [1, u_n]} n e^{-\eta \min\{n\epsilon_n^2, \sqrt{n}\epsilon_n x\}} / (1 - \Phi(x)) = 0.$$

On the other hand, it follows from (1.10) that

$$\begin{aligned}
P[S_n/V_n \geq (1 - \epsilon_n)x] &\leq \left\{1 - \Phi[(1 - \epsilon_n)x]\right\} \left[1 + O(1)x^3/\sqrt{n}\right] \\
&\leq [1 - \Phi(x)] \left\{1 + \frac{|\Phi[(1 - \epsilon_n)x] - \Phi(x)|}{1 - \Phi(x)}\right\} \left[1 + O(1)x^3/\sqrt{n}\right] \\
&= [1 - \Phi(x)] [1 + o(1)], \quad (2.11)
\end{aligned}$$

where we have used the result:

$$|\Phi[(1 - \epsilon_n)x] - \Phi(x)| \leq \epsilon_n x e^{-(1 - \epsilon_n)^2 x^2/2} = o[1 - \Phi(x)],$$

uniformly in $[1, o(n^{1/6})]$, since $\epsilon_n x^2 \leq \epsilon'_n x^3/\sqrt{n} = o(1)$.

By virtue of (2.10)–(2.11) and the upper bound of (1.11), we obtain $P(T_n \geq x) \leq [1 - \Phi(x)][1 + o(1)]$. Similarly we have $P(T_n \geq x) \geq [1 - \Phi(x)][1 + o(1)]$. This proves (1.8).

In a similar matter, by choosing $\epsilon_n = \max\{n^{-1/8}, \epsilon'_n\}$ where ϵ'_n is a sequence of constants such that $\epsilon'_n \rightarrow 0$ so slowly that $n\epsilon_n'^2/x_n^2 \rightarrow \infty$, we have

$$\begin{aligned} n &= o(1)e^{0.5\eta \min(n\epsilon_n^2, \sqrt{n}\epsilon_n x)}, \\ x^2 &= o(\min(n\epsilon_n^2, \sqrt{n}\epsilon_n x)) \end{aligned}$$

and therefore

$$ne^{-\eta \min(n\epsilon_n^2, \sqrt{n}\epsilon_n x)} = o(1)e^{-x^2/4},$$

which together with (1.11) and (1.9) proves (1.7).

The proof of Theorem 1.2 is now complete.

3. PROOF OF PROPOSITION 2.1

In this section, we give the proof of Proposition 2.1. Lemma 3.1 is interesting in itself as it provides an exponential bound for martingale difference under finite moment conditions.

Lemma 3.1. *Let $\{\xi_i, \mathcal{F}_i, i \geq 1\}$ be a sequence of martingale difference with $E\xi_i^2 < \infty$ and put $d_i^2 = E(\xi_i^2|\mathcal{F}_{i-1})$. Then*

$$P\left(\frac{|\sum_{i=1}^n \xi_i|}{(\sum_{i=1}^n (\xi_i^2 + 2d_i^2 + 3E\xi_i^2))^{1/2}} \geq x\right) \leq \sqrt{2} \exp(-x^2/8) \tag{3.1}$$

for all $x > 0$.

Proof. We first show that

$$e^{x-x^2} \leq 1 + x1_{\{x \geq -1/2\}} \tag{3.2}$$

for all $x \in R$.

It is easy to see that (3.2) holds for $x < -1/2$. For $x \geq -1/2$ let $f(x) = x - x^2 - \ln(1 + x)$. Observe that

$$\begin{aligned} f'(x) &= 1 - 2x - \frac{1}{1+x} = \frac{-x(1+2x)}{1+x} \\ &\begin{cases} > 0 & \text{for } -1/2 < x < 0, \\ = 0 & \text{for } x = 0, \\ < 0 & \text{for } x > 0. \end{cases} \end{aligned}$$

Therefore f achieves maximum at $x = 0$, that is, $f(x) \leq f(0) = 0$ for $x > -1/2$. This proves (3.2).

