

Berry-Esseen Bound for Independent Random Sum via Stein's Method

N. Chaidee and K. Neammanee

Department of Mathematics, Faculty of Science
Chulalongkorn University, Bangkok 10330, Thailand
nattakarn.c@chula.ac.th

Abstract

Let (X_n) be a sequence of independent identically distributed random variables with zero means and finite third moments and N a positive integral-valued random variable such that N, X_1, X_2, \dots are independent. In this work, we apply a new technique, which is called Stein's method, to give uniform and non-uniform bounds in normal approximation of a random sum $X_1 + X_2 + \dots + X_N$.

Mathematics Subject Classification: 60F05, 60G05

Keywords: normal approximation, Stein's method, random sum, concentration inequality, uniform and non-uniform bounds

1 Introduction

One of the most important theorem in probability theory and statistics is the central limit theorem (CLT) which is stated as follows.

Let X_1, X_2, \dots, X_n be independent random variables with $EX_i = 0$ and $EX_i^2 = \sigma_i^2 < \infty$. Then for

$$W_n = \frac{X_1 + X_2 + \dots + X_n}{\sqrt{s_n^2}} \quad \text{and} \quad s_n^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$$

we have

$$P(W_n \leq x) \rightarrow \Phi(x) \quad \text{as} \quad n \rightarrow \infty$$

where $\Phi(x)$ is the standard normal distribution function, that is,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

To apply the CLT we always need to know a bound between $P(W_n \leq x)$ and $\Phi(x)$, i.e., $|P(W_n \leq x) - \Phi(x)|$. Berry([2]) and Esseen([6]) are the first two mathematicians who gave this bound. Their bound is of the form

$$\sup_{x \in \mathbb{R}} |P(W_n \leq x) - \Phi(x)| \leq \frac{C}{s_n^3} \sum_{i=1}^n E|X_i|^3. \quad (1)$$

Many authors([2], [6], [20]) find and improve the constant C in (1) and the best C is 0.7655 which is obtained by Shiganov in 1986([12]) in case of X_1, X_2, \dots, X_n are identically and 0.7915 in case of non-identically([12]).

In many situations, we need to approximation the probability of random sum. Let X_1, X_2, \dots be independent identically distributed random variables with $EX_i = 0$ and $E|X_i|^3 = \gamma < \infty$. Let N be a positive integral-valued random variable with $EN < \infty$ and N, X_1, X_2, \dots are independent.

Let

$$W = \frac{X_1 + X_2 + \dots + X_N}{\sqrt{EN}\sigma}$$

where $\sigma^2 = EX_i^2$ for all $i = 1, 2, \dots$.

The CLT for W are obtained by many mathematicians(see, [14], [1], [22], [9], [13], [15], [11] and [5] for examples). In this work, we concerns on uniform and non-uniform bounds between $P(W \leq x)$ and $\Phi(x)$ by using a new method, which is called Stein method.

Normally, the method which is used in our literatures is Fourier transform while in this paper we will use a new method which introduced by Stein in 1972([7]) to obtain bounds of $|P(W \leq x) - \Phi(x)|$. Stein's method was introduced by Stein in 1972 for normal approximation of dependent random variables. This method use differential equation which always call Stein's equation in stead of Fourier transform. There are three approaches of Stein's method for normal approximation, namely, inductive approach(for example, see [4], [10]), coupling approach(for example, see [16], [8]) and concentration inequality approach(for example, see [21]). The concentration inequality approach was originally used by Stein for independent and identically distributed random variable([21]) and it was extended by Chen in([18], [19], [17]). In this paper, we use concentration inequality approach which used in [18] to find bounds of $|P(W \leq x) - \Phi(x)|$. A uniform bound in normal approximation is in section 2 while a non-uniform bound is in section 3.

2 Uniform Bound

Let X_1, X_2, \dots be independent identically distributed random variables with $EX_i = 0$, $EX_i^2 = \sigma^2$ and $E|X_i|^3 = \gamma < \infty$. Let N be an integral-valued

random variable with $EN^{3/2} < \infty$ and N, X_1, X_2, \dots are independent. Define

$$Y_i = \frac{X_i}{\sqrt{EN}\sigma}$$

and

$$W = Y_1 + Y_2 + \dots + Y_N.$$

Then $EY_i = 0$ and

$$EW^2 = E\left(\sum_{i=1}^N Y_i\right)^2 = \sum_{n=1}^{\infty} P(N = n) \sum_{i=1}^n EY_i^2 = 1.$$

In this section, we provide a uniform bound between $P(W \leq x)$ and $\Phi(x)$. Our main result is in Theorem 2.1.

Theorem 2.1. *There exists a constant C such that*

$$\sup_{x \in \mathbb{R}} |P(W \leq x) - \Phi(x)| \leq 10.32\delta + \frac{0.125}{\sqrt{EN}} + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}} + \frac{E|N - EN|}{EN}$$

where $\delta = \frac{\gamma}{\sqrt{EN}\sigma^3}$.

