ขอบเขตค่าคลาดเคลื่อนของการประมาณความน่าจะเป็นของจำนวนจุดยอดที่มีดีกรีตามกำหนดในกราฟเชิงสุ่ม

นางสาวอังคณา สันทัดการ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2550

BOUNDS ON ERRORS OF PROBABILITY APPROXIMATION OF NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH

Miss Angkana Suntadkarn

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Mathematics

Department of Mathematics
Faculty of Science
Chulalongkorn University

Academic Year 2007

Thesis Title	BOUNDS ON ERRORS OF PROBABILITY APPROXIMATION OF NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH
By	Miss Angkana Suntadkarn
Field of Study	Mathematics
Thesis Advisor	Professor Kritsana Neammanee, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.... Dean of the Faculty of Science

(Professor Supot Hannongbua, Ph.D.)

THESIS COMMITTEE

.... Chairman

(Associate Professor Wanida Hemakul, Ph.D.)

.... Thesis Advisor

(Professor Kritsana Neammanee, Ph.D.)

.... Member

(Associate Professor Imchit Termwuttipong, Ph.D.)

.... Member

(Kittipat Wong, Ph.D.)

.... Member

(Associate Professor Virool Boonyasombat, Ph.D.)

อังคณา สันทัดการ : ขอบเขตค่าคลาดเคลื่อนของการประมาณความน่าจะเป็นของจำนวนจุดยอดที่มีดีกรีตามกำหนดในกราฟเชิงสุ่ม (BOUNDS ON ERRORS OF PROBABILITY APPROXIMATION OF NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH) อ.ที่ปรึกษา: ศ. ดร. กฤษณะ เนียมมณี, 66 หน้า

เราหาขอบเขตค่าคลาดเคลื่อนของการประมาณด้วยการแจกแจงปกติและการประมาณด้วยการแจกแจงปัวซงของจำนวนจุดยอดที่มีดีกรีตามกำหนดในกราฟเชิงสุ่มโดยวิธีของสไตน์-เชน

ภาควิชา ...**คณิตศาสตร์**...	ลายมือชื่อนิสิต Angkana Santadkarn
สาขาวิชา ...**คณิตศาสตร์**...	ลายมือชื่ออาจารย์ที่ปรึกษา Kritsana Neammanee
ปีการศึกษา ......2550......	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม..............................

##4773848723 : MAJOR MATHEMATICS
KEY WORDS : RANDOM GRAPH/VERTEX OF DEGREE d/STEIN-CHEN METHOD/ NORMAL APPROXIMATION/POISSON APPROXIMATION

ANGKANA SUNTADKARN : BOUNDS ON ERRORS OF PROBABILITY APPROXIMATION OF NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH. THESIS ADVISOR : PROF. KRITSANA NEAMMANEE, Ph.D., 66 pp.

We investigate bounds on errors of normal approximation and Poisson approximation of the number of vertices of a fixed degree in a random graph by using Stein-Chen method.

Department	...**Mathematics**...	Student's signature	Angkana Suntadkarn
Field of study	...**Mathematics**...	Advisor's signature	Kritsana Neammanee
Academic year	.........**2007**............	Co-advisor's signature	....

ACKNOWLEDGEMENTS

I am greatly indebted to Professor Dr. Kritsana Neammanee, my thesis advisor, for his suggestions and helpful advice in preparing and writing this thesis. I would like to thank Associate Professor Dr. Wanida Hemakul, Associate Professor Dr. Imchit Termwuttipong, Dr. Kittipat Wong and Associate Professor Dr. Virool Boonyasombat, my thesis committee, Professor Andrew D. Barbour and Professor Martin Raic for their suggestions to this thesis. I am also sincerely grateful to all of my teachers for my knowledge and skill.

I would like to thank you the Institute for Mathematical Sciences, National University of Singapore, for giving me the scholarship to attend the Random Graphs and Large-Scale Real-World Networks Workshop (1 May - 30 Jun 2006), especially Professor Louis H.Y. Chen.

In particular, thank you my dear friends for giving me good experiences at Chulalongkorn university.

Finally, I would like to express my deep gratitude to my beloved family for their love and encouragement throughout my graduate study.

CONTENTS

CHAPTER I

INTRODUCTION

A random graph is a collection of points, or vertices, with lines, or edges, connecting pairs of them at random. The notion of a random graph was first introduced in 1947 by Erdős ([21]) to show the existence of a graph with a certain Ramsey property. A decade later, the theory of random graph began with the paper entitled *On Random Graphs I* by Erdős and Rényi ([22]), and the theory had been developed by a series ([23], [24], [25], [26], [27]) of papers of them.

The theory of random graphs has developed into one of the mainstays of modern discrete mathematics, and has produced a prodigious number of results, many of them highly ingenious, describing statistical properties of graphs, such as distribution of component sizes, existence and size of a giant component, and typical vertex-vertex distances, that are of interest to practitioners in the field mentioned above.

Random graphs have been used as models of networks in diverse areas of science, engineering, and sociology. These include ecological food webs ([17], [60]), epidemiology ([2], [3], [38], [41], [53]) metabolic pathway ([11], [12], [19], [36], [37], [43]), electric power grids([40]), telephone call network ([1]), networks of social contacts and scientific collaboration ([29], [30], [48], [59]), and, of particular interest to computer science, the internet ([4], [13], [18], [28]). Many additional citations to these topics may be found in ([49]) and the review article ([56]). There are many situations in which the theory tells us that the distribution of a random variable may be approximated by normal distribution or Poisson distribution.

In 1972, Stein([54]) introduced a new powerful technique for obtaining the rate of convergence to the standard normal distribution. His approach was subsequently extended to cover a convergence to Poisson distributions by Chen ([14]). Both methods were illustrated, in the context of random graph theory, in Barbour ([8]). The method for proving Poisson convergence has since been widely taken up (Karo$\acute{n}$ski [34], Karo$\acute{n}$ski and Ruci$\acute{n}$ski [35], Nowicki [50], Janson[32]), but results for random graphs subsequently obtained by the method for normal convergence seem to be limited to examples in Barbour and Eagleson ([5], [6]).

Let $\mathbb{G}(n,p)$ be a random graph on n labeled vertices $\{1,2,\ldots,n\}$ where possible edge $\{i,j\}$ is present randomly and independently with the probability p, $0<p<1$. In our work, we will approximate the distribution functions of the number of vertices of a fixed degree by normal and Poisson distribution functions.

1.1 Normal approximation

Stein's method was given by Stein([54]) in 1972. His technique was relied on the elementary differential equation

$$f'(w) - wf(w) = I_z(w) - \Phi(z) \tag{1.1}$$

where $\Phi(z) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{z} e^{-\frac{t^2}{2}}dt$ and the test function $I_z : \mathbb{R} \to \mathbb{R}$ be defined by

$$I_z(w) = \begin{cases} 1 & ; w \le z \\ 0 & ; w > z. \end{cases}$$

The unique solution, f_z, of the Stein's equation (1.1) is

$$f_z(w) = \begin{cases} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(z)] & \text{if } w \le z \\ \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(z)[1-\Phi(w)] & \text{if } w > z. \end{cases} \tag{1.2}$$

Stein's method was applied to random graphs by Barbour[8] in 1982. In 1989, Barbour, Karo$\acute{n}$ski and Ruci$\acute{n}$ski[9] used Stein's method to show that the distributions of the number of vertices of a fixed degree in $\mathbb{G}(n,p)$ can be approximated by the standard normal distribution function. The followings are their results.

Theorem 1.1. ([9], pp.141)

Let S_n be the number of vertices of a fixed degree, d, in $\mathbb{G}(n,p)$. If $d \geq 1$ then there exists a constant C such that

$$d_1\Big(\frac{S_n - ES_n}{\sqrt{VarS_n}}, \mathcal{N}(0,1)\Big) \leq \frac{C}{\sqrt{VarS_n}},$$

where a metric d_1 is defined by, for any random variables X and Y

$$d_1(X,Y) = \sup\Big\{|Eh(X) - Eh(Y)| : \sup_{x\in\mathbb{R}}|h(x)| + \sup_{x\in\mathbb{R}}|h'(x)| \leq 1$$
$$\textit{for all bounded test functions } h \textit{ with bounded derivative}\Big\}.$$

Inparticular, if $ES_n \to \infty$ and $\log n + d\log\log n - np \to \infty$ then

$$\sup_{z\in\mathbb{R}}\Big|P\Big(\frac{S_n - ES_n}{\sqrt{VarS_n}} \leq z\Big) - \Phi(z)\Big| \to 0 \quad \textit{as } n \to \infty.$$

A Berry-Esseen bound between distribution functions of S_n and Φ in the form $\delta_n := \sup_{z\in\mathbb{R}}\Big|P\Big(\frac{S_n - ES_n}{\sqrt{VarS_n}} \leq z\Big) - \Phi(z)\Big|$ and the metric d_1 are different (see Barbour and Hall [10]). In general, $\delta_n = O(\varepsilon_n^{\frac{1}{2}})$ where ε_n is an upper estimate in metric d_1.

In Theorem 1.1 they gave only sufficient conditions in order to the distribution function of $\frac{S_n - ES_n}{\sqrt{VarS_n}}$ converge to Φ. But, in case of isolated vertices, i.e. $d = 0$, Theorem 1.2 gave both necessary and sufficient conditions for the convergence of S_n but they did not give a Berry-Esseen bound between the normalized S_n and Φ.

Theorem 1.2. ([9], pp.143)

Let S_n be the number of isolated vertices in $\mathbb{G}(n,p)$. Then

$$\sup_{z\in\mathbb{R}}\Big|P\Big(\frac{S_n - ES_n}{\sqrt{VarS_n}} \leq z\Big) - \Phi(z)\Big| \to 0$$

as $n \to \infty$ if and only if $n^2p \to \infty$ and $\log n - np \to \infty$.

In the work of Barbour, Karoński and Ruciński[9], they used the assumption that f_z (in (1.2)) and f_z' can be expanded by Taylor's formula. In 2003, Martin[44] found that this fact is not true. He corrected this mistake by using another test function in stead of I_z. The new test function used by Martin is Lipschitz test function, $I_{z,\varepsilon} : \mathbb{R} \to \mathbb{R}$ which is defined by

$$I_{z,\varepsilon}(w) = \begin{cases} 1 & ;\ w < z - \varepsilon \\ -\dfrac{1}{2\varepsilon}(w - z - \varepsilon) & ;\ z - \varepsilon \leq w < z + \varepsilon \\ 0 & ;\ w \geq z + \varepsilon \end{cases}$$

where $\varepsilon > 0$ is fixed.

In this work we use an idea of Martin to give a uniform bound of the normal approximation of the number of vertices of a fixed degree in a random graph. The followings are our main results.

Theorem 1.3. *Let S_n be defined as in Theorem 1.1 where $d \geq 1$ and $p = \frac{1}{n^\gamma}$ where $\gamma \in \left[1, 1 + \frac{1}{d}\right)$. Then there exists a constant $C(d)$ such that for $0 < \beta < 1$,*

$$\sup_{z \in \mathbb{R}} |P(W_n \leq z) - \Phi(z)| \leq \frac{C(d)}{\sigma_n^\beta}$$

where $W_n := \dfrac{S_n - ES_n}{\sqrt{VarS_n}}$ and $\sigma_n^2 = VarS_n > 0$.

Theorem 1.4. *Let W_n be defined as in Theorem 1.3 and $d = 0$. If $p = \frac{1}{n}$, then there exists a constant $C(d)$ such that for $0 < \beta < 1$,*

$$\sup_{z \in \mathbb{R}} |P(W_n \leq z) - \Phi(z)| \leq \frac{C(d)}{\sigma_n^\beta}.$$

1.2 Poisson approximation

In 1992, Barbour, Holst and Janson[7] proved that the distribution of W, the number of vertices of a fixed degree d in $\mathbb{G}(n, p)$, can be approximated by Poisson distribution, Poi_λ, with parameter

$$\lambda := EW = n\binom{n-1}{d}p^d(1-p)^{n-1-d}$$

and the uniform bound is given by the following theorems.

