Berry-Esseen Bounds for Studentized Statistics by Stein's Method

by

Zhang Kan

A Thesis Submitted to

The Hong Kong University of Science and Technology

in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy

in Mathematics

September 2012, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Zhang Kan

Berry-Esseen Bounds for Studentized Statistics by Stein's Method

by

Zhang Kan

This is to certify that I have examined the above Ph.D thesis
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by
the thesis examination committee have been made.

Prof. SHAO, Qi-Man, Supervisor

Prof. Allen Moy, Head of Department

Department of Mathematics

September 2012

Acknowledgment

Firstly, I'd like to take this opportunity to sincerely thank my supervisor, Prof. Shao, Qi-Man, who offered me the opportunity to study in the beautiful campus of the Hong Kong University of Science and Technology. During my PhD study in the past years, Prof. Shao has given me a lot of valuable advice and help in my study. I really appreciate it for Prof. Shao's great support and patience for me in such a long time. In addition, I would like to thank the Department of Mathematics for providing the postgraduate studentship in the first four years to support my study and life in Hong Kong. My heartfelt gratitude also goes to Prof. Jing Bing-Yi, Prof. Ling Shi-Qing, and Prof. Chen Kani, who not only taught me knowledge in statistics but also gave me all kinds of helps when I had difficulties in my study.

Secondly, I am eager to express my great thanks to my family. Life would be different without their love. I want to thank my parents. When I was a little girl, they let me know the most important thing in our lives is to be a nice person. Although we have different opinions in many aspects about my PhD study, I can always feel their love. I would also like to extend my deepest thanks to my husband, Yarui, who is now a Postdoc in Division of Life Science in HKUST. We came to Hong Kong together six years ago. He makes my life rich and colorful. In the past years, he witnessed all of my happiness, tensions and depressions, and always gave me the biggest support and understanding. My thanks also extended to my parents-in-law. I am so lucky to become a family with them, who are always there when we were facing challenges in the life. Without the support and encouragement from all my family, it would be impossible for me to complete my PhD study.

During the period of my study, I have obtained innumerous help from my colleagues and friends. I give my special thank to Mr. Zhou, Wenxin, who is also

Prof. Shao's PhD student, for his unselfish help, valuable discussions and suggestions on my study, especially about the Studentized U-statistics. I want to thank Dr. Liu Dongwen and Dr. Zhu Ke. They gave me lots of help and useful suggestions for my study and we became really good friends. I also want to thank Dr. Li Dong, Dr. Liu Zhi, Dr. Kong Xinbing, Dr. Yao Yuan, Dr. Lin Yuanyuan, Dr. Yang Yingsi, Mr. Zhou Chaoxu, Dr. Cong Xuri, Dr. Cai Tao, Dr. Yang Liang, Mr. Zhu Muyan, who helped me in different ways. All my dearest friends whom I have forgotten to mention here are also greatly appreciated.

Contents

Title Page i

Authorization Page ii

Signature Page iii

Acknowledgments iv

Table of Contents vi

Abstract viii

1 Introduction 1

2 Stein's Method 5

2.1 A Brief History of Stein's Method 5

2.2 The Main Idea of Stein's Method 7

3 Main Theorem 12

4 Proof of Main Theorem **15**

4.1 Main Idea of the Proof 15

4.2 Proof 16

4.3 Proof of Lemmas 22

5 Applications **35**

5.1 Application to Student's t-Statistics 35

5.1.1 Introduction 35

5.1.2 Main Result 36

5.2 Application to Studentized U-Statistics 37

5.2.1 Introduction 37

5.2.2 Main Result 38

5.2.3 Proof of Theorem 5.2. 39

5.3 Application to Studentized L-Statistics 49

5.3.1 Introduction 49

5.3.2 Main Result 49

5.3.3 Proof of Theorem 5.5. 50

6 Conclusion and Future Work **59**

Bibliography **60**

Berry-Esseen Bounds for Studentized Statistics by Stein's Method

Zhang Kan

Department of Mathematics

Abstract

The Stein's method, originally introduced by Stein (1972), gives us a novel way to investigate normal approximation. This method has since been developed in different areas and can be applied to different distributional approximations. It is a very powerful tool to estimate the absolute error of approximations and also the relative error.

In this thesis, we establish the uniform and non-uniform Berry-Esseen bound for a class of Studentized statistics via Stein's method. In particular, we recover the optimal results for Student's t-statistics, Studentized U-statistics and Studentized L-statistics.

Chapter 1

Introduction

Let $X_1, X_2, \ldots, X_n$ be independent random variables and let $T := T(X_1, \ldots, X_n)$ be a general sampling statistic of interest. Assume that when n goes to infinity, T converges to a standard normal distribution so that the normal law can be used to approximate a p-value of a hypothesis test. However, even though in practice sample sizes may be large, the normal approximation may not be accurate caused by other factors. Thus, it is important for us to evaluate the quality of the normal approximation by estimating the Kolmogorov distance between the distribution function of T and normal distribution function or the ratio of the two tail probabilities.

The main contribution of this thesis is to evaluate the absolute error $\sup_{z\in\mathbb{R}} |P(T \leq z) - \Phi(z)|$ via Berry-Esseen type bounds. Suppose T is a linear statistic plus an error term, say, $T = W + \Delta$, where

$$W = \sum_{i=1}^{n} \xi_i = \sum_{i=1}^{n} g_i(X_i), \qquad \Delta := \Delta(X_1, \ldots, X_n) = T - W.$$

Here $g_i := g_{n,i}$ are Borel measurable functions.

Assume that

$$E\xi_i = 0 \text{ for } i = 1, 2, \ldots, n, \qquad \text{and} \quad \sum_{i=1}^{n} E\xi_i^2 = 1. \tag{1.1}$$

It is clear that if $\Delta \to 0$ in probability as $n \to \infty$, then the central limit theorem

$$\sup_{z\in\mathbb{R}} |P(T \leq z) - \Phi(z)| \to 0,$$

holds provided that the Lindeberg condition is satisfied, i.e.

$$\forall \epsilon > 0, \qquad \sum_{i=1}^{n} E\xi_i^2 I(|\xi_i| > \epsilon) \to 0.$$

Here and throughout the following, we let Φ denote the standard normal distribution function, that is,

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} dt, \quad \forall x \in \mathbb{R}.$$

In the realm of normal approximation for independent random variables, the classical approach is the so-called Fourier transform method. However, when it comes to dependent random variables, this approach may not work well. Stein (1972) introduced a novel method to get the bounds on distance between a s-tandard normal distribution and the distribution of a sum of dependent random variables. This method provides a new way to prove the central limit theorem, and more importantly, it gives the convergence rate at the same time. This method is now known as Stein's method, which has been extensively studied and well developed during the last two decades. In addition, the method has been extended to Poisson and compound Poisson approximations, exponential approximation, multivariate, combinatorial and discretized normal approximations.

Due to statistical considerations, Chen and Shao (2007) proved the following uniform and non-uniform Berry-Esseen bounds for a class of non-linear statistics. Recall that $\xi_i = g_{n,i}(X_i)$, put

$$\beta_2 = \sum_{i=1}^{n} E\{\xi_i^2 I(|\xi_i| > 1)\}, \qquad \beta_3 = \sum_{i=1}^{n} \{E|\xi_i|^3 I(|\xi_i| \leq 1)\},$$

and let $\delta > 0$ satisfy

$$\sum_{i=1}^{n} E\{|\xi_i| \min(\delta, |\xi_i|)\} \geq 1/2. \tag{1.2}$$

For each $1 \leq i \leq n$, let Δ_i be a random variable such that X_i and $(\Delta_i, W - \xi_i)$ are independent. Then for any $p \geq 2$,

$$\sup_{x \in \mathbb{R}} |P(T \leq x) - \Phi(x)| \leq 6.1(\beta_2 + \beta_3) + E|W\Delta| + \sum_{i=1}^{n} E|\xi_i(\Delta - \Delta_i)| \tag{1.3}$$

and

$$|P(T \leq x) - \Phi(x)| \leq \gamma_x + e^{-|x|/3}\tau, \quad \text{for all } x \in \mathbb{R}, \tag{1.4}$$

where

$$\begin{aligned}
\gamma_x &= P(|\Delta| > (|x|+1)/3) + 2\sum_{i=1}^{n} P\big(|\xi_i| > (|x|+1)/(6p)\big) \\
&\quad + e^p\big(1 + x^2/(36p)\big)^{-p}\beta_2, \\
\tau &= 22\delta + 8.6\|\Delta\|_2 + 3.6\sum_{i=1}^{n} \|\xi_i\|_2 \|\Delta - \Delta_i\|_2.
\end{aligned}$$

The general results have been successfully applied to obtain optimal Berry-Esseen bounds for U-statistics, multi-sample U-statistics, L-statistics, and random sums.

However, the general results of standardized statistics are not enough. It is also important to obtain similar Berry-Esseen type bounds for Studentized version of T, under the assumption of asymptotic normality. This is because in statistical inference studentized statistics are commonly used since non-studentized statistics often involve some unknown nuisance parameters. A prototypical example is Student's t-statistic, whose high degree of robustness against heavy-tailed data has been quantified in recent years. The main purpose of this thesis is to establish uniform and non-uniform Berry-Essend bounds for a class of Studentized non-linear statistics, including Studentized U-statistics and L-statistics.

This thesis is organized as follows. Chapter 2 contains a brief introduction of Stein's method, including Stein's equation, properties of the solutions and construction of Stein identities for sums of independent random variables. Chapter

3 presents the main theorem, Theorem 3.1, on uniform and non-uniform Berry-Essend bounds for Studentized non-linear statistics. A detailed proof of the main theorem is given in Chapter 4. In Chapter 5, we specialize Theorem 3.1 to some well-known and widely applicable statistics, such as Student's t-statistics, Studentized U-statistics and Studentized L-statistics. Finally, we end up in Chapter 6 with a conclusion on both current and future work.

Chapter 2

Stein's Method

2.1 A Brief History of Stein's Method

The classical approach to study Central Limit Theorem relied heavily on Fourier method. The Fourier methods work well for sums of independent random variables. Stein (1972) initiated a novel method to obtain a bound between the distribution of a sum of m-dependent sequence of random variables and a standard normal distribution in the Kolmogorov metric. This method is now called Stein's method. Stein's method is a powerful tool because it works not only for independent random variables but also for dependent random variables. It can prove the central limit theorem and give bounds for accuracy of approximations at the same time. Extensive applications of Stein's method to obtain uniform and non-uniform Berry-Esseen bounds for independent and dependent random variables can be found in, for example, Diaconis (1977), Baldi, Rinott and Stein (1989), Barbour (1990), Dembo and Rinott (1996), Goldstein and Reinert (1997), Chen and Shao (2001, 2004, 2007), Chatterjee (2008), and Nourdin and Peccati (2009).

After Stein (1972), Louis Chen developed this method to investigate Poisson approximation for sums of dependent indicator random variables, see Chen (1975). From then on, statisticians put their efforts to extend the applications of Stein's idea to many probability approximations other than Normal and Poisson, say, Poisson process, compound Poisson and binomial approximations, which can be found in Diaconis and Holmes (2004), Barbour and Chen (2005) and Chen, Goldstein and Shao (2010). Other than the study of more probability approximations, intensive efforts have been made to apply Stein's method to a wide range of areas to examine different distributional approximations, for example, in Arratia, Goldstein and Gordon (1990), Barbour, Holst and Janson (1992), and Chen (1993).

Recently, Stein's method has some new developments. Firstly, it is used to get the Cramér type moderate deviation, which basically considers the ratio of two tail probabilities of T and Z. Chen, Fang and Shao (2012) uses Stein's method to establish the optimal Cramér type moderate deviations for a lot of cases. Secondly, exchangeable pairs approach is well developed using Stein's method. Chatterjee and Shao (2011) uses exchangeable pairs approach to investigate non-normal approximation and obtains a Berry-Esseen bound of order $O(1/\sqrt{n})$ in the noncentral limit theorem for the magnetization in the Curie-Weiss ferromagnet at the critical temperature. Thirdly, the concentration inequality approach has become a powerful technique. It can be used to prove the normal approximation by Stein's method. For instance, Chen and Shao (2007) uses the concentration inequalities to give us the optimal uniform and non-uniform Berry-Esseen bound for a kind of non-linear statistics. Also the concentration inequalities are well developed to give us the Berry-Esseen bound for independent random variables and dependent random variables under local dependence. Shao (2010) gives a new exponential concentration inequality and it can be used to get Berry-Esseen bounds for other statistics. In this thesis, we will also use this exponential concentration inequality in our proof.

2.2 The Main Idea of Stein's Method

Let Z be a standard normal random variable and let $\mathcal{C}_{bd}$ be the class of bounded, continuous and piecewise differentiable functions $f : R \to R$ with $E|f'(Z)| < \infty$. Stein's method rests on the following characterization.

Lemma 2.1. *([22] Lemma 2.1) Let W be a real valued random variable. Then W has a standard normal distribution if and only if*

$$Ef'(W) = E(Wf(W)), \tag{2.1}$$

for all $f \in \mathcal{C}_{bd}$.

The proof of necessity is essentially a direct consequence of integration by parts. For the sufficiency, for fixed $x \in R$, let f_x be the unique bounded solution of the Stein equation

$$f'(w) - wf(w) = I(w \leq x) - \Phi(x). \tag{2.2}$$

The solution f_x is given by

$$f_x(w) \quad = \quad \begin{cases} \sqrt{2\pi}e^{w^2/2}(1-\Phi(w))\Phi(x), & w \geq x \\ \sqrt{2\pi}e^{w^2/2}(1-\Phi(x))\Phi(w), & w < x \end{cases} \tag{2.3}$$

and $f_x \in \mathcal{C}_{bd}$. Moreover, f_x has the following properties.

Lemma 2.2. *([22] Lemma 2.2) For the function f_x defined by (2.3), we have*

$$wf_x(w) \text{ is an increasing function of } w, \tag{2.4}$$

$$|wf_x(w)| \leq 1, \quad |wf_x(w) - uf_x(u)| \leq 1, \tag{2.5}$$

$$|f_x'(w)| \leq 1, \quad |f_x'(w) - f_x'(v)| \leq 1, \tag{2.6}$$

$$0 \leq f_x(w) \leq \min(1, 1/x) \tag{2.7}$$

and

$$|(w+u)f_x(w+u)-(w+v)f_x(w+v)| \le (|w|+\sqrt{2\pi}/4)(|u|+|v|) \tag{2.8}$$

for all real w,u, and v.

Let W be the random variable of interest. We aim to estimate $P(W \le x)-\Phi(x)$. By (2.2), it is equivalent to consider $Ef'(W)-EWf_x(W)$, which is often much easier to handle than the original one. For the most simple case, we consider when W is the standardized sum of independent random variables. Let $\xi_1, \xi_2, ..., \xi_n$ be independent random variables satisfying $E\xi_i = 0$ for $1 \le i \le n$ and $\sum_{i=1}^n E\xi_i^2 = 1$. Let

$$W = \sum_{i=1}^n \xi_i \qquad \text{and} \qquad W^{(i)} = W - \xi_i, \tag{2.9}$$

and define

$$K_i(t) = E(\xi_i(I(0 \le t \le \xi_i) - I(\xi_i \le t < 0))).$$

It is clear that $K_i(t) \ge 0$ for all real t, and

$$\int_{-\infty}^{\infty} K_i(t)dt = E\xi_i^2, \quad \text{and} \quad \int_{-\infty}^{\infty} |t|K_i(t)dt = \frac{1}{2}E|\xi_i|^3. \tag{2.10}$$

Then we can write

$$\begin{aligned} EWf(W) &= \sum_{i=1}^n E(\xi_i f(W)) \\ &= \sum_{i=1}^n E(\xi_i(f(W)-f(W^{(i)})) \\ &= \sum_{i=1}^n E(\xi_i \int_0^{\xi_i} f'(W^{(i)}+t)dt) \\ &= \sum_{i=1}^n E(\int_{-\infty}^{\infty} f'(W^{(i)}+t)\xi_i(I(0 \le t \le \xi_i) - I(\xi_i \le t < 0))dt) \\ &= \sum_{i=1}^n \int_{-\infty}^{\infty} E(f'(W^{(i)}+t))K_i(t)dt, \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} Ef'(W) &= \sum_{i=1}^{n} Ef'(W)E\xi_i^2 \\ &= \sum_{i=1}^{n}\int_{-\infty}^{\infty} Ef'(W)K_i(t)dt. \end{aligned} \tag{2.12}$$

Therefore by (2.11) and (2.12), our goal is to estimate

$$\begin{aligned} P(W \le x) - \Phi(x) &= Ef'(W) - EWf(W) \\ &= \sum_{i=1}^{n}\int_{-\infty}^{\infty} E(f'(W) - f'(W^{(i)} + t))K_i(t)dt. \end{aligned} \tag{2.13}$$

Now using the properties of $f'(x)$, we can get the Berry-Esseen bound. This is the most simple application of Stein's method. Next, we list some properties that will be used later.