Corollary 2.2. *If $N = n$, then there exists a constants C such that*

$$\sup_{x \in \mathbb{R}} |P(W \leq x) - \Phi(x)| \leq \frac{11.82\gamma}{\sigma^3\sqrt{n}} \left(1 + \frac{1}{\sqrt{n}}\right).$$

To prove Theorem 2.1, we need the following concentration inequality.

Proposition 2.3. *(Uniform Concentration Inequality). Let*

$$W_n = \sum_{i=1}^n Y_i \quad \text{and} \quad W_n^{(i)} = W_n - Y_i \quad \text{for } i = 1, 2, \dots, n.$$

If $\delta < \frac{1}{20}$, then for $a < b$

$$P(a \leq W_n^{(i)} \leq b) \leq \frac{2.5EN}{n} \left[\left\{ \frac{1}{2}(b - a) + \delta \right\} \sqrt{\frac{n}{EN}} + 2\delta + \frac{\delta}{\sqrt{EN}} \right].$$

Proof. In [18], Chen and Shao defined $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(t) = \begin{cases} -\frac{1}{2}(b-a) - \delta & \text{for } t < a - \delta, \\ t - \frac{1}{2}(b+a) & \text{for } a - \delta \leq t \leq b + \delta, \\ \frac{1}{2}(b-a) + \delta & \text{for } t \geq b + \delta \end{cases}$$

and showed that

$$EW_n^{(i)} f(W_n^{(i)}) \geq E\{I(a \leq W_n^{(i)} \leq b)S\} - P(a \leq W_n^{(i)} \leq b)E(|Y_i| \min\{\delta, |Y_i|\}), \quad (2)$$

$$ES \geq \sum_{j=1}^n \frac{EX_j^2}{EN\sigma^2} - \frac{1}{2\delta} \sum_{j=1}^n \frac{E|X_j|^3}{(EN\sigma^2)^{3/2}} \quad (3)$$

and

$$|EW_n^{(i)} f(W_n^{(i)})| \leq \left\{ \frac{1}{2}(b-a) + \delta \right\} E|W_n^{(i)}| \quad (4)$$

where

$$S = \sum_{j=1}^n |Y_j| \min\{\delta, |Y_j|\},$$

$$I(A)(w) = \begin{cases} 1 & \text{if } w \in A, \\ 0 & \text{if } w \notin A. \end{cases}$$

By (3),

$$ES \geq \frac{n}{EN} - \frac{n\gamma}{2\delta(EN\sigma^2)^{3/2}} = \frac{n}{2EN}. \quad (5)$$

From this fact and the fact that

$$\begin{aligned} ES^2 &= E\left(\sum_{j=1}^n |Y_j| \min\{\delta, |Y_j|\}\right)^2 \\ &= \sum_{j=1}^n E(|Y_j| \min\{\delta, |Y_j|\})^2 + \sum_{\substack{i,j \\ i \neq j}} E[|Y_i| \min\{\delta, |Y_i|\}][|Y_j| \min\{\delta, |Y_j|\}] \\ &\leq \sum_{j=1}^n E(|Y_j| \min\{\delta, |Y_j|\})^2 + \left(\sum_{j=1}^n E|Y_j| \min\{\delta, |Y_j|\}\right)^2 \\ &\leq \delta^2 \sum_{j=1}^n E|Y_j|^2 + E^2 S \\ &= \frac{n\delta^2}{EN} + E^2 S, \end{aligned}$$

we have

$$\begin{aligned}
 P(S \leq \frac{0.4n}{EN}) &= P(E(S) - S \geq E(S) - \frac{0.4n}{EN}) \\
 &\leq P(E(S) - S \geq \frac{n}{10EN}) \\
 &\leq \frac{100(EN)^2}{n^2} E(S - E(S))^2 \\
 &= \frac{100(EN)^2}{n^2} (ES^2 - E^2S) \\
 &\leq \frac{5\delta EN}{n}.
 \end{aligned}$$

Hence, by (2) and the fact that for $s < t, y \geq 0, c > 0$,

$$I(s \leq w \leq t)y \geq c(I(s \leq w \leq t) - (1 - y/c)I(y \leq c)), \tag{6}$$

we have

$$\begin{aligned}
 EW_n^{(i)} f(W_n^{(i)}) &\geq \frac{0.4n}{EN} \{P(a \leq W_n^{(i)} \leq b) - E[(1 - \frac{2.5SEN}{n})I(S \leq \frac{0.4n}{EN})]\} \\
 &\quad - P(a \leq W_n^{(i)} \leq b)E(|Y_i| \min\{\delta, |Y_i|\}) \\
 &\geq \frac{0.4n}{EN} \{P(a \leq W_n^{(i)} \leq b) - EI(S \leq \frac{0.4n}{EN})\} - \delta E|Y_i| \\
 &\geq \frac{0.4n}{EN} \{P(a \leq W_n^{(i)} \leq b) - P(S \leq \frac{0.4n}{EN})\} - \frac{\delta}{\sqrt{EN}} \\
 &\geq \frac{0.4n}{EN} P(a \leq W_n^{(i)} \leq b) - 2\delta - \frac{\delta}{\sqrt{EN}}.
 \end{aligned}$$

This implies,

$$\begin{aligned}
 P(a \leq W_n^{(i)} \leq b) &\leq \frac{2.5EN}{n} \left\{ EW_n^{(i)} f(W_n^{(i)}) + 2\delta + \frac{\delta}{\sqrt{EN}} \right\} \\
 &\leq \frac{2.5EN}{n} \left\{ (\frac{1}{2}(b - a) + \delta) \sqrt{E|W_n^{(i)}|^2} + 2\delta + \frac{\delta}{\sqrt{EN}} \right\} \\
 &\leq \frac{2.5EN}{n} \left\{ (\frac{1}{2}(b - a) + \delta) \sqrt{\frac{n}{EN}} + 2\delta + \frac{\delta}{\sqrt{EN}} \right\}
 \end{aligned}$$

where we have used (4) in the second inequality. □

Proof of Theorem 2.1.