It follows from (3.2) that for $t \in R$

$$\begin{aligned} E\left(\exp(t\xi_i - t^2(\xi_i^2 + 2d_i^2))|\mathcal{F}_{i-1}\right) &= e^{-2t^2 d_i^2} E\left(\exp(t\xi_i - t^2\xi_i^2)|\mathcal{F}_{i-1}\right) \\ &\leq e^{-2t^2 d_i^2} \left(1 + E(t\xi_i 1_{\{t\xi_i \geq -1/2\}}|\mathcal{F}_{i-1})\right) \\ &= e^{-2t^2 d_i^2} \left(1 - E(t\xi_i 1_{\{t\xi_i < -1/2\}}|\mathcal{F}_{i-1})\right) \\ &\leq e^{-2t^2 d_i^2} \left(1 + 2E((t\xi_i)^2|\mathcal{F}_{i-1})\right) \\ &\leq 1. \end{aligned}$$

This shows that $\left\{ \exp\left(t \sum_{j=1}^i \xi_j - t^2 \sum_{j=1}^i (\xi_j^2 + 2d_j^2)\right), \mathcal{F}_i, i \geq 1 \right\}$ is a super-martingale and hence

$$E \exp\left(t \sum_{j=1}^n \xi_j - t^2 \sum_{j=1}^n (\xi_j^2 + 2d_j^2)\right) \leq 1. \tag{3.3}$$

By (3.3) and Theorem 2.1 of,

$$E \exp \left(\frac{a |\sum_{i=1}^n \xi_i|}{(2 \sum_{i=1}^n (\xi_i^2 + 2d_i^2 + 3E\xi_i^2))^{1/2}} \right) \leq \sqrt{2} \exp(a^2) \quad (3.4)$$

for all $a > 0$. Letting $a = x/(2\sqrt{2})$ together with Markov's inequality yields

$$\begin{aligned} P \left(\frac{|\sum_{i=1}^n \xi_i|}{(\sum_{i=1}^n (\xi_i^2 + 2d_i^2 + 3E\xi_i^2))^{1/2}} \geq x \right) &\leq e^{-ax/\sqrt{2}} E \exp \left(\frac{a |\sum_{i=1}^n \xi_i|}{(2 \sum_{i=1}^n (\xi_i^2 + 2d_i^2 + 3E\xi_i^2))^{1/2}} \right) \\ &\leq \sqrt{2} \exp(-ax/\sqrt{2} + a^2) = \sqrt{2} \exp(-x^2/8). \end{aligned}$$

This proves (3.1). \square

Lemma 3.2. *Let $\{\xi_i, i \geq 1\}$ be independent random variables with zero means and finite variances. Put*

$$S_n = \sum_{i=1}^n \xi_i, \quad V_n^2 = \sum_{i=1}^n \xi_i^2, \quad B_n^2 = \sum_{i=1}^n E\xi_i^2.$$

Then

$$P(|S_n| \geq x(V_n^2 + 5B_n^2)^{1/2}) \leq \sqrt{2} \exp(-x^2/8) \quad \text{for } x > 0 \quad (3.5)$$

and

$$ES_n^2 I(|S_n| \geq x(V_n + 4B_n)) \leq 23B_n^2 e^{-x^2/4} \quad (3.6)$$

Proof. Result (3.5) follows from (3.1) directly because $E(\xi_i^2 | \mathcal{F}_{i-1}) = E\xi_i^2$ by independence of random variables.

When $0 < x < 3$, we have $23e^{-x^2/4} \geq 1$ and $ES_n^2 I(|S_n| \geq x(V_n + 4B_n)) \leq ES_n^2 = B_n^2$ and hence (3.6) holds. When $x > 3$, let $\{\eta_i, 1 \leq i \leq n\}$ be an independent copy of $\{\xi_i, 1 \leq i \leq n\}$. Set

$$S_n^* = \sum_{i=1}^n \eta_i, \quad V_n^{*2} = \sum_{i=1}^n \eta_i^2.$$

By the Chebyshev inequality,

$$P(|S_n^*| \leq 2B_n, V_n^{*2} \leq 4B_n^2) \geq 1 - P(|S_n^*| > 2B_n) - P(V_n^{*2} > 4B_n^2) \geq 1 - 1/4 - 1/4 = 1/2.$$