Lemma 2.3. *We have the following properties*

$$f_x(w) \le \begin{cases} Ce^{-x/2}, & w \le x/2 \\ 1, & w > x/2, \end{cases}$$

$$f_x'(w) \le \begin{cases} Ce^{-x/2}, & w \le x/2 \\ 1, & w > x/2, \end{cases}$$

and

$$(wf_x(w))' \le \begin{cases} Ce^{-x/2}, & w \le 11x/12 \\ Cx, & w > 11x/12. \end{cases}$$

Proof of Lemma 2.3 For the function f_x defined by (2.3), we directly have

$$f_x(w) \le \begin{cases} \sqrt{2\pi}e^{x^2/8}(1-\Phi(x))\Phi(x/2), & w \le x/2 \\ 1, & w > x/2 \end{cases} \tag{2.14}$$

$$\le \begin{cases} Ce^{-x/2}, & w \le x/2 \\ 1, & w > x/2. \end{cases} \tag{2.15}$$

By (2.2), we get

$$f_x'(w) = \begin{cases} (\sqrt{2\pi} w e^{w^2/2}\Phi(w) + 1)(1 - \Phi(x)), & \text{if } w \le x \\ (\sqrt{2\pi} w e^{w^2/2}(1 - \Phi(w)) - 1)\Phi(x), & \text{if } w > x. \end{cases} \tag{2.16}$$

Then we get

$$f_x'(w) \le \begin{cases} (1 + \sqrt{2\pi}(x/2)e^{x^2/8}(1 - \Phi(x)), & w \le x/2 \\ 1, & w > x/2 \end{cases} \tag{2.17}$$

$$\le \begin{cases} Ce^{-x/2}, & w \le x/2 \\ 1, & w > x/2. \end{cases} \tag{2.18}$$

Let $G(w) = wf_x(w)$ and $g(w) = (wf_x(w))'$, we get

$$g(w) = \begin{cases} \left(\sqrt{2\pi}(1 + w^2)e^{w^2/2}(1 - \Phi(w)) - w\right)\Phi(x), & \text{if } w \ge x \\ \left(\sqrt{2\pi}(1 + w^2)e^{w^2/2}\Phi(w) + w\right)(1 - \Phi(x)), & \text{if } w < x. \end{cases} \tag{2.19}$$

A direct calculation (cf. Lemma 6.5 in [22]) shows that

$$\begin{aligned} 0 \le \sqrt{2\pi}(1 + w^2)e^{w^2/2}(1 - \Phi(w)) - w &\le \frac{2}{1 + w^3}, && \text{for } w \ge 0, \\ \sqrt{2\pi}(1 + w^2)e^{w^2/2}\Phi(w) + w &\le 2, && \text{for } w \le 0. \end{aligned}$$

This implies that $g \ge 0$, $g(w) \le 2(1 - \Phi(x)) \le 2$ for $w \le 0$ and $g(w) \le 2/(1+w^3)$ for $w \ge x$; furthermore, g is clearly increasing for $0 \le w < x$. Thus, we have

$$g(w) \le \begin{cases} 4(1 + x^2)e^{x^2/2}(1 - \Phi(x)), & \text{if } w \le 11x/12 \\ 4(1 + x^2)e^{x^2/2}(1 - \Phi(x)), & \text{if } w > 11x/12. \end{cases} \tag{2.20}$$

$$\le \begin{cases} Ce^{-x/2}, & w \le 11x/12 \\ Cx, & w > 11x/12. \end{cases} \tag{2.21}$$

Lemma 2.4. *For any* $|a| \le \frac{x}{4}$, $|b| \le \frac{x}{4}$, *we have*

$$|f_x(w + a) - f_x(w + b)| \le |b - a|e^{-\frac{x^2}{4}} + 4|b - a|I(w \ge \frac{x}{4}). \tag{2.22}$$

Proof of Lemma 2.4 Using Lemma 2.3 and noting that $xf(x)$ is increasing, it follows that

$$
\begin{aligned}
&|f(w+a)-f(w+b)| \\
&= |\int_{w+a}^{w+b} f'(t)dt| \\
&= |\int_{w+a}^{w+b} (tf(t)+I(t\le x)-\Phi(x))dt| \\
&= |\int_{w+a}^{w+b} (1-\Phi(x)+tf(t)+I(t>x))dt| \\
&\le |(b-a)(1-\Phi(x))| + |\int_{w+a}^{w+b} tf(t)dt| + |\int_{w+a}^{w+b} I(t\ge x)dt| \\
&\le |(b-a)(1-\Phi(x))| + |(a-b)(w+a)f(w+a)| + |(b-a)(w+b)f(w+b)| \\
&\quad +|(a-b)I(w+a>x)| + |(b-a)I(w+b>x)| \\
&\le |(b-a)(1-\Phi(x)| + |(a-b)(w+a)f(w+a)I(w+a<x/2)| \\
&\quad +|(a-b)(w+a)f(w+a)I(w+a\ge x/2)| + |(a-b)I(w+a>x)| \\
&\quad +|(b-a)(w+b)f(w+b)I(w+b<x/2)| \\
&\quad +|(b-a)(w+b)f(w+b)I(w+b\ge x/2)| + |(b-a)I(w+b>x)| \\
&\le |(b-a)(1-\Phi(x)| + C|a-b|e^{-x^2/4} + |a-b|I(w+a\ge x/2) \\
&\quad +|a-b|I(w+a>x) + C|b-a|e^{-x^2/4} \\
&\quad +|b-a|I(w+b\ge x/2) + |b-a|I(w+b>x) \\
&\le C|a-b|e^{-x^2/4} + 2|a-b|I(w+a>x/2) + 2|b-a|I(w+b>x/2) \\
&\le C|a-b|e^{-x^2/4} + 2|a-b|I(w\ge x/4) + 2|b-a|I(w\ge x/4) \\
&\le C|a-b|e^{-x^2/4} + 4|a-b|I(w\ge x/4).
\end{aligned}
$$

The proof of Lemma 2.4 is thus complete.

Chapter 3

Main Theorem

The main objective of interest in this thesis is the following class of abstract non-linear statistics of the form

$$T_s = \frac{W + \Delta_1}{\sqrt{1 + \Delta_2}}, \tag{3.1}$$

where in general we assume that both random variables Δ_1 and Δ_2 converge to zero in probability and W is sum of independent random variables as defined in (2.9). Many Studentized statistics can be written in terms of (3.1), including Studentized U-statistics, L-statistics, etc. The main purpose is to establish uniform and non-uniform Berry Esseen bounds of optimal order on the closeness of normality for T_s via Kolmogorov distance, that is,

$$\sup_{x \in \mathbb{R}} \big| P(T_s \leq x) - \Phi(x) \big|.$$

NOTATION. Throughout this thesis, C will denote an absolute positive constant whose value may be different at each appearance. We shall use $\|X\|_p$ to denote the L_p norm of a random variable X that is defined by $\|X\|_p := (E|X|^p)^{1/p}$, for every $p \geq 1$.

We now present the following main results.

Theorem 3.1. *Assume that $\xi_1, ..., \xi_n$ are independent random variables satisfying (1.1) and $T_s \equiv T_s(\xi_1, ..., \xi_n)$ is given by (3.1), where $W = \sum_{i=1}^n \xi_i$ and $\Delta_j \equiv \Delta_j(\xi_1, ..., \xi_n)$, $j = 1, 2$ are measurable functions of $\xi_1, ..., \xi_n$. Moreover, assume that for some $2 < r \le 3$, $E|\xi_i|^r < \infty$, $1 \le i \le n$. For any $x \in \mathbb{R}$, set*

$$\Delta_x = \Delta_1 - x\Delta_2/2. \tag{3.2}$$

Then, for every $p > 1$ with (p, q) satisfying $1/p + 1/q = 1$, we have

$$\begin{aligned} &|P(T_s \le x) - \Phi(x)| \\ \le\ & C\Big\{E(\Delta_1^2 \wedge 1) + (1+|x|)^{-1}E(\Delta_2^2 \wedge 1) + e^{-|x|/2}E(\Delta_2^2 e^{\bar{W}}) \\ & +|E[\Delta_x f_x(W)]| + (1+|x|^r)^{-1}\sum_{i=1}^n E|\xi_i|^r \\ & +e^{-|x|/40}\sum_{j=1}^n \big(\|\xi_j\|_p\|\Delta_1 - \Delta_1^{(j)}\|_q + \|\xi_j\|_p\|\Delta_2 - \Delta_2^{(j)}\|_q\big)\Big\}, \end{aligned}$$

where for every $x \in \mathbb{R}$, $f_x(\cdot)$ denotes the unique solution of the ordinary differential equation

$$f'(w) - wf(w) = I(w \le x) - \Phi(x), \quad w \in \mathbb{R},$$

that is,

$$f_x(w) \;=\; \begin{cases} \sqrt{2\pi}e^{w^2/2}(1-\Phi(w))\Phi(x), & w \ge x \\ \sqrt{2\pi}e^{w^2/2}(1-\Phi(x))\Phi(w), & w < x, \end{cases}$$

$\bar{W} = \sum_{i=1}^n \xi_i I(|\xi_i| \le 1)$ and $\Delta_1^{(j)}$, $\Delta_2^{(j)}$, for each $1 \le j \le n$ are arbitrary measurable functions of $\{\xi_i, 1 \le i \le n, i \ne j\}$.

The non-uniform Berry-Esseen bound obtained in Theorem 3.1 is optimal for many statistics.

Corollary 3.2. *Suppose the assumptions in Theorem 1 hold. If we assume that*

$$\Delta_2 = \sum_{i=1}^n \xi_i^2 - 1 + \Delta_3 = \sum_{i=1}^n (\xi_i^2 - E\xi_i^2) + \Delta_3,$$

where Δ_3 converges to zero in probability, then there exists a constant $C > 0$ such that

$$
\begin{aligned}
&|P(T_s \le x) - \Phi(x)| \\
\le\ & C\Big\{E(\Delta_1^2 \wedge 1) + (1+|x|)^{-1}E(\Delta_3^2 \wedge 1) + e^{-|x|/2}E(\Delta_3^2 e^{\bar{W}}) \\
& + |E[\Delta_1 f_x(W)]| + |E[\Delta_3 f_x(W)]| + \sum_{i=1}^{n} E|\xi_i|^r \\
& + e^{-|x|/40}\sum_{j=1}^{n}\big(\|\xi_j\|_p\|\Delta_1 - \Delta_1^{(j)}\|_q + \|\xi_j\|_p\|\Delta_3 - \Delta_3^{(j)}\|_q\big)\Big\}.
\end{aligned}
$$

Remark 3.3. *In Theorem 3.1 and Corollary 3.2, the choice of $\Delta_1^{(j)}$ and $\Delta_2^{(j)}$ are flexible. For example, for $k = 1, 2$, one can choose $\Delta_k^{(j)} = \Delta_k(\xi_1, ..., \xi_{j-1}, 0, \xi_{j+1}, ..., \xi_n)$ or $\Delta_k(\xi_1, ..., \xi_{j-1}, \hat{\xi}_j, \xi_{j+1}, ..., \xi_n)$ where $\hat{\xi}_j$ is an independent copy of ξ_j.*

Chapter 4

Proof of Main Theorem

4.1 Main Idea of the Proof

First, without loss of generality, assume $x \geq 0$ as we can simply apply the result to $-T_s$. By Taylor's expansion, we know, when $\Delta_2 \geq -1$,

$$\sqrt{1+\Delta_2} \leq 1 + \frac{1}{2}\Delta_2 \quad \text{and} \quad \sqrt{1+\Delta_2} \geq 1 + \frac{1}{2}\Delta_2 - \Delta_2^2.$$

Then it follows that

$$\begin{aligned} P(W + \Delta_x \geq x) &\leq P(\frac{W+\Delta_1}{\sqrt{1+\Delta_2}} \geq x) = P(W + \Delta_1 \geq x\sqrt{1+\Delta_2}) \\ &\leq P(W + \Delta_x \geq x) + P(x(1 + \frac{1}{2}\Delta_2 - \frac{1}{2}\Delta_2^2) \leq W + \Delta_1 \leq x(1 + \frac{1}{2}\Delta_2)), \end{aligned}$$

where we let $\Delta_x = \Delta_1 - x\Delta_2/2$.

Therefore, to estimate

$$P\Big(\frac{W+\Delta_1}{\sqrt{1+\Delta_2}} \leq x\Big) - \Phi(x),$$

it is sufficient to estimate

$$P(W + \Delta_x \leq x) - \Phi(x)$$

together with an error term

$$P\Big\{x\Big(1+\frac{1}{2}\Delta_2-\frac{1}{2}\Delta_2^2\Big)\le W+\Delta_1\le x\Big(1+\frac{1}{2}\Delta_2\Big)\Big\}.$$

Hence the proof is mainly formulated into two parts.

The first part is to get the estimation of $P(W+\Delta_x\le x)-\Phi(x)$. To use the Stein's equation, we add a term $E\Delta_x f(W)$ and consider $P(W+\Delta_x\le x)-\Phi(x)+E\Delta_x f(W)$ instead. The term $E\Delta_x f(W)$ will be kept in the result and be calculated when the result is applied to different statistics. Then we follow three steps. Firstly, we truncate ξ, W, $K_i(t)$ and Δ_x and prove it suffices to consider $P(\bar{W}+\bar{\Delta}_x\le x)-\Phi(x)+E\bar{\Delta}_x f(\bar{W})$. Secondly, we use Stein's equation to rewrite it and group them into four terms. Thirdly, we estimate these four terms one by one, which will be shown in Lemma 4.2.

In the second part, we will use a randomized concentration inequality to estimate the error term

$$P(x(1+\frac{1}{2}\Delta_2-\frac{1}{2}\Delta_2^2)\le W+\Delta_1\le x(1+\frac{1}{2}\Delta_2)).$$

To do this, we also do the truncations first, and then apply the randomized concentration inequality.

To complete the proof, we will combine all estimations together and make the result neat.

4.2 Proof

Before we use Stein's method to evaluate

$$P(W+\Delta_x\le x)-\Phi(x)+E\Delta_x f_x(W),$$

we will first introduce truncated variables $\bar{\xi}_i=\xi_i I(|\xi_i|\le 1)$, $1\le i\le n$ and set

$$\bar{W}=\sum_{i=1}^n\bar{\xi}_i,\quad \bar{K}_i(t)=E\{\bar{\xi}_i I(0\le t\le\bar{\xi}_i)-\bar{\xi}_i I(\bar{\xi}_i\le t\le 0)\},\tag{4.1}$$

and

$$\bar{\Delta}_x = \begin{cases} (x+1)/4, & \text{if } \Delta^* \geq (x+1)/4, \\ \Delta^*, & \text{if } |\Delta^*| \leq (x+1)/4, \\ -(x+1)/4, & \text{if } \Delta^* \leq -(x+1)/4, \end{cases} \tag{4.2}$$

where $\Delta^* \equiv \Delta(\bar{\xi}_1, \ldots, \bar{\xi}_n)$. With the above notations, we have

Lemma 4.1. *For* $\Delta_x = \Delta_1 - \frac{x}{2}\Delta_2$,

$$\begin{aligned} & |P(W + \Delta_x \leq x) - \Phi(x) + E\Delta_x f_x(W)| \\ \leq \; & 2P(|\Delta_x| \geq \frac{x+1}{4}) + \sum_{i=1}^n P(|W^{(i)}| \geq \frac{x-1}{2})P(|\xi_i| > 1) + \sum_{i=1}^n P(|\xi_i| \geq \frac{x+1}{4}) \\ & + |P(\bar{W} + \bar{\Delta}_x \leq x) - \Phi(x) + E\Delta_x f_x(W)|, \end{aligned}$$

where $W^{(i)}$ *is independent of* ξ_i*. In particular, for* $2 < r \leq 3$*, we have*

$$\begin{aligned} & |P(W + \Delta_x \leq x) - \Phi(x) + E\Delta_x f_x(W)| \\ \leq \; & |P(\bar{W} + \bar{\Delta}_x \leq x) - \Phi(x) + E\Delta_x f_x(W)| + 2P(|\Delta_1| \geq \frac{1}{4}) + 2P(|\Delta_2| \geq \frac{1}{2}) \\ & + C\sum_{i=1}^n \frac{E|\xi_i|^r}{1+|x|^r}. \end{aligned}$$

The proof of Lemma 4.1 will be presented in Section 4.3. By Lemma 4.1, it suffices to consider $P(\bar{W}+\bar{\Delta}_x \le x)-\Phi(x)+E\bar{\Delta}_x f_x(\bar{W})$. Write, by (2.2),

$$\begin{aligned}
& P(\bar{W}+\bar{\Delta}_x \le x)-\Phi(x)+E\bar{\Delta}_x f_x(\bar{W}) \\
= \quad & Ef_x'(\bar{W}+\bar{\Delta}_x)-E(\bar{W}+\bar{\Delta}_x)f_x(\bar{W}+\bar{\Delta}_x)+E\bar{\Delta}_x f_x(\bar{W}+\bar{\Delta}_x) \\
& -E\bar{\Delta}_x(f_x(\bar{W}+\bar{\Delta}_x)-f_x(\bar{W})) \\
= \quad & Ef_x'(\bar{W}+\bar{\Delta}_x)-E\bar{W}f_x(\bar{W}+\bar{\Delta}_x)-E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f_x'(\bar{W}+t)dt \\
= \quad & Ef_x'(\bar{W}+\bar{\Delta}_x)-\sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x)-f_x(\bar{W}+\bar{\Delta}_x^{(i)})] \\
& -\sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x^{(i)})-f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)})] \\
& +\sum_{i=1}^n E(\xi_i-\bar{\xi}_i)f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)})-E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f_x'(\bar{W}+t)dt \\
= \quad & \sum_{i=1}^n E\int_{-1}^1 f_x'(\bar{W}+\bar{\Delta}_x)\bar{K}_i(t)dt+\beta_2 Ef_x'(\bar{W}+\bar{\Delta}_x) \\
& -\sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x^{(i)})-f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)})] \\
& -\sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x)-f_x(\bar{W}+\bar{\Delta}_x^{(i)})]+\sum_{i=1}^n E(\xi_i-\bar{\xi}_i)f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)}) \\
& -E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f_x'(\bar{W}+t)dt
\end{aligned}$$

Then we have

$$\begin{aligned}
& |P(\bar{W}+\bar{\Delta}_x \le x)-\Phi(x)+E\bar{\Delta}_x f_x(\bar{W})| \\
\le \quad & |R_1|+|R_2|+|R_3|+|R_4|
\end{aligned}$$

where

$$\begin{aligned}
R_1 &= \sum_{i=1}^n E\int_{-1}^1 f_x'(\bar{W}+\bar{\Delta}_x)\bar{K}_i(t)dt-\sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x^{(i)})-f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)})] \\
R_2 &= \sum_{i=1}^n E\bar{\xi}_i[f_x(\bar{W}+\bar{\Delta}_x)-f_x(\bar{W}+\bar{\Delta}_x^{(i)})] \\
R_3 &= \sum_{i=1}^n E\xi_i^2 I(|\xi_i|>1)Ef_x'(\bar{W}+\bar{\Delta}_x)+\sum_{i=1}^n E(\xi_i-\bar{\xi}_i)f_x(\bar{W}^{(i)}+\bar{\Delta}_x^{(i)}) \\
R_4 &= E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f_x'(\bar{W}+t)dt.
\end{aligned}$$

The estimations of each term are given in the following lemma.

Lemma 4.2. *There exists a universal constant $C > 0$, such that, for any $p \geq 1$ and $q = q(p)$ satisfying $1/p + 1/q = 1$,*

$$\begin{aligned}
|R_1| &\leq Ce^{-x/40}\Big\{\sum_{i=1}^n E|\xi_i|^r + \sum_{i=1}^n \|\xi_i\|_p\big(\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q\big)\Big\},\\
|R_2| &\leq Ce^{-x/8}\sum_{i=1}^n \|\xi_i\|_p\big(\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q\big),\\
|R_3| &\leq Ce^{-x/4}\sum_{i=1}^n E|\xi_i|^r
\end{aligned}$$

and

$$|R_4| \leq C\Big(E\bar{\Delta}_1^2 + \frac{1}{1+x}E\bar{\Delta}_2^2\Big) + Ce^{-x}E(\bar{\Delta}_2^2 e^{\bar{W}}).$$

For the second part, we will estimate the error term. We will state the result in the following lemma and give the proof in Section 4.3.

Lemma 4.3. *We have*

$$\begin{aligned}
&P(x(1+\frac{1}{2}\Delta_2 - \frac{1}{2}\Delta_2^2) \leq W + \Delta_1 \leq x(1+\frac{1}{2}\Delta_2))\\
&\leq Ce^{-\frac{17}{32}x}[\sum_{i=1}^n E|\xi_i|^r + E(e^{\bar{W}}\bar{\Delta}_2^2) + \sum_{i=1}^n \|\bar{\xi}_i\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \sum_{i=1}^n \|\bar{\xi}_i\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q]\\
&\quad + P(|\Delta_1| \geq \frac{1}{4}) + P(|\Delta_2| \geq \frac{1}{2}) + \sum_{i=1}^n P(W^{(i)} \geq \frac{x-1}{2}, |\xi_i| > 1) + \sum_{i=1}^n P(|\xi_i| \geq \frac{x+1}{4}).
\end{aligned}$$

Hence, from Lemma 4.1, Lemma 4.2 and Lemma 4.3, we get the result

$$\begin{aligned}
&|P(\frac{W+\Delta_1}{\sqrt{1+\Delta_2}} \leq x) - \Phi(x)|\\
&\leq C\Big(|E(\Delta_x f_x(W))| + (E(\Delta_1^2 \wedge 1) + \frac{1}{1+x}E(\Delta_2^2 \wedge 1)) + Ce^{-x}E(\Delta_2^2 e^{\bar{W}})\\
&\quad + \frac{1}{1+|x|^r}\sum_{i=1}^n E|\xi_i|^r + e^{-x/40}(\sum_{j=1}^n \|\xi_j\|_p\|\Delta_1 - \Delta_1^{(j)}\|_q + \sum_{j=1}^n \|\xi_j\|_p\|\Delta_2 - \Delta_2^{(j)}\|_q)\Big).
\end{aligned}$$

The proof of Theorem 3.1 is then complete. Next we prove the corollary.