Theorem 1.5. ([7], pp.99)

Let W *be the number of vertices of degree* d, $d \geq 1$, *in a random graph* $\mathbb{G}(n,p)$ *and* $A \subseteq \{0, 1, \ldots, n\}$. *Then*

$$\sup_A |P(W \in A) - Poi_\lambda(A)| \leq \mu\Big(1 + R_1 + R_2\Big) \tag{1.3}$$

where $\mu = \binom{n-1}{d} p^d (1-p)^{n-1-d}$

$$R_1 = \Big[\frac{(n-1-d)}{(n-1)(1-p)} + \frac{d}{(n-1)p}\Big] E(d - deg(i))^+$$

$$R_2 = \frac{(n-1-d)}{(n-1)(1-p)}\Big[1 + \frac{(n-d-2)p}{(d+1)(1-p)}\Big] E(deg(i) - d)^+ \quad \textit{and}$$

$deg(i)$ *is the degree of vertex* i.

Inparticular, a bound in (1.3) *converges to* 0 *as* $n \to \infty$ *if either*

1. $np \to 0$ *and* $d \geq 2$, *or*
2. np *is bounded away from* 0 *and* $(np)^{-\frac{1}{2}}|d - np| \to \infty$.

In this work, we give bounds of this approximation of the number of vertices of a fixed degree d, $d \geq 0$ in $\mathbb{G}(n,p)$ by using Stein-Chen method. The followings are our main results.

Theorem 1.6. *Let* W *be the number of vertices of degree* d, $d \geq 1$, *in a random graph* $\mathbb{G}(n,p)$ *and* $A \subseteq \{0, 1, \ldots, n\}$. *Then*

1. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq C(\lambda, A)\mu(1 + R_1 + R_2)$,
2. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq (1 - e^{-\lambda})\mu(1 + R_1 + R_2)$

where $C(\lambda, A)$ *is a constant which defined by*

$$C(\lambda, A) = \min\Big\{1, \lambda, \frac{\Delta(\lambda)}{M_A + 1}\Big\},$$

$$\Delta(\lambda) = \begin{cases} e^\lambda + \lambda - 1 & \textit{if} \quad \lambda^{-1}(e^\lambda - 1) \leq M_A, \\ 2(e^\lambda - 1) & \textit{if} \quad \lambda^{-1}(e^\lambda - 1) > M_A, \end{cases}$$

$$M_A = \begin{cases} \max\{w|C_w \subseteq A\} & \textit{if} \quad 0 \in A, \\ \min\{w|w \in A\} & \textit{if} \quad 0 \notin A, \end{cases} \qquad \textit{and}$$

$$C_w = \{0, 1, \ldots, w\}.$$

Furthermore, we know from [46],[57] *that*

$$C(\lambda, \{0, 1, \ldots, w_0\}) \leq (1 - e^{-\lambda}) \min\Big\{1, \frac{e^{\lambda}}{w_0 + 1}\Big\},$$

where $w_0 = 0, 1, \ldots, n$ *and*

$$C(\lambda, \{w_0\}) \leq \min\Big\{1, \frac{\lambda}{w_0}\Big\}$$

where $w_0 = 1, 2, \ldots, n$.

Corollary 1.7. *Let* W *be the number of vertices of degree* d, $d \geq 1$, *in a random graph* $\mathbb{G}(n, p)$ *and* $p = \dfrac{1}{n^{\gamma}}$ *for any* $\gamma \in \mathbb{R}^+$. *Then for* $A \subseteq \{0, 1, \ldots, n\}$

1. *if* $\delta > 1$ *then*

$$1.1 \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq \frac{C(\lambda, A, d)}{n^{(\gamma-1)(d-1)}},$$

$$1.2 \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq \frac{6d^2}{d!q^{d+3}} \frac{(1 - e^{-\lambda})}{n^{(\gamma-1)(d-1)}},$$

where $C(\lambda, A, d) = \dfrac{6d^2}{d!q^{d+3}} C(\lambda, A)$,

2. *if* $0 < \gamma < 1$ *then*

$$2.1 \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq \frac{C(\lambda, A, d)}{n^{d(1-\gamma)}},$$

$$2.2 \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq \frac{6d^2(2d+2)!}{d!q^{3+d}} \frac{(1 - e^{-\lambda})}{n^{d(1-\gamma)}},$$

where $C(\lambda, A, d) = \dfrac{6d^2(2d+2)!}{d!q^{3+d}} C(\lambda, A)$.

Theorem 1.8. *Let* W *be the number of isolated vertices, i.e.,* $d = 0$, *in a random graph* $\mathbb{G}(n, p)$. *Then, for* $A \subseteq \{0, 1, 2, \ldots, n\}$,

$$1. \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq C(\lambda, A)[(n-2)p + 1](1-p)^{n-2}$$

$$2. \quad \Big|P(W \in A) - Poi_{\lambda}(A)\Big| \leq (1 - e^{-\lambda})[(n-2)p + 1](1-p)^{n-2}$$

where Poi_{λ} *is a Poisson distribtuion with parameter* $\lambda = nq^{n-1}$.

Using the fact that $(1-p) < \frac{1}{e^p}$, we see that the bounds in Theorem 1.8 converge to 0 when $np \to \infty$, that is $p = \frac{1}{n^\gamma}$ for $0 < \gamma < 1$.

Corollary 1.9. *Let W be the number of isolated vertices, i.e., $d = 0$, in a random graph $\mathbb{G}(n,p)$ and $p = \frac{1}{n^\gamma}$ for any $0 < \gamma < 1$. Then, for $A \subseteq \{0,1,2,\ldots,n\}$,*

1. $\left|P(W \in A) - Poi_\lambda(A)\right| \leq \frac{3C(\lambda, A)}{q^2 n^{1-\gamma}}$
2. $\left|P(W \in A) - Poi_\lambda(A)\right| \leq \frac{3(1-e^{-\lambda})}{q^2 n^{1-\gamma}}$.

According to our work, we can conclude as the followings.

1. The number of vertices of a fixed degree d, $d \geq 1$, in a random graph $\mathbb{G}(n,p)$.

$p = \frac{1}{n^\gamma}$	Normal approximation	Poisson approximation
$0 < \gamma < 1$		✓
$1 \leq \gamma < 1 + \frac{1}{d}$	✓	✓ $(\gamma \neq 1)$
$\gamma \geq 1 + \frac{1}{d}$		✓

2. The number of isolated vertices in a random graph $\mathbb{G}(n,p)$.

$p = \frac{1}{n^\gamma}$	Normal approximation	Poisson approximation
$0 < \gamma < 1$		✓
$\gamma = 1$	✓	
$\gamma > 1$	open problem	

This thesis is organized as follows. Preliminaries are in Chapter 2. In chapter 3 we give a uniform bound of normal approximation of the number of vertices of a fixed degree in a random graph. While Poisson approximation of the number of vertices of a fixed degree in a random graph is considered in Chapter 4. In chapter 5, we give an open problem on normal approximation of the number of isolated tree in a random graph. Throughout this work, a constant C stands for an absolute constant with possible different values in different places.

CHAPTER II

PRELIMINARIES

In this chapter, we review some basic knowledges in probability, graph and a model of random graph which will be used in our work. The proof is omited but can be found in [20], [42] and [51].

2.1 Probability Space and Random Variables

A **probability space** is a measure space $(\Omega, \mathcal{F}, P)$ for which $P(\Omega) = 1$. The measure P is called a **probability measure**. The set Ω will be refered as a **sample space** and its elements are called **points** or **elementary events**. The elements of $\mathcal{F}$ are called **events**. For any event A, the value $P(A)$ is called the **probabilty of** A.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is called a **random variable** if for every Borel set B in $\mathbb{R}$, $X^{-1}(B)$ belongs to $\mathcal{F}$. We shall use the notation $P(X \in B)$ in place of $P(\omega \in \Omega | X(\omega) \in B)$. In the case where $B = (-\infty, a]$ or $[a, b]$, $P(X \in B)$ is denoted by $P(X \leq a)$ or $P(a \leq X \leq b)$, respectively.

Let X be a random variable. A function $F : \mathbb{R} \to [0, 1]$ which is defined by

$$F(x) = P(X \leq x)$$

is called the **distribution function** of X.

A random variable X with the distribution function F is said to be a **discrete random variable** if the image of X is countable and it is called a **continuous random variable** if F can be written in the form

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

for some nonnegative integrable function f on $\mathbb{R}$.

Now we will give some examples of random variables.

A random variable X, taking on one of the values $0, 1, 2, \ldots$, is said to be a **Poisson** random variable with parameter λ, $\lambda > 0$, written as $X \sim Poi(\lambda)$, if

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots.$$

We say that X is a **normal** random variable with parameter μ and σ^2, written as $X \sim N(\mu, \sigma^2)$, if its probability function is defined by

$$f(x) = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2\sigma^2}(x-\mu)^2).$$

Moreover, if $X \sim N(0, 1)$ then X is said to be a **standard normal** random variable.

2.2 Independence

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_\alpha$ is a sub-σ algebra of $\mathcal{F}$ for every $\alpha \in \Lambda$. We say that $\{\mathcal{F}_\alpha | \alpha \in \Lambda\}$ is **independent** if and only if for any subset $J = \{1, 2, \ldots, k\}$ of Λ,

$$P(\bigcap_{m=1}^{k} A_m) = \prod_{m=1}^{k} P(A_m)$$

where $A_m \in \mathcal{F}_m$ for $m = 1, 2, \ldots, k$.

Let $\mathcal{E}_\alpha \subseteq \mathcal{F}$ for all $\alpha \in \Lambda$. We say that $\{\mathcal{E}_\alpha | \alpha \in \Lambda\}$ is **independent** if and only if $\{\sigma(\mathcal{E}_\alpha) | \alpha \in \Lambda\}$ is independent where $\sigma(\mathcal{E}_\alpha)$ is the smallest σ-algebra with $\mathcal{E}_\alpha \subseteq \sigma(\mathcal{E}_\alpha)$.

We say that the set of random variables $\{X_\alpha | \alpha \in \Lambda\}$ is **independent** if $\{\sigma(X_\alpha) | \alpha \in \Lambda\}$ is independent, where $\sigma(X) = \{X^{-1}(B) | B \text{ is a Borel subset of } \mathbb{R}\}$.

Theorem 2.1.

Random variables $X_1, X_2, \ldots, X_n$ *are* **independent** *if for any Borel sets* $B_1, B_2, \ldots, B_n$, *we have*

$$P\Big(\bigcap_{i=1}^{n} \{X_i \in B_i\}\Big) = \prod_{i=1}^{n} P(X_i \in B_i).$$

Proposition 2.2.

If X_{ij} ; $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m_i$ *are independent and* $f_i : \mathbb{R}^{m_i} \to \mathbb{R}$ *are measurable, then* $\{f_i(X_{i1}, X_{i2}, \ldots, X_{im_i}),\ \ i = 1, 2, \ldots, n\}$ *is independent.*

2.3 Expectation, Variance and Conditional Expectation

Let X be any random variable on a probability space $(\Omega, \mathcal{F}, P)$. If $\int_\Omega |X| dP < \infty$, then we define its **expected value** to be

$$E(X) = \int_\Omega X dP.$$

Proposition 2.3.

1. If X *is a discrete random variable, then* $E(X) = \sum_{x \in ImX} xP(X = x)$.

2. If X *is a continuous random variable with probability function* f, *then*

$$E(X) = \int_\mathbb{R} x f(x) dx.$$

Proposition 2.4.

Let X *and* Y *be random variables such that* $E(|X|) < \infty$ *and* $E(|Y|) < \infty$ *and* $a, b \in \mathbb{R}$. *Then we have the followings:*

1. $E(aX + bY) = aE(X) + bE(Y)$.

2. If $X \leq Y$, *then* $E(X) \leq E(Y)$.

3. $|E(X)| \leq E(|X|)$.

Let X be a random variable which $E(|X|^k) < \infty$. Then $E(|X|^k) < \infty$ is called the k**-th moment** of X about the origin and call $E[(X - E(X))^k]$ or $E[X - E(X)]^k$ the k**-th moment** of X about the mean.