Noting that

$$\begin{aligned} &\{|S_n| \geq x(4B_n + V_n), |S_n^*| \leq 2B_n, V_n^{*2} \leq 4B_n^2\} \\ &\subset \{|S_n - S_n^*| \geq x(4B_n + \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2} - V_n^*) - 2B_n, |S_n^*| \leq 2B_n, V_n^{*2} \leq 4B_n^2\} \\ &\subset \{|S_n - S_n^*| \geq x(2B_n + \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}) - 2B_n, |S_n^*| \leq 2B_n\} \\ &\subset \{|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}, |S_n^*| \leq 2B_n\}, \end{aligned}$$

we have

$$\begin{aligned}
 ES_n^2 I(|S_n| \geq x(V_n + 4B_n)) &= \frac{ES_n^2 I(|S_n| \geq x(V_n + 4B_n)) I(|S_n^*| \leq 2B_n, V_n^{*2} \leq 4B_n^2)}{P(|S_n^*| \leq 2B_n, V_n^{*2} \leq 4B_n^2)} \\
 &\leq 2ES_n^2 I\left(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}, |S_n^*| \leq 2B_n\right) \\
 &\leq 4E(S_n - S_n^*)^2 I\left(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}, |S_n^*| \leq 2B_n\right) \\
 &\quad + 4ES_n^{*2} I\left(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}, |S_n^*| \leq 2B_n\right) \\
 &\quad \quad \quad \text{[by the fact that } S_n^2 \leq 2(S_n - S_n^*)^2 + 2S_n^{*2}] \\
 &\leq 4E(S_n - S_n^*)^2 I\left(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}\right) \\
 &\quad + 16B_n^2 P\left(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}\right). \tag{3.7}
 \end{aligned}$$

Let $\{\varepsilon_i, 1 \leq i \leq n\}$ be a Rademacher sequence independent of $\{\xi_i, 1 \leq i \leq n\}$ and $\{\eta_i, 1 \leq i \leq n\}$. Noting that $\{\xi_i - \eta_i, 1 \leq i \leq n\}$ is a sequence of independent symmetric random variables, $\{\varepsilon_i(\xi_i - \eta_i), 1 \leq i \leq n\}$ and $\{\xi_i - \eta_i, 1 \leq i \leq n\}$ have the same joint distribution. It is known that

$$P\left(\left|\sum_{i=1}^n a_i \varepsilon_i\right| \geq x \left(\sum_{i=1}^n a_i^2\right)^{1/2}\right) \leq 2e^{-x^2/2} \tag{3.8}$$

for any real numbers $\{a_i\}$. Hence with $Y = \left|\sum_{i=1}^n a_i \varepsilon_i\right| / \left(\sum_{i=1}^n a_i^2\right)^{1/2}$

$$\begin{aligned}
 \left(\sum_{i=1}^n a_i^2\right)^{-1} E\left(\sum_{i=1}^n a_i \varepsilon_i\right)^2 I\left(\left|\sum_{i=1}^n a_i \varepsilon_i\right| \geq x \left(\sum_{i=1}^n a_i^2\right)^{1/2}\right) &= EY^2 I(Y \geq x) \\
 &= x^2 P(Y \geq x) + 2 \int_x^\infty t P(Y \geq t) dt \\
 &\leq 2x^2 e^{-x^2/2} + 4 \int_x^\infty t e^{-t^2/2} dt \\
 &= 2(2 + x^2) e^{-x^2/2} \leq 2.4 e^{-x^2/4} \tag{3.9}
 \end{aligned}$$

for $x > 3$. Thus by (3.8) and (3.9) for $x > 3$

$$P(|S_n - S_n^*| \geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2\right)^{1/2}) \leq 2e^{-x^2/2} \leq 0.22e^{-x^2/4} \tag{3.10}$$

and

$$\begin{aligned}
 E(S_n - S_n^*)^2 I(|S_n - S_n^*| \geq x) &\geq x \left(\sum_{i=1}^n (\xi_i - \eta_i)^2 \right)^{1/2} I(|S_n^*| \leq 2B_n) = E \left(\sum_{i=1}^n \varepsilon_i (\xi_i - \eta_i) \right)^2 I \left(\left| \sum_{i=1}^n \varepsilon_i (\xi_i - \eta_i) \right| \right. \\
 &\geq x \left. \left(\sum_{i=1}^n (\xi_i - \eta_i)^2 \right)^{1/2} \right) \\
 &\leq 2.4 e^{-x^2/4} E \sum_{i=1}^n (\xi_i - \eta_i)^2 \\
 &= 4.8 B_n^2 e^{-x^2/4}.
 \end{aligned} \tag{3.11}$$

This proves (3.6) by (3.7), (3.10) and (3.11). □

In the following two lemmas we continue to use the notations given in Section 2.