To bound $P(W \leq x) - \Phi(x)$, it suffices to consider $x \geq 0$ as we can apply result to $-W$ when $x < 0$. By [18], p. 246, we have

$$|P(W \leq x) - \Phi(x)| \leq 0.55 \quad \text{for } x \geq 0,$$

hence we only need to consider the case where

$$\delta < \frac{1}{20}.$$

Let f be the unique bounded solution f_x of the Stein equation

$$f'(w) - wf(w) = I(w \leq x) - \Phi(x). \quad (7)$$

If we replace w in (7) by W_n , then

$$Ef'(W_n) - EW_n f(W_n) = P(W_n \leq x) - \Phi(x).$$

Set $p_n = P(N = n)$. Then, by the fact that

$$EW_n f(W_n) = \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n^{(i)} + t) M_i(t) dt$$

(eq. (2.17), p.12 of [3]) where

$$M_i(t) = E[Y_i(I(0 \leq t \leq Y_i) - I(Y_i \leq t < 0))],$$

yields

$$\begin{aligned} & P(W \leq x) - \Phi(x) \\ &= \sum_{n=1}^{\infty} p_n P(W_n \leq x) - \sum_{n=1}^{\infty} p_n \Phi(x) \\ &= \sum_{n=1}^{\infty} p_n [P(W_n \leq x) - \Phi(x)] \\ &= \sum_{n=1}^{\infty} p_n [Ef'(W_n) - EW_n f(W_n)] \\ &= \sum_{n=1}^{\infty} p_n \left(\sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n) M_i(t) dt - \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n^{(i)} + t) M_i(t) dt \right) \\ &\quad + \sum_{n=1}^{\infty} p_n \left(Ef'(W_n) - \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n) M_i(t) dt \right) \\ &= A_1 + A_2 \end{aligned} \quad (8)$$

where

$$\begin{aligned} A_1 &= \sum_{n=1}^{\infty} p_n \left(\sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n) M_i(t) dt - \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n^{(i)} + t) M_i(t) dt \right), \\ A_2 &= \left| \sum_{n=1}^{\infty} p_n \left(Ef'(W_n) - \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n) M_i(t) dt \right) \right|. \end{aligned}$$

By noting that

$$f'(w + s) - f'(w + t) \leq \begin{cases} 1 & \text{if } w + s \leq x, w + t > x, \\ (|w| + 0.63)(|s| + |t|) & \text{if } s \geq t, \\ 0 & \text{otherwise} \end{cases}$$

(see [18], p. 247), we have

$$\begin{aligned} A_1 &= \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \left(\int_{-\infty}^{\infty} \left\{ f'(W_n^{(i)} + Y_i) - f'(W_n^{(i)} + t) \right\} M_i(t) dt \right) \\ &\leq \sum_{n=1}^{\infty} p_n \left(\sum_{i=1}^n E \int_{\substack{W_n^{(i)} + t > x \\ W_n^{(i)} + Y_i \leq x}} M_i(t) dt + \sum_{i=1}^n E \int_{Y_i \geq t} (|W_n^{(i)}| + 0.63)(|Y_i| + |t|) M_i(t) dt \right) \\ &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{t > Y_i} I(x - t < W_n^{(i)} \leq x - Y_i) M_i(t) dt \\ &\quad + \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{Y_i \geq t} (|W_n^{(i)}| + 0.63)(|Y_i| + |t|) M_i(t) dt \\ &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E E^{Y_i} \int_{t > Y_i} I(x - t < W_n^{(i)} \leq x - Y_i) M_i(t) dt \\ &\quad + \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E (|W_n^{(i)}| + 0.63) E \int_{-\infty}^{\infty} (|Y_i| + |t|) M_i(t) dt \\ &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{t > Y_i} E^{Y_i} I(x - t < W_n^{(i)} \leq x - Y_i) M_i(t) dt \\ &\quad + \sum_{n=1}^{\infty} p_n \sum_{i=1}^n (\sqrt{E(W_n^{(i)})^2} + 0.63)(E|Y_i|E|Y_i|^2 + \frac{1}{2}E|Y_i|^3) \\ &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{t > Y_i} P(x - t < W_n^{(i)} \leq x - Y_i | Y_i) M_i(t) dt \\ &\quad + \frac{3}{2} \sum_{n=1}^{\infty} p_n \sum_{i=1}^n (\sqrt{\frac{n}{EN}} + 0.63) E|Y_i|^3 \\ &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{-\infty}^{\infty} P(x - t < W_n^{(i)} \leq x - Y_i | Y_i) M_i(t) dt \\ &\quad + \frac{3\gamma}{2(EN)^{3/2}\sigma^3} \sum_{n=1}^{\infty} p_n n (\sqrt{\frac{n}{EN}} + 0.63) \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \int_{t>Y_i} \left[\frac{2.5EN}{n} \left(\left\{ \frac{1}{2}(|t| + |Y_i|) + \delta \right\} \sqrt{\frac{n}{EN}} + 2\delta + \frac{\delta}{\sqrt{EN}} \right) M_i(t) dt \right] \\
 &\quad + \frac{1.5\gamma EN^{3/2}}{(EN)^2\sigma^3} + \frac{0.945\gamma EN}{(EN)^{3/2}\sigma^3} \\
 &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n \left[\frac{2.5EN}{n} \left(\left\{ \frac{1}{2} \left(\frac{E|Y_i|^3}{2} + E|Y_i|E|Y_i|^2 \right) + \delta E|Y_i|^2 \right\} \sqrt{\frac{n}{EN}} + 2\delta E|Y_i|^2 \right. \right. \\
 &\quad \left. \left. + \frac{\delta E|Y_i|^2}{\sqrt{EN}} \right) \right] + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}} + 0.945\delta \\
 &= 9.375\delta + \frac{2.5\delta}{\sqrt{EN}} + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}} + 0.945\delta \\
 &\leq 10.32\delta + \frac{0.125}{\sqrt{EN}} + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}} \tag{9}
 \end{aligned}$$