We call the second moment of X about the mean, the **variance** of X, denoted by $Var(X)$. Then

$$Var(X) = E[X - E(X)]^2.$$

We note that

1. $Var(X) = E(X^2) - E^2(X)$.
2. If $X \sim N(\mu, \sigma^2)$, then $E(X) = \mu$ and $Var(X) = \sigma^2$.
3. If $X \sim Poi(\lambda)$, then $E(X) = \lambda$ and $Var(X) = \lambda$.

Proposition 2.5.

If $X_1, X_2, \ldots, X_n$ are independent and $E|X_i| < \infty$ for $i = 1, 2, \ldots, n$, then

1. $E(X_1 X_2 \cdots X_n) = E(X_1)E(X_2) \cdots E(X_n)$,

2. $Var(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \cdots + a_n^2 Var(X_n)$ for any real numbers $a_1, a_2, \ldots, a_n$.

The following inequalities are useful in our work.

1. **Hölder's inequality**:

$$E(|XY|) \leq E^{\frac{1}{p}}(|X|^p) E^{\frac{1}{q}}(|Y|^q)$$

where $1 \leq p, q \leq \infty$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$ and $E(|X|^p) < \infty$, $E(|X|^q) < \infty$.

2. **Chebyshev's inequality**:

$$P(\{|X - E(X)| \geq \varepsilon\}) \leq \frac{Var(X)}{\varepsilon^2} \quad \text{for all } \varepsilon > 0$$

where $E(X^2) < \infty$.

Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $E|X| < \infty$ and $\mathcal{D}$ a sub σ-algebra of $\mathcal{F}$. Define a probability measure $P_{\mathcal{D}} : \mathcal{D} \to [0, 1]$ by

$$P_{\mathcal{D}}(E) = P(E)$$

and a sign-measure $\mathcal{Q}_X : \mathcal{D} \to \mathbb{R}$ by

$$\mathcal{Q}_X(E) = \int_E X dP.$$

Then, by Radon-Nikodym theorem we have $\mathcal{Q}_X \ll P_{\mathcal{D}}$ and there exists a unique measurable function $E(X|\mathcal{D})$ on $(\Omega, \mathcal{F}, P_{\mathcal{D}})$ such that

$$\int_E E(X|\mathcal{D}) dP_{\mathcal{D}} = \mathcal{Q}_X(E) = \int_E X dP \qquad \text{for any } \; E \in \mathcal{D}.$$

We call $E(X|\mathcal{D})$ the **conditional expectation** of X with respect to $\mathcal{D}$.

Moreover, for any random variables X and Y on the same probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|) < \infty$, we will denote $E(X|\sigma(Y))$ by $E(X|Y)$.

Theorem 2.6.

Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $E(|X|) < \infty$, then the followings hold for any sub σ-algebra $\mathcal{D}$ of $\mathcal{F}$.

1. *If X is a random variable on $(\Omega, \mathcal{D}, P_{\mathcal{D}})$, then $E(X|\mathcal{D}) = X$ a.s.$[P_{\mathcal{D}}]$.*

2. *$E(X|\mathcal{F}) = X$ a.s.$[P]$.*

3. *If $\sigma(X)$ and $\mathcal{D}$ are independent, then $E(X|\mathcal{D}) = E(X)$ a.s.$[P_{\mathcal{D}}]$.*

2.4 Graph Theory

A **graph** G consists a non-empty set of elements, called **vertices**, and a list of unordered pairs of these distinct elements, called **edges**. The set of vertices of the graph G is called the **vertex set** of G, denoted by $V(G)$, and the set of edges is called the **edge set** of G, denoted by $E(G)$. If $\{v, w\}$ is an edge e, for some vertices v and w in G then v and w are said to be **adjacent** or edge e is said to be **incident** v **and** w.

Definition 2.7.

The **degree** *of a vertex v in graph G, denoted by $deg(v)$, is the number of edges incident to v.*

Any vertex of degree zero is called an **isolated vertex**.

2.5 Models of Random Graphs

The notion of a random graph originated in a paper of Erdős(1947)[21], which is considered by some as the first conscious application of the probabilistic method. It was used there to prove the existence of a graph with a specific Ramsey property.

The model introduced by Erdős is very natural and can be described as choosing a graph at random, with equal probabilities, from the set of all $2^{\binom{n}{2}}$ graphs whose vertex set is $\{1,2,\ldots,n\}$. In other words, it can be described as the probability space $(\Omega,\mathcal{F},P)$, where Ω is the set of all graphs with vertex set $\{1,2,\ldots,n\}$, $\mathcal{F}$ is the family of all subsets of Ω, and for every $\omega\in\Omega$

$$P(\omega)=2^{-\binom{n}{2}}.$$

Generally speaking, a random graph is a graph constructed by a random procedure. In accordance with standard definitions in probability theory, this is formalized by representing the "random procedure" by a probability space $(\Omega,\mathcal{F},P)$ and the "construction" by a function from the probability space into a suitable family of graphs. The *distribution* of a random graph is the induced probability distribution on the family of graphs: for many purpose this is the only relevant feature of the construction and we usually do not distinguish between different random graphs with the same distribution. Indeed, it is often convenient to define a random graph by specifying its distribution.

The word "model" is used rather loosely in theory of random graphs. It may refer to a specific class of random graph, defined as above, or perhaps to a specific distribution. Nowadays, among several models of random graphs, there are two basic ones, the binomial model and the uniform model, both originating in the simple model introduced by Erdős(1947).

Given a real number p, $0\le p\le 1$, the *binomial random graph*, denoted by $\mathbb{G}(n,p)$, is defined by taking as Ω the set of all graphs on vertex set $\{1,2,\ldots,n\}$ and setting

$$P(G)=p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|},$$

where $|E(G)|$ stands for the number of edges of a graph G. For $p=\dfrac{1}{2}$ this is the model of 1947. However, most of the random graph literature is devoted to cases in which $p=p(n)$ as $n\to\infty$.

Given an integer M, $0\le M\le\dbinom{n}{2}$, the *uniform random graph*, denoted by $\mathbb{G}(n,M)$, is defined by taking as Ω the family of all graphs on the vertex set $\{1,2,\ldots,n\}$

with exactly M edges, and the uniform probability on Ω,

$$P(G) = \binom{\binom{n}{2}}{M}^{-1}, \quad G \in \Omega.$$

In this work, we are interested in one of two models which is the *binomial random graph*, is shorthand for a random graph.

CHAPTER III

NORMAL APPROXIMATION OF THE NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH

In this chapter, we use Stein method to give a uniform bound on normal approximation of number of vertices of a fixed degree in a random graph with n vertices.

Let $\mathbb{G}(n,p)$ be a random graph on n labeled vertices $\{1,2,\ldots,n\}$ where possible edge $\{i,j\}$ is present randomly and independently with the probability p, $0<p<1$. Let S_n be the number of vertices of a fixed degree d, where $d \geq 0$, in $\mathbb{G}(n,p)$. Then $S_n = Y_1 + Y_2 + \cdots + Y_n$ where

$$Y_i = \begin{cases} 1 & \text{if vertex } i \text{ has degree } d \text{ in } \mathbb{G}(n,p), \\ 0 & \text{otherwise,} \end{cases}$$

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

Note that the expectation of Y_i for $i=1,2,\ldots,n$ is

$$\mu = P(Y_i = 1) = \binom{n-1}{d} p^d q^{n-1-d} \text{ and } E(S_n) = n\mu \tag{3.1}$$

where $q = 1-p$ and

$$VarS_n = \frac{n}{n-1}\binom{n-1}{d}^2 (d-(n-1)p)^2 p^{2d-1}(1-p)^{2(n-d)-3} + E(S_n) - \frac{(E(S_n))^2}{n} \tag{3.2}$$

([9], pp.142).

Let

$$W_n := \frac{S_n - E(S_n)}{\sqrt{VarS_n}}.$$

In this chapter, we give a uniform bound between $P(W_n \leq z)$ and $\Phi(z)$. The followings are our main results.

Theorem 3.1. *Let* $p = \dfrac{1}{n^\gamma}$ *where* $\gamma \in \left[1, 1+\dfrac{1}{d}\right)$ *for* $d \geq 1$. *Then there exists a constant* $C(d)$, *such that for* $0 < \beta < 1$,

$$\sup_{z\in\mathbb{R}} |P(W_n \leq z) - \Phi(z)| \leq \frac{C(d)}{\sigma_n^\beta}$$

$\sigma_n^2 = VarS_n > 0$.

Theorem 3.2. *Let* $d = 0$. *If* $p = \dfrac{1}{n}$, *then there exists a constant* $C(d)$, *such that for* $0 < \beta < 1$,

$$\sup_{z\in\mathbb{R}} |P(W_n \leq z) - \Phi(z)| \leq \frac{C(d)}{\sigma_n^\beta}.$$

This chapter is organized as follows. In section 3.1, we prove auxiliary results for the proof of main results and in section 3.2 we introduce Stein's method for normal approximation which is used in the proof of main results in section 3.3.

3.1 Auxiliary Results

In this section, we give auxiliary results for proving Theorem 3.1 and Theorem 3.2. For each $i \in \{1, 2, \ldots, n\}$, let

$$X_i = \frac{Y_i - E(Y_i)}{\sqrt{VarS_n}}.$$

Then

$$E(X_i) = 0, \quad W_n = \sum_{i=1}^n X_i \quad \text{and} \quad E(W_n^2) = 1.$$

For any $\Lambda \subset \{1, 2, \ldots, n\}$ and $i, j \in \{1, 2, \ldots, n\}$, we define

$$Y_j^{(\Lambda)} = \begin{cases} 1 & \text{if the vertex } j \text{ has degree } d \text{ in } \mathbb{G}(n,p) - \{\Lambda\}, \\ 0 & \text{otherwise.} \end{cases}$$

where the random graph $\mathbb{G}(n,p)-\{\Lambda\}$ is obtained from $\mathbb{G}(n,p)$ by removing the vertex in Λ. For $i,j=1,2,\ldots,n$, let

$$Z_{ij} = \begin{cases} \frac{1}{\sigma_n} Y_i & ;\ i=j, \\ \frac{1}{\sigma_n}(Y_j - Y_j^{(i)}) & ;\ i \neq j, \end{cases}$$

$$Z_i = \sum_{j=1}^{n} Z_{ij}, \tag{3.3}$$

$$W^{(i)} = \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{1}{\sigma_n}(Y_j^{(i)} - EY_j^{(i)}) - E(Z_i) = W_n - Z_i, \tag{3.4}$$

$$V_{ij} = \begin{cases} 0 & ; i=j, \\ \frac{1}{\sigma_n}\Big\{ Y_j^{(i)} + \sum_{\substack{l=1 \\ l\neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)}) \Big\} & ; i \neq j, \end{cases}$$

$$W_{ij} = \sum_{\substack{l=1 \\ l\neq i,j}}^{n} \frac{1}{\sigma_n}(Y_l^{(i,j)} - E(Y_l^{(i,j)})) - E(V_{ij}) - E(Z_i), = W^{(i)} - V_{ij} \tag{3.5}$$

where $Y_j^{(i)} := Y_j^{(\{i\})}$ and $Y_l^{(i,j)} := Y_l^{(\{i,j\})}$.

Note that

$$W^{(i)} \text{ is independent of } X_i \text{ and } W_{ij} \text{ is independent of the pair } (X_i, Z_{ij}) \tag{3.6}$$

([9], pp.137).

Proposition 3.3.

1. For $d \geq 1$ and $p = \frac{1}{n^{\gamma}}$ for $\gamma \geq 1$, we have

$$\sigma_n^2 \geq \frac{n\mu}{2}.$$

2. For $d = 0$ and $p = \frac{1}{n}$, we have

$$\sigma_n^2 = n\mu.$$

Proof. 1. From the fact that $q = 1 - p < \dfrac{1}{e^p}$ and $p = \dfrac{1}{n^\gamma}$ for $\gamma \geq 1$ we have

$$\begin{aligned} \mu = P(Y_i = 1) &= \binom{n-1}{d} p^d q^{n-1-d} \\ &= \frac{(n-1)(n-2)\cdots(n-d)p^d}{d! e^{(n-1-d)p}} \\ &< \frac{n^d p^d}{d! e^{(n-1-d)p}} \\ &\leq \frac{1}{d! e} \end{aligned} \tag{3.7}$$

for large n.