Proof of Corollary 3.2 Recall that for $1 \le i \le n$, ξ_i has mean zero, variance $1/n$ and $E|\xi_i|^r < \infty$ with some $2 < r \le 3$. As we assumed

$$\Delta_2 = \sum_{i=1}^n \xi_i^2 - 1 + \Delta_3 = \sum_{i=1}^n (\xi_i^2 - E\xi_i^2) + \Delta_3,$$

where $\Delta_3 \to 0$ in probability. Noting that $|\bar{\Delta}_2| \le 1$, we can also assume $|\bar{\Delta}_3| \le 1$. Now we calculate each term.

$$\begin{aligned} & E\bar{\Delta}_2^2 \\ &= E(\sum_{i=1}^n (\xi_i^2 - E\xi_i^2) + \Delta_3)^2 \\ &\le C\sum_{i=1}^n E|\sum_{i=1}^n (\xi_i^2 - E\xi_i^2)|^2 + CE\bar{\Delta}_3^2 \\ &\le C\sum_{i=1}^n E|\xi_i|^r + CE\bar{\Delta}_3^2, \end{aligned} \tag{4.3}$$

where $|\xi_i| \le 1$ gives us $E|\xi_i|^4 \le E|\xi_i|^r$.

$$\begin{aligned} & Ee^{\bar{W}}\bar{\Delta}_2^2 \\ &\le Ee^{\bar{W}}(\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2) + \bar{\Delta}_3)^2 \\ &\le Ee^{\bar{W}}[\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2)^2 + \sum_{i \ne j} (\bar{\xi}_i^2 - E\bar{\xi}_i^2)(\bar{\xi}_j^2 - E\bar{\xi}_j^2)] + Ee^{\bar{W}}\bar{\Delta}_3^2 \\ &= \sum_{i=1}^n Ee^{\bar{W}^{(i)}} Ee^{\bar{\xi}_i}(\bar{\xi}_i^2 - E\bar{\xi}_i^2)^2 + \sum_{i \ne j} Ee^{\bar{W}^{(i,j)}} Ee^{\bar{\xi}_i}(\bar{\xi}_i^2 - E\bar{\xi}_i^2) Ee^{\bar{\xi}_j}(\bar{\xi}_j^2 - E\bar{\xi}_j^2) + Ee^{\bar{W}}\bar{\Delta}_3^2 \\ &\le C\sum_{i=1}^n E|\xi_i|^r + Ee^{\bar{W}}\bar{\Delta}_3^2. \end{aligned} \tag{4.4}$$

$$\begin{aligned} & \sum_{i=1}^n \|\bar{\xi}_i\|_p \|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q \\ &= \sum_{i=1}^n \|\bar{\xi}_i\|_p \|\xi_j^2 - E\xi_j^2 + \bar{\Delta}_3 - \bar{\Delta}_3^{(j)}\|_q \\ &\le \sum_{i=1}^n E|\xi_i|^r + \sum_{i=1}^n \|\bar{\xi}_i\|_p \|\bar{\Delta}_3 - \bar{\Delta}_3^{(j)}\|_q. \end{aligned} \tag{4.5}$$

$$
\begin{aligned}
&P(|\Delta_2| \geq 1/4) \\
&\leq \; P(|\Delta_3| \geq 1/8) + P(|\sum_{i=1}^n (\xi_i^2 - E\xi_i^2)| \geq 1/8) \\
&\leq \; P(|\Delta_3| \geq 1/8) + C\sum_{i=1}^n E|\xi_i^2|^{r/2}, \\
&\leq \; E|\Delta_3|^2 + C\sum_{i=1}^n E|\xi_i|^r. \qquad (4.6)
\end{aligned}
$$

And

$$
|E\bar{\Delta} f_x(\bar{W})| \;\leq\; |E\bar{\Delta}_1 f_x(\bar{W})| + \frac{x}{2}|E\bar{\Delta}_3 f_x(\bar{W})| + \frac{x}{2}|E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2) f_x(\bar{W})|,
$$

where

$$
\begin{aligned}
&E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2) f_x(\bar{W}) \\
&= \; E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2)[f(\bar{W}) - f(\bar{W}^{(i)})] \\
&= \; E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2)[\int_0^{\bar{\xi}_i} f'(\bar{W}^{(i)} + t)dt] \\
&= \; E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2)[\int_{-1}^1 (I(0 \leq t \leq \bar{\xi}_i) - I(\bar{\xi}_i \leq t \leq 0)) f'(\bar{W}^{(i)} + t)dt] \\
&= \; \sum_{i=1}^n E\int_{-1}^1 f'(\bar{W}^{(i)} + t)(\bar{\xi}_i^2 - E\bar{\xi}_i^2)(I(0 \leq t \leq \bar{\xi}_i) - I(\bar{\xi}_i \leq t \leq 0))dt.
\end{aligned}
$$

Then for $2 < r \leq 3$, we get

$$
\begin{aligned}
&\frac{x}{2}|E\sum_{i=1}^n (\bar{\xi}_i^2 - E\bar{\xi}_i^2) f_x(\bar{W})| \\
&\leq \; \frac{x}{2}\sum_{i=1}^n \int_{-1}^1 |Ef'(\bar{W}^{(i)} + t)||E(\bar{\xi}_i^2 - E\bar{\xi}_i^2)(I(0 \leq t \leq \bar{\xi}_i) - I(\bar{\xi}_i \leq t \leq 0))|dt \\
&\leq \; Ce^{-x/8}\sum_{i=1}^n E|\xi_i|^r.
\end{aligned}
$$

Therefore we get

$$\begin{aligned}
&|P(\frac{W+\Delta_1}{\sqrt{1+\Delta_2}} \le x) - \Phi(x)| \\
\le\ & CE(\Delta_1^2 \wedge 1) + \frac{C}{1+x}E(\Delta_3^2 \wedge 1) + Ce^{-x/2}E(\Delta_3^2 e^{\bar{W}}) + \frac{C}{1+|x|^r}\sum_{i=1}^{n} E|\xi_i|^r \\
& + Ce^{-x/40}\sum_{j=1}^{n}(\|\xi_j\|_p\|\Delta_1 - \Delta_1^{(j)}\|_q + \|\xi_j\|_p\|\Delta_3 - \Delta_3^{(j)}\|_q) \\
& + C|E\Delta_1 f_x(W)| + Cx|E\Delta_3 f_x(W)|.
\end{aligned}$$

The proof of Corollary 3.2 is then complete. For the reminder of this section, we shall give the proof of each lemma.

4.3 Proof of Lemmas

First, we state two properties we will use in the proof of the lemmas.

Lemma 4.4. *For $\bar{W} = \sum_{i=1}^{n} \bar{\xi}_i$, we have*

$$Ee^{\bar{W}} \le e^2. \tag{4.7}$$

Proof of Lemma 4.4 To prove this lemma, we need to use Bennett-Hoeffding Inequality (cf. Lemma 6.2 in [22]), which is as follows. Assume that $EX_i \le 0$, $X_i \le a$ for each $1 \le i \le n$, and $\sum_{i=1}^{n} EX_i^2 \le B_n^2$. Then for $t > 0$,

$$Ee^{tS_n} \le \exp\left(a^{-2}(e^{ta} - 1 - ta)B_n^2\right). \tag{4.8}$$

Now let $a = 1$, $B_n^2 = 1$, $t = 1$, then it follows from (4.8) that

$$\begin{aligned}
Ee^{\bar{W}} &= Ee^{(\bar{W} - E\bar{W})}e^{E\bar{W}} \\
&= e^{E\bar{W}} E\exp\{\sum_{i=1}^{n}(\bar{\xi}_i - E\bar{\xi}_i)\} \\
&\le e^2.
\end{aligned}$$

Proposition 4.5. *For $\bar{W} = \sum_{i=1}^n \bar{\xi}_i$, by using Lemma 5.2 in Chen and Shao(2007), it follows that*

$$\begin{aligned} & P(\Delta_1 \le \bar{W} \le \Delta_2, \Delta_1 \ge a) \\ \le\ & e^{-a/2}\{7.5\|\Delta_2 - \Delta_1\|_q + 10(\beta_2 + \beta_3) + 4.2\sum_{l=1}^{2}\sum_{i=1}^{n} \|\xi_i\|_p \|\Delta_l - \Delta_{l,i}\|_q\} \end{aligned}$$

where a is a real number, $\Delta_{l,i}$ and $\Delta_{2,i}$ are Borel measurable functions of $(\xi_j, 1 \le j \le n, j \ne i)$, $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Next we prove the lemmas.

Proof of Lemma 4.1 From the definitions (4.1) and (4.2), we get

$$\begin{aligned} & P(W + \Delta_x \ge x) \\ =\ & P(W + \Delta_x \ge x, \max|\xi_i| \le 1) + P(W + \Delta_x \ge x, \max|\xi_i| > 1) \\ \le\ & P(\bar{W} + \Delta_x^* \ge x, |\Delta_x^*| \le \frac{x+1}{4}) + P(\bar{W} + \Delta_x^* \ge x, |\Delta_x^*| \ge \frac{x+1}{4}, \max|\xi_i| \le 1) \\ & + P(W + \Delta_x \ge x, |\Delta_x| \le \frac{x+1}{4}, \max|\xi_i| > 1) \\ & + P(W + \Delta_x \ge x, |\Delta_x| \ge \frac{x+1}{4}, \max|\xi_i| > 1) \\ \le\ & P(\bar{W} + \bar{\Delta}_x \ge x) + P(|\Delta_x| \ge \frac{x+1}{4}) + P(W \ge \frac{3x-1}{4}, \max|\xi_i| > 1) \\ \le\ & P(\bar{W} + \bar{\Delta}_x \ge x) + P(|\Delta_x| \ge \frac{x+1}{4}) + \sum_{i=1}^{n} P(W \ge \frac{3x-1}{4}, |\xi_i| > 1) \\ \le\ & P(\bar{W} + \bar{\Delta}_x \ge x) + P(|\Delta_x| \ge \frac{x+1}{4}) + \sum_{i=1}^{n} P(W \ge \frac{3x-1}{4}, 1 < |\xi_i| < \frac{x+1}{4}) \\ & + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}) \\ \le\ & P(\bar{W} + \bar{\Delta}_x \ge x) + P(|\Delta_x| \ge \frac{x+1}{4}) + \sum_{i=1}^{n} P(W - \xi_i \ge \frac{x-1}{2}, |\xi_i| > 1) \\ & + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}) \\ \le\ & P(\bar{W} + \bar{\Delta}_x \ge x) + P(|\Delta_x| \ge \frac{x+1}{4}) + \sum_{i=1}^{n} P(W^{(i)} \ge \frac{x-1}{2}) P(|\xi_i| > 1) \\ & + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}), \end{aligned}$$

which gives us

$$
\begin{aligned}
&|P(W+\Delta_x \le x) - P(\bar{W}+\bar{\Delta}_x \le x)| \\
&= |P(W+\Delta_x \ge x) - P(\bar{W}+\bar{\Delta}_x \ge x)| \\
&\le P(|\Delta| \ge \frac{x+1}{4}) + \sum_{i=1}^{n} P(W^{(i)} \ge \frac{x-1}{2})P(|\xi_i|>1) + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}).
\end{aligned}
$$

Therefore we get

$$
\begin{aligned}
&|P(W+\Delta_x \le x) - \Phi(x) + E\Delta_x f(W)| \\
&\le |P(\bar{W}+\bar{\Delta}_x \le x) - \Phi(x) + E\Delta_x f(W)| + |P(W+\Delta_x \le x) - P(\bar{W}+\bar{\Delta}_x \le x)| \\
&\le |P(\bar{W}+\bar{\Delta}_x \le x) - \Phi(x) + E\Delta_x f(W)| + P(|\Delta_x| \ge \frac{x+1}{4}) \\
&\quad + \sum_{i=1}^{n} P(|W^{(i)}| \ge \frac{x-1}{2})P(|\xi_i|>1) + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}).
\end{aligned}
$$

For $2 < r \le 3$, using Chebyshev's inequality and Rosenthal inequality, we have

$$
\begin{aligned}
&\sum_{i=1}^{n} P(|W^{(i)}| \ge \frac{x-1}{2})P(|\xi_i|>1) + \sum_{i=1}^{n} P(|\xi_i| \ge \frac{x+1}{4}) \\
&\le \sum_{i=1}^{n} \frac{E(|W^{(i)}|+1)^r}{(\frac{x+1}{2})^r} P(|\xi_i|>1) + \sum_{i=1}^{n} \frac{E|\xi_i|^r}{(\frac{x+1}{4})^r} \\
&\le \sum_{i=1}^{n} \frac{C[(\sum_{j\ne i} E|\xi_j|^2)^r + \sum_{j\ne i} E|\xi_j|^r]}{(\frac{x+1}{2})^r} P(|\xi_i|>1) + \sum_{i=1}^{n} \frac{E|\xi_i|^r}{(\frac{x+1}{4})^r} \\
&\le \frac{C\sum_{j=1}^{n}(1+E|\xi_j|^r)}{1+|x|^r} \sum_{i=1}^{n} P(|\xi_i|>1) + \frac{C\sum_{i=1}^{n} E|\xi_i|^r}{1+|x|^r} \\
&\le \frac{C\sum_{j=1}^{n} E|\xi_j|^r}{1+|x|^r} \sum_{i=1}^{n} E|\xi_i|^r + \frac{C\sum_{i=1}^{n} E|\xi_i|^r}{1+|x|^r} \\
&\le C\sum_{i=1}^{n} \frac{E|\xi_i|^r}{1+|x|^r}. \qquad (4.9)
\end{aligned}
$$

Noting that $\Delta_x = \Delta_1 - \frac{x}{2}\Delta_2$, we have

$$
\begin{aligned}
P(|\Delta_x| \ge \frac{x+1}{4}) &\le P(|\Delta_1 - \frac{x}{2}\Delta_2| \ge \frac{x+1}{4}) \\
&\le P(|\Delta_1| \ge \frac{1}{4}) + P(|\frac{1}{2}x\Delta_2| \ge \frac{x}{4}) \\
&\le P(|\Delta_1| \ge \frac{1}{4}) + P(|\Delta_2| \ge \frac{1}{2}). \qquad (4.10)
\end{aligned}
$$

The proof of Lemma 4.1 is complete.

Proof of Lemma 4.2 For the first part, we prove the bound of R_1. Let $G(w) = wf(w)$ and $g(w) = (wf(w))'$, we get

$$\begin{aligned}
& f'(\bar{W}+\bar{\Delta}) - f'(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t) \\
= \; & G(\bar{W}+\bar{\Delta}) - G(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t) + I(\bar{W}+\bar{\Delta}\le x) - I(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t\le x) \\
= \; & \int_t^{\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)du + I(\bar{W}+\bar{\Delta}\le x) - I(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t\le x).
\end{aligned}$$

Then recalling $\bar{K}_i(t) = E\{\bar{\xi}_i I(0\le t\le \bar{\xi}_i) - \bar{\xi}_i I(\bar{\xi}_i \le t \le 0)\}$ in (4.1), we can get

$$\begin{aligned}
R_1 \; = \; & \sum_{i=1}^n E\int_{-1}^1 f'(\bar{W}+\bar{\Delta})\bar{K}_i(t)dt - \sum_{i=1}^n E\bar{\xi}_i[f(\bar{W}+\bar{\Delta}^{(i)}) - f(\bar{W}^{(i)}+\bar{\Delta}^{(i)})] \\
= \; & \sum_{i=1}^n E\int_{-1}^1 [(f'(\bar{W}+\bar{\Delta}) - f'(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t)]\bar{K}_i(t)dt \\
= \; & \sum_{i=1}^n \int_{-1}^1 E(\int_t^{\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)du)\bar{K}_i(t)dt \\
& + \sum_{i=1}^n E\int_{-1}^1 (I(\bar{W}+\bar{\Delta}\le x) - I(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t\le x))\bar{K}_i(t)dt \\
= \; & \sum_{i=1}^n \int_{-1}^1 E(\int_t^{\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)du)\bar{K}_i(t)dt \\
& + \sum_{i=1}^n \int_{-1}^1 \{P(\bar{W}+\bar{\Delta}\le x) - P(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t\le x)\}\bar{K}_i(t)dt \\
= \; & R_{1,1} + R_{1,2},
\end{aligned}$$

where

$$\begin{aligned}
R_{1,1} \; &= \; \sum_{i=1}^n \int_{-1}^1 E(\int_t^{\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)du)\bar{K}_i(t)dt, \\
R_{1,2} \; &= \; \sum_{i=1}^n \int_{-1}^1 \{P(\bar{W}+\bar{\Delta}\le x) - P(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+t\le x)\}\bar{K}_i(t)dt.
\end{aligned}$$

It suffices to show that for some $2 < r \le 3$,

$$\begin{aligned}
& |R_{1,1}| \\
& \quad \le \; Ce^{-x/40}(\sum_{i=1}^n E|\xi_i|^r + \sum_{i=1}^n \|\bar{\xi}_i\|_p(\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q)), \quad (4.11) \\
& |R_{1,2}| \\
& \quad \le \; Ce^{-x/40}(\sum_{i=1}^n E|\xi_i|^r + \sum_{i=1}^n \|\xi_i\|_p(\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q)), \quad (4.12)
\end{aligned}$$

Proof of (4.11) We first write

$$\begin{aligned}
R_{1,1} &= \sum_{i=1}^{n}\int_{-1}^{1} E\int_{t}^{\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)du\bar{K}_i(t)dt \\
&= \sum_{i=1}^{n}\int_{-1}^{1} E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(t<u<\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)})du\bar{K}_i(t)dt \\
&\quad -\sum_{i=1}^{n}\int_{-1}^{1} E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}<u<t)du\bar{K}_i(t)dt \\
&:= \sum_{i=1}^{n}\int_{-1}^{1} E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_1<u<\eta_2)du\bar{K}_i(t)dt \\
&\quad -\sum_{i=1}^{n}\int_{-1}^{1} E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_2<u<\eta_1)du\bar{K}_i(t)dt,
\end{aligned}$$

where we let $\eta_1=t$ and $\eta_2=\bar{\xi}_i+\bar{\Delta}-\bar{\Delta}^{(i)}$. Then we have

$$\begin{aligned}
|R_{1,1}| &\leq \sum_{i=1}^{n}\int_{-\infty}^{\infty}|E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_1<u<\eta_2)du|\bar{K}_i(t)dt \\
&\quad +\sum_{i=1}^{n}\int_{-\infty}^{\infty}|E\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_2<u<\eta_1)du|\bar{K}_i(t)dt.
\end{aligned}$$

Claim that for $t\leq 1$ and $\bar{\xi}_i\leq 1$, we have

$$\begin{aligned}
&E(\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_1<u<\eta_2)du|\bar{\xi}_i) \\
&\leq Ce^{-x/40}(\|\bar{\xi}_i\|_q+\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q+|t|). \qquad (4.13)
\end{aligned}$$

And similarly, we have

$$\begin{aligned}
&E(\int_{-\infty}^{\infty} g(\bar{W}^{(i)}+\bar{\Delta}^{(i)}+u)I(\eta_2<u<\eta_1)du|\bar{\xi}_i) \\
&\leq Ce^{-x/40}(\|\bar{\xi}_i\|_q+\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q+|t|). \qquad (4.14)
\end{aligned}$$

Then note that $|\bar{\xi}_i|\leq 1$,we get the estimation of $R_{1,1}$.

$$\begin{aligned}
|R_{1,1}| &\leq Ce^{-x/40}\sum_{i=1}^{n}\int_{-1}^{1}(|t|\wedge 1+\|\bar{\xi}_i\|_q+\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q)\bar{K}_i(t)dt \\
&\leq Ce^{-x/40}(\sum_{i=1}^{n}E|\bar{\xi}_i|^r+\sum_{i=1}^{n}\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q E\bar{\xi}_i^2) \\
&\leq Ce^{-x/40}(\sum_{i=1}^{n}E|\bar{\xi}_i|^r+\sum_{i=1}^{n}\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q\|\bar{\xi}_i\|_p) \\
&\leq Ce^{-x/40}(\sum_{i=1}^{n}E|\xi_i|^r+\sum_{i=1}^{n}(\|\bar{\xi}_i\|_p\|\bar{\Delta}_1-\bar{\Delta}_1^{(i)}\|_q+\|\bar{\xi}_i\|_p\|\bar{\Delta}_2-\bar{\Delta}_2^{(i)}\|_q)),
\end{aligned}$$

where $2 < r \leq 3$.