where we have used Proposition 2.3 in the seventh inequality. Similar to (9), we can use the fact that

$$f'(w + s) - f'(w + t) \geq \begin{cases} -1 & \text{if } w + s > x, w + t \leq x, \\ -(|w| + 0.63)(|s| + |t|) & \text{if } s < t, \\ 0 & \text{otherwise,} \end{cases}$$

(see [18], p. 247) to prove that

$$A_1 \geq -10.32\delta - \frac{0.125}{\sqrt{EN}} - \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}}.$$

Hence

$$|A_1| \leq 10.32\delta + \frac{0.125}{\sqrt{EN}} + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}}.$$

By [18], p. 247,

$$|f'(w)| \leq 1 \text{ for all } w, \tag{10}$$

then

$$\begin{aligned}
 |A_2| &= \left| \sum_{n=1}^{\infty} p_n E f'(W_n) \left(1 - \sum_{i=1}^n E \int_{-\infty}^{\infty} M_i(t) dt \right) \right| \\
 &\leq \sum_{n=1}^{\infty} p_n \left| 1 - \sum_{i=1}^n E Y_i^2 \right| \\
 &= \sum_{n=1}^{\infty} p_n \left| 1 - \frac{n}{EN} \right|
 \end{aligned}$$

$$\begin{aligned}
 &= E\left|1 - \frac{N}{EN}\right| \\
 &= \frac{E|N - EN|}{EN}.
 \end{aligned} \tag{11}$$

By combining (8)-(11), we have

$$|P(W \leq x) - \Phi(x)| \leq 10.32\delta + \frac{0.125}{\sqrt{EN}} + \frac{1.5\delta EN^{3/2}}{(EN)^{3/2}} + \frac{E|N - EN|}{EN}.$$

□

3 Non-uniform Bound

In this section, we will give a non-uniform bound in normal approximation of

$$W = Y_1 + Y_2 + \dots + Y_N$$

and further assume that $EN^5 < \infty$.

Theorem 3.1. *There exists a constant C such that for every real number x ,*

$$|P(W \leq x) - \Phi(x)| \leq \frac{C\delta}{(1 + |x|)^3} \left(1 + \frac{EN^5}{(EN)^5}\right) + \frac{C\sqrt{(\text{Var}N)EN^2}}{(1 + |x|)^2(EN)^2}.$$

Corollary 3.2. *If $N = n$, then there exists a constants C such that for every real number x ,*

$$|P(W \leq x) - \Phi(x)| \leq \frac{C\delta}{(1 + |x|)^3}.$$

The following proposition concerns a concentration inequality for a non-uniform bound.

Proposition 3.3. *(Non-uniform Concentration Inequality) For $0 \leq a < b$, we have*

$$P(a \leq W_n^{(i)} \leq b) \leq \frac{C(b - a + \delta)}{(1 + a)^3} \left(\frac{EN}{n} + \frac{n^4}{(EN)^4}\right).$$

where C is a constant.