Hence by (3.2) and (3.7),

$$\begin{aligned} \sigma_n^2 &= n\binom{n-1}{d} p^d q^{n-1-d} \Big\{ \frac{1}{(n-1)} \binom{n-1}{d} (d-(n-1)p)^2 p^{d-1} q^{n-2-d} + 1 - \frac{E(S_n)}{n} \Big\} \\ &= n\mu \Big\{ \binom{n-1}{d} p^d q^{n-1-d} \frac{(d-(n-1)p)^2}{(n-1)} \frac{1}{pq} + 1 - \frac{n\mu}{n} \Big\} \\ &= n\mu \Big\{ \mu \frac{(d-(n-1)p)^2}{(n-1)} \frac{1}{pq} + 1 - \mu \Big\} \\ &\geq n\mu \Big\{ 1 - \mu \Big\} \\ &\geq \Big(1 - \frac{1}{d!e}\Big) n\mu \\ &\geq \frac{n\mu}{2}. \end{aligned}$$

2. For $d = 0$ and $i = 1, 2, \ldots, n$ we have

$$\mu = P(Y_i = 1) = q^{n-1}.$$

From this fact, (3.2) and $p = \dfrac{1}{n}$, we have

$$\begin{aligned}
\sigma_n^2 &= n(n-1)pq^{2n-3} + E(S_n) - \frac{(E(S_n))^2}{n} \\
&= n(n-1)pq^{2n-3} + nq^{n-1} - nq^{2n-2} \\
&= nq^{n-1}\Big\{(n-1)pq^{n-2} + 1 - q^{n-1}\Big\} \\
&= nq^{n-1}\Big\{npq^{n-2} - pq^{n-2} + 1 - q^{n-2}q\Big\} \\
&= nq^{n-1}\Big\{npq^{n-2} - q^{n-2}(p+q) + 1\Big\} \\
&= nq^{n-1}\Big\{npq^{n-2} - q^{n-2} + 1\Big\} \\
&= nq^{n-1}\Big\{1 + nq^{n-2}(p - \frac{1}{n})\Big\} \\
&= nq^{n-1} \\
&= n\mu.
\end{aligned}$$

□

Proposition 3.4. *Let $d \geq 0$ and $p = \dfrac{1}{n^{\gamma}}$ where $\gamma \geq 1$. For $n \geq 3$, $r_1, r_2, r_3 \in \mathbb{N}$ and $i, j \in \{1, 2, \ldots, n\}$, there exists a positive constant $C_1(d, r_1, r_2, r_3)$ such that*

$$E(|X_i^{r_1} Z_i^{r_2} V_{ij}^{r_3}|) \leq \frac{C_1(d, r_1, r_2, r_3)\mu}{\sigma_n^{r_1+r_2+r_3}}$$

where

$$\begin{aligned}
C_1(d, r_1, r_2, r_3) \leq{}& 2^{r_2 - \frac{2(r_1+r_2)}{r_1+r_2+r_3}} \Big[2 + \frac{3r_2(r_1+r_2+r_3)(d+2)}{r_1+r_2}\Big]^{\frac{r_1+r_2}{r_1+r_2+r_3}} \\
&+ 2^{r_3 - \frac{2r_3}{r_1+r_2+r_3}} [4 + 3(r_1+r_2+r_3)(d+4)]^{\frac{r_3}{r_1+r_2+r_3}}.
\end{aligned}$$

Proof. Note that

$$|Y_j - Y_j^{(i)}| \leq E_{ij} I[deg(j) = d \text{ or } d+1] \tag{3.8}$$

where

$$E_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent in } \mathbb{G}(n,p), \\ 0 & \text{otherwise} \end{cases}$$

and $deg(j)$ is a degree of a vertex j for all $j \in \{1, 2, \ldots, n\}$.

Hence, for $r_1, r_2, \ldots, r_m \in \mathbb{N}$ and for distinct $j_1, j_2, \ldots, j_m$ which are not equal to i we have

$$\begin{aligned}
&E(|Y_{j_1} - Y_{j_1}^{(i)}|^{r_1}|Y_{j_2} - Y_{j_2}^{(i)}|^{r_2} \cdots |Y_{j_m} - Y_{j_m}^{(i)}|^{r_m}) \\
&\leq E(E_{ij_1}^{r_1}(I[deg(j_1) = d \text{ or } d+1])^{r_1} E_{ij_2}^{r_2} \cdots E_{ij_m}^{r_m}) \\
&= P(E_{ij_1} = 1, I[deg(j_1) = d \text{ or } d+1] = 1, E_{ij_2} = 1, \ldots, E_{ij_m} = 1) \\
&= p^{m-1}\Big[\binom{n-2}{d-1} p^d q^{n-1-d} + \binom{n-2}{d} p^{d+1} q^{n-2-d}\Big] \\
&= p^{m-1}\binom{n-1}{d} p^d q^{n-1-d}\Big[\frac{d}{n-1} + \frac{(n-1-d)p}{(n-1)q}\Big] \\
&\leq \frac{p^{m-1}\mu}{(n-1)}\Big[d + \frac{np}{q}\Big] \\
&\leq \frac{p^{m-1}\mu}{n}\Big[\frac{n}{(n-1)}\Big]\Big[d+2\Big] \\
&\leq \frac{3(d+2)}{2}\frac{p^{m-1}\mu}{n}
\end{aligned} \tag{3.9}$$

where $\binom{n}{k} = 0$ for any $k < 0$.

From this fact and $np \leq 1$ we have,

$$\begin{aligned}
&E\Big(\Big|\sum_{\substack{j=1 \\ j \neq i}}^{n} (Y_j - Y_j^{(i)})\Big|^r\Big) \\
&\leq \sum_{\substack{j=1 \\ j \neq i}}^{n} E(|Y_j - Y_j^{(i)}|^r) + \sum_{\substack{j_1=1 \\ j_1 \neq i \\ r_1+r_2=r}}^{n} \sum_{\substack{j_2=1 \\ j_2 \neq i, j_1}}^{n} E(|(Y_{j_1} - Y_{j_1}^{(i)})^{r_1}(Y_{j_2} - Y_{j_2}^{(i)})^{r_2}|) \\
&\quad + \cdots + \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{n} \sum_{\substack{j_2=1 \\ j_2 \neq i, j_2}}^{n} \cdots \sum_{\substack{j_r=1 \\ j_r \neq i, j_1, \ldots, j_{r-1}}}^{n} E(|(Y_{j_1} - Y_{j_1}^{(i)}) \cdots (Y_{j_r} - Y_{j_r}^{(i)})|) \\
&\leq \frac{3(d+2)}{2}\Big(\frac{n\mu}{n} + \frac{n^2 p\mu}{n} + \cdots + \frac{n^r p^{r-1}\mu}{n}\Big) \\
&\leq \frac{3r(d+2)}{2}\mu
\end{aligned}$$

which implies that

$$\begin{aligned}
E|Z_i^r| &= \frac{1}{\sigma_n^r} E\Big|\Big(Y_i + \sum_{\substack{j=1\\ j\neq i}}^{n}(Y_j - Y_j^{(i)})\Big)^r\Big| \\
&\leq \frac{2^{r-1}}{\sigma_n^r}\Big\{E|Y_i^r| + E\Big|\sum_{\substack{j=1\\ j\neq i}}^{n}(Y_j - Y_j^{(i)})\Big|^r\Big\} \\
&\leq \frac{2^{r-1}}{\sigma_n^r}\Big(E(Y_i) + \frac{3r(d+2)}{2}\mu\Big) \\
&= 2^{r-2}(2+3r(d+2))\frac{\mu}{\sigma_n^r}. \qquad (3.10)
\end{aligned}$$

From (3.10) and the fact that

$$|X_i| = \Big|\frac{(Y_i-\mu)}{\sigma_n}\Big| \leq \frac{1}{\sigma_n}, \qquad (3.11)$$

we have

$$E|X_i^{r_1}Z_i^{r_2}| \leq \frac{1}{\sigma_n^{r_1}}E|Z_i^{r_2}| \leq 2^{r_2-2}(2+3r_2(d+2))\frac{\mu}{\sigma_n^{r_1+r_2}}. \qquad (3.12)$$

Similarly to (3.8), we observe that

$$|Y_l^{(i)} - Y_l^{(i,j)}| \leq E_{jl}I^{(i)}[\deg(l) = d \text{ or } d+1] \qquad (3.13)$$

where

$$I^{(i)}[\deg(l) = d \text{ or } d+1] = \begin{cases} 1 & \text{if } \deg(l) = d \text{ or } d+1 \text{ in } G(n,p)-\{i\}, \\ 0 & \text{otherwise.} \end{cases}$$

From (3.13), for $r_1, r_2, \ldots, r_m \in \mathbb{N}$ and for distinct $l_1, l_2, \ldots, l_m$ which are not equal to i, j we have

$$
\begin{aligned}
&E(|Y_{l_1}^{(i)}-Y_{l_1}^{(i,j)}|^{r_1}|Y_{l_2}^{(i)}-Y_{l_2}^{(i,j)}|^{r_2}\cdots|Y_{l_m}^{(i)}-Y_{l_m}^{(i,j)}|^{r_m})\\
&\le E(E_{jl_1}^{r_1}(I^{(i)}[\deg(l_1)=d \text{ or } d+1])^{r_1}E_{jl_2}^{r_2}\cdots E_{jl_m}^{r_m})\\
&=P(E_{jl_1}=1, I^{(i)}[\deg(l_1)=d \text{ or } d+1]=1, E_{jl_2}=1,\dots,E_{jl_m}=1)\\
&=p^{m-1}\Big[\binom{n-3}{d-1}p^dq^{n-2-d}+\binom{n-3}{d}p^{d+1}q^{n-3-d}\Big]\\
&=p^{m-1}\binom{n-1}{d}p^dq^{n-1-d}\Big[\frac{(n-1-d)d}{(n-1)(n-2)}+\frac{(n-1-d)(n-2-d)p}{(n-1)(n-2)q^2}\Big]\\
&=\frac{p^{m-1}\mu}{(n-1)}\Big[\frac{(n-1-d)d}{(n-2)}+\frac{(n-1-d)(n-2-d)p}{(n-2)q^2}\Big]\\
&=\frac{p^{m-1}\mu}{n}\Big[\frac{n}{(n-1)}\Big]\Big[\frac{(n-1-d)d}{(n-2)}+\frac{(n-1-d)(n-2-d)p}{(n-2)q^2}\Big]\\
&\le\frac{3}{2}(d+4)\frac{p^{m-1}\mu}{n} \qquad (3.14)
\end{aligned}
$$

and from (3.14) and the fact that $np\le 1$ we have

$$
\begin{aligned}
&E\Big|\sum_{\substack{l=1\\ l\neq i,j}}^{n}(Y_l^{(i)}-Y_l^{(i,j)})^r\Big|\\
&\le\sum_{\substack{l=1\\ l\neq i,j}}^{n}E|Y_l^{(i)}-Y_l^{(i,j)}|^r+\sum_{\substack{l_1=1\\ l_1\neq i,j\\ r_1+r_2=r}}^{n}\ \sum_{\substack{l_2=1\\ l_2\neq i,j,l_1}}^{n}E|(Y_{l_1}^{(i)}-Y_{l_1}^{(i,j)})^{r_1}(Y_{l_2}^{(i)}-Y_{l_2}^{(i,j)})^{r_2}|\\
&\quad+\cdots+\sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n}\cdots\sum_{\substack{l_r=1\\ l_r\neq i,j,l_1,\dots,l_{r-1}}}^{n}E|(Y_{l_1}^{(i)}-Y_{l_1}^{(i,j)})\cdots(Y_{l_r}^{(i)}-Y_{l_r}^{(i,j)})|\\
&\le\frac{3}{2}(d+4)\Big(\frac{n\mu}{n}+\frac{n^2p\mu}{n}+\cdots+\frac{n^rp^{r-1}\mu}{n}\Big)\\
&\le\frac{3r}{2}(d+4)\mu.
\end{aligned}
$$