Proof of (4.13) and (4.14). By Lemma 2.3 and Lemma 4.4, we get

$$
\begin{aligned}
& E\int_{-\infty}^{\infty} g(\bar W^{(i)}+\bar\Delta^{(i)}+u)I(\eta_1<u<\eta_2)du \\
\leq\ & E\int_{-\infty}^{\infty} g(\bar W^{(i)}+\bar\Delta^{(i)}+u)I(\eta_1<u<\eta_2)I(\bar W^{(i)}+\bar\Delta^{(i)}+u\leq 11x/12)du \\
& +CxE\int_{-\infty}^{\infty} I(\eta_1<u<\eta_2)I(\bar W^{(i)}+\bar\Delta^{(i)}+u>11x/12, u\leq 7x/12)du \\
& +CxE\int_{-\infty}^{\infty} I(\eta_1<u<\eta_2)I(\bar W^{(i)}+\bar\Delta^{(i)}+u>11x/12, u> 7x/12)du \ \ \text{(by (2.21))} \\
\leq\ & Ce^{-x/2}E|\eta_2-\eta_1| + CxE[I(\bar W^{(i)}+\bar\Delta^{(i)}>x/3)|\eta_2-\eta_1|] \\
& +CxE\int_{-\infty}^{\infty} I(\eta_1<u<\eta_2)I(u>7x/12)du \\
\leq\ & Ce^{-x/2}E|\eta_2-\eta_1| + Cx[P(\bar W^{(i)}+\bar\Delta^{(i)}>x/3)]^{1/p}\|\eta_2-\eta_1\|_q \\
& +Cx[P(\eta_2>7x/12)]^{1/p}\|\eta_2-\eta_1\|_q \\
\leq\ & Ce^{-x/2}E|\eta_2-\eta_1| + [Cx(e^{-x/3}Ee^{\bar W^{(i)}+\bar\Delta^{(i)}})^{1/p} + Cx(e^{-7x/12}Ee^{\bar\xi_i+\bar\Delta-\bar\Delta^{(i)}})^{1/p}]\|\eta_2-\eta_1\|_q \\
\leq\ & Ce^{-x/2}E|\eta_2-\eta_1| + Cx(e^{-x/12})^{1/p}\|\eta_2-\eta_1\|_q \\
\leq\ & Cx(e^{-x/12})^{1/p}\|\bar\xi_i+\bar\Delta-\bar\Delta^{(i)}-t\|_q \\
\leq\ & e^{-x/40}(\|\bar\xi_i\|_q+\|\bar\Delta-\bar\Delta^{(i)}\|_q+|t|),
\end{aligned}
$$

which gives (4.13). The proof of (4.14) is similar and omitted here.

Proof of (4.12) Note that

$$
\begin{aligned}
& |P(\bar W+\bar\Delta\leq x)-P(\bar W^{(i)}+\bar\Delta^{(i)}+t\leq x)| \\
\leq\ & P(x-\bar\Delta\leq \bar W\leq x-\bar\Delta^{(i)}+t-\bar\xi_i)+P(x-\bar\Delta^{(i)}-t+\bar\xi_i\leq \bar W\leq x-\bar\Delta),
\end{aligned}
$$

where for $|t|\leq 1$,

$$x-\bar\Delta\geq x-\frac{x+1}{4}=\frac{3x}{4}-\frac{1}{4},\ \text{and}\ x-\bar\Delta^{(i)}-t+\bar\xi_i\geq x-\frac{x+1}{4}-1-1=\frac{3x}{4}-\frac{9}{4}.$$

Then using Proposition 4.5, it follows that

$$\begin{aligned}
&P(x - \bar{\Delta} \le \bar{W} \le x - \bar{\Delta}^{(i)} + t - \bar{\xi}_i) \\
\le\ & Ce^{-\frac{3x}{8}}\{7.5\|\bar{\xi}_i\|_2 + 7.5|t| + 10(\beta_2 + \beta_3) + 8.4\sum_{j=1}^{n}\|\xi_j\|_p\|\bar{\Delta}^{(j)} - \bar{\Delta}\|_q\},
\end{aligned}$$

and

$$\begin{aligned}
&P(x - \bar{\Delta}^{(i)} - t + \bar{\xi}_i \le \bar{W} \le x - \bar{\Delta}) \\
\le\ & Ce^{-\frac{3x}{8}}\{7.5\|\bar{\xi}_i\|_2 + 7.5|t| + 10(\beta_2 + \beta_3) + 8.4\sum_{j=1}^{n}\|\xi_j\|_p\|\bar{\Delta}^{(j)} - \bar{\Delta}\|_q\}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&|R_{1,2}| \\
\le\ & \sum_{i=1}^{n}\int_{-1}^{1}|P(\bar{W} + \bar{\Delta} \le x) - P(\bar{W}^{(i)} + \bar{\Delta}^{(i)} + t \le x)|\bar{K}_i(t)dt \\
\le\ & Ce^{-\frac{3x}{8}}\beta_3 + Ce^{-\frac{3x}{8}}\sum_{j=1}^{n}\|\xi_j\|_p\|\bar{\Delta}^{(j)} - \bar{\Delta}\|_q + Ce^{-\frac{3x}{8}}(\beta_2 + \beta_3) + Ce^{-\frac{3x}{8}}\sum_{i=1}^{n}\int_{-1}^{1}|t|\bar{K}_i(t)dt \\
\le\ & Ce^{-\frac{3x}{8}}(\sum_{j=1}^{n}\|\xi_j\|_p\|\bar{\Delta}^{(j)} - \bar{\Delta}\|_q + \beta_2 + \beta_3) \\
\le\ & Ce^{-\frac{3x}{8}}(\sum_{i=1}^{n}E|\xi_i|^r + \sum_{i=1}^{n}(\|\xi_i\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \|\xi_i\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q)),
\end{aligned}$$

where $2 < r \le 3$.

For the second part, we prove the bound of R_2. Using Lemma 2.4, we get

$$|f(\bar{W} + \bar{\Delta}) - f(\bar{W} + \bar{\Delta}^{(i)})| \le C|\bar{\Delta} - \bar{\Delta}^{(i)}|e^{-x^2/4} + 4|\bar{\Delta} - \bar{\Delta}^{(i)}|I(\bar{W} \ge x/4).$$

Then it follows that

$$\begin{aligned}
&E|\bar{\xi}_i||f(\bar{W} + \bar{\Delta}) - f(\bar{W} + \bar{\Delta}^{(i)})| \\
\le\ & Ce^{-x^2/4}E|\bar{\xi}_i||\bar{\Delta} - \bar{\Delta}^{(i)}| + 4E|\bar{\xi}_i||\bar{\Delta} - \bar{\Delta}^{(i)}|I(\bar{W} \ge x/4) \\
\le\ & Ce^{-x^2/4}\|\bar{\xi}_i\|_p\|\bar{\Delta} - \bar{\Delta}^{(i)}\|_q + 4\|\bar{\xi}_i I(\bar{W} \ge x/4)\|_p\|\bar{\Delta} - \bar{\Delta}^{(i)}\|_q,
\end{aligned}$$

where $p \geq 1, 1/p + 1/q = 1$ and by Lemma 4.4, we get

$$\begin{aligned}
E|\bar{\xi}_i|^p I(\bar{W} \geq x/4) &\leq e^{-x/4} E|\bar{\xi}_i|^p e^{\bar{W}} \\
&= e^{-x/4} E|\bar{\xi}_i|^p e^{\bar{\xi}_i} E e^{\bar{W}-\bar{\xi}_i} \\
&\leq e^{e-2+1} e^{-x/4} E|\bar{\xi}_i|^p \\
&\leq C e^{-x/4} E|\bar{\xi}_i|^p.
\end{aligned}$$

Then we have

$$\begin{aligned}
&E|\bar{\xi}_i||f(\bar{W}+\bar{\Delta}) - f(\bar{W}+\bar{\Delta}^{(i)})| \\
&\leq Ce^{-x^2/4}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q + 4(Ce^{-x/4}E|\bar{\xi}_i|^p)^{1/p}\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q \\
&\leq Ce^{-x^2/4}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q + Ce^{-x/4}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q,
\end{aligned}$$

Now we get the estimation

$$\begin{aligned}
|R_2| &\leq \sum_{i=1}^{n} E|\bar{\xi}_i||f(\bar{W}+\bar{\Delta}) - f(\bar{W}+\bar{\Delta}^{(i)})| \\
&\leq \sum_{i=1}^{n}(Ce^{-x^2/4}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q + Ce^{-x/4}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q) \\
&\leq Ce^{-x/4}\sum_{i=1}^{n}\|\bar{\xi}_i\|_p\|\bar{\Delta}-\bar{\Delta}^{(i)}\|_q \\
&\leq Ce^{-x/8}\sum_{i=1}^{n}(\|\xi_i\|_p\|\bar{\Delta}_1-\bar{\Delta}_1^{(i)}\|_q + \|\xi_i\|_p\|\bar{\Delta}_2-\bar{\Delta}_2^{(i)}\|_q)).
\end{aligned}$$

For the third part, we prove the bound of R_3. First we claim that

$$E|f_x'(\bar{W}+\bar{\Delta})| \leq Ce^{-x/4}, \tag{4.15}$$

$$E|f_x(\bar{W}^{(i)}+\bar{\Delta}^{(i)})| \leq Ce^{-x/4}. \tag{4.16}$$

These can be proved by Lemma, Lemma 4.4, $|\bar{\Delta}| \leq 1/4$ and $|\bar{\Delta}^{(i)}| \leq 1/4$.

$$\begin{aligned}
&E|f_x'(\bar{W}+\bar{\Delta})| \\
&= E|f_x'(\bar{W}+\bar{\Delta})|I(\bar{W}+\bar{\Delta} \leq x/2) + E|f_x'(\bar{W}+\bar{\Delta})|I(\bar{W}+\bar{\Delta} \geq x/2) \\
&\leq Ce^{-x/2} + P(\bar{W}+\bar{\Delta} \geq x/2) \\
&\leq Ce^{-x/2} + e^{-x/2}Ee^{\bar{W}+\bar{\Delta}} \\
&\leq Ce^{-x/4}.
\end{aligned}$$

$$
\begin{aligned}
&E|f_x(\bar{W}^{(i)}+\bar{\Delta}^{(i)})| \\
&= E|f_x(\bar{W}^{(i)}+\bar{\Delta}^{(i)})|I(\bar{W}^{(i)}+\bar{\Delta}^{(i)}\le x/2)+E|f_x(\bar{W}^{(i)}+\bar{\Delta}^{(i)})|I(\bar{W}^{(i)}+\bar{\Delta}^{(i)}\ge x/2) \\
&\le Ce^{-x/2}+P(\bar{W}^{(i)}+\bar{\Delta}^{(i)}\ge x/2) \\
&\le Ce^{-x/2}+e^{-x/2}Ee^{\bar{W}^{(i)}+\bar{\Delta}^{(i)}} \\
&\le Ce^{-x/4}.
\end{aligned}
$$

Therefore we get

$$
\begin{aligned}
|R_3| &\le |\sum_{i=1}^n E\xi_i^2 I(|\xi_i|>1)Ef'(\bar{W}+\bar{\Delta})|+|\sum_{i=1}^n E(\xi_i-\bar{\xi}_i)f(\bar{W}^{(i)}+\bar{\Delta}^{(i)})| \\
&\le |\sum_{i=1}^n E\xi_i^2 I(|\xi_i|>1)Ef'(\bar{W}+\bar{\Delta})|+|\sum_{i=1}^n E\xi_i^2 I(|\xi_i|\ge 1)Ef(\bar{W}^{(i)}+\bar{\Delta}^{(i)})| \\
&\le \sum_{i=1}^n E|\xi_i|^r E|f_x'(\bar{W}+\bar{\Delta})|+\sum_{i=1}^n E|\xi_i|^r E|f_x(\bar{W}+\bar{\Delta})| \\
&\le Ce^{-x/4}\sum_{i=1}^n E|\xi_i|^r,
\end{aligned}
$$

where $2<r\le 3$.

For the last part, we prove the bound of R_4. Noting the explicit form of f in (2.16), we get $f'(w)\le\frac{1}{1+w^3}$ when $w>x$, and $f'(w)\le e^{-x/2}$ when $w<x/2$. Here $\Delta_x=\Delta_1-\frac{x}{2}\Delta_2$ and $|\bar{\Delta}_x|\le\frac{x+1}{4}$, we have

$$
\begin{aligned}
&|E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f'(\bar{W}+t)dt| \\
&= |E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f'(\bar{W}+t)I(\bar{W}<x/2)dt+E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f'(\bar{W}+t)I(x/2<\bar{W}<2x)dt \\
&\quad +E\bar{\Delta}_x\int_0^{\bar{\Delta}_x} f'(\bar{W}+t)I(\bar{W}>2x)dt| \\
&\le \frac{1}{1+x^3}E|\bar{\Delta}_x|^2+E|\bar{\Delta}_x|^2 I(x/2<\bar{W}<2x)+e^{-x/2}E|\bar{\Delta}_x|^2 \\
&\le 2(\frac{1}{1+x^3}+e^{-x/2})(E\bar{\Delta}_1^2+\frac{x^2}{4}E\bar{\Delta}_2^2)+2E\bar{\Delta}_1^2+\frac{x^2}{2}E(\bar{\Delta}_2^2 I(x/2<\bar{W}<2x) \\
&\le 2(\frac{1}{1+x^3}+e^{-x/2})(E\bar{\Delta}_1^2+\frac{x^2}{4}E\bar{\Delta}_2^2)+2E\bar{\Delta}_1^2+Cx^2E(\frac{\bar{\Delta}_2^2 e^{\bar{W}}}{e^x}) \\
&\le C(E\bar{\Delta}_1^2+\frac{1}{1+x}E\bar{\Delta}_2^2)+Ce^{-x}E(\bar{\Delta}_2^2 e^{\bar{W}}).
\end{aligned}
$$

Now the proof of Lemma 4.2 is complete.

Proof of Lemma 4.3 To estimate $P(x(1+\frac{1}{2}\Delta_2-\frac{1}{2}\Delta_2^2) \le W+\Delta_1 \le x(1+\frac{1}{2}\Delta_2))$, we first do the truncation, then use the randomized concentration inequality (Shao, 2010). As before, we let

$$\Delta_x = \Delta_1 - \frac{x}{2}\Delta_2, \quad \Delta_x' = \Delta_1 - \frac{x}{2}\Delta_2 + \frac{x}{2}\Delta_2^2,$$

then it follows that

$$\begin{aligned}
&P(x(1+\frac{1}{2}\Delta_2 - \frac{1}{2}\Delta_2^2) \le W + \Delta_1 \le x(1+\frac{1}{2}\Delta_2)) \\
&= P(x - \Delta_x' \le W \le x - \Delta_x) \\
&= P(x - \Delta_x' \le W \le x - \Delta_x, \max|\xi_i| \le 1) + P(x - \Delta_x' \le W \le x - \Delta_x, \max|\xi_i| > 1) \\
&\le P(x - \Delta_x' \le W \le x - \Delta_x, \max|\xi_i| \le 1, |\Delta_x| \le \frac{x+1}{4}, |\Delta_x'| \le \frac{x+1}{4}) \\
&\quad + P(\max|\xi_i| \le 1, |\Delta_x| > \frac{x+1}{4}) + P(\max|\xi_i| \le 1, |\Delta_x'| > \frac{x+1}{4}) \\
&\quad + P(x - \Delta_x' \le W \le x - \Delta_x, \max|\xi_i| > 1, |\Delta_x| \le \frac{x+1}{4}, |\Delta_x'| \le \frac{x+1}{4}) \\
&\quad + P(\max|\xi_i| > 1, |\Delta_x| > \frac{x+1}{4}) + P(\max|\xi_i| > 1, |\Delta_x'| > \frac{x+1}{4}) \\
&\le P(x - \bar{\Delta}_x' \le \bar{W} \le x - \bar{\Delta}_x) + P(W \ge \frac{3x-1}{4}, \max|\xi_i| > 1) \\
&\quad + P(|\Delta_x| > \frac{x+1}{4}) + P(|\Delta_x'| > \frac{x+1}{4}) \\
&\le P(x(1+\frac{1}{2}\bar{\Delta}_2 - \frac{1}{2}\bar{\Delta}_2^2) \le \bar{W} + \bar{\Delta}_1 \le x(1+\frac{1}{2}\bar{\Delta}_2)) + P(|\Delta_1| \ge \frac{1}{4}) + P(\frac{x}{2}|\Delta_2| \ge \frac{x}{4}) \\
&\quad + P(\frac{x}{2}|\Delta_2|^2 \ge \frac{x+1}{4}) + \sum_{i=1}^n P(W^{(i)} \ge \frac{x-1}{2}, |\xi_i| > 1) + \sum_{i=1}^n P(|\xi_i| \ge \frac{x+1}{4}) \\
&\le P(x(1+\frac{1}{2}\bar{\Delta}_2 - \frac{1}{2}\bar{\Delta}_2^2) \le \bar{W} + \bar{\Delta}_1 \le x(1+\frac{1}{2}\bar{\Delta}_2)) + P(|\Delta_1| \ge \frac{1}{4}) + P(|\Delta_2| \ge \frac{1}{2}) \\
&\quad + \sum_{i=1}^n P(W^{(i)} \ge \frac{x-1}{2}, |\xi_i| > 1) + \sum_{i=1}^n P(|\xi_i| \ge \frac{x+1}{4}),
\end{aligned}$$

where all terms can be included in the main theorem except

$$P(x(1+\frac{1}{2}\bar{\Delta}_2 - \frac{1}{2}\bar{\Delta}_2^2) \le \bar{W} + \bar{\Delta}_1 \le x(1+\frac{1}{2}\bar{\Delta}_2)).$$

We will use the following lemma which can be found in Shao (2010).

Lemma 4.6. *Let*

$$\begin{aligned}\bar{\xi}_i &= \xi_i I(|\xi_i| \le 1), \quad W^* = \sum_{i=1}^n \bar{\xi}_i, \quad \Delta^* = \Delta(\bar{\xi}_1, ..., \bar{\xi}_n), \\ \Delta_1^* &= \Delta_1(\bar{\xi}_1, ..., \bar{\xi}_n), \quad \Delta_2^* = \Delta_2(\bar{\xi}_1, ..., \bar{\xi}_n).\end{aligned}$$

Assume that there exists $c_1 > c_2 > 0$, $\delta > 0$ such that

$$\sum_{i=1}^n E\xi_i^2 \le c_1 \tag{4.17}$$

and

$$\sum_{i=1}^n E|\bar{\xi}_i| \min(\delta, |\bar{\xi}_i|/2) \ge c_2. \tag{4.18}$$

Then the randomized concentration inequality gives us for $\lambda \ge 0$

$$\begin{aligned}&Ee^{\lambda(W^*+\Delta^*)} I(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*) \\ \le\ & (Ee^{2\lambda(W^*+\Delta^*)})^{1/2} \exp(-\frac{c_2^2}{8c_1\delta^2}) + \frac{2e^{\lambda\delta}}{c_2}[Ee^{\lambda(W^*+\Delta^*)}|W^*|(|\Delta_2^* - \Delta_1^*| + 2\delta) \\ & +2\sum_{i=1}^n Ee^{\lambda(W^{*(i)}+\Delta^{*(i)})}|\bar{\xi}_i|(|\Delta_1^* - \Delta_1^{*(i)}| + |\Delta_2^* - \Delta_2^{*(i)}|) \\ & +8\sum_{i=1}^n E|\Delta^* - \Delta^{*(i)}| \min(|\bar{\xi}_i|, |\Delta^* - \Delta^{*(i)}|)(1 + \lambda(\Delta_2^* - \Delta_1^* + 2\delta)) \\ & \max(e^{\lambda(W^*+\Delta^*)}, e^{\lambda(W^{*(i)}+\Delta^{*(i)})})],\end{aligned} \tag{4.19}$$

for any measurable functions $\Delta^{*(i)}$, $\Delta_1^{*(i)}$, $\Delta_2^{*(i)}$ such that ξ_i is independent of $(W^{*(i)}, \Delta^{*(i)}, \Delta_1^{*(i)}, \Delta_2^{*(i)})$, where $W^{*(i)} = W^* - \bar{\xi}_i$.