Lemma 3.3. *Assume $\sigma_1^2 = 1$. Then for all $y > 0$,*

$$P(|S_n| \geq y(V_n + \sqrt{5n})) \leq 2e^{-y^2/8} \tag{3.12}$$

and

$$P(V_n^2 \leq n/2) \leq e^{-\eta_0 n}. \tag{3.13}$$

where $\eta_0 = 1/(32a_0^2)$ and a_0 satisfied

$$Eg(X_1)^2 I\{|g(X_1)| \geq a_0\} \leq 1/4. \tag{3.14}$$

Proof. Recall $Eg(X_1) = 0$ and $Eg(X_1)^2 = 1$. (3.12) is a special case of (3.5). We next prove (3.13). Let $Y_k = g(X_k)I_{(|g(X_k)| \leq a_0)}$. Since $e^{-x} \leq 1 - x + x^2/2$ for $x > 0$, we have with $t = 1/(4a_0^2)$

$$\begin{aligned}
 P(V_n^2 \leq n/2) &\leq P\left(\sum_{k=1}^n Y_k^2 \leq n/2\right) \\
 &\leq e^{tn/2} E e^{-t \sum_{k=1}^n Y_k^2} = e^{tn/2} (E e^{-tY_1^2})^n \\
 &\leq e^{tn/2} \left(1 - tEY_1^2 + t^2EY_1^4/2\right)^n \\
 &\leq e^{tn/2} \left(1 - (3/4)t + t^2a_0^2/2\right)^n \\
 &\leq \exp\left(- (t/4 - t^2a_0^2/2)n\right) = \exp\left(- \frac{n}{32a_0^2}\right),
 \end{aligned}$$

as desired. □

Lemma 3.4. *Assume $\sigma_1^2 = 1$. Then, for all $y \geq 0$,*

$$P[\Lambda_n^2 \geq a_0 y^2 n (7V_n^2 + 11n)] \leq \sqrt{2} n e^{-y^2/8} \tag{3.15}$$

where $a_0 = 2(c_0 + 4)$, and

$$P\left[\left|\sum_{1 \leq i < j \leq n} \psi(X_i, X_j)\right| \geq a_1 y^2 \sqrt{n} (V_n^2 + 106n)^{1/2}\right] \leq \sqrt{2} (n+2) e^{-y^2/8}, \tag{3.16}$$

where $a_1^2 = 46(c_0 + 4)$.

Proof. First prove (3.15). Note that, given X_i , $W_n^{(i)}$ is a sum of i.i.d. random variables with zero means. It follows from (3.5) that

$$P\left\{|W_n^{(i)}| \geq y[V_n^{(i)2} + 5(n-1)\tau^2(X_i)]^{1/2}\right\} \leq \sqrt{2}e^{-y^2/8} \tag{3.17}$$

where $V_n^{(i)2} = \sum_{\substack{j=1 \\ j \neq i}}^n \psi^2(X_i, X_j)$ and $\tau^2(x) = E(\psi^2(X_1, X_j)|X_j = x)$. Note that $\psi^2(x_1, x_2) \leq 2(c_0 + 4)[1 + g^2(x_1) + g^2(x_2)]$. We have

$$V_n^{(i)2} + 5(n-1)\tau^2(X_i) \leq 2(c_0 + 4) \left[11n + 6ng^2(X_i) + \sum_{i=1}^n g^2(X_i) \right].$$

This, together with (3.17) and the fact that

$$\sum_{i=1}^n \left[11n + 6ng^2(X_i) + \sum_{i=1}^n g^2(X_i) \right] = n(7V_n^2 + 11n),$$

yields that

$$\begin{aligned} P\left[\Lambda_n^2 \geq a_0 y^2 n (7V_n^2 + 11n)\right] &\leq \sum_{i=1}^n P\left\{|W_n^{(i)}| \geq y[V_n^{(i)2} + 5(n-1)\tau^2(X_i)]^{1/2}\right\} \\ &\leq \sqrt{2}n e^{-y^2/8}, \end{aligned}$$

as required.