From this fact and the fact that

$$
E(Y_j^{(i)})=P(Y_j^{(i)}=1)=\binom{n-2}{d}p^dq^{n-2-d}=\binom{n-1}{d}p^dq^{n-1-d}\Big[\frac{(n-1-d)}{(n-1)q}\Big]\le 2\mu,
$$

we have,

$$
\begin{aligned}
E|V_{ij}^r| &= \frac{1}{\sigma_n^r} E\Big|\Big\{Y_j^{(i)} + \sum_{\substack{l=1\\ l\neq i,j}}^{n}(Y_l^{(i)} - Y_l^{(i,j)})\Big\}^r\Big| \\
&\leq \frac{2^{r-1}}{\sigma_n^r}\Big\{E|Y_j^{(i)}|^r + E\Big|\sum_{\substack{l=1\\ l\neq i,j}}^{n}(Y_l^{(i)} - Y_l^{(i,j)})\Big|^r\Big\} \\
&\leq 2^{r-2}(4+3r(d+4))\frac{\mu}{\sigma_n^r} \qquad (3.15)
\end{aligned}
$$

for any $r \in \mathbb{N}$. Thus, from (3.12) and (3.15), we have

$$
\begin{aligned}
E|X_i^{r_1} Z_i^{r_2} V_{ij}^{r_3}| &\leq \Big\{E|X_i^{r_1} Z_i^{r_2}|^{\frac{r_1+r_2+r_3}{r_1+r_2}}\Big\}^{\frac{r_1+r_2}{r_1+r_2+r_3}} \Big\{E|V_{ij}^{r_3}|^{\frac{r_1+r_2+r_3}{r3}}\Big\}^{\frac{r_3}{r_1+r_2+r_3}} \\
&\leq \Big\{2^{\frac{r_2(r_1+r_2+r_3)}{r_1+r_2}-2}\Big[2+\frac{3r_2(r_1+r_2+r_3)(d+2)}{r_1+r_2}\Big]\frac{\mu}{\sigma_n^{r_1+r_2+r_3}}\Big\}^{\frac{r_1+r_2}{r_1+r_2+r_3}} \\
&\quad \Big\{2^{r_1+r_2+r_3-2}(4+3(r_1+r_2+r_3)(d+4))\frac{\mu}{\sigma_n^{r_1+r_2+r_3}}\Big\}^{\frac{r_3}{r_1+r_2+r_3}} \\
&= C_1(d,r_1,r_2,r_3)\frac{\mu}{\sigma^{r_1+r_2+r_3}}
\end{aligned}
$$

where

$$
\begin{aligned}
C_1(d,r_1,r_2,r_3) \leq\ & 2^{r_2-\frac{2(r_1+r_2)}{r_1+r_2+r_3}}\Big[2+\frac{3r_2(r_1+r_2+r_3)(d+2)}{r_1+r_2}\Big]^{\frac{r_1+r_2}{r_1+r_2+r_3}} \\
&+ 2^{r_3-\frac{2r_3}{r_1+r_2+r_3}}\left[4+3(r_1+r_2+r_3)(d+4)\right]^{\frac{r_3}{r_1+r_2+r_3}}
\end{aligned}
$$

for $r_1, r_2, r_3 \in \mathbb{N}$. □

Proposition 3.5. *Let $d \geq 0$ and $p = \frac{1}{n^\gamma}$ where $\gamma \geq 1$. For $n \geq 3$, $r_1, r_2, r_3 \in \mathbb{N}$ and $i, j \in \{1, 2, \ldots, n\}$, there exists a positive constant $C_2(d, r_3)$ such that*

$$E|X_i^{r_1} Z_{ij}^{r_2} V_{ij}^{r_3}| \leq \frac{C_2(d,r_3)\mu}{n\sigma_n^{r_1+r_2+r_3}}$$

where $C_2(d, r_3) = 2^{r_3-2}(4 + 3r_3(d + 4))$.

Proof. Note from (3.8) and (3.13) that, for any $i, j \in \{1, 2, \ldots, n\}, i \neq j, l_1, \ldots, l_m, \in \{1, 2, \ldots, n\} - \{i, j\}$ and $r_1, r_2, \ldots, r_{m+1} \in \mathbb{N}$,

$$\begin{aligned}
E|(Y_j^{(i)})^{r_1}(Y_j - Y_j^{(i)})^{r_2}| &\leq E(|Y_j^{(i)}|^{r_1} E_{ij}^{r_2} (I[deg(j) = d \text{ or } d+1])^{r_2}) \\
&= P(Y_j^{(i)} = 1, E_{ij} = 1, I[deg(j) = d \text{ or } d+1] = 1) \\
&= \binom{n-2}{d} p^{d+1} q^{n-2-d} \\
&= \binom{n-1}{d} p^d q^{n-1-d} \frac{(n-1-d)p}{(n-1)q} \\
&\leq \frac{\mu}{nq} \\
&\leq \frac{2\mu}{n}
\end{aligned} \tag{3.16}$$

and

$$\begin{aligned}
&E|(Y_j - Y_j^{(i)})^{r_1}(Y_{l_1}^{(i)} - Y_{l_1}^{(i,j)})^{r_2}(Y_{l_2}^{(i)} - Y_{l_2}^{(i,j)})^{r_3} \cdots (Y_{l_m}^{(i)} - Y_{l_m}^{(i,j)})^{r_{m+1}}| \\
&\leq E(E_{ij}^{r_1} E_{jl_1}^{r_2} (I^{(i)}[\deg(l_1) = d \text{ or } d+1])^{r_2} E_{jl_2}^{r_3} \cdots E_{jl_m}^{r_{m+1}}) \\
&= P(E_{ij} = 1, E_{jl_1} = 1, I^{(i)}[\deg(l_1) = d \text{ or } d+1] = 1, E_{jl_2} = 1, \ldots, E_{jl_m} = 1) \\
&= pp^{m-1}\Big[\binom{n-3}{d-1} p^d q^{n-2-d} + \binom{n-3}{d} p^{d+1} q^{n-3-d}\Big] \\
&= p^m \binom{n-1}{d} p^d q^{n-1-d} \Big[\frac{(n-1-d)d}{(n-1)(n-2)} + \frac{(n-1-d)(n-2-d)p}{(n-1)(n-2)q^2}\Big] \\
&= \frac{p^m \mu}{(n-1)} \Big[\frac{(n-1-d)d}{(n-2)} + \frac{(n-1-d)(n-2-d)p}{(n-2)q^2}\Big] \\
&\leq \frac{3}{2}(d+4)\frac{p^m \mu}{n}.
\end{aligned} \tag{3.17}$$

From (3.17) we see that

$$E\Big|(Y_j - Y_j^{(i)})^{r_2}\Big[\sum_{\substack{l=1 \\ l \neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)})\Big]^{r_3}\Big| \leq \frac{3r_3}{2}(d+4)\frac{\mu}{n}.$$

From this fact, (3.11) and (3.16), we have

$$
\begin{aligned}
&E|X_i^{r_1}Z_{ij}^{r_2}V_{ij}^{r_3}| \\
&\leq \frac{1}{\sigma_n^{r_1+r_2+r_3}}E\Big|(Y_j-Y_j^{(i)})^{r_2}\Big\{Y_j^{(i)}+\sum_{\substack{l=1\\ l\neq i,j}}^{n}(Y_l^{(i)}-Y_l^{(i,j)})\Big\}^{r_3}\Big| \\
&\leq \frac{2^{r_3-1}}{\sigma^{r_1+r_2+r_3}}\Big\{E|(Y_j-Y_j^{(i)})^{r_2}(Y_j^{(i)})^{r_3}|+E\Big|(Y_j-Y_j^{(i)})^{r_2}\Big[\sum_{\substack{l=1\\ l\neq i,j}}^{n}(Y_l^{(i)}-Y_l^{(i,j)})\Big]^{r_3}\Big|\Big\} \\
&\leq 2^{r_3-2}(4+3r_3(d+4))\frac{\mu}{n\sigma^{r_1+r_2+r_3}}.
\end{aligned}
$$

□

3.2 Stein's method for normal approximation

Stein's method was given by Stein[54] in 1972. His technique was relied on the elementary differential equation

$$f'(w)-wf(w)=h(w)-\mathcal{N}h \tag{3.18}$$

where $f:\mathbb{R}\to\mathbb{R}$ is a continuous and piecewise continuously differentiable function, h is a bounded test function with bounded derivative and $\mathcal{N}h$ is defined by

$$\mathcal{N}h=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}h(z)e^{-\frac{1}{2}z^2}dz.$$

To use (3.18) to find a bound of normal approximation, many authors([9], [15], [16], [39], [45]) choose the test function $h=I_z$ where $I_z:\mathbb{R}\to\mathbb{R}$ is defined by

$$I_z(w)=\begin{cases}1 & ;w\leq z\\ 0 & ;w>z.\end{cases} \tag{3.19}$$

It is well-known that the solution f_z of (3.18) with test function I_z is of the form

$$f_z(w)=\begin{cases}\sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(z)] & \text{if } w\leq z\\ \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(z)[1-\Phi(w)] & \text{if } w>z.\end{cases} \tag{3.20}$$

where $\Phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt$([55], pp.22).
Observe that

$$f_z'(w)=\begin{cases}[1-\Phi(z)][1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)] & ;\ w<z\\ \Phi(z)[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))] & ;\ w>z\end{cases}$$

and the first derivative of f_z does not exists at $w=z$. But from (3.18) and (3.20) we have to define

$$f_z'(z)=[1-\Phi(z)][1+\sqrt{2\pi}ze^{\frac{1}{2}z^2}\Phi(z)].$$

Thus

$$f_z'(w)=\begin{cases}[1-\Phi(z)][1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)] & ;\ w\le z\\ \Phi(z)[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))] & ;\ w>z.\end{cases} \tag{3.21}$$

By substituting any random variable W for w in (3.18) with $h=I_z$, we yields

$$E[f_z'(W)-Wf_z(W)]=P(W\le z)-\Phi(z).$$

Hence, to bound $|P(W\le z)-\Phi(z)|$, it suffices to bound $E[f_z'(W)-Wf_z(W)]$.

Barbour, Karo$\acute{n}$ski and Ruci$\acute{n}$ski([9],1989) used Taylor expansion of f_z and f_z' to show that

$$|P(W_n\le z)-\Phi(z)|\le\frac{C}{\sigma_n}$$

where W_n is the number of vertices of a fixed degree in a random graph. Unfortunately, since f_z' is not continuous at $w=z$, we can not use the Taylor's expansion of f_z'. In 2003, Martin[14] found that this fact is not true. He corrected this mistake by use another test function. In this chapter we will correct the idea of Barbour, Karo$\acute{n}$ski and Ruci$\acute{n}$ski by using another test function instead of I_z. The new test function is $I_{z,\varepsilon}:\mathbb{R}\to\mathbb{R}$ which is defined by

$$I_{z,\varepsilon}(w)=\begin{cases}1 & ;\ w<z-\varepsilon\\ -\frac{1}{2\varepsilon}(w-z-\varepsilon) & ;\ z-\varepsilon\le w<z+\varepsilon\\ 0 & ;\ w\ge z+\varepsilon\end{cases}$$

where $\varepsilon>0$ is fixed. This function is introduced by Martin([44] pp.84, 2003).