To use the lemma, now we let

$$\begin{aligned}W^* &= \bar{W},\ \Delta^* = \bar{\Delta}_1,\ \Delta_1^* = x(1 + \frac{1}{2}\bar{\Delta}_2 - \frac{1}{2}\bar{\Delta}_2^2),\ \Delta_2^* = x(1 + \frac{1}{2}\bar{\Delta}_2), \\ \Delta^{*(i)} &= \bar{\Delta}_1^{(i)},\ \Delta_1^{*(i)} = x(1 + \frac{1}{2}\bar{\Delta}_2^{(i)} - \frac{1}{2}(\bar{\Delta}_2^{(i)})^2),\ \Delta_2^{*(i)} = x(1 + \frac{1}{2}\bar{\Delta}_2^{(i)}).\end{aligned}$$

Let $c_1 = 1$, $c_2 = 1/2$, $\lambda = 1$. Noting $|\bar{\Delta}_2| \le \frac{1}{4}$, $|\bar{\Delta}_1| \le \frac{x}{8}$, it follows that

$$W^* + \Delta^* \ge \Delta_1^* = x(1 + \frac{1}{2}\bar{\Delta}_2 - \frac{1}{2}\bar{\Delta}_2^2) \ge \frac{27x}{32}.$$

Then we have

$$
\begin{aligned}
& Ee^{(W^*+\Delta^*)}I(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*) \\
\ge \ \ & Ee^{\frac{27x}{32}}I(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*) \\
= \ \ & e^{\frac{27}{32}x}P(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*).
\end{aligned}
$$

Hence, from (4.19), we obtain

$$
\begin{aligned}
& P(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*) \\
\le \ \ & e^{-\frac{27}{32}x}Ee^{(W^*+\Delta^*)}I(\Delta_1^* \le W^* + \Delta^* \le \Delta_2^*) \\
\le \ \ & e^{-\frac{27}{32}x}(Ee^{2(W^*+\Delta^*)})^{1/2}\exp(-\frac{(1/2)^2}{8\delta^2}) + \frac{2e^{\delta}e^{-\frac{23}{32}x}}{\frac{1}{2}}[Ee^{(W^*+\Delta^*)}|W^*|(|\frac{x}{2}\bar{\Delta}_2^2| + 2\delta) \\
& +2\sum_{i=1}^{n} Ee^{(W^{*(i)}+\Delta^{*(i)})}|\bar{\xi}_i|(|\Delta_1^* - \Delta_1^{*(i)}| + |\Delta_2^* - \Delta_2^{*(i)}|) \\
& +8\sum_{i=1}^{n} E|\Delta^* - \Delta^{*(i)}|\min(|\bar{\xi}_i|, |\Delta^* - \Delta^{*(i)}|)(1 + (\frac{x}{2}\bar{\Delta}_2^2 + 2\delta)) \\
& \max(e^{(W^*+\Delta^*)}, e^{(W^{*(i)}+\Delta^{*(i)})})] \\
\le \ \ & Ce^{-\frac{23}{32}x-\frac{1}{32\delta^2}} + 4e^{\delta}e^{-\frac{23}{32}x}[Ee^{W^*}|W^*|(|\frac{x}{2}\bar{\Delta}_2^2| + 2\delta) \\
& +2\sum_{i=1}^{n}(Ee^{W^{*(i)}}|\bar{\xi}_i||\Delta_1^* - \Delta_1^{*(i)}| + Ee^{W^{*(i)}}|\bar{\xi}_i||\Delta_2^* - \Delta_2^{*(i)}|) \\
& +8(1 + \frac{x}{32} + 2\delta)\sum_{i=1}^{n} E|\Delta^* - \Delta^{*(i)}||\bar{\xi}_i|(e^{W^*} + e^{W^{*(i)}})] \\
\le \ \ & Ce^{-\frac{23}{32}x-\frac{1}{32\delta^2}} + 8e^{\delta}e^{-\frac{23}{32}x}\delta + 2xe^{\delta}e^{-\frac{23}{32}x}Ee^{\bar{W}}|\bar{W}||\bar{\Delta}_2^2| \\
& +8e^{\delta-\frac{23}{32}x}\sum_{i=1}^{n}\|Ee^{\bar{W}^{(i)}}|\bar{\xi}_i|\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q + 8e^{\delta-\frac{23}{32}x}\sum_{i=1}^{n}\|Ee^{\bar{W}^{(i)}}|\bar{\xi}_i|\|_p\|\bar{\Delta}_2^2 - \bar{\Delta}_2^{(i)2}\|_q \\
& +C(1+\delta)e^{\delta-\frac{17}{32}x}\sum_{i=1}^{n}\|Ee^{\bar{W}^{(i)}}|\bar{\xi}_i|\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q \\
\le \ \ & Ce^{-\frac{17}{32}x}\delta + Ce^{-\frac{17}{32}x}E(e^{\bar{W}}|\bar{W}|\bar{\Delta}_2^2) \\
& +Ce^{-\frac{17}{32}x}(\sum_{i=1}^{n}\|\bar{\xi}_i\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \sum_{i=1}^{n}\|\bar{\xi}_i\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q) \\
\le \ \ & Ce^{-\frac{17}{32}x}\delta + Ce^{-\frac{17}{32}x}E(e^{\bar{W}}\bar{\Delta}_2^2) \\
& +Ce^{-\frac{17}{32}x}(\sum_{i=1}^{n}\|\bar{\xi}_i\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \sum_{i=1}^{n}\|\bar{\xi}_i\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q),
\end{aligned}
$$

where we use the fact

$$e^{\bar{W}}|\bar{W}| \leq C(e^{2\bar{W}} + e^{\frac{1}{2}\bar{W}}).$$

Note that $E|\bar{\xi}_i|^3 < \infty$, using Remark 2.1 in [23], we can let $\delta = 1/2\sum_{i=1}^n E|\bar{\xi}_i|^3 \leq 1/2\sum_{i=1}^n E|\bar{\xi}_i|^r$. Then we get

$$\begin{aligned}
&P(\Delta_1^* \leq W^* + \Delta^* \leq \Delta_2^*) \\
&\leq \; Ce^{-\frac{17}{32}x}\sum_{i=1}^n E|\xi_i|^r + Ce^{-\frac{17}{32}x}E(e^{\bar{W}}\bar{\Delta}_2^2) \\
&\quad +Ce^{-\frac{17}{32}x}(\sum_{i=1}^n \|\bar{\xi}_i\|_p\|\bar{\Delta}_1 - \bar{\Delta}_1^{(i)}\|_q + \sum_{i=1}^n \|\bar{\xi}_i\|_p\|\bar{\Delta}_2 - \bar{\Delta}_2^{(i)}\|_q).
\end{aligned}$$

Chapter 5

Applications

Theorem 3.1 can be applied to a wide range of studentized statistics and provide bounds of the best possible order in many cases. In this section, we give three applications to illustrate the usefulness of Theorem 3.1.

5.1 Application to Student's t-Statistics

5.1.1 Introduction

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent real-valued random variables with $EX_i = 0$ and $EX_i^2 < \infty$. Put

$$S_n = \sum_{i=1}^n X_i, \qquad V_n^2 = \sum_{i=1}^n X_i^2, \qquad B_n^2 = \sum_{i=1}^n EX_i^2. \qquad \text{for } n = 1, 2, \ldots$$

Then the Student's t-statistic is defined by

$$T_n = \frac{S_n}{\sqrt{n} s_n},$$

where $s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - S_n/n)^2$.

Student's t-statistic is one of the most important statistics. When σ is known, we know that $S_n/\sqrt{n}\sigma$ is asymptotically normal distributed, then we can construct a confidence interval or do the hypothesis testing for μ. But when σ is unknown, which is in most cases, we replace σ by s_n and use this student's t-statistic to do the hypothesis testing or construct the confidence interval for μ.

The student's t-statistic follows a t distribution if we assume the sampling is from a normal distribution. But even if the data is not normally distributed, when the sample size is large enough, we can use the standard normal distribution to describe the student's t-statistics since s_n converges to σ almost surely as n goes to infinity. A lot of statisticians then naturally consider the behavior of the approximation.

There are two ways to consider the approximation. One is to consider the ratio of the tail probabilities, say, Shao (1999) gives a Cramer type large deviation result for Student's t-statistic. The other is to consider the difference of the two distribution, using Berry-Esseen bound. The Berry-Esseen bounds of Student's t-statistics have been investigated by different authors, such as, Slavova (1985), Hall (1988) and Bentkus and Gotze (1996).

5.1.2 Main Result

It is known that the Student's t-statistic is closely related to the self-normalized sum S_n/V_n via the following identity

$$T_n = \frac{S_n}{V_n}\left(\frac{n-1}{n-(S_n/V_n)^2}\right)^{1/2}. \tag{5.1}$$

Since $s/(n-s^2)^{1/2}$ is increasing on $(-\sqrt{n}, \sqrt{n})$, (5.1) follows that for $x > 0$

$$\{T_n \geq x\} = \{\frac{S_n}{V_n} \geq x(\frac{n}{n+x^2-1})^{1/2}\}. \tag{5.2}$$

Therefore it is sufficient to state our main result in terms of the self-normalized sum. Noting that

$$\begin{aligned}\frac{S_n}{V_n} &= \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n X_i^2}} \\ &= \frac{\sum_{i=1}^n X_i/B_n}{\sqrt{\sum_{i=1}^n (X_i/B_n)^2}} \\ &:= \frac{\sum_{i=1}^n \xi_i}{\sqrt{\sum_{i=1}^n \xi_i^2}},\end{aligned}$$

where $\xi_i = X_i/B_n$ satisfying $E\xi_i = 0$ and $\sum_{i=1}^n E\xi_i^2 = 1$, it follows the form

$$T_n = \frac{W + \Delta_1}{\sqrt{1+\Delta_2}},$$

where $W = \sum_{i=1}^n \xi_i$, $\Delta_1 = 0$, and $\Delta_2 = \sum_{i=1}^n \xi_i^2 - 1 = \sum_{i=1}^n (\xi_i^2 - E\xi_i^2)$. It is a special case with $\Delta_1 = \Delta_3 = 0$. Therefore we can directly get

Theorem 5.1. *There exists a constant $C > 0$, such that, for some $2 < r \leq 3$*

$$|P(T_n \leq x) - \Phi(x)| \leq C \sum_{i=1}^n E|\xi_i|^r.$$

The result Theorem 5.1 got is known before, see Bentkus, V. and Gotze, F (1996) for instance.

5.2 Application to Studentized U-Statistics

5.2.1 Introduction

Let $X, X_1, X_2, ..., X_n$ be a sequence of i.i.d. random variables, and let $h(x, y)$ be a real-valued Borel measurable function, symmetric in its arguments with $Eh(X_1, X_2) = \theta$. Then the U-statistic of degree 2 for estimation of θ with kernel $h(x, y)$ is defined to be

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

For a non-degenerate U-statistic, if $Eh^2(X_1, X_2) < \infty$ and $\sigma_g^2 = \mathrm{Var}(g(X_1)) > 0$, where $g(x) = Eh(x, X)$, then the central limit theorem holds, i.e.,

$$\sup_{x\in\mathbb{R}} \left| P\Big(\frac{\sqrt{n}}{2\sigma_g}(U_n - \theta) \le x\Big) - \Phi(x) \right| \to 0. \tag{5.3}$$

However, since σ_g is typically unknown, it is necessary to estimate σ_g first and then substitute it into (5.3). Indeed, what commonly used in practice is the following Studentized U-statistic

$$T_n = \frac{\sqrt{n}}{2r_n}(U_n - \theta), \tag{5.4}$$

where $n^{-1}r_n^2$ denotes the jackknife estimator of σ_g^2,

$$r_n^2 = (n-1)(n-2)^{-2}\sum_{i=1}^{n}(q_i - U_n)^2 \quad \text{with} \quad q_i = \frac{1}{n-1}\sum_{j=1, j\neq i}^{n} h(X_i, X_j). \tag{5.5}$$

Berry-Esseen bounds for Studentized U-statistics have been obtained by various authors. For example, Callaert and Veraverbeke (1981) gave the Berry-Esseen bounds among others. Zhao (1983) sharpened the work of Callaert and Veraverbeke (1981) and obtained the optimal rate of convergence, $O(n^{-1/2})$, when $E|h(X_1, X_2)|^4 < \infty$ and $\sigma_g^2 > 0$. Wang, Jing and Zhao (2000) further weakened the former condition to $E|h(X_1, X_2)|^3 < \infty$.

In particular, when $h(x, y) = (x + y)/2$, the Studentized U-statistic reduces to Student's t-statistic.

5.2.2 Main Result

We now study the rate of convergence for the distribution of the Studentized U-statistic T_n, i.e.

$$P(T_n \le x) = P\Big(\frac{\sqrt{n}}{2r_n}(U_n - \theta) \le x\Big), \tag{5.6}$$

to its normal limit in Kolmogorov distance. Under the current setup, Theorem 3.1 specializes to the following result.

Theorem 5.2. *Assume that $\theta = Eh(X_1, X_2) = 0$ and $E|h(X_1, X_2)|^3 < \infty$. Let $g(x) = Eh(x, X_1)$, $\sigma_g^2 = Eg^2(X_1)$ and assume also that $\sigma_g^2 > 0$, $E|g(X_1)|^3 < \infty$. Then*

$$\begin{aligned} & |P(T_n \le x) - \Phi(x)| \\ \le \ & \frac{C}{\sqrt{n}}\Big(\sigma_g^{-3} E|g(X_1)|^3 + \sigma_g^{-2}(E|g(X_2)|^3)^{\frac{1}{3}}(E|h(X_1, X_2)|^3)^{\frac{1}{3}}\Big). \end{aligned}$$

5.2.3 Proof of Theorem 5.2.

First, if we put $\tilde{h} = h/\sigma_g$ and $\tilde{g} = g/\sigma_g$, then $\tilde{g}(x) = E\{\tilde{h}(x, X_1)\}$ and $\tilde{g}(X_1)$, ... , $\tilde{g}(X_n)$ are i.i.d. random variables with zero mean and unit variance. From the scaling invariance property of Studentized U-statistic, we can replace h and g with $\tilde{h}$ and $\tilde{g}$ respectively, which does not change the definition of T_n. For brevity of notation, we will still use h and g but assume without loss of generality that $\sigma_g^2 = 1$.

We begin with some standard truncations by letting, for any $1 \le i \le n$,

$$\bar{g}(X_i) = g(X_i) I(|g(X_i)| \le n^{1/2}), \qquad \bar{\theta} = E\bar{g}(X_i), \qquad \bar{\sigma}_g^2 = \mathrm{Var}(\bar{g}(X_i)),$$

$$\begin{aligned} h^*(X_i, X_j) &= h(X_i, X_j)/\bar{\sigma}_g, \qquad g^*(x) = (\bar{g}(x) - \bar{\theta})/\bar{\sigma}_g, \\ \psi^*(x_1, x_2) &= h^*(x_1, x_2) - g^*(x_1) - g^*(x_2), \end{aligned}$$

so that $g^*(X_1), ..., g^*(X_n)$ are i.i.d. random variables with mean zero and variance one. Put $\xi_i = g^*(X_i)/\sqrt{n}$, $1 \le i \le n$,

$$\begin{aligned} & S_n = \sum_{i=1}^n \xi_i, \qquad V_n^2 = \sum_{i=1}^n \xi_i^2, \qquad \Delta_n = \frac{1}{n-1} \sum_{1 \le i < j \le n} \frac{\psi^*(X_i, X_j)}{\sqrt{n}}, \\ & \Lambda_n^2 = \sum_{i=1}^n (W_n^{(i)})^2, \qquad W_n^{(i)} = \sum_{j=1, j \ne i}^n \frac{\psi^*(X_i, X_j)}{\sqrt{n}}, \end{aligned}$$

and define

$$U_n^* = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} h^*(X_i, X_j), \qquad q_i^* = \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} h^*(X_i, X_j),$$
$$r_n^{*2} = \frac{n-1}{(n-2)^2} \sum_{i=1}^{n} (q_i^* - U_n^*)^2, \qquad s_n^{*2} = \frac{n-1}{(n-2)^2} \sum_{i=1}^{n} q_i^{*2},$$

and note that

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

On the event $\{ \max_{1\le i\le n} |g(X_i)| \le n^{1/2} \}$, it is easy to see that

$$U_n/r_n = U_n^*/r_n^*,$$

which further implies

$$\begin{aligned} T_n &= \sqrt{n}U_n^*/(2r_n^*) \\ &= \frac{\sqrt{n}U_n^*/2}{\{s_n^{*2} - (n^2-n)U_n^{*2}/(n-2)^2\}^{1/2}} \\ &= \frac{T_n^*}{\{1 - 4(n-1)T_n^{*2}/(n-2)^2\}^{1/2}}, \end{aligned} \tag{5.7}$$

where $T_n^* := \sqrt{n}U_n^*/(2s_n^*)$, and

$$\{T_n \ge x\} = \{T_n^* \ge x/(1 + 4x^2(n-1)/(n-2)^2)^{1/2}\}. \tag{5.8}$$

Therefore we only need to consider T_n^* instead of T_n. In order to apply Theorem 3.1 to T_n^*, first observe that

$$q_i^* = \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} h^*(X_i, X_j) = \frac{n^{1/2}}{n-1} \Big\{ (n-2)\xi_i + S_n + W_n^{(i)} \Big\},$$

such that by routine calculations,

$$\begin{aligned} \frac{(n-2)^2(n-1)}{n} s_n^{*2} &= \sum_{i=1}^{n} \Big\{ (n-2)\xi_i + S_n + W_n^{(i)} \Big\}^2 \\ &= (n-2)^2 V_n^2 + \Lambda_n^2 + (3n-4)S_n^2 \\ &\quad + 2(n-2) \sum_{i=1}^{n} \xi_i W_n^{(i)} + 2S_n \sum_{i=1}^{n} W_n^{(i)}. \end{aligned}$$

By Cauchy-Schwarz inequality, the last term can be bounded by

$$|S_n \sum_{i=1}^{n} W_n^{(i)}| \le \sqrt{n} S_n \Lambda_n,$$

so that if we write

$$s_n^{*2} = \frac{n}{n-1}\big(V_n^2 + \delta_n\big), \tag{5.9}$$

then $\delta_n \equiv \delta_n(X_1, ..., X_n)$ satisfies

$$\begin{aligned} |\delta_n| &\le 4S_n^2/n + 2\Lambda_n^2/n^2 + 2\big(\sum_{i=1}^n \xi_i W_n^{(i)}\big)/n \\ &\le \delta_{n,1} + \delta_{n,2}, \end{aligned} \tag{5.10}$$

where

$$\delta_{n,1} := 4S_n^2/n + 2\Lambda_n^2/n^2, \qquad \delta_{n,2} := 2\big(\sum_{i=1}^n \xi_i W_n^{(i)}\big)/n.$$