We next prove (3.16). Let $\mathcal{F}_j = \sigma(X_i, i \leq j)$ and rewrite

$$\sum_{1 \leq i < j \leq n} \psi(X_i, X_j) = \sum_{j=2}^n Y_j,$$

where $Y_j = \sum_{i=1}^{j-1} \psi(X_i, X_j)$. Then $\{Y_j, \mathcal{F}_j, j \geq 2\}$ is a martingale difference sequence. By (3.1), we have

$$P\left(\left|\sum_{j=2}^n Y_j\right| \geq y \left\{\sum_{j=2}^n [Y_j^2 + 3EY_j^2 + 2E(Y_j^2|\mathcal{F}_{j-1})]\right\}^{1/2}\right) \leq \sqrt{2}e^{-y^2/8}. \tag{3.18}$$

Note that $EY_j^2 \leq (j-1)Eh^2(X_1, X_2) \leq 3(j-1)$ by (1.6) and $Eg^2(X_1) = 1$. The result (3.16) follows if we prove

$$P\left[T_{1n}^2 \geq a_2 y^2 n (V_n^2 + n)\right] \leq \sqrt{2}n e^{-y^2/8}, \tag{3.19}$$

where $T_{1n}^2 = \sum_{j=2}^n Y_j^2$ and $a_2 = 14(c_0 + 4)$, and

$$P\left[T_{2n}^2 \geq a_3 y^2 n (V_n^2 + 50n)\right] \leq \sqrt{2}e^{-y^2/4}, \tag{3.20}$$

where $T_{2n}^2 = \sum_{j=2}^n E(Y_j^2|\mathcal{F}_{j-1})$ and $a_3 = 16(c_0 + 4)$.

We only prove (3.20). The proof of (3.19) is similar to (3.15). We omit the details. Without loss of generality, assume $y \geq 1$. Otherwise (3.20) is obvious. Write $V'_j = V_{\psi,j} + 4(j-1)^{1/2}\tau(X_j)$, where $V_{\psi,j}^2 = \sum_{i=1}^{j-1} \psi^2(X_i, X_j)$. Observe that

$$\begin{aligned}
 P\{T_{2n}^2 &\geq 2y^2[4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1)]\} \leq P\left\{\sum_{j=2}^n E[Y_j^2 I(|Y_j| \leq yV'_j) | \mathcal{F}_{j-1}]\right\} \\
 &\geq y^2[4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1)] + P\left\{\sum_{j=2}^n E[Y_j^2 I(|Y_j| > yV'_j) | \mathcal{F}_{j-1}]\right\} \\
 &\geq y^2[4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1)] \\
 &:= J_1 + J_2.
 \end{aligned} \tag{3.21}$$

Note that

$$\begin{aligned}
 J_1 &\leq P\left\{\sum_{j=2}^n y^2 E[V_j'^2 | \mathcal{F}_{j-1}] \geq y^2 \left[4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1)\right]\right\} \\
 &= P\left\{\sum_{j=2}^n \sum_{i=1}^{j-1} 2\tau^2(X_i) + 32 \sum_{j=2}^n (j-1)E\tau^2(X_1) \geq 4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1)\right\} \\
 &= 0
 \end{aligned} \tag{3.22}$$

and that (recall $y \geq 1$)

$$\begin{aligned}
 J_2 &\leq \frac{1}{64y^2n^2E\tau^2(X_1)} \sum_{j=2}^n E[Y_j^2 I(|Y_j| > yV'_j)] \\
 &= \frac{1}{64y^2n^2E\tau^2(X_1)} \sum_{j=2}^n E\{E[Y_j^2 I(|Y_j| > yV'_j) | X_j]\} \\
 &\leq \frac{16}{64y^2n^2E\tau^2(X_1)} \sum_{j=2}^n E[j\tau^2(X_1)]e^{-y^2/4} \quad \text{by (3.6)} \\
 &\leq e^{-y^2/4}.
 \end{aligned} \tag{3.23}$$

The result (3.20) now follows from (3.21)–(3.23) and the fact that

$$4n \sum_{j=2}^n \tau^2(X_j) + 64n^2 E\tau^2(X_1) \leq 8(c_0 + 4)n(50n + V_n^2),$$

as $\tau^2(x) \leq 2(c_0 + 4)[2 + g(x)]$. This also completes the proof of Lemma 3.4. □

We are now ready to prove Proposition 2.1. Without loss of generality, assume $\sigma_1^2 = 1$. Otherwise, consider h/σ_1 in the place of h . We only prove (2.8). The proof of (2.9) is given in a similar manner except we use (3.16) in the place of (3.15).