Proposition 3.6. *The solution of Stein's equation*

$$f'(w) - wf(w) = I_{z,\varepsilon}(w) - \mathcal{N}I_{z,\varepsilon}$$

is of the form

$$f_{z,\varepsilon} = \begin{cases} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)\Big[1 - \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big] & ; w < z - \varepsilon \\ \\ \frac{1}{2\varepsilon}\sqrt{2\pi}e^{\frac{w^2}{2}}[1-\Phi(w)]\int_{z-\varepsilon}^{w}\Phi(t)dt & \\ +\sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)\Big[\frac{(z+\varepsilon-w)}{2\varepsilon} - \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon}\Phi(t)dt\Big] & ; z-\varepsilon \leq w < z+\varepsilon \\ \\ \frac{1}{2\varepsilon}\sqrt{2\pi}e^{\frac{w^2}{2}}[1-\Phi(w)]\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt & ; w \geq z+\varepsilon. \end{cases}$$

Proof. Note that

$$\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} I_t(w)dt = \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} 1dt = 1, \qquad \text{for } w < z-\varepsilon,$$

$$\begin{aligned} \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} I_t(w)dt &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{w} I_t(w)dt + \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon} I_t(w)dt \\ &= \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon} 1dt \\ &= -\frac{1}{2\varepsilon}(w-z-\varepsilon), \qquad \text{for } z-\varepsilon \leq w < z+\varepsilon, \end{aligned}$$

and

$$\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} I_t(w)dt = 0, \qquad \text{for } w \geq z+\varepsilon.$$

Hence

$$I_{z,\varepsilon}(w) = \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} I_t(w)dt \tag{3.22}$$

for all $w \in \mathbb{R}$. From this fact we can see that

$$\mathcal{N}I_{z,\varepsilon} = \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} \mathcal{N}I_t dt = \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} \Phi(t)dt. \tag{3.23}$$

By (3.22), (3.23) and Stein's equation we have

$$\begin{aligned} I_{z,\varepsilon} - \mathcal{N}I_{z,\varepsilon} &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} (I_t(w) - \Phi(t))dt \\ &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} (f_t'(w) - wf_t(w))dt. \end{aligned}$$

From this fact and (3.18), the solution of Stein equation,

$$f'(w) - wf(w) = I_{z,\varepsilon}(w) - \mathcal{N}I_{z,\varepsilon}$$

is $f_{z,\varepsilon} : \mathbb{R} \to \mathbb{R}$ which is defined by

$$f_{z,\varepsilon}(w) = \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} f_t(w)dt. \tag{3.24}$$

We use (3.20) and (3.24) to give the form of $f_{z,\varepsilon}$.

Case $w < z - \varepsilon$.

$$\begin{aligned} f_{z,\varepsilon}(w) &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(t)]dt \\ &= \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)\Big[1 - \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]. \end{aligned}$$

Case $z - \varepsilon \leq w < z + \varepsilon$.

$$\begin{aligned} f_{z,\varepsilon} &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{w} f_t(w)dt + \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon} f_t(w)dt \\ &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{w} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(t)[1-\Phi(w)]dt + \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(t)]dt \\ &= \frac{1}{2\varepsilon}\sqrt{2\pi}e^{\frac{w^2}{2}}[1-\Phi(w)]\int_{z-\varepsilon}^{w}\Phi(t)dt \\ &\quad + \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)\Big[\frac{(z+\varepsilon-w)}{2\varepsilon} - \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon}\Phi(t)dt\Big]. \end{aligned}$$

Case $w \geq z + \varepsilon$.

$$\begin{aligned} f_{z,\varepsilon}(w) &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(t)[1-\Phi(w)]dt \\ &= \frac{1}{2\varepsilon}\sqrt{2\pi}e^{\frac{w^2}{2}}[1-\Phi(w)]\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt. \end{aligned}$$

This completes the proof. □

Proposition 3.7. *For any $z \in \mathbb{R}$ and $\varepsilon > 0$, the first derivative of $f_{z,\varepsilon}$ exists and is continuous.*

Proof. First, we will use (3.20) and (3.21) to show that $f'_{z,\varepsilon}$ exists.

Case $w < z - \varepsilon$.

$$\begin{aligned} f'_{z,\varepsilon}(w) &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} f'_t(w)dt \\ &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} [1-\Phi(t)][1+\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)]dt \\ &= \frac{1}{2\varepsilon}[1+\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)]\Big\{2\varepsilon - \int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big\} \\ &= [1+\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)]\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]. \end{aligned}$$

Case $z - \varepsilon < w < z + \varepsilon$

$$\begin{aligned} f'_{z,\varepsilon}(w) &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{w} f'_t(w)dt + \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon} f'_t(w)dt \\ &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{w}\Phi(t)[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]dt \\ &\quad + \frac{1}{2\varepsilon}\int_{w}^{z+\varepsilon}[1-\Phi(t)][1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)]dt \\ &= \frac{1}{2\varepsilon}[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]\int_{z-\varepsilon}^{w}\Phi(t)dt \\ &\quad + \frac{1}{2\varepsilon}[1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)]\Big\{(z+\varepsilon-w) - \int_{w}^{z+\varepsilon}\Phi(t)dt\Big\}. \end{aligned}$$

Case $w > z + \varepsilon$

$$\begin{aligned} f'_{z,\varepsilon}(w) &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon} f'_t(w)dt \\ &= \frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]dt \\ &= \frac{1}{2\varepsilon}[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt. \end{aligned}$$

Next, we will find the first derivative of $f_{z,\varepsilon}(w)$ at points, $w = z-\varepsilon$ and $w = z+\varepsilon$. To find $f'_{z,\varepsilon}(z-\varepsilon)$, we use Proposition 3.6 and L'Hopital's rule to consider the left and the right derivatives of $f_{z,\varepsilon}(z-\varepsilon)$ as follows.

$$\begin{aligned}
&\lim_{h\to 0^-}\frac{f_{z,\varepsilon}(z-\varepsilon+h)-f_{z,\varepsilon}(z-\varepsilon)}{h}\\
&=\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\lim_{h\to 0^-}\frac{1}{h}\Big\{\sqrt{2\pi}e^{\frac{(z-\varepsilon+h)^2}{2}}\Phi(z-\varepsilon+h)-\sqrt{2\pi}e^{\frac{(z-\varepsilon)^2}{2}}\Phi(z-\varepsilon)\Big\}\\
&=\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\lim_{h\to 0^-}\sqrt{2\pi}\Big[\frac{1}{\sqrt{2\pi}}+(z-\varepsilon+h)\Phi(z-\varepsilon+h)e^{\frac{(z-\varepsilon+h)^2}{2}}\Big]\\
&=\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\Big[1+\sqrt{2\pi}(z-\varepsilon)\Phi(z-\varepsilon)e^{\frac{(z-\varepsilon)^2}{2}}\Big]
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{h\to 0^+}\frac{f_{z,\varepsilon}(z-\varepsilon+h)-f_{z,\varepsilon}(z-\varepsilon)}{h}\\
&=\lim_{h\to 0^+}\frac{1}{h}\Big\{\frac{1}{2\varepsilon}\sqrt{2\pi}e^{\frac{(z-\varepsilon+h)^2}{2}}[1-\Phi(z-\varepsilon+h)]\int_{z-\varepsilon}^{z-\varepsilon+h}\Phi(t)dt\\
&\qquad+\sqrt{2\pi}e^{\frac{(z-\varepsilon+h)^2}{2}}\Phi(z-\varepsilon+h)\Big[1-\frac{h}{2\varepsilon}-\frac{1}{2\varepsilon}\int_{z-\varepsilon+h}^{z+\varepsilon}\Phi(t)dt\Big]\\
&\qquad-\sqrt{2\pi}e^{\frac{(z-\varepsilon)^2}{2}}\Phi(z-\varepsilon)\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\Big\}\\
&=\lim_{h\to 0^+}\frac{\sqrt{2\pi}}{2\varepsilon}\Big\{e^{\frac{(z-\varepsilon+h)^2}{2}}[1-\Phi(z-\varepsilon+h)]\Phi(z-\varepsilon+h)\\
&\qquad+\int_{z-\varepsilon}^{z-\varepsilon+h}\Phi(t)dt\Big[\frac{-1}{\sqrt{2\pi}}+[1-\Phi(z-\varepsilon+h)](z-\varepsilon+h)e^{\frac{(z-\varepsilon+h)^2}{2}}\Big]\Big\}\\
&\qquad+\sqrt{2\pi}\Big\{e^{\frac{(z-\varepsilon+h)^2}{2}}\Phi(z-\varepsilon+h)\Big[\frac{-1}{2\varepsilon}+\frac{1}{2\varepsilon}\Phi(z-\varepsilon+h)\Big]+\\
&\qquad+\Big[1-\frac{h}{2\varepsilon}-\frac{1}{2\varepsilon}\int_{z-\varepsilon+h}^{z+\varepsilon}\Phi(t)dt\Big]\Big[\frac{1}{\sqrt{2\pi}}+(z-\varepsilon+h)e^{\frac{(z-\varepsilon+h)^2}{2}}\Phi(z-\varepsilon+h)\Big]\Big\}\\
&=\frac{\sqrt{2\pi}}{2\varepsilon}e^{\frac{(z-\varepsilon)^2}{2}}[1-\Phi(z-\varepsilon)]\Phi(z-\varepsilon)-\frac{\sqrt{2\pi}}{2\varepsilon}e^{\frac{(z-\varepsilon)^2}{2}}[1-\Phi(z-\varepsilon)]\Phi(z-\varepsilon)\\
&\qquad+\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\Big[1+\sqrt{2\pi}(z-\varepsilon)e^{\frac{(z-\varepsilon)^2}{2}}\Phi(z-\varepsilon)\Big]\\
&=\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\Big[1+\sqrt{2\pi}(z-\varepsilon)e^{\frac{(z-\varepsilon)^2}{2}}\Phi(z-\varepsilon)\Big].
\end{aligned}$$

Thus

$$f'_{z,\varepsilon}(z-\varepsilon)=\Big[1-\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big]\Big[1+\sqrt{2\pi}(z-\varepsilon)e^{\frac{(z-\varepsilon)^2}{2}}\Phi(z-\varepsilon)\Big].$$

Similarly to $f'_{z,\varepsilon}(z-\varepsilon)$, we have

$$f'_{z,\varepsilon}(z+\varepsilon)=\frac{1}{2\varepsilon}\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt\Big[-1+\sqrt{2\pi}(1-\Phi(z+\varepsilon))(z+\varepsilon)e^{\frac{(z+\varepsilon)^2}{2}}\Big].$$

Thus we have

$$
f'_{z,\varepsilon}(w)
= \begin{cases}
[1+\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)] - \frac{1}{2\varepsilon}[1+\sqrt{2\pi}we^{\frac{w^2}{2}}\Phi(w)]\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt & ;\ w \le z-\varepsilon \\[2em]
\frac{1}{2\varepsilon}[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]\int_{z-\varepsilon}^{w}\Phi(t)dt \\
+\frac{1}{2\varepsilon}[1+\sqrt{2\pi}we^{\frac{1}{2}w^2}\Phi(w)]\Big\{(z+\varepsilon-w)-\int_{w}^{z+\varepsilon}\Phi(t)dt\Big\} & ; z-\varepsilon \le w \le z+\varepsilon \\[2em]
\frac{1}{2\varepsilon}[-1+\sqrt{2\pi}we^{\frac{1}{2}w^2}(1-\Phi(w))]\int_{z-\varepsilon}^{z+\varepsilon}\Phi(t)dt & ; w \ge z+\varepsilon.
\end{cases}
$$

Observe that $f'_{z,\varepsilon}$ is continuous. □

Remark. The idea of the proof of Proposition 3.6 and Proposition 3.7 is introduced by Martin.

Proposition 3.8. *For the function $f_{z,\varepsilon}$ defined in Proposition 3.6 we have*

1. $\sup\limits_{\substack{x\in\mathbb{R}\\ \varepsilon>0}} |f_{z,\varepsilon}(x)| \le \frac{\sqrt{2\pi}}{4}$
2. $\sup\limits_{\substack{x,y\in\mathbb{R}\\ x\neq y\\ \varepsilon>0}} \Big|\frac{f_{z,\varepsilon}(x)-f_{z,\varepsilon}(y)}{x-y}\Big| \le 1$
3. $I_{z,\varepsilon}(x) - I_{z,\varepsilon}(y) = \frac{y-x}{2\varepsilon}\int_0^1 I_{[z-\varepsilon\le(1-\theta)x+\theta y\le z+\varepsilon]}d\theta$ *for every $x, y \in \mathbb{R}$*

where

$$
I_A(w) = \begin{cases} 1 & ;\ w \in A \\[1em] 0 & \textit{otherwise} \end{cases}
$$

where $A \subseteq \mathbb{R}$

Proof. See Lemma 2 in Chapter II of Stein([55], 1986) for the proof of (1) and (2). The proof of (3), see page 86 of Martin([44],2003). □

Theorem 3.9 is one of the main results in [44] that we will apply in our work.