Moreover, by Hoeffding's decomposition,

$$\sqrt{n}U_n^*/2 = S_n + \Delta_n,$$

which together with (5.9) gives

$$T_n^* = \frac{S_n + \Delta_n}{d_n\sqrt{1 + (V_n^2 - 1) + \delta_n}}, \tag{5.11}$$

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

By Markov's inequality and (5.10), together with the fact (recall that $Eg(X_1) = 0$, $Eg^2(X_1) = 1$ and thus $Eh^2(X_1, X_2) \ge 1$)

$$|\bar{\theta}| \le n^{-1/2}, \qquad \bar{\sigma}_g^2 \ge 1/2, \tag{5.12}$$

we have

$$\begin{aligned} &P(|\delta_{n,1}| > n^{-1/2}) \le n^{1/2} E\delta_{n,1} \\ &\le n^{1/2}\Big\{4/n + 2(E\Lambda_n^2)/n^2\Big\} \\ &\le n^{1/2}\Big\{4/n + \frac{2}{n}Eh^2(X_1, X_2)\Big\} \\ &\le \frac{C}{\sqrt{n}}Eh^2(X_1, X_2), \end{aligned} \tag{5.13}$$

where

$$E\Lambda_n^2 = \frac{1}{n}\sum_{i=1}^n E\Big\{\sum_{j\ne i}\psi^*(X_i, X_j)\Big\}^2 \le nE(\psi^*(X_1, X_2))^2 \le CnEh^2(X_1, X_2),$$

and

$$ES_n^{*2} = \sum_{i=1}^{n} E\xi_i^2 = 1.$$

By (5.11), (5.13) and noting that $|d_n - 1| = O(n^{-1})$, it is sufficient to consider

$$\begin{aligned}
& P\Big(\frac{S_n + \Delta_n}{\sqrt{1 + (V_n^2 - 1) + \delta_n}} \le x\Big) \\
\le\ & P\Big(\frac{S_n + \Delta_n}{\sqrt{1 + (V_n^2 - 1) + \delta_{n,1} + \delta_{n,2}}} \le x, |\delta_{n,1}| \le n^{-1/2}\Big) + P\big(|\delta_{n,1}| > n^{-1/2}\big) \\
\le\ & P\Big(\frac{S_n + \Delta_n}{\sqrt{1 + (V_n^2 - 1) + n^{-1/2} + \delta_{n,2}}} \le x\Big) + \frac{C}{\sqrt{n}} E h^2(X_1, X_2). \qquad (5.14)
\end{aligned}$$

In Theorem 3.1, let

$$W = S_n = \sum_{i=1}^{n} \xi_i = \sum_{i=1}^{n} \frac{g^*(X_i)}{\sqrt{n}}, \qquad (5.15)$$

$$\Delta_1 = \Delta_n = \frac{1}{n-1} \sum_{1 \le i \ne j \le n} \frac{\psi^*(X_i, X_j)}{\sqrt{n}}, \qquad (5.16)$$

$$\begin{aligned}
\Delta_2 &= (V_n^2 - 1) + n^{-1/2} + 2\big(\sum_{i=1}^{n} \xi_i W_n^{(i)}\big)/n \\
&= \sum_{i=1}^{n} (\xi_i^2 - E\xi_i^2) + n^{-1/2} + \frac{1}{n^{3/2}} \sum_{i=1}^{n} \big(\xi_i \sum_{j \ne i} \psi^*(X_i, X_j)\big). \qquad (5.17)
\end{aligned}$$

However, since what we really dealt with in Theorem 3.1 is a truncated version of Δ_2, we assume in the following without loss of generality that $|\Delta_2| \le 1$. For each $1 \le l \le n$, set

$$\Delta_1^{(l)} = \frac{1}{(n-1)\sqrt{n}} \sum_{1 \le i \ne j (\ne l) \le n} \psi^*(X_i, X_j)$$

and

$$\Delta_2^{(l)} = \sum_{\substack{i=1 \\ i \ne l}}^{n} (\xi_i^2 - E\xi_i^2) + \frac{1}{\sqrt{n}} + \frac{1}{n^{3/2}} \sum_{i=1, i \ne l}^{n} \big(\xi_i \sum_{j \ne i, l} \psi^*(X_i, X_j)\big),$$

so that

$$\Delta_1 - \Delta_1^{(l)} = \frac{2}{(n-1)\sqrt{n}} \sum_{i \ne l} \psi^*(X_i, X_l)$$

and

$$\Delta_2 - \Delta_2^{(l)} = \xi_l^2 - E\xi_l^2 + \frac{1}{n^{3/2}} \Big\{ \xi_l \sum_{j \ne l} \psi^*(X_j, X_l) + \sum_{i \ne l} \xi_i \psi^*(X_i, X_l) \Big\}.$$

Lemma 5.3. *There exists an absolute constant $C > 0$, such that*

$$E|\Delta_1 - \Delta_1^{(l)}|^2 \leq Cn^{-2}Eh^2(X_1, X_2), \tag{5.18}$$

and

$$\begin{aligned} & E|\Delta_2 - \Delta_2^{(l)}|^2 \\ \leq\ & E\xi_1^4 + \frac{C}{n^2}\Big(1 + Eh^2(X_1, X_2) + E|g(X_2)h(X_1, X_2)| \\ & +(E|g(X_2)|^3)^{2/3}(E|h(X_1, X_2)|^3)^{2/3}\Big). \end{aligned} \tag{5.19}$$

As a direct consequence of above lemma, we have

$$\sum_{l=1}^{n} \|\xi_l\|_2 \|\Delta_1 - \Delta_1^{(l)}\|_2 \leq Cn^{-1/2}\big\{Eh^2(X_1, X_2)\big\}^{1/2} \tag{5.20}$$

and

$$\begin{aligned} & \sum_{l=1}^{n} \|\xi_l\|_2 \|\Delta_2 - \Delta_2^{(l)}\|_2 \\ \leq\ & Cn^{-1/2}(E|g(X_1)|^3)^{1/3}(E|h(X_1, X_2)|^3)^{1/3}. \end{aligned} \tag{5.21}$$

Now we calculate $E\bar{\Delta}_x f_x(\bar{W})$, where $\Delta_x = \Delta_1 - x\Delta_2/2$. As proved before, we have $E|f_x(\bar{W})| \leq Ce^{-x/4}$, and

$$E|\Delta_1|^2 \leq \frac{Eh^2(X_1, X_2)}{2(n-1)}, \tag{5.22}$$

therefore

$$E|\Delta_1 f_x(\bar{W})| \leq C\big(e^{-x/4}n^{-1}Eh^2(X_1, X_2)\big)^{1/2} \leq Ce^{-x/8}n^{-1/2}\big(Eh^2(X_1, X_2)\big)^{1/2} \tag{5.23}$$

For

$$\begin{aligned} & -\frac{x}{2}E\Delta_2 f_x(\bar{W}) \\ =\ & -\frac{x}{2}E\Big\{\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2) + \frac{1}{\sqrt{n}} + \frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i \sum_{j\neq i}\psi^*(X_i, X_j)\big)\Big\} f_x(\bar{W}) \\ \leq\ & -\frac{x}{2}E(\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2))f_x(\bar{W}) - \frac{x}{2}\frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i \sum_{j\neq i}\psi^*(X_i, X_j)\big)f_x(\bar{W}) - \frac{x}{2}\frac{1}{\sqrt{n}}Ef_x(\bar{W}) \\ \leq\ & H_1 + H_2 + H_3, \end{aligned}$$

it is easy to see that

$$|H_3| \le \frac{Cxe^{-x/4}}{\sqrt{n}}. \tag{5.24}$$

Lemma 5.4. *There exists an absolute constant $C > 0$, such that*

$$|H_1| \le Cxe^{-x/4}n^{-1/2}E|g(X_1)|^3, \tag{5.25}$$

and

$$|H_2| \le Cxe^{-x/8}n^{-1/2}\Big((E|g(X_2)|^3)^{1/3}(E|h(X_1,X_2)|^3)^{1/3} + E|g(X_1)|^3\Big). \tag{5.26}$$

By (5.23), (5.24) and Lemma 5.4, we get

$$\begin{aligned}
&|E\bar{\Delta}_x f_x(\bar{W})| \\
&\le |E\bar{\Delta}_1 f_x(\bar{W})| + \frac{x}{2}|E\bar{\Delta}_2 f_x(\bar{W})| \\
&\le \frac{Cxe^{-x/8}}{\sqrt{n}}\Big((E|g(X_2)|^3)^{1/3}(E|h(X_1,X_2)|^3)^{1/3} + E|g(X_1)|^3\Big). \qquad (5.27)
\end{aligned}$$

Next we calculate

$$E\bar{\Delta}_1^2 + \frac{1}{1+x}E\bar{\Delta}_2^2 + e^{-x/2}E\bar{\Delta}_2^2 e^{\bar{W}},$$

where $E\Delta_1^2$ is given in (5.22). Using $|\bar{\Delta}_2| \le 1$, by (5.34), (5.35) and (5.36), we get

$$\begin{aligned}
&E\bar{\Delta}_2^2 \\
&\le CE\Big|\sum_{i=1}^n (\xi_i^2 - E\xi_i^2) + \frac{1}{n^{3/2}}\sum_{i=1}^n \big(\xi_i \sum_{j\ne i}\psi^*(X_i,X_j)\big) + \frac{1}{\sqrt{n}}\Big|^2 \\
&\le CnE|\xi_i|^3 + \|\frac{C}{n^{3/2}}\sum_{i=1}^n \big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\big)\|_2 + \frac{C}{\sqrt{n}} \\
&\le Cn^{-1/2}E|g(X_1)|^3 + Cn^{-1/2}\Big((E|g(X_2)|^3)^{\frac{1}{3}}(E|h(X_1,X_2)|^3)^{\frac{1}{3}} + E|g(X_1)|^3\Big) \\
&\le Cn^{-1/2}\Big(E|g^3(X_1)| + (E|g(X_2)|^3)^{1/3}(E|h(X_1,X_2)|^3)^{1/3}\Big). \qquad (5.28)
\end{aligned}$$

and

$$
\begin{aligned}
&E\bar{\Delta}_2^2 e^{\bar{W}} \\
&= CE\Big|\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2) + \frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i\sum_{j\neq i}\psi^*(X_i,X_j)\big) + \frac{1}{\sqrt{n}}\Big|^2 e^{\bar{W}} \\
&\le CE|\frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i\sum_{j\neq i}\psi^*(X_i,X_j)\big) + \frac{1}{\sqrt{n}}|^2 e^{\bar{W}} + CE|\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2)|^2 e^{\bar{W}} \\
&\le C(E|\frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i\sum_{j\neq i}\psi^*(X_i,X_j)\big) + \frac{1}{\sqrt{n}}|^2)^{1/2}(Ee^{2\bar{W}})^{1/2} + C\sum_{i=1}^{n}E|\xi_i|^3 \\
&\le \frac{C}{\sqrt{n}}\Big((Eh^2(X_1,X_2))^{1/2} + (E|g(X_2)|^3)^{\frac{1}{3}}(E|h(X_1,X_2)|^3)^{\frac{1}{3}} + E|g(X_1)|^3\Big) + C\sum_{i=1}^{n}E|\xi_i|^3 \\
&\le \frac{C}{\sqrt{n}}\Big((E|g(X_2)|^3)^{\frac{1}{3}}(E|h(X_1,X_2)|^3)^{\frac{1}{3}} + E|g(X_1)|^3\Big). \qquad (5.29)
\end{aligned}
$$

Now, combining (5.14), (5.20), (5.21), (5.22), (5.27), (5.28) and (5.29), we get Theorem5.2

$$
\begin{aligned}
&|P(T_n \le x) - \Phi(x)| \\
&\le \frac{C}{\sqrt{n}}\Big(E|g(X_1)|^3 + (E|g(X_2)|^3)^{\frac{1}{3}}(E|h(X_1,X_2)|^3)^{\frac{1}{3}}\Big). \qquad (5.30)
\end{aligned}
$$

For the reminder of this section, we prove the lemmas we used.

Proof of Lemma 5.3 For (5.18), using the fact (5.12), we get

$$
\begin{aligned}
E|\Delta_1 - \Delta_1^{(l)}|^2 &\le \frac{4}{n(n-1)^2}\Big\{\sum_{j\neq i}E\Big(\psi^*(X_i,X_j)\Big)^2\Big\} \\
&\le Cn^{-2}Eh^2(X_1,X_2).
\end{aligned}
$$

For (5.19), we get

$$
\begin{aligned}
&E|\Delta_2 - \Delta_2^{(l)}|^2 \\
&\le E\xi_l^4 + \frac{1}{n^3}\Big\{E\Big(\xi_l\sum_{j\neq l}\psi^*(X_j,X_l)\Big)^2 + E\Big(\sum_{i\neq l}\xi_i\psi^*(X_i,X_l)\Big)^2\Big\} \\
&\le E\xi_l^4 + \frac{C}{n^2}\Big(Eh^2(X_1,X_2) + Eg^2(X_3)Eg(X_2)h(X_1,X_2) + (Eg^2(X_1))^2 \\
&\quad +(Eg(X_2)E(h(X_1,X_2)|X_1))^2\Big),
\end{aligned}
$$

where we use

$$
\begin{aligned}
& E\Big(\xi_l \sum_{j\neq l} \psi^*(X_j, X_l)\Big)^2 \\
&= E(\xi_l^2 E((\sum_{j\neq l} \psi^*(X_j, X_l))^2 | X_l)) \\
&= E(\xi_1^2 E((n-1)\psi^{*2}(X_1, X_2)|X_1)) \\
&= (n-1)E(\xi_1^2 \psi^{*2}(X_1, X_2)) \\
&\leq (n-1)E\xi_1^2 h^2(X_1, X_2) \\
&\leq nEh^2(X_1, X_2), \qquad (5.31)
\end{aligned}
$$

and

$$
\begin{aligned}
& E\Big(\sum_{i\neq l} \xi_i \psi^*(X_i, X_l)\Big)^2 \\
&= E(\sum_{i\neq l} \xi_i^2 \psi^{*2}(X_i, X_l)) + \sum_{i\neq j\neq l} E(\xi_i \xi_j \psi^*(X_i, X_l)\psi^*(X_j, X_l)) \\
&\leq nEh^2(X_1, X_2) + n^2 E\Big(\xi_2 \xi_3 \psi^*(X_1, X_2)\psi^*(X_1, X_3)\Big) \\
&\leq nEh^2(X_1, X_2) + n^2 E\Big(E\big(\xi_2 \psi^*(X_1, X_2)|X_1\big)E\big(\xi_3 \psi^*(X_1, X_3)|X_1\big)\Big) \\
&\leq nEh^2(X_1, X_2) \\
&\quad + CnE\Big\{E\Big(\bar{g}(X_2)(h(X_1, X_2) - \bar{g}(X_2))|X_1\Big)E\Big(\bar{g}(X_3)(h(X_1, X_3) - \bar{g}(X_3))|X_1\Big)\Big\} \\
&\leq nEh^2(X_1, X_2) + Cn\Big(1 + E|g(X_2)h(X_1, X_2)|\Big)\Big(1 + E|g(X_3)h(X_1, X_3)|\Big) \\
&\leq nEh^2(X_1, X_2) + Cn\Big\{1 + E|g(X_2)h(X_1, X_2)| + E|g(X_2)g(X_3)h(X_1, X_2)h(X_1, X_3)|\Big\} \\
&\leq Cn\Big((E|g(X_2)g(X_3)|^3)^{1/3}(E|h(X_1, X_2)h(X_1, X_3)|^{3/2})^{2/3}\Big) \\
&\leq Cn\Big((E|g(X_2)|^3 E|g(X_3)|^3)^{1/3}((E|h(X_1, X_2)|^3)^{1/2}(E|h(X_1, X_3)|^3)^{1/2})^{2/3}\Big) \\
&\leq Cn\Big((E|g(X_2)|^3)^{2/3}(E|h(X_1, X_2)|^3)^{2/3}\Big). \qquad (5.32)
\end{aligned}
$$

Proof of Lemma 5.4 For H_1, we have

$$
\begin{aligned}
H_1 &= -\frac{x}{2}E(\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2))f_x(\bar{W}) \\
&= -\frac{x}{2}E\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2)[f(\bar{W}) - f(\bar{W}^{(i)})] \\
&= -\frac{x}{2}E\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2)[\int_0^{\xi_i} f'(\bar{W}^{(i)}+t)dt] \\
&= -\frac{x}{2}E\sum_{i=1}^{n}(\xi_i^2 - E\xi_i^2)[\int_{-1}^{1}(I(0 \le t \le \xi_i) - I(\xi_i \le t \le 0))f'(\bar{W}^{(i)}+t)dt] \\
&= -\frac{x}{2}\sum_{i=1}^{n}E\int_{-1}^{1} f'(\bar{W}^{(i)}+t)(\xi_i^2 - E\xi_i^2)(I(0 \le t \le \xi_i) - I(\xi_i \le t \le 0))dt,
\end{aligned}
$$

therefore

$$
\begin{aligned}
|H_1| &\le \frac{x}{2}\sum_{i=1}^{n}\int_{-1}^{1}|Ef'(\bar{W}^{(i)}+t)||E(\xi_i^2 - E\xi_i^2)(I(0 \le t \le \xi_i) - I(\xi_i \le t \le 0))|dt \\
&\le Cxe^{-x/4}\sum_{i=1}^{n}E|\xi_i|^3 \\
&\le Cxe^{-x/4}\frac{Eg^3(X_1)}{\sqrt{n}}.
\end{aligned}
$$

For H_2, first write

$$
\begin{aligned}
|H_2| &\le \Big|\frac{x}{2}\frac{1}{n^{3/2}}\sum_{i=1}^{n}\big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\big)f_x(\bar{W})\Big| \\
&\le \frac{Cx}{n^{3/2}}\Big\{E\Big(\sum_{i=1}^{n}\big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\big)\Big)^2\Big\}^{1/2}\Big(E|f_x(\bar{W})|^2\Big)^{1/2} \\
&\le \frac{Cxe^{-x/8}}{n^{3/2}}\Big\{E\Big(\sum_{i=1}^{n}\big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\big)\Big)^2\Big\}^{1/2}, \qquad (5.33)
\end{aligned}
$$

therefore we need to calculate

$$
\begin{aligned}
&E\Big(\sum_{i=1}^{n}\big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\big)\Big)^2 \\
&= \sum_{i=1}^{n}E\Big(\xi_i\sum_{j\ne i}\psi^*(X_i,X_j)\Big)^2 + \sum_{i\ne j}E\Big(\xi_i\sum_{k\ne i}\psi^*(X_i,X_k)\xi_j\sum_{m\ne j}\psi^*(X_j,X_m)\Big) \\
&= nE\Big(\xi_1\sum_{j\ne 1}\psi^*(X_1,X_j)\Big)^2 + n(n-1)E\Big(\xi_1\sum_{k\ne 1}\psi^*(X_1,X_k)\xi_2\sum_{m\ne 2}\psi^*(X_2,X_m)\Big)
\end{aligned}
$$

(5.34)

By (5.31), we get

$$E\Big(\xi_1 \sum_{j\neq 1} \psi^*(X_1, X_j)\Big)^2 \leq nEh^2(X_1, X_2). \tag{5.35}$$

By (5.32), we get

$$\begin{aligned}
& E\Big(\xi_1 \sum_{k\neq 1} \psi^*(X_1, X_k)\xi_2 \sum_{m\neq 2} \psi^*(X_2, X_m)\Big) \\
\leq\ & (n-2)E\xi_1\psi^*(X_1, X_3)\xi_2\psi^*(X_2, X_3) + n^2\big\{E[\xi_1\psi^*(X_1, X_3)]\big\}^2 \\
& +2(n-2)E\xi_1\psi^*(X_1, X_3)\xi_2\psi^*(X_1, X_2) + E\xi_1\xi_2\psi^{*2}(X_1, X_2) \\
\leq\ & C\big\{(E|g(X_2)|^3)^{2/3}(E|h(X_1, X_2)|^3)^{2/3} + (E|g(X_1)|^3)^2\big\},
\end{aligned} \tag{5.36}$$

where

$$\begin{aligned}
& E[\xi_1\psi^*(X_1, X_3)] \\
=\ & E[\xi_1 E(\psi^*(X_1, X_3)|X_1)] \\
=\ & \bar{\sigma}_g^{-1} E[\xi_1 E(h(X_1, X_3) - \bar{g}(X_1) + \bar{\theta})|X_1] \\
\leq\ & \bar{\sigma}_g^{-1} E[\xi_1 g(X_1) I(|g(X_1)| > \sqrt{n})] + \bar{\sigma}_g^{-1} n^{-1/2} E|\xi_i| \\
\leq\ & Cn^{-1}E|g(X_1)|^3.
\end{aligned} \tag{5.37}$$

Therefore we get

$$\begin{aligned}
|H_2| \leq\ & \frac{Cxe^{-x/8}}{n^{3/2}}(n^2Eh^2(X_1, X_2))^{1/2} + \frac{Cxe^{-x/8}n}{n^{3/2}}\Big((E|g(X_2)|^3)^{1/3}(E|h(X_1, X_2)|^3)^{1/3} \\
& +E|g(X_1)|^3\Big) \\
\leq\ & \frac{Cxe^{-x/8}}{\sqrt{n}}\Big((E|g(X_2)|^3)^{1/3}(E|h(X_1, X_2)|^3)^{1/3} + E|g(X_1)|^3\Big).
\end{aligned} \tag{5.38}$$

5.3 Application to Studentized L-Statistics

5.3.1 Introduction

Let $X_1, ..., X_n$ be i.i.d. real random variables with distribution function F. Define F_n be the empirical distribution, which is,

$$F_n(x) = n^{-1} \sum_{j=1}^{n} I\{X_i \leq x\}.$$

Let $J(t)$ be a real-valued function on $[0, 1]$ and

$$T(G) = \int_{-\infty}^{\infty} xJ(G(x))dG(x).$$

The statistic $T(F_n)$ is called an L-statistic. Write

$$\sigma^2 \equiv \sigma^2(J, F) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(s))J(F(t))F(\min(s, t))(1 - F(\max(s, t)))dsdt.$$

Clearly, $\hat{\sigma}^2 \equiv \sigma^2(J, F_n)$ is a natural estimate of σ^2.