By (3.12) and (3.13), for any $x > 0$

$$\begin{aligned} P(|S_n| \geq 5xV_n) &\leq P(V_n^2 \leq n/2) + P[|S_n| \geq x(V_n + \sqrt{5n})] \\ &\leq 2e^{-x^2/8} + e^{-\eta_0 n}. \end{aligned}$$

By (3.15) and (3.13), for any $x > 0$,

$$\begin{aligned} P(\Lambda_n \geq \sqrt{7a_0 + 22}x\sqrt{n}V_n) &\leq P(V_n^2 \leq n/2) + P[\Lambda_n^2 \geq a_0 x^2 n (7V_n^2 + 11n)] \\ &\leq \sqrt{2}n e^{-x^2/8} + e^{-\eta_0 n}. \end{aligned}$$

These facts imply that, for any $y > 0$,

$$\begin{aligned} P(|\delta_n| \geq y) &\leq 2P(|S_n| \geq \sqrt{y(n-2)}V_n/3) + 2P(\Lambda_n \geq y(n-2)V_n/4) + P(\Lambda_n \geq \sqrt{y}(n-2)V_n/\sqrt{3}) \\ &\leq 2\sqrt{2}(n+1)e^{-\delta'_0 y n} + 2\sqrt{2}n e^{-\delta''_0 y^2 n} + 5e^{-\eta_0 n} \\ &\leq 4\sqrt{2}(n+2) \exp(-\delta_0 n \min\{1, y, y^2\}), \end{aligned}$$

where δ_0 , δ'_0 and δ''_0 are constants depending only on σ_1^2 and c_0 . This proves (2.8) and hence completes the proof of Proposition 2.1.

Acknowledgements. The authors thank the referee for his/her valuable comments.

REFERENCES

- [1] I.B. Alberink and V. Bentkus, Berry-Esseen bounds for von-Mises and U -statistics. *Lith. Math. J.* **41** (2001) 1–16.
- [2] I.B. Alberink and V. Bentkus, Lyapunov type bounds for U -statistics. *Theory Probab. Appl.* **46** (2002) 571–588.
- [3] J.N. Arvesen, Jackknifing U -statistics. *Ann. Math. Statist.* **40** (1969) 2076–2100.
- [4] Y.V. Borovskikh and N.C. Weber, Large deviations of U -statistics I. *Lietuvos Matematikos Rinkiny*s **43** (2003) 13–37.
- [5] Y.V. Borovskikh and N.C. Weber, Large deviations of U -statistics I. *Lietuvos Matematikos Rinkiny*s **43** (2003) 294–316.
- [6] H. Callaert and N. Veraverbeke, The order of the normal approximation for a studentized U -statistics. *Ann. Statist.* **9** (1981) 194–200.
- [7] W. Hoeffding, A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** (1948) 293–325.
- [8] B.-Y. Jing, Q.M. Shao and Q. Wang, Self-normalized Cramér-type large deviation for independent random variables. *Ann. Probab.* **31** (2003) 2167–2215.
- [9] B.-Y. Jing, Q.M. Shao, W. Zhou, Saddlepoint approximation for Student’s t -statistic with no moment conditions. *Ann. Statist.* **32** (2004) 2679–2711.
- [10] V.S. Koroljuk and V. Yu. Borovskikh, *Theory of U -statistics*. Kluwer Academic Publishers, Dordrecht (1994).
- [11] Q.M. Shao, Self-normalized large deviations. *Ann. Probab.* **25** (1997) 285–328.
- [12] Q.M. Shao, Cramér-type large deviation for Student’s t statistic. *J. Theoret. Probab.* **12** (1999) 387–398.
- [13] V.H. De La Pena, M.J. Klass and T.L. Lai, Self-normalized processes: exponential inequalities, moment bound and iterated logarithm laws. *Ann. Probab.* **32** (2004) 1902–1933.
- [14] M. Vardemaele and N. Veraverbeke, Cramer type large deviations for studentized U -statistics. *Metrika* **32** (1985) 165–180.
- [15] Q. Wang, Bernstein type inequalities for degenerate U -statistics with applications. *Ann. Math. Ser. B* **19** (1998) 157–166.
- [16] Q. Wang, B.-Y. Jing and L. Zhao, The Berry-Esséen bound for studentized statistics. *Ann. Probab.* **28** (2000) 511–535.
- [17] Q. Wang and N.C. Weber, Exact convergence rate and leading term in the central limit theorem for U -statistics. *Statist. Sinica* **16** (2006) 1409–1422.
- [18] L. Zhao, The rate of the normal approximation for a studentized U -statistic. *Science Exploration* **3** (1983) 45–52.