Proof. In the proof of non-uniform concentration inequality in [18], p. 241, Chen and Shao defined

$$\begin{aligned}
 Z_j &= Y_j I(|Y_j| \leq 1 + a) - EY_j I(|Y_j| \leq 1 + a), \\
 T_n^{(i)} &= \sum_{\substack{j=1 \\ j \neq i}}^n Z_j \quad \text{and} \\
 r &= \sum_{\substack{j=1 \\ j \neq i}}^n EY_j I(|Y_j| > 1 + a),
 \end{aligned}$$

and showed that

$$P(a \leq W_n^{(i)} \leq b) \leq P(a+r \leq T_n^{(i)} \leq b+r) + P(\max_{1 \leq j \leq n} |Y_j| > 1+a).$$

Hence, by the fact that

$$P(\max_{1 \leq j \leq n} |Y_j| > 1+a) \leq \frac{1}{(1+a)^3} \sum_{j=1}^n E|Y_j|^3 = \frac{n\delta}{EN(1+a)^3},$$

we have

$$P(a \leq W_n^{(i)} \leq b) \leq P(a+r \leq T_n^{(i)} \leq b+r) + \frac{n\delta}{EN(1+a)^3}. \quad (12)$$

To prove this proposition, it suffices to bound

$$P(a+r \leq T_n^{(i)} \leq b+r).$$

By Rosenthal inequality (see, for example, [23], p.59) and the fact that

$$E|T_n^{(i)}|^2 \leq \frac{n}{EN}, \quad (13)$$

we have

$$\begin{aligned} E|T_n^{(i)}|^4 &\leq C \left\{ (E|T_n^{(i)}|^2)^2 + \sum_{\substack{j=1 \\ j \neq i}}^n E|Y_j|^4 I(|Y_j| \leq 1+a) \right\} \\ &\leq C \left\{ (nEY_1^2)^2 + n(1+a)E|Y_1|^3 \right\} \\ &\leq C \left\{ \frac{n^2}{(EN)^2} + \frac{n(1+a)\delta}{EN} \right\}. \end{aligned} \quad (14)$$

Case 1. $(1+a)\delta \geq 1$.

By the fact that $r \leq \frac{n}{EN}$ and $\frac{1}{1+a} < \delta$, we have

$$\begin{aligned} P(a+r \leq T_n^{(i)} \leq b+r) &\leq P(1+a \leq 1-r+T_n^{(i)}) \\ &\leq \frac{E|1-r+T_n^{(i)}|^4}{(1+a)^4} \\ &\leq \frac{C(1+r^4+E|T_n^{(i)}|^4)}{(1+a)^4} \\ &\leq \frac{C}{(1+a)^4} + \frac{Cn^4}{(EN)^4(1+a)^4} + \frac{C}{(1+a)^4} \left\{ \frac{n^2}{(EN)^2} + \frac{n(1+a)\delta}{EN} \right\} \\ &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{n}{EN} + \frac{n^2}{(EN)^2} + \frac{n^4}{(EN)^4} \right\} \\ &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{n^4}{(EN)^4} \right\} \\ &\leq \frac{C\delta}{(1+a)^3} \left\{ \frac{EN}{n} + \frac{n^4}{(EN)^4} \right\}. \end{aligned} \quad (15)$$

Case 2. $(1 + a)\delta < 1$.

Follows the argument in [18], p. 243 and choose C in (6) is $\frac{n-1}{4EN}$, we can show that

$$E(T_n^{(i)} f(T_n^{(i)})) \geq \frac{(n-1)(1+a)^3}{4EN} \left\{ P(a+r \leq T_n^{(i)} \leq b+r) - P(U \leq \frac{n-1}{4EN}) \right\} \tag{16}$$

where $U = \sum_{\substack{j=1 \\ j \neq i}}^n \eta_j$ and $\eta_j = |Z_j| \min(\delta, |Z_j|)$ and

$$f(x) = \begin{cases} 0 & \text{for } x < a+r-\delta, \\ (1+x-r+\delta)^3(x-a-r+\delta) & \text{for } a+r-\delta \leq x \leq b+r+\delta, \\ (1+x-r+\delta)^3(b-a+2\delta) & \text{for } x > b+r+\delta. \end{cases}$$

By the fact that for any $c, d > 0$, $\min(c, d) \geq d - \frac{d^2}{4c}$ and $EY_i = 0$, we have

$$\begin{aligned} EU &\geq \sum_{\substack{j=1 \\ j \neq i}}^n E \left\{ Z_j^2 - \frac{|Z_j|^3}{4\delta} \right\} \\ &= \sum_{\substack{j=1 \\ j \neq i}}^n \left\{ EY_j^2 I(|Y_j| \leq 1+a) - (EY_j I(|Y_j| \leq 1+a))^2 \right\} \end{aligned}$$

$$\begin{aligned}
& - \frac{1}{4\delta} \sum_{\substack{j=1 \\ j \neq i}}^n E \left(\left| Y_j I(|Y_j| \leq 1+a) - EY_j I(|Y_j| \leq 1+a) \right| \right)^3 \\
& = \sum_{\substack{j=1 \\ j \neq i}}^n \left\{ EY_j^2 I(|Y_j| \leq 1+a) - (EY_j I(|Y_j| > 1+a))^2 \right\} \\
& \quad - \frac{1}{4\delta} \sum_{\substack{j=1 \\ j \neq i}}^n E \left(\left| Y_j I(|Y_j| \leq 1+a) - EY_j I(|Y_j| \leq 1+a) \right| \right)^3 \\
& \geq \sum_{\substack{j=1 \\ j \neq i}}^n \left\{ EY_j^2 I(|Y_j| \leq 1+a) - EY_j^2 I(|Y_j| > 1+a) \right\} \\
& \quad - \frac{1}{\delta} \sum_{\substack{j=1 \\ j \neq i}}^n E \left(|Y_j|^3 I(|Y_j| \leq 1+a) + |EY_j I(|Y_j| \leq 1+a)|^3 \right) \\
& \geq \sum_{\substack{j=1 \\ j \neq i}}^n EY_j^2 - \frac{2}{\delta} \sum_{\substack{j=1 \\ j \neq i}}^n E|Y_j|^3 \\
& = \frac{(n-1)\sigma^2}{EN\sigma^2} - \frac{2(n-1)\gamma}{\delta(EN\sigma^2)^{3/2}} \\
& = \frac{n-1}{2EN}.
\end{aligned}$$