Theorem 3.9. *Let W_n be a decomposed random variable defined by*

$$W_n = \sum_{i \in I} X_i$$

$$E(X_i) = 0, i \in I; \quad E(W_n^2) = 1;$$

$$W_n = W^{(i)} + Z_i, i \in I, \text{ where } W^{(i)} \text{ is independent of } X_i;$$

$$Z_i = \sum_{j \in K_i} Z_{ij}, i \in I, K_i \subset I;$$

$$W^{(i)} = W_{ij} + V_{ij}, i \in I, j \in K_i,$$

$$\text{where } W_{ij} \text{ is independent of the pair } (X_i, Z_{ij}).$$

Suppose that

$$|X_i| \leq A_i, |Z_{ik}| \leq B_{ik}, |V_{ik}| \leq C_{ik}, |Z_i + V_{ik}| \leq C'_{ik}$$

for some constants A_i, B_{ik}, C_{ik} and C'_{ik}. Then

$$\sup_{z \in \mathbb{R}} |P(W_n \leq z) - \Phi(z)| \leq 13.7 \sum_{i \in I} A_i B_i^2 + \sum_{i \in I} \sum_{k \in K_i} A_i B_{ik}(6.8 C_{ik} + 9.3 C'_{ik})$$

where $B_i := \sum_{k \in K_i} B_{ik}$.

To prove this theorem, Martin showed that for all $\varepsilon > 0$

$$|P(W_n \leq z) - \Phi(z)| \leq A_1(\varepsilon) + A_2(\varepsilon) + A_3(\varepsilon) + B_1(\varepsilon) + B_2(\varepsilon) + B_3(\varepsilon) + \frac{\varepsilon}{\sqrt{2\pi}}, \quad (3.25)$$

where

$$A_1(\varepsilon) := \sum_{i \in I} E(|g_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i) - g_{z,\varepsilon}(W^{(i)})||X_i Z_i|)$$

$$A_2(\varepsilon) := \sum_{i \in I} \sum_{k \in K_i} E(|g_{z,\varepsilon}(W^{(i)}) - g_{z,\varepsilon}(W_{ik})||X_i Z_{ik}|)$$

$$A_3(\varepsilon) := \sum_{i \in I} \sum_{k \in K_i} E(|g_{z,\varepsilon}(W_n) - g_{z,\varepsilon}(W_{ik})|)E(|X_i Z_{ik}|)$$

$$B_1(\varepsilon) := \sum_{i \in I} E(|I_{z,\varepsilon}(W^{(i)}) - I_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i)||X_i Z_i|)$$

$$B_2(\varepsilon) := \sum_{i \in I} \sum_{k \in K_i} E(|I_{z,\varepsilon}(W_{ij}) - I_{z,\varepsilon}(W_{ij} + V_{ij})||X_i Z_{ij}|)$$

$$B_3(\varepsilon) := \sum_{i \in I} \sum_{k \in K_i} E(|I_{z,\varepsilon}(W_n) - I_{z,\varepsilon}(W_{ij})|)E(|X_i Z_{ij}|)$$

and

$$g_{z,\varepsilon}(x) = f_{z,\varepsilon}(x)x, \tag{3.26}$$

$\theta_1 \in [0,1]$ ([44], pp.84-86).

Then he used the boundedness of X_i, Z_{ik}, V_{ik} and $Z_i + V_{ik}$ to bound every term on the right handside of (3.25). In fact we can not apply Theorem 3.9 to our work, since our random variable $Z_i = \sum_{j=1}^{n} \frac{1}{\sigma_n}\Big\{Y_i + (Y_j - Y_j^{(i)})\Big\}$ is not bounded. But in our work, we also prove our result by using equation (3.25).

3.3 Proof of main results

In this section, we give the proof of Theorem 3.1 and Theorem 3.2.

Proof of Theorem 3.1. In Theorem 3.1 we need to find δ where

$$\delta := \sup_{z\in\mathbb{R}} |P(W_n \le z) - \Phi(z)|.$$

We note that for any $a, b \in \mathbb{R}$,

$$\begin{aligned}
P(a \le W_n \le b) &= P(W_n \le b) - P(W < a) - \Phi(b) + \Phi(a) + \Phi(b) - \Phi(a) \\
&= [P(W_n \le b) - \Phi(b)] - [P(W < a) - \Phi(a)] + \Phi(b) - \Phi(a) \\
&\le [P(W_n \le b) - \Phi(b)] + [P(W_n \le a) - \Phi(a)] + \Phi(b) - \Phi(a) \\
&\le 2\delta + \frac{b-a}{\sqrt{2\pi}}.
\end{aligned} \tag{3.27}$$

To find δ, we divide the proof into 4 steps as follows.

Step 1. We will show that

$$A_1(\varepsilon) + A_2(\varepsilon) + A_3(\varepsilon) \le \frac{C}{\sigma_n}$$

for every $\varepsilon > 0$.

From (3.4), (3.6), (3.10), (3.12) and the fact that

$$E(X_i^2) = \frac{1}{\sigma_n^2} E((Y_i - \mu)^2) \le \frac{\mu}{\sigma_n^2}$$

we have

$$
\begin{aligned}
E(|W_nX_i|Z_i^2) &= E(|(W^{(i)}+Z_i)X_i|Z_i^2) \\
&= E(|W^{(i)}X_i|Z_i^2) + E(|X_iZ_i^3|) \\
&\le \Big\{E((W^{(i)})^2X_i^2)\Big\}^{\frac{1}{2}}\Big\{E(Z_i^4)\Big\}^{\frac{1}{2}} + \frac{C\mu}{\sigma_n^4} \\
&= \Big\{E((W^{(i)})^2)E(X_i^2)\Big\}^{\frac{1}{2}}\Big\{E(Z_i^4)\Big\}^{\frac{1}{2}} + \frac{C\mu}{\sigma_n^4} \\
&\le C\Big\{E(W_n^2+Z_i^2)E(X_i^2)\Big\}^{\frac{1}{2}}\Big\{E(Z_i^4)\Big\}^{\frac{1}{2}} + \frac{C\mu}{\sigma_n^4} \\
&\le C\Big\{2E(X_i^2)\Big\}^{\frac{1}{2}}\Big\{E(Z_i^4)\Big\}^{\frac{1}{2}} + \frac{C\mu}{\sigma_n^4} \\
&\le \frac{C\mu}{\sigma_n^3}.
\end{aligned}
\tag{3.28}
$$

From (3.10) and (3.15) we have

$$
E|W_{ij}| = E\Big|\sum_{\substack{l=1\\ l\neq i,j}}^{n}\frac{1}{\sigma_n}(Y_l^{(i,j)} - EY_l^{(i,j)}) - E(V_{ij}) - E(Z_i)\Big| \le E|V_{ij}| + E|Z_i| \le \frac{C}{\sigma_n}
\tag{3.29}
$$

and from (3.9) we have

$$
E|X_iZ_{ij}| \le \frac{1}{\sigma_n^2}E|Y_j - Y_j^{(i)}| \le \frac{3(d+2)\mu}{2n\sigma_n^2}.
\tag{3.30}
$$

From Proposition 3.3, Proposition 3.8, (3.12), (3.26) and (3.28) we have

$$
\begin{aligned}
A_1(\varepsilon) &= \sum_{i=1}^{n} E(|g_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i) - g_{z,\varepsilon}(W^{(i)})||X_i Z_i|) \\
&= \sum_{i=1}^{n} E(|f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i)(W^{(i)}+\theta_1 Z_i - W_n) - f_{z,\varepsilon}(W^{(i)})(W^{(i)} - W_n) \\
&\quad + (f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i) - f_{z,\varepsilon}(W^{(i)}))W_n||X_i Z_i|) \\
&= \sum_{i=1}^{n} E(|f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i)(\theta_1 - 1)Z_i - f_{z,\varepsilon}(W^{(i)})Z_i) \\
&\quad + \Big(\frac{f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i) - f_{z,\varepsilon}(W^{(i)})}{\theta_1 Z_i}\Big)\theta_1 Z_i W_n||X_i Z_i|) \\
&\le \sum_{i=1}^{n} \Big\{ E(|f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i)||X_i|Z_i^2) + E(|f_{z,\varepsilon}(W^{(i)})||X_i|Z_i^2) \\
&\quad + E(\Big|\frac{f_{z,\varepsilon}(W^{(i)}+\theta_1 Z_i) - f_{z,\varepsilon}(W^{(i)})}{\theta_1 Z_i}\Big||W_n X_i|Z_i^2) \Big\} \\
&\le \sum_{i=1}^{n} \Big\{ \frac{\sqrt{2\pi}}{4} E(|X_i|Z_i^2) + \frac{\sqrt{2\pi}}{4} E(|X_i|Z_i^2) + E(|W_n X_i|Z_i^2) \Big\} \\
&= \sum_{i=1}^{n} \Big\{ \frac{\sqrt{2\pi}}{2} E(|X_i|Z_i^2) + E(|W_n X_i|Z_i^2) \Big\} \\
&\le C\Big\{\frac{n\mu}{\sigma_n^3}\Big\} \\
&\le \frac{C}{\sigma_n}. \qquad (3.31)
\end{aligned}
$$

From Proposition 3.3, Proposition 3.5, Proposition 3.8, (3.4), (3.5), (3.6), (3.26), (3.29) and (3.30), we have

$$\begin{aligned}
A_2(\varepsilon) &= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|g_{z,\varepsilon}(W^{(i)}) - g_{z,\varepsilon}(W_{ij})||X_i Z_{ij}|) \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|f_{z,\varepsilon}(W^{(i)})(W^{(i)} - W_n) - f_{z,\varepsilon}(W_{ij})(W_{ij} - W_n) \\
&\quad + (f_{z,\varepsilon}(W^{(i)}) - f_{z,\varepsilon}(W_{ij}))W_n||X_i Z_{ij}|) \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|f_{z,\varepsilon}(W^{(i)})W^{(i)} - f_{z,\varepsilon}(W^{(i)})(W_{ij} + V_{ij} + Z_i) - f_{z,\varepsilon}(W_{ij})W_{ij} \\
&\quad + f_{z,\varepsilon}(W_{ij})(W^{(i)} + Z_i) + f_{z,\varepsilon}(W^{(i)})W_n - f_{z,\varepsilon}(W_{ij})W_n||X_i Z_{ij}|) \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|[f_{z,\varepsilon}(W^{(i)}) + f_{z,\varepsilon}(W_{ij})](W^{(i)} - W_{ij}) - f_{z,\varepsilon}(W^{(i)})V_{ij} - f_{z,\varepsilon}(W^{(i)})Z_i \\
&\quad + f_{z,\varepsilon}(W_{ij})Z_i + f_{z,\varepsilon}(W^{(i)})W_n - f_{z,\varepsilon}(W_{ij})W_n||X_i Z_{ij}|) \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|[f_{z,\varepsilon}(W^{(i)}) + f_{z,\varepsilon}(W_{ij})]V_{ij} - f_{z,\varepsilon}(W^{(i)})V_{ij} + f_{z,\varepsilon}(W^{(i)})[W_n - Z_i] \\
&\quad - f_{z,\varepsilon}(W_{ij})[W_n - Z_i]||X_i Z_{ij}|) \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|[f_{z,\varepsilon}(W^{(i)}) + f_{z,\varepsilon}(W_{ij})]V_{ij} - f_{z,\varepsilon}(W^{(i)})V_{ij} + f_{z,\varepsilon}(W^{(i)})[W_{ij} + V_{ij}] \\
&\quad - f_{z,\varepsilon}(W_{ij})[W_{ij} + V_{ij}]||X_i Z_{ij}|) \\
&\le \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{5\sqrt{2\pi}}{4}E|X_i Z_{ij} V_{ij}| + \frac{\sqrt{2\pi}}{2}E|W_{ij} X_i Z_{ij}| \\
&\le \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{5\sqrt{2\pi}}{4}E|X_i Z_{ij} V_{ij}| + \frac{\sqrt{2\pi}}{2}E|W_{ij}|E|X_i Z_{ij}| \\
&\le C\Big\{\frac{n^2\mu}{n\sigma_n^3} + \frac{n^2\mu}{n\sigma_n^3}\Big\} \\
&\le \frac{C}{\sigma_n}. \qquad (3.32)
\end{aligned}$$