For Studentized L-statistics, the rates of convergence to a normal distribution have been studied by different authors. Helmers (1982) gave us a Berry-Esseen bound of rate $O(n^{-1/2})$ when assuming that $E|X_1|^{4.5} < \infty$, $\sigma^2 > 0$ and some smoothness conditions on $J(t)$. Wang, Jing, and Zhao (2000) weakened the moment condition to $E|X_1|^3 < \infty$.

5.3.2 Main Result

Define the distribution of the Studentized L-statistic by

$$H(x) \quad = \quad P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \leq x).$$

Applying Theorem 3.1, we get the following result.

Theorem 5.5. *Assume that $E|X_1|^3 < \infty$ and $\sigma^2 > 0$. If the weight function $J(t)$ satisfying that $J''(t)$ is bounded on $t \in [0,1]$, then we have*

$$|H(x) - \Phi(x)| \le \frac{C}{\sqrt{n}} \max\Big\{\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{1/3}, \Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{3}\Big\}.$$

5.3.3 Proof of Theorem 5.5.

We first truncate X_i and then apply the theorem to the truncated sum. Let

$$\bar{X}_i = X_i I(|X_i| \le n^{1/2}\sigma),$$

then

$$\begin{aligned}
&P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \ge x) \\
&= P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \ge x, \text{ all } |X_i| \le n^{1/2}\sigma) \\
&\quad + P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \ge x, |X_i| > n^{1/2}\sigma \text{ for some } i) \\
&\le P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \ge x, \text{ all } |X_i| \le n^{1/2}\sigma) + \sum_{i=1}^{n} P(|X_i| > n^{1/2}\sigma) \\
&\le P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) \ge x, \text{ all } |X_i| \le n^{1/2}\sigma) + \frac{E|X_1|^3}{n^{1/2}\sigma^3}. \qquad (5.39)
\end{aligned}$$

Therefore, we only need to consider $\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F))$ when all $|X_i| \le n^{1/2}\sigma$. Therefore we let $\bar{X}_1, ..., \bar{X}_n$ be i.i.d. real r.v.'s with distribution F and define $F_n = \frac{1}{n}\sum_{i=1}^{n} I(\bar{X}_i \le x)$. To apply Theorem 3.1, we rewrite $\sqrt{n}\hat{\sigma}^{-1}(T(F_n) - T(F)) = T/S$, where

$$T = \frac{\sqrt{n}(T(F_n) - T(F))}{\sigma} \quad \text{and} \quad S^2 = \frac{\hat{\sigma}^2}{\sigma^2}.$$

Let $\psi(t) = \int_0^t J(u)du$. From Lemma B of Serfling [(1980), p.265], we have

$$T(F_n) - T(F) = -\int_{-\infty}^{\infty} \Big\{\psi(F_n(x)) - \psi(F(x))\Big\}dx,$$

and hence we can write

$$T = \sqrt{n}(T(F_n) - T(F))/\sigma = W + \Delta_1,$$

where

$$
\begin{aligned}
W &= -\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n}\int_{-\infty}^{\infty}(I(X_i \le x) - F(x))J(F(x))dx := \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\alpha(X_j), \\
\Delta_1 &= -\frac{\sqrt{n}}{\sigma}\int_{-\infty}^{\infty}[\psi(F_n(x)) - \psi(F(x)) - (F_n(x) - F(x))J(F(x))]dx.
\end{aligned}
$$

And following the results in Wang, Jing and Zhao (2000), we write

$$
S^2 = 1 + n^{-3}\sum_{i\neq j\neq k}\gamma(X_i, X_j, X_k) + V_{2n},
$$

where

$$
\begin{aligned}
\gamma(X_i, X_j, X_k) &= \xi(X_i, X_j) + \varphi(X_i, X_j, X_k), \\
\xi(X_i, X_j) &= \sigma^{-2}\int\int J(F(s))J(F(t))(I(X_i \le \min(s,t))I(X_j > \max(s,t)) \\
&\quad -F(\min(s,t))(1 - F(\max(s,t))))dsdt, \\
\varphi(X_i, X_j, X_k) &= \sigma^{-2}\int\int J'(F(s))J(F(t))(I(X_i \le s) - F(s)) \\
&\quad I(X_j \le \min(s,t))I(X_k > \max(s,t))dsdt, \\
V_{2n} &= Q_1 + Q_2 + Q_3,
\end{aligned}
$$

which satisfies

$$
P\Big(\frac{\sqrt{n}\sigma^3}{E|X_1|^3}|Q_1 + Q_2| \ge 1\Big) \le \frac{A(J)}{\sqrt{n}}\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^2 \le \frac{1}{\sqrt{n}}\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{8/3}, \tag{5.40}
$$

$$
P\Big(\frac{\sqrt{n}\sigma^3}{E|X_1|^3}|Q_3| \ge 1\Big) \le \frac{A(J)}{\sqrt{n}}\frac{E|X_1|^2}{\sigma^2}\frac{E|X_1|^3}{\sigma^3} \le \frac{1}{\sqrt{n}}\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{7/3}, \tag{5.41}
$$

and

$$
|\alpha(X_j)| \le A(J)\sigma^{-1}(|X_j| + E|X_1|), \tag{5.42}
$$

$$
|\xi(X_j, X_k)| \le A(J)\sigma^{-2}(X_j^2 + X_k^2 + EX_1^2), \tag{5.43}
$$

$$
|\varphi(X_i, X_j, X_k)| \le A(J)\sigma^{-2}(X_j^2 + X_k^2). \tag{5.44}
$$

Noting $E\alpha^2(X_1) = 1$, it follows from (5.42) that

$$
A(J) \le \frac{EX_1^2}{\sigma^2} \le \Big(\frac{EX_1^3}{\sigma^3}\Big)^{2/3}.
$$

Therefore

$$
\begin{aligned}
&P(\sqrt{n}\hat{\sigma}^{-1}(T(F_n)-T(F))\geq x)\\
&= P(\frac{W+\Delta_1}{\sqrt{1+n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+V_{2n}}}\geq x)\\
&\leq P\Big(\frac{W+\Delta_1}{\sqrt{1+n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}}}\geq x\Big)+P\Big(|V_{2n}|\geq\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\Big)\\
&\leq P\Big(\frac{W+\Delta_1}{\sqrt{1+n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}}}\geq x\Big)\\
&\quad+\frac{1}{\sqrt{n}}\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{7/3}+\frac{1}{\sqrt{n}}\Big(\frac{E|X_1|^3}{\sigma^3}\Big)^{8/3}.
\end{aligned}
$$

Let

$$
\begin{aligned}
\Delta_2 &= n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3},\\
\Delta_2^{(l)} &= n^{-3}\sum_{i\neq j\neq k, i\neq l, j\neq l, k\neq l}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}.
\end{aligned}
$$

To complete the proof, we need to prove the following results:

$$
\sum_{l=1}^n\|\xi_l\|_p\|\Delta_1-\Delta_1^{(l)}\|_q \leq \frac{CA(J)}{\sqrt{n}\sigma^2}E|X_1|^2, \tag{5.45}
$$

$$
\sum_{l=1}^n\|\xi_l\|_p\|\Delta_2-\Delta_2^{(l)}\|_q \leq \frac{CA(J)^2}{\sqrt{n}\sigma^3}E|X_1|^3, \tag{5.46}
$$

$$
\begin{aligned}
|-\frac{x}{2}E\Delta_2 f_x(\bar{W})| &\leq \frac{xe^{-x/4}}{2}\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\\
&\quad+\frac{Cxe^{-x/12}}{n^{1/2}}\Big(\frac{A(J)}{\sigma^2}(E|X_1|^3)^{\frac{2}{3}}+\frac{A(J)^2}{\sigma^3}E|X_1|^3\Big).
\end{aligned} \tag{5.47}
$$

To prove (5.45), let

$$
\Delta_1^{(l)}=-\frac{\sqrt{n}}{\sigma}\int_{-\infty}^{\infty}\Big[\psi(F_{n,l}(x))-\psi(F(x))-(F_{n,l}(x)-F(x))J(F(x))\Big]dx,
$$

where $F_{n,l}(x)=\frac{1}{n}\Big(F(x)+\sum_{1\leq j\leq n, j\neq l}I(X_j\leq x)\Big)$. Then by Chen and Shao (2007), we get

$$
E|\Delta_1|^2\leq\frac{C}{n}\frac{EX_1^2}{\sigma^2},
$$

$$
E|\Delta_1-\Delta_1^{(l)}|^2\leq\frac{C}{n^2}\frac{EX_1^2}{\sigma^2}.
$$

Therefore

$$\sum_{l=1}^{n} \|\xi_l\|_p \|\Delta_1 - \Delta_1^{(l)}\|_q \leq \sum_{l=1}^{n} \frac{1}{n^{1/2}} \frac{A(J)}{\sigma} E(|X_1|^2)^{1/2} \frac{C}{n} \frac{(E|X_1|^2)^{1/2}}{\sigma} = \frac{CA(J)}{\sqrt{n}\sigma^2} E|X_1|^2,$$

where $p = 2$, $q = 2$.

To prove (5.46), take $l = 1$ as example, we get

$$\begin{aligned}
\Delta_2 - \Delta_2^{(1)} &= \frac{1}{n^3}\Big(\sum_{i \neq j \neq k} \gamma(X_i, X_j, X_k) - \sum_{i \neq j \neq k, i \neq 1, j \neq 1, k \neq 1} \gamma(X_i, X_j, X_k) \Big) \\
&= \frac{1}{n^3}\Big(\sum_{j \neq k \neq 1} \gamma(X_1, X_j, X_k) + \sum_{i \neq k \neq 1} \gamma(X_i, X_1, X_k) + \sum_{i \neq j \neq 1} \gamma(X_i, X_j, X_1) \Big) \\
&= \frac{1}{n^3}\Big(\sum_{j \neq k \neq 1} \xi(X_1, X_j) + \sum_{j \neq k \neq 1} \varphi(X_1, X_j, X_k) + \sum_{i \neq k \neq 1} \xi(X_i, X_1) \\
&\quad + \sum_{i \neq k \neq 1} \varphi(X_i, X_1, X_k) + \sum_{i \neq j \neq 1} \xi(X_i, X_j) + \sum_{i \neq j \neq 1} \varphi(X_i, X_j, X_1) \Big).
\end{aligned}$$

Therefore

$$\begin{aligned}
&E\big|\Delta_2 - \Delta_2^{(1)}\big|^{3/2} \\
&\leq \frac{C}{n^{9/2}}\Big(E\big| \sum_{j \neq k \neq 1} \xi(X_1, X_j) \big|^{3/2} + E\big| \sum_{j \neq k \neq 1} \varphi(X_1, X_j, X_k) \big|^{3/2} + E\big| \sum_{i \neq k \neq 1} \xi(X_i, X_1) \big|^{3/2} \\
&\quad + E\big| \sum_{i \neq k \neq 1} \varphi(X_i, X_1, X_k) \big|^{3/2} + E\big| \sum_{i \neq j \neq 1} \xi(X_i, X_j) \big|^{3/2} + E\big| \sum_{i \neq j \neq 1} \varphi(X_i, X_j, X_1) \big|^{3/2} \Big) \\
&\leq \frac{Cn^3}{n^{9/2}}\Big(E|\xi(X_1, X_j)|^{3/2} + E|\varphi(X_1, X_j, X_k)|^{3/2} + E|\xi(X_i, X_1)|^{3/2} \\
&\quad + E|\varphi(X_i, X_1, X_k)|^{3/2} + E|\xi(X_i, X_j)|^{3/2} + E|\varphi(X_i, X_j, X_1)|^{3/2} \Big) \\
&\leq \frac{C}{n^{3/2}}\Big(E|\xi(X_1, X_2)|^{3/2} + E|\varphi(X_1, X_2, X_3)|^{3/2} \Big) \\
&\leq \frac{C}{n^{3/2}} \frac{A(J)^{3/2}}{\sigma^3} E|X_1|^3,
\end{aligned}$$

then

$$\big(E\big|\Delta_2 - \Delta_2^{(1)}\big|^{3/2}\big)^{2/3} \leq \frac{C}{n} \frac{A(J)}{\sigma^2} (E|X_1|^3)^{2/3}.$$

Then note that

$$\xi_i = -\frac{1}{\sqrt{n}\sigma} \int_{-\infty}^{\infty} (I(X_i \leq x) - F(x)) J(F(x)) dx = \frac{1}{\sqrt{n}} \alpha(X_i),$$

we have

$$\begin{aligned}\sum_{l=1}^{n}\|\xi_l\|_p\|\Delta_2-\Delta_2^{(l)}\|_q &\le \sum_{l=1}^{n}\frac{1}{n^{1/2}}\frac{A(J)}{\sigma}E(|X_1|^3)^{1/3}\frac{C}{n}\frac{A(J)}{\sigma^2}(E|X_1|^3)^{2/3}\\ &= \frac{CA(J)^2}{\sqrt{n}\sigma^3}E|X_1|^3,\end{aligned}$$

where $p=3$, $q=3/2$.

To prove (5.47), now we calculate $E\Delta_x f_x(\bar{W})$. Note that here $\bar{\Delta}_x=\Delta_1-\frac{x}{2}\Delta_2$, where

$$\begin{aligned}\Delta_1 &= -\frac{\sqrt{n}}{\sigma}\int_{-\infty}^{\infty}[\psi(F_n(x))-\psi(F(x))-(F_n(x)-F(x))J(F(x))]\,dx,\\ \Delta_2 &= n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}.\end{aligned}$$

As proved before, we have $E|f_x(\bar{W})|\le Ce^{-x/4}$, and

$$E|\Delta_1|^2\le\frac{C}{n}\frac{EX_1^2}{\sigma^2}, \tag{5.48}$$

therefore

$$\begin{aligned}E|\Delta_1 f_x(\bar{W})| &\le \left(Ce^{-x/4}\frac{C}{n}\frac{EX_1^2}{\sigma^2}\right)^{1/2},\\ &\le \frac{Ce^{-x/8}}{\sqrt{n}}\frac{(EX_1^2)^{1/2}}{\sigma}.\end{aligned} \tag{5.49}$$

Note that

$$\begin{aligned}-\frac{x}{2}E\Delta_2 f_x(\bar{W}) &= -\frac{x}{2}E\left(n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right)f_x(\bar{W})\\ &= -\frac{x}{2}E\left(\frac{1}{n^3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)\right)f_x(\bar{W})-\frac{x}{2}\frac{E|X_1|^3}{\sqrt{n}\sigma^3}Ef_x(\bar{W})\\ &:= H_1+H_2,\end{aligned}$$

where

$$|H_2|\le E\left|\frac{x}{2\sqrt{n}}Ef_x(\bar{W})\right|\le\frac{xe^{-x/4}}{2}\frac{E|X_1|^3}{\sqrt{n}\sigma^3}.$$

To estimate H_1, we first write

$$\begin{aligned}
& n^{-3} \sum_{i \neq j \neq k} \gamma(X_i, X_j, X_k) \\
= & \frac{6}{n^3} \sum_{i<j<k} \gamma(X_i, X_j, X_k) \\
= & \frac{1}{n^3} \sum_{i<j<k} [\gamma(X_i, X_j, X_k) + \gamma(X_i, X_k, X_j) + \gamma(X_j, X_i, X_k) \\
& + \gamma(X_j, X_k, X_i) + \gamma(X_k, X_i, X_j) + \gamma(X_k, X_j, X_i)] \\
:= & \frac{1}{n^3} \sum_{i<j<k} h(X_i, X_j, X_k),
\end{aligned}$$

where $h(X_i, X_j, X_k)$ is a symmetric function satisfying $Eh(X_i, X_j, X_k) = 0$. Then

$$\frac{n^3}{\binom{n}{3}} n^{-3} \sum_{i \neq j \neq k} \gamma(X_i, X_j, X_k) = \frac{1}{\binom{n}{3}} \sum_{i<j<k} h(X_i, X_j, X_k)$$

is a U-statistic with rank $r = 1$. Then by Hoeffding's representation, we let

$$\begin{aligned}
g_1(x_1) &= Eh(x_1, X_2, X_3), \\
g_2(x_1, x_2) &= Eh(x_1, x_2, X_3) - g_1(x_1) - g_1(x_2), \\
g_3(x_1, x_2, x_3) &= Eh(x_1, x_2, x_3) - \sum_{i=1}^{3} g_1(x_i) - \sum_{1 \leq i < j \leq 3} g_2(x_i, x_j), \\
U_{nc} &= \binom{n}{c}^{-1} \sum_{1 \leq i_1 < \ldots < i_c \leq n} g_c(X_{i1}, ..., X_{ic}), \text{ for } c = 1, 2, 3.
\end{aligned}$$

Then we get

$$\begin{aligned}
& \frac{1}{\binom{n}{3}} \sum_{i<j<k} h(X_i, X_j, X_k) \\
= & \sum_{c=1}^{3} \binom{3}{c} U_{nc} \\
= & \frac{3}{n} \sum_{j=1}^{n} g_1(X_j) + 3 \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} g_2(X_i, X_j) + \binom{n}{3}^{-1} \sum_{1 \leq i < j < k \leq n} g_3(X_i, X_j, X_k) \\
:= & \frac{3}{n} \sum_{j=1}^{n} g_1(X_j) + \Delta_2^*.
\end{aligned}$$