From this fact and Rosenthal inequality,

$$\begin{aligned}
P\left(U \leq \frac{n-1}{4EN}\right) & \leq P\left(EU - U \geq \frac{n-1}{2EN} - \frac{n-1}{4EN}\right) \\
& = P\left(EU - U \geq \frac{n-1}{4EN}\right) \\
& \leq \frac{16(EN)^2}{(n-1)^2} E|U - E(U)|^4 \\
& \leq \frac{C(EN)^2}{(n-1)^2} \left\{ \left(\sum_{\substack{j=1 \\ j \neq i}}^n E\eta_j^2 \right)^2 + \sum_{\substack{j=1 \\ j \neq i}}^n E\eta_j^4 \right\} \\
& \leq \frac{C(EN)^2}{(n-1)^2} \left\{ \left((n-1)\delta E|Z_1|^3 \right)^2 + (n-1)\delta^4 E|Z_1|^4 \right\} \\
& \leq \frac{C(EN)^2}{(n-1)^2} \left\{ \left((n-1)\delta E|Y_1|^3 \right)^2 + (n-1)(1+a)\delta^4 E|Y_1|^3 \right\}
\end{aligned}$$

$$\begin{aligned}
 &= \frac{C(EN)^2}{(n-1)^2} \left\{ \left(\frac{(n-1)\delta^2}{EN} \right)^2 + \frac{(n-1)(1+a)\delta^5}{EN} \right\} \\
 &= C \left\{ \delta^4 + \frac{(1+a)\delta^5 EN}{n-1} \right\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n-1} \right\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\}
 \end{aligned} \tag{17}$$

where we have used the fact that $(1+a)\delta < 1$ in the last inequality.

By (13), (14), (16) and (17),

$$\begin{aligned}
 &P(a+r \leq T_n^{(i)} \leq b+r) \\
 &\leq P(U \leq \frac{n-1}{4EN}) + \frac{4EN}{(n-1)(1+a)^3} E(T_n^{(i)} f(T_n^{(i)})) \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\} + \frac{CEN}{n(1+a)^3} \{(b-a+2\delta)E|T_n^{(i)}(1+T_n^{(i)}-r+\delta)^3|\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\} + \frac{CEN}{n(1+a)^3} \{(b-a+2\delta)((1+r^3)E|T_n^{(i)}| + E|T_n^{(i)}|^4)\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\} \\
 &\quad + \frac{CEN}{n(1+a)^3} \left\{ (b-a+\delta) \left\{ 1 + \frac{n^3}{(EN)^3} \sqrt{E|T_n^{(i)}|^2} + \frac{n^2}{(EN)^2} + \frac{n(1+a)\delta}{EN} \right\} \right\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\} \\
 &\quad + \frac{CEN}{n(1+a)^3} \left\{ (b-a+2\delta) \left\{ \frac{\sqrt{n}}{\sqrt{EN}} + \frac{n}{EN} + \frac{n^2}{(EN)^2} + \frac{n^{7/2}}{(EN)^{7/2}} \right\} \right\} \\
 &\leq \frac{C\delta}{(1+a)^3} \left\{ 1 + \frac{EN}{n} \right\} + \frac{CEN(b-a+\delta)}{n(1+a)^3} \left\{ 1 + \frac{n^4}{(EN)^4} \right\} \\
 &\leq \frac{C(b-a+\delta)}{(1+a)^3} \left(\frac{EN}{n} + \frac{n^3}{(EN)^3} \right).
 \end{aligned} \tag{18}$$

By (12), (15) and (18), we have

$$P(a \leq W_n^{(i)} \leq b) \leq \frac{C(b-a+\delta)}{(1+a)^3} \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right).$$

□

Lemma 3.4. Let f_x be the solution of (7) and $g(w) = (wf_x(w))'$.

1. For $x > 0$, there exists a constant C such that

$$E|f'_x(W_n)| \leq \frac{C}{(1+x)^2} \left(1 + \frac{n}{EN}\right).$$

2. For $x \geq 4$ and $|u| \leq 1 + \frac{x}{4}$, we have

$$Eg(W_n^{(i)} + u) \leq \frac{C}{(1+x)^3} \left\{ (1 + \delta x) \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \right\}$$

for $i = 1, 2, \dots, n$.

Proof.

1. Using the argument of Lemma 5.1 in [18], p.248 and the fact that

$$EW_n^2 = \frac{n}{EN},$$

we have

$$E|f'_x(W_n)| \leq \frac{C}{(1+x)^2} \left(1 + \frac{n}{EN}\right).$$

2. Use the argument of Lemma 5.2 in [18], p.248 and Proposition 3.3.

□

Proof of Theorem 3.1.

In view of Theorem 2.1, we may without loss of generality, assume $x \geq 4$. If $(1+x)\delta \geq 1$, then, the arguments in (14) and (15) yield

$$P(W \geq x) \leq \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^4}{(EN)^4}\right)$$

which implies

$$\begin{aligned} |F(x) - \Phi(x)| &\leq P(W \geq x) + 1 - \Phi(x) \\ &\leq \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^4}{(EN)^4}\right) + \frac{C}{(1+x)^4} \\ &\leq \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^4}{(EN)^4}\right). \end{aligned}$$

Hence, we assume $(1+x)\delta < 1$.

By the same argument as (8), we have

$$\begin{aligned}
 & F(x) - \Phi(x) \\
 &= \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \left(I(|Y_i| \leq 1 + \frac{x}{4}) \int_{-\infty}^{\infty} (f'(W_n^{(i)} + Y_i) - f'(W_n^{(i)} + t)) M_i(t) dt \right) \\
 &+ \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \left(I(|Y_i| > 1 + \frac{x}{4}) \int_{-\infty}^{\infty} (f'(W_n^{(i)} + Y_i) - f'(W_n^{(i)} + t)) M_i(t) dt \right) \\
 &+ \sum_{n=1}^{\infty} p_n \left(E f'(W_n) - \sum_{i=1}^n E \int_{-\infty}^{\infty} f'(W_n) M_i(t) dt \right) \\
 &:= A_1 + A_2 + A_3.
 \end{aligned}$$

One can follow the argument of [18] on p.251-252 to show that

$$\begin{aligned}
 |A_1| &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n \left| E \left(I(|Y_i| \leq 1 + \frac{x}{4}) \int_{-\infty}^{\infty} M_i(t) \int_t^{Y_i} E g(W_n^{(i)} + u) du dt \right) \right| \\
 &+ \sum_{n=1}^{\infty} p_n \sum_{i=1}^n E \left(I(|Y_i| \leq 1 + \frac{x}{4}) \right. \\
 &\quad \times \left. \int_{-\infty}^{\infty} P(x - \max(Y_i, t) < W_n^{(i)} \leq x - \min(Y_i, t) | Y_i) M_i(t) dt \right) \\
 &= A_{11} + A_{12}.
 \end{aligned}$$

Then Proposition 3.3 and Lemma 3.4(2) yields

$$\begin{aligned}
 |A_{11}| &\leq \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \sum_{i=1}^n \left| E \left(I(|Y_i| \leq 1 + \frac{x}{4}) \right. \right. \\
 &\quad \times \left. \left. \int_{-\infty}^{\infty} M_i(t) \int_t^{Y_i} (1 + \delta x) \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) du dt \right) \right| \\
 &\leq \frac{C(1+\delta x)}{(1+x)^3} \sum_{n=1}^{\infty} p_n \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \sum_{i=1}^n E |Y_i|^3 \\
 &\leq \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \frac{n\delta}{EN} \\
 &= \frac{C\delta}{(1+x)^3 EN} \sum_{n=1}^{\infty} p_n \left(EN + \frac{n^5}{(EN)^4} \right) \\
 &= \frac{C\delta}{(1+x)^3 EN} \left(EN + \frac{EN^5}{(EN)^4} \right) \\
 &= \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^5}{(EN)^5} \right)
 \end{aligned}$$

where we have used the fact that $(1+x)\delta < 1$ in the third inequality, and

$$\begin{aligned}
 |A_{12}| &\leq \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \\
 &\quad \times \sum_{i=1}^n E \left(I(|Y_i| \leq 1 + \frac{x}{4}) \int_{|t| \leq 1 + \frac{x}{4}} (|t| + |Y_i| + \delta) K_i(t) dt \right) \\
 &\leq \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \sum_{i=1}^n (E|Y_i|^3 + \delta E|Y_i|^2) \\
 &= \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \left(\frac{EN}{n} + \frac{n^4}{(EN)^4} \right) \left(\frac{n\delta}{EN} + \frac{n\delta}{EN} \right) \\
 &= \frac{C\delta}{(1+x)^3 EN} \sum_{n=1}^{\infty} p_n \left(EN + \frac{n^5}{(EN)^4} \right) \\
 &= \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^5}{(EN)^5} \right).
 \end{aligned}$$

Hence

$$|A_1| \leq \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^5}{(EN)^5} \right).$$

By (10), we have

$$\begin{aligned}
 |A_2| &\leq \sum_{n=1}^{\infty} p_n \sum_{i=1}^n P(|Y_i| > 1 + \frac{x}{4}) \\
 &\leq C \sum_{n=1}^{\infty} p_n \sum_{i=1}^n \frac{E|Y_i|^3}{(1+x)^3} \\
 &\leq \frac{C}{(1+x)^3} \sum_{n=1}^{\infty} p_n \frac{n\delta}{EN} \\
 &\leq \frac{C\delta}{(1+x)^3}.
 \end{aligned}$$