Note that

$$E \sup_{|t| \leq 1} |f'(\bar{W}^{(j)} + t)| \leq e^{-x/4},$$

we get

$$\begin{aligned}
& E\left(\frac{3}{n}\sum_{j=1}^{n} g_1(X_j) f(\bar{W})\right) \\
= \quad & \frac{3}{n} E \sum_{j=1}^{n} g_1(X_j)\left(f(\bar{W}) - f(\bar{W}^{(j)})\right) \\
\leq \quad & \frac{3}{n}\sum_{j=1}^{n} E|g_1(X_j)\bar{\xi}_j| E\left(\sup_{|t|\leq 1}\left|f'\left(\bar{W}^{(j)}+t\right)\right|\right) \\
\leq \quad & \frac{Ce^{-x/4}}{n}\sum_{j=1}^{n} E|g_1(X_j)\bar{\xi}_j| \\
\leq \quad & \frac{Ce^{-x/4}}{n}\sum_{j=1}^{n} (E|g_1(X_j)|^{3/2})^{2/3}(E|\bar{\xi}_j|^3)^{1/3} \\
\leq \quad & \frac{Ce^{-x/4}A(J)^2}{\sqrt{n}\sigma^3} E|X_1|^3.
\end{aligned}$$

Using Theorem 2.1.3 in Koroljuk and Borovskich (1994), we get

$$\begin{aligned}
& E\left|\frac{1}{\binom{n}{3}}\sum_{i<j<k} h(X_i, X_j, X_k)\right|^{3/2} \\
\leq \quad & \sqrt{3}\sum_{c=1}^{3}\binom{3}{c}^{3/2}\binom{n}{c}^{-1/2}(\sqrt{2})^{c+1}E|g_c|^{3/2}. \qquad (5.50)
\end{aligned}$$

Therefore

$$\begin{aligned}
E|\Delta_2^*|^{3/2} \quad \leq \quad & \frac{36}{\sqrt{n(n-1)}}E|g_2(X_1,X_2)|^{3/2} + \frac{12\sqrt{2}}{\sqrt{n(n-1)(n-2)}}E|g_3(X_1,X_2,X_3)|^{3/2} \\
\leq \quad & C\left(\frac{1}{n}E|g_2(X_1,X_2)|^{3/2} + \frac{1}{n^{3/2}}E|g_3(X_1,X_2,X_3)|^{3/2}\right) \\
\leq \quad & C\left(\frac{1}{n}+\frac{1}{n^{3/2}}\right)E|\gamma(X_1,X_2,X_3)|^{3/2} \\
\leq \quad & \frac{CA(J)^{3/2}}{n\sigma^3}E|X_1|^3.
\end{aligned}$$

Then we get

$$E\left|x\Delta_2^* f_x(\bar{W})\right| \leq x\left(E|\Delta_2^*|^{\frac{3}{2}}\right)^{\frac{2}{3}}\left(E|f_x(\bar{W})|^3\right)^{\frac{1}{3}} \leq \frac{Cxe^{-x/12}A(J)}{n^{2/3}\sigma^2}\left(E|X_1|^3\right)^{2/3}.$$

Then we get

$$\begin{aligned}|H_1| &\leq \left|\frac{x}{2}E\left(\frac{1}{n^3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)\right)f_x\left(\bar{W}\right)\right| \\ &\leq \left|Cx\left(\frac{3}{n}\sum_{j=1}^{n}g_1(X_j)+\Delta_2^*\right)f_x\left(\bar{W}\right)\right| \\ &\leq \frac{Cxe^{-x/12}}{n^{1/2}}\left(\frac{A(J)}{\sigma^2}(E|X_1|^3)^{2/3}+\frac{A(J)^2}{\sigma^3}E|X_1|^3\right).\end{aligned}$$

Then we get the proof of (5.47). Next we estimate

$$E\left|\bar{\Delta}_1\right|^2+\frac{1}{1+x}E\left|\bar{\Delta}_2\right|^2+e^{-x/2}E\bar{\Delta}_2^2e^{\bar{W}}.$$

Using (5.50) again, we get

$$\begin{aligned}E\left|\Delta_2\right|^{3/2} &= E\left|n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right|^{3/2} \\ &\leq CE\left|\frac{1}{\binom{n}{3}}\sum_{i<j<k}h(X_i,X_j,X_k)\right|^{3/2}+\frac{C}{n^{3/4}}\frac{E|X_1|^3}{\sqrt{n}\sigma^3} \\ &\leq C\left(\frac{1}{\sqrt{n}}+\frac{1}{n}+\frac{1}{n^{3/2}}\right)E\left|\gamma(X_1,X_2,X_3)\right|^{3/2}+\frac{C}{n^{3/4}}\frac{E|X_1|^3}{\sqrt{n}\sigma^3} \\ &\leq \frac{CA(J)^{3/2}}{\sqrt{n}\sigma^3}E|X_1|^3.\end{aligned}$$

For $|\bar{\Delta}_1|\leq\frac{x+1}{8}$ and $|\bar{\Delta}_2|\leq\frac{1}{4}$, we get

$$\begin{aligned}&E\left|\bar{\Delta}_1\right|^2+\frac{1}{1+x}E\left|\bar{\Delta}_2\right|^2 \\ &\quad\leq E|\Delta_1|^2+\frac{C}{1+x}E|\Delta_2|^{3/2} \\ &\quad\leq \frac{C}{n}\frac{EX_1^2}{\sigma^2}+\frac{CA(J)^{3/2}}{(1+x)\sqrt{n}\sigma^3}E|X_1|^3. \qquad (5.51)\end{aligned}$$

Noting

$$\Delta_2=n^{-3}\sum_{i\neq j\neq k}\gamma(X_i,X_j,X_k)+1/\sqrt{n}=\frac{C}{n}\sum_{j=1}^{n}g_1(X_j)+\Delta_2^*+\frac{E|X_1|^3}{\sqrt{n}\sigma^3}, \quad (5.52)$$

it follows from (5.50) and $|X_i| \le \sqrt{n}\sigma$ that

$$
\begin{aligned}
&E\bar{\Delta}_2^2 e^{\bar{W}} \\
= \;& E\left|\frac{C}{n}\sum_{j=1}^n g_1(X_j) + \Delta_2^* + \frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right|^2 I\left(\left|\frac{C}{n}\sum_{j=1}^n g_1(X_j) + \Delta_2^* + \frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right| \le 1/4\right) e^{\bar{W}} \\
\le \;& CE\left|\frac{C}{n}\sum_{j=1}^n g_1(X_j)\right|^2 e^{\bar{W}} + CE\left|\Delta_2^*\right| e^{\bar{W}} + C\left(\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right)^2 Ee^{\bar{W}} \\
\le \;& \frac{C}{n^2}E\left|\sum_{j=1}^n g_1(X_j)\right|^2 e^{\bar{W}} + C\|\Delta_2^*\|_{3/2}\|e^{\bar{W}}\|_3 + C\left(\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right)^2 \\
\le \;& \frac{C}{n^2}E\left(\sum_{j=1}^n g^2(X_j) + \sum_{j\neq k} g(X_j)g(X_k)\right) e^{\bar{W}} + \left(\frac{CA(J)^{3/2}E|X_1|^3}{n\sigma^3}\right)^{2/3} + C\left(\frac{E|X_1|^3}{\sqrt{n}\sigma^3}\right)^2 \\
\le \;& \frac{C}{n^2}\sum_{j=1}^n Eg^2(X_j)e^{\bar{\xi}_j}Ee^{\bar{W}^{(j)}} + \frac{C}{n^{2/3}}\left(\frac{E|X_1|^3}{\sigma^3}\right)^{4/3} + \frac{C}{n}\left(\frac{E|X_1|^3}{\sigma^3}\right)^2 \\
\le \;& \frac{C}{\sqrt{n}}\left(\frac{E|X_1|^3}{\sigma^3}\right)^{7/3} + \frac{C}{n^{2/3}}\left(\frac{E|X_1|^3}{\sigma^3}\right)^{4/3} + \frac{C}{n}\left(\frac{E|X_1|^3}{\sigma^3}\right)^2. \qquad (5.53)
\end{aligned}
$$

Combining (5.39) (5.40), (5.41), (5.45), (5.46), (5.47), (5.49), (5.51) and (5.53), we get

$$
\begin{aligned}
&|H(x) - \Phi(x)| \\
\le \;& \frac{E|X_1|^3}{n^{1/2}\sigma^3} + \frac{C}{\sqrt{n}}(\frac{E|X_1|^3}{\sigma^3})^{8/3} + \frac{C}{\sqrt{n}}(\frac{E|X_1|^3}{\sigma^3})^{7/3} + \frac{Ce^{-x/8}}{\sqrt{n}}(\frac{E|X_1|^3}{\sigma^3})^{4/3} + \frac{C}{\sqrt{n}}(\frac{E|X_1|^3}{\sigma^3})^{7/3} \\
&+\frac{xe^{-x/4}}{2}\frac{E|X_1|^3}{\sqrt{n}\sigma^3} + \frac{Cxe^{-x/12}}{n^{1/2}}((\frac{E|X_1|^3}{\sigma^3})^{4/3} + (\frac{E|X_1|^3}{\sigma^3})^{7/3}) + \frac{C}{n}(\frac{E|X_1|^3}{\sigma^3})^{2/3} \\
&+\frac{C}{\sqrt{n}(1+|x|^3)}(\frac{E|X_1|^3}{\sigma^3})^3 + \frac{C}{\sqrt{n}}(\frac{E|X_1|^3}{\sigma^3})^{7/3} + \frac{C}{n^{2/3}}(\frac{E|X_1|^3}{\sigma^3})^{4/3} + \frac{C}{n}(\frac{E|X_1|^3}{\sigma^3})^2 \\
\le \;& \frac{C}{\sqrt{n}}\max\left\{(\frac{E|X_1|^3}{\sigma^3})^{1/3}, (\frac{E|X_1|^3}{\sigma^3})^3\right\}.
\end{aligned}
$$

Chapter 6

Conclusion and Future Work

In this thesis, we extend the use of Stein's method. We obtains a non-uniform Berry-Esseen bound for a class of Studentized statistics via Stein's method and a randomized concentration inequality. In Chapter 5, we apply the Berry-Esseen bound we obtained to three studentized statistics. The result of Student's t-statistics, Studentized U-statistics and L-statistics are as good as the existing ones.

In the future, we wish to apply our result to some other studentized statistics to get more applications. For example, we will try to apply it to a studentized random sum and this is still under study. We are also interested in applying our result to a U-statistics $U_n = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} h(X_i, X_j)$ to check whether we can get $|P(U_n \le x) - \Phi(x)| \le \frac{C}{\sqrt{n}}$ under conditions $E|h(X_1, X_2)|^{5/3} < \infty$ and $E|g(X)|^3 < \infty$, which are the weakest moment conditions for U-statistics. In addition, we used a randomized concentration inequality in our proof. This inequality may have more applications in proving Berry-Esseen bounds for some other class of statistics in the future.

Bibliography

[1] Arratia, R., Goldstein, L. and Gordon, L. (1990). Poisson approximation and the Chen-Stein method. *Statist. Sci.* **5**, 403-434.

[2] Arvesen, J.N. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* **40**, 2076-2100.

[3] Baldi, P., Rinott, Y. and Stein, C. (1989). A normal approximation for the number of local maxima of a random function on a graph. In: *Probability, statistics, and mathematics. Papers in honor of Samuel Karlin*, Eds: Anderson, T. W., Athreya, K. B., Iglehart, D. L., pp. 59-81. Academic Press, Boston.

[4] Barbour, A.D. and Hall, P. (1984). Stein's method and the Berry-Esseen theorem. *Austral. J. Statist.* **26**, 8-15.

[5] Barbour, A. D. (1988). Stein's method and Poisson process convergence. *J. Appl. Probab.* **25A**, 175-184.

[6] Barbour, A. D. (1990). Stein's method for diffusion approximations. *Probability Theory and Related Fields* **84**, 297-322.

[7] Barbour, A. D., Holst, L. and Janson, S. (1992). Poisson approximation. Oxford Studies in Probability. 2. The Clarendon Press Oxford University Press.

[8] Barbour, A.D. and Chen, L.H.Y (2005). An Introduction to Stein method. Lecture Notes Series **4**, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific.

[9] Barbour, A.D. and Chen, L.H.Y (2005). Stein's Method and Application. Lecture Notes Series **5**, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific.

[10] P. Van Beeck (1972). An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality. *Z. Wahrsch. Verw. Geb.* **23**, 187-196.

[11] Bentkus, V., Bloznelis, M. and Götze, F. (1996). A Berry-Esseen bound for Student's statistic in the non-i.i.d. case. *J. Theoret. Probab.* **9**, 765-796.

[12] Bentkus, V., and Götze, F. (1996). The Berry-Esseen bound for Student's statistic. *Ann. Prob.* **24**, 491-503.

[13] Callaert, H. and Veraverbeke, N. (1981). The order of the normal approximation for a Studentized statistic. *Ann. Statist.* **9**, 194-200.

[14] Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **35**, 1584-1610.

[15] Chatterjee and Shao, Q.M. (2011). Stein's method of exchangeable pairs with application to the Curie-Weiss model.

[16] Chen, L.H.Y. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3**, 534-545.

[17] Chen, L.H.Y. (1993). Extending the Poisson approximation. *Science.* **262**, 379-380.

[18] Chen, L.H.Y. (1998) Stein's method: some perspectives with applications. In: *Probability towards 2000,* Eds: L. Accardi and C.C. Heyde, Lecture Notes in Statistics **128**, 97-122. Springer-Verlag, Berlin.

[19] Chen, L.H.Y. (2000) Non-uniform bounds in probability approximations using Stein's method. In: *Probability and statistical models with applications: a volume in honor of Theophilos Cacoullos*, Eds: N. Balakrishnan, Ch. A. Charalambides and M.V. Koutras, pp. 3-14. Chapman and Hall/CRC Press, Boca Raton, Florida.

[20] Chen, L.H.Y., Fang, X. and Shao, Q.M. (2012) From Stein identities to moderate deviations. *Ann. Probab.* To appear.

[21] Chen, L.H.Y. and Shao, Q.M. (2001). A non-uniform Berry-Esseen bound via Stein's method. *Probab. Theory Relat. Fields* **120**, 236-254.

[22] Chen, L.H.Y. and Shao, Q.M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32**, 1985-2028.

[23] Chen, L.H.Y. and Shao, Q.M. (2005). Stein's method for normal approximation. *An Introduction to Stein's Method* (A.D. Barbour and L.H.Y. Chen eds). Lecture Notes Series, Institute for Mathematical Sciences, NUS, Vol. 4, p. 1-59. World Scientific.

[24] Chen, L.H.Y. and Shao, Q.M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13**, 581-599.

[25] Chen, L.H.Y., Goldstein, L. and Shao, Q.M. (2011). Normal Approximation by Stein's Method. Springer-verlag, Berlin Heidelberg.

[26] Dembo, A. and Y. Rinott, Y. (1996). Some examples of normal approximations by Stein's method. Random discrete structures. IMA Vol. Math. Appl., **76**, 25-44.

[27] Diaconis, P. (1977). The distribution of leading digits and uniform idstribution mod 1. *Ann. Probab.* **5**, 72-81.

[28] Diaconis, P. (1989). An example for Stein's method. Stanford Stat. Dept. Technical Report.

[29] Diaconis, P. (2004). Stein's method for Markov chains: first examples. In: *Stein's Method: Expository Lectures and Applications*, Eds: Diaconis, P. and Holmes, S., pp. 27-43. IMS Lecture Notes **46**, Beachwood, Ohio.

[30] Diaconis, P. and Holmes, S. (2004). Stein's method: expository lectures and applications. IMS Lecture Notes, Vol. **46**, Hayward, CA.

[31] Gine, E., Gotze, F. and Mason, D.M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25**, 1514-1531.

[32] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Probab.* **25**, 1514-1531.

[33] Goldstein L. (2005) Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **42**.

[34] Hall, P. (1988). On the effect of random norming on the rate of convergence in the central limit theorem. *Ann. Probab.* **16**, 1264-1280.

[35] Helmers, R. (1982). Edgeworth Expansions for Linear Combinations of Order Statistics. *Math. Centre Tracts* **105**. Math. Centrum, Amsterdam.

[36] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19**, 293-325.

[37] Koroljuk, V.S. and Borovskich, Yu. V. (1994). Theory of U-statistics. Kluwer Academic Publishers, Dordrecht.

[38] Lai, T.L., de la Pena, V. and Shao, Q. M. (2009). Self-normalized Processes: Limit Theory and Statistical Applications. Springer Series in Probability and its Applications, Springer-Verlag, New York.

[39] Lai, T.L, Shao, Q.M. and Wang, Q. (2009). Cramér type moderate deviations for Studentized U-statistics. ESAIM: Probab. Statist.

[40] Nourdin, I. and Pecatti, G. (2009). Stein's method on Weiner Chaos. Probab. Th. Related Fields **145**, 75-118.

[41] Petrov, V.V. (1975). Sums of Independent Random Variables. Springer-Verlag, New York.

[42] Petrov, V.V. (1995). Limit theorems of probability theory: sequences of independent random variables. Oxford studies in Probability 4, Clarendon Press, Oxford.

[43] Raic, M. (1975). CLT-related large deviation bounds based on Stein's method. Adv. Appl. Probab. **39**, 731-752.

[44] Rinott, Y. and Rotar, V. (1996). A multivariate CLT for local dependence with $n^{1/2} \log n$ rate and pplications to multivate graph related statistics. *J. Multiv. Anal.* **56**, 333-350.

[45] Rinott, Y. and Rotar, V. (2000). Normal approximations by Stein's method. *Decis. Econ. Finance.* **23**, 15-29.

[46] Robert J. Serfling. (1980). Approximation Theorem of Mathematical Statistics.

[47] Slavova, V.V. (1985). On the Berry-Esseen bound for Student's statistics. *Lecture Notes in Math.* 1155 355-390. Springer, Berlin.

[48] Shao, Q.M. (1997). Self-normalized large deviations. Ann. Probab. **25**, 285-328.

[49] Shao, Q.M. (1999). Cramer-type large deviation for Student's t statistic. J. Theorect. Probab. **12**, 387-398.

[50] Shao, Q.M. (2005). An explicit Berry - Esseen bound for the Student t-statistic via Stein's method. *Stein's Method and Applications* (A.D. Barbour and L.H.Y. Chen eds). Lecture Notes Series, Institute for Mathematical Sciences, NUS, Vol. 5, p. 143-155. World Scientific.

[51] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proc. Sixth Berkeley Symp. Math. Stat. Prob. **2**, 583-602, Univ. California Press. Berkeley, Calif.

[52] Stein, C. (1986). Approximate computation of expectations. Institute of Mathematical Statistics Lecture Notes, Monograph Series, 7. Hayward, California.

[53] Wang, Q. and Jing, B.Y. (1999). An exponential non-uniform Berry-Esseen bound for self-nromalized sums. *Ann. Probab.* **27**, 2068-2088.

[54] Wang, Q., Jing, B.Y. and Zhao, L. (2000). The Berry-Esseen bound for studentized statistics. *Ann. Probab.* **28**, 511-535.

[55] Wang, Q. and Weber, N.C. (2006). Exact convergence rate and leading term in the central limit theorem for U-statistics. *Statist. Sinica.* **16**, 1409-1422.

[56] Zhao, L. (1983). The rate of the normal approximation for a Studentized U-statistic. *Science Exploration* **3** 45-52.