Follows the argument of (11) and Lemma 3.4(1), we have

$$\begin{aligned}
 |A_3| &\leq \frac{C}{(1+x)^2} \sum_{i=1}^n p_n \left(1 + \frac{n}{EN} \right) \left| 1 - \frac{n}{EN} \right| \\
 &= \frac{CE|N^2 - (EN)^2|}{(1+x)^2(EN)^2} \\
 &\leq \frac{C\sqrt{\text{Var}N}EN^2}{(1+x)^2(EN)^2}.
 \end{aligned}$$

Hence,

$$|P(W \leq x) - \Phi(x)| \leq \frac{C\delta}{(1+x)^3} \left(1 + \frac{EN^5}{(EN)^5}\right) + \frac{C\sqrt{(\text{Var}N)EN^2}}{(1+x)^2(EN)^2}.$$

□

References

- [1] A. Krajka and Z. Rychilk, Necessary and sufficient conditions for weak convergence of random sums of independent random variables. *Comment.Math.Univ.Carolin.*, **34**(1993), 465-482.
- [2] A.C. Berry, The accuracy of the Gaussian approximation to the sum of independent variables. *Trans.Amer.Math.Soc.*, **49**(1941), 122-136.
- [3] A.D. Barbour and L.H.Y.Chen, An introduction to Stein's method, *Lecture Notes Series*, Institute for Mathematical Sciences, National University of Singapore, **4**(2005), Singapore university press and world scientific publishing, Singapore.
- [4] A.D. Barbour and P. Hall, Stein's method and the Berry-Esseen theorem. *Austral. J. Statist.*, **26**(1984), 8-15.
- [5] A.V. Pecinkin, The convergence to normal law of sums of random number of random summands. *Theor.Verojatnost. i Primenen.*, **18**(1973), 380-382.
- [6] C.G. Esseen, Fourier analysis of distribution functions: a mathematical study of the Laplace-Gaussian law, *Acta Math.*, **77**(1945), 1-125.
- [7] C. Stein, A bounded on the error in the normal approximation to a distribution of sums of dependent random variables. *Proc.Sixth Berkeley Symp.Math.Statist.Prob.*, **7**(1972), 583-602.
- [8] C. Stein, Approximate computation of expectations. *Institute of Mathematical Statistics, Lecture notes-monograph series 7*, Hayward, California, 1986.
- [9] D. Landers and L. Rogge, The exact approximation order in the central limit theorem for random summation. *Z.Wahrsch.Verw.Gebiete.*, **36**(1976), 269-283.
- [10] E. Bolthausen, An estimate of the remainder in a combinatorial central limit theorem. *z.Wahrsch.Verw.Gebiete.*, **66**(1984), 379-386.

- [11] G. Orazov, Refinement of theorems on the asymptotic distribution of sums of a random number of random summands under various normalizations. *Taskent.Gos.Univ.Nauch.Trudy Vyp.*, **402**(1972), 87-93.
- [12] I.S. Shiganov, A refinement of the upper bound of the constant in the remainder term of the central limit theorem. *J. Soviet Math.*,**3**(1986), 2545-2550.
- [13] J. Mogyorodi, A central limit theorem for the sums of a random number of random variables. *Ann.Univ.Sci.Budapest.Eotvos Sect.Math.*, **10**(1967), 171-182.
- [14] J.R. Blum, D.L. Handon and J.I. Rosenblatt, On the central limit theorem for the sum of a random number of independent random variables. *Z.Wahrsch.Verw.Gebiete.*,**1**(1962), 389-393.
- [15] K. Neammanee, A limit theorem for random sums of independent random variables with finite variance. Doctoral dissertation, Department of Mathematics, Graduate School, Chulalongkorn University, 1993.
- [16] K. Neammanee, On the constant in the non-uniform version of the Berry-Esseen theorem. *Int.J.Math. and Math.Sci.*,**12**(2005), 1951-1967.
- [17] L.H.Y. Chen, Stein's method: some perspectives with applications, *Probability Towards 2000*(L.Accardi and C.C. Heyde, eds.), *Lecture Notes in Statistics*, **128**(1998), Springer Verlag, 97-122.
- [18] L.H.Y. Chen and Q.M. Shao, A non-uniform Berry-Esseen bound via Stein's method. *Probab. Theory Related Fields*, **120**(2001), 236-254.
- [19] L.H.Y. Chen and Q.M. Shao, Normal approximation under local dependence. *Ann.Probab.*, **32**(2004), 1985-2028.
- [20] P. Van Beeck, An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality. *Z.Wahrsch.Verw.Gebiete.*, **23**(1972), 187-196.
- [21] S.T. Ho and L.H.Y. Chen, An L_p bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.*, **6**(1978), 231-249.
- [22] V.M. Kruglov, The convergence of the distributions of sums of a random number of independent random variables to the normal distribution. *Vestnik Moskov.Univ.Ser.I Mat.Meh.*, **31**(1976), 5-12.
- [23] V.V. Petrov, *Limit theorems of probability theory : sequences of independent random variables.* Oxford Studies in Probability 4, Clarendon Press, Oxford, 1995.