From $EW_n = 0$, Proposition 3.8, (3.10), (3.15), (3.26) and (3.30), we have

$$\begin{aligned}
A_3(\varepsilon) &= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E|g_{z,\varepsilon}(W_n) - g_{z,\varepsilon}(W_{ij})|E|X_i Z_{ij}| \\
&= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E|f_{z,\varepsilon}(W_n)(W_n - W_n) - f_{z,\varepsilon}(W_{ij})(W_{ij} - W_n) \\
&\quad + (f_{z,\varepsilon}(W_n) - f_{z,\varepsilon}(W_{ij}))W_n|E|X_i Z_{ij}| \\
&\leq \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E|f_{z,\varepsilon}(W_{ij})(Z_i + V_{ij})|E|X_i Z_{ij}| + E|(f_{z,\varepsilon}(W_n) - f_{z,\varepsilon}(W_{ij}))W_n||E|X_i Z_{ij}| \\
&\leq \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{\sqrt{2\pi}}{4}E|Z_i + V_{ij}|E|X_i Z_{ij}| + \frac{\sqrt{2\pi}}{2}E|W_n||E|X_i Z_{ij}| \\
&\leq \frac{Cn^2\mu}{n\sigma_n^3} \\
&\leq \frac{C}{\sigma_n}. \qquad (3.33)
\end{aligned}$$

Therefore, from (3.31), (3.32) and (3.33),

$$A_1(\varepsilon) + A_2(\varepsilon) + A_3(\varepsilon) \leq \frac{C}{\sigma_n}.$$

To proof step 2 - 4 we let

$$\varepsilon = 6\Big[\frac{2(2 + 6(d+2))}{\sigma_n^\beta} + \frac{3(d+2)}{\sigma_n^\beta} + \frac{3(d+2)}{\sigma_n^\beta}\Big]. \qquad (3.34)$$

Step 2. We will show that, for $0 < \beta < 1$

$$B_1(\varepsilon) \leq \frac{\delta}{6} + \frac{C(d)}{\sigma_n^\beta}.$$

From Proposition 3.8 (3) we have for $0 < \beta < 1$

$$\begin{aligned}
B_1(\varepsilon) &= \sum_{i=1}^{n} E(|I_{z,\varepsilon}(W^{(i)}) - I_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i)||X_i Z_i|) \\
&\le \frac{1}{2\varepsilon} \sum_{i=1}^{n} E(\Big| \int_0^1 I_{[z-\varepsilon \le W^{(i)} + \theta\theta_1 Z_i \le z+\varepsilon]} d\theta \Big| |X_i Z_i^2|) \\
&\le \frac{1}{2\varepsilon} \sum_{i=1}^{n} E(|X_i Z_i^2| \Big| \int_0^1 I_{[z-\varepsilon \le W^{(i)} + \theta\theta_1 Z_i \le z+\varepsilon, |V_{ij}| > \frac{1}{\sigma_n^\beta}]} d\theta \Big|) \\
&\quad + \frac{1}{2\varepsilon} \sum_{i=1}^{n} E(|X_i Z_i^2| \Big| \int_0^1 I_{[z-\varepsilon \le W^{(i)} + \theta\theta_1 Z_i \le z+\varepsilon, |V_{ij}| \le \frac{1}{\sigma_n^\beta}]} d\theta \Big|) \\
&:= B_{11}(\varepsilon) + B_{12}(\varepsilon).
\end{aligned}$$

By Proposition 3.4 we have

$$\begin{aligned}
B_{11}(\varepsilon) &\le \frac{\sigma_n^{r\beta}}{2\varepsilon} \sum_{i=1}^{n} E(|X_i Z_i^2 V_{ij}^r| \Big| \int_0^1 I_{[z-\varepsilon \le W^{(i)} + \theta\theta_1 Z_i \le z+\varepsilon, |V_{ij}| > \frac{1}{\sigma_n^\beta}]} d\theta \Big|) \\
&\le \frac{\sigma_n^{r\beta}}{2\varepsilon} \sum_{i=1}^{n} E|X_i Z_i^2 V_{ij}^r| \\
&\le \frac{Cn\mu}{\varepsilon \sigma_n^{r(1-\beta)+3}} \qquad (3.35)
\end{aligned}$$

for every $r > 0$.

We will use (3.27) to bound $B_{12}(\varepsilon)$. Note that from (3.4), (3.5), (3.27) and Chebyshev's inequality we have

$$\begin{aligned}
P(a \le W_{ij} \le b) &= P(a \le W_n - (Z_i + V_{ij}) \le b, |Z_i + V_{ij}| \le \frac{1}{\sigma_n^\beta}) \\
&\quad + P(a \le W_n - (Z_i + V_{ij}) \le b, |Z_i + V_{ij}| > \frac{1}{\sigma_n^\beta}) \\
&\le P(a - \frac{1}{\sigma_n^\beta} \le W_n \le b + \frac{1}{\sigma_n^\beta}) + P(|Z_i + V_{ij}| > \frac{1}{\sigma_n^\beta}) \\
&\le 2\delta + \frac{b-a}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta} + P(|Z_i + V_{ij}| > \frac{1}{\sigma_n^\beta}) \\
&\le 2\delta + \frac{b-a}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta} + E|Z_i + V_{ij}|^s \sigma_n^{s\beta} \qquad (3.36)
\end{aligned}$$

for any $s > 0$.

From (3.6), (3.10), (3.12), (3.15) and (3.36)(for $s = \dfrac{\beta}{1-\beta}$) we have

$$\begin{aligned}
B_{12}(\varepsilon) &= \frac{1}{2\varepsilon}\sum_{i=1}^{n} E(|X_iZ_i^2|\Big|\int_0^1 I_{[z-\varepsilon\le W_{ij}+V_{ij}+\theta\theta_1 Z_i\le z+\varepsilon,|Z_i|\le\frac{1}{\sigma_n^\beta},|V_{ij}|\le\frac{1}{\sigma_n^\beta}]}d\theta \\
&\quad+\int_0^1 I_{[z-\varepsilon\le W_{ij}+V_{ij}+\theta\theta_1 Z_i\le z+\varepsilon,|Z_i|>\frac{1}{\sigma_n^\beta},|V_{ij}|\le\frac{1}{\sigma_n^\beta}]}d\theta\Big|) \\
&\le \frac{1}{2\varepsilon}\sum_{i=1}^{n} E(|X_iZ_i^2|\int_0^1 I_{[z-\varepsilon-\frac{2}{\sigma_n^\beta}\le W_{ij}\le z+\varepsilon+\frac{2}{\sigma_n^\beta}]}d\theta)+\frac{\sigma_n^{r\beta}}{2\varepsilon}\sum_{i=1}^{n}E|X_iZ_i^{r+2}| \\
&= \frac{1}{2\varepsilon}\sum_{i=1}^{n} E(|X_iZ_i^2|\int_0^1 P\Big(z-\varepsilon-\frac{2}{\sigma_n^\beta}\le W_{ij}\le z+\varepsilon+\frac{2}{\sigma_n^\beta}\Big)d\theta)+\frac{C\sigma_n^{r\beta}n\mu}{\varepsilon\sigma_n^{r+3}} \\
&\le \frac{(2+6(d+2))n\mu}{2\varepsilon\sigma_n^3}\Big[2\delta+\frac{2\varepsilon}{\sqrt{2\pi}}+\frac{6}{\sqrt{2\pi}\sigma_n^\beta}\Big] \\
&\quad+\frac{(2+6(d+2))n\mu}{2\varepsilon\sigma_n^3}E|Z_i+V_{ij}|^{\frac{\beta}{1-\beta}}\sigma_n^{\beta(\frac{\beta}{1-\beta})}+\frac{Cn\mu}{\varepsilon\sigma_n^{r(1-\beta)+3}} \\
&\le \frac{(2+6(d+2))n\mu}{2\varepsilon\sigma_n^3}\Big[2\delta+\frac{2\varepsilon}{\sqrt{2\pi}}+\frac{6}{\sqrt{2\pi}\sigma_n^\beta}\Big]+\frac{Cn\mu}{\varepsilon\sigma_n^3}\Big[\frac{\sigma_n^{\frac{\beta^2}{1-\beta}}}{\sigma_n^{\frac{\beta}{1-\beta}}}\Big]+\frac{Cn\mu}{\varepsilon\sigma_n^{r(1-\beta)+3}}.
\end{aligned} \tag{3.37}$$

Thus by Proposition 3.3, (3.34), (3.35), (3.37) we can choose $r>0$ be such that $r(1-\beta)+1-\beta>1$ where $0<\beta<1$ so

$$\begin{aligned}
B_1(\varepsilon) &\le \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}}+\frac{(2+6(d+2))}{\varepsilon\sigma_n}\Big[2\delta+\frac{2\varepsilon}{\sqrt{2\pi}}+\frac{6}{\sqrt{2\pi}\sigma_n^\beta}\Big]+\frac{C}{\varepsilon\sigma_n^{1+\frac{\beta-\beta^2}{1-\beta}}} \\
&= \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}}+\frac{2(2+6(d+2))\delta}{\varepsilon\sigma_n}+\frac{2(2+6(d+2))}{\sqrt{2\pi}\sigma_n}+\frac{6(2+6(d+2))}{\varepsilon\sigma_n\sigma_n^\beta\sqrt{2\pi}}+\frac{C}{\varepsilon\sigma_n^{1+\beta}} \\
&\le \frac{C}{\sigma_n^{r(1-\beta)+1-\beta}}+\frac{2(2+6(d+2))\delta\sigma_n^\beta}{\sigma_n^\beta 6[2(2+6(d+1))]}+\frac{C}{\sigma_n^\beta} \\
&\le \frac{\delta}{6}+\frac{C(d)}{\sigma_n^\beta}
\end{aligned}$$

for $0<\beta<1$.

Step 3. We will show that there exists a constant $C>0$, for $0<\beta<1$

$$B_2(\varepsilon)\le\frac{\delta}{6}+\frac{C(d)}{\sigma_n^\beta}.$$

By using Proposition 3.8(3), Proposition 3.3, Proposition 3.5, (3.6), (3.10), (3.15), (3.30) and (3.36) (for $s=\dfrac{1}{1-\beta}$) we have for $r>0$ and $0<\beta<1$

$$
\begin{aligned}
B_2(\varepsilon) &= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|I_{z,\varepsilon}(W_{ij}) - I_{z,\varepsilon}(W_{ij}+V_{ij})||X_iZ_{ij}|)\\
&\le \frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|X_iZ_{ij}V_{ij}|\int_0^1 I_{[z-\varepsilon\le W_{ij}+\theta V_{ij}\le z+\varepsilon]}d\theta)\\
&\le \frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|X_iZ_{ij}V_{ij}|\int_0^1 I_{[z-\varepsilon\le W_{ij}+\theta V_{ij}\le z+\varepsilon, |V_{ij}|>\frac{1}{\sigma_n^\beta}]}d\theta)\\
&\quad + \frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|X_iZ_{ij}V_{ij}|\int_0^1 I_{[z-\varepsilon\le W_{ij}+\theta V_{ij}\le z+\varepsilon, |V_{ij}|\le\frac{1}{\sigma_n^\beta}]}d\theta)\\
&\le \frac{\sigma_n^{r\beta}}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E|X_iZ_{ij}V_{ij}^{r+1}|\\
&\quad + \frac{1}{2\varepsilon\sigma_n^\beta}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|X_iZ_{ij}|\int_0^1 I_{[z-\varepsilon-\frac{1}{\sigma_n^\beta}\le W_{ij}\le z+\varepsilon+\frac{1}{\sigma_n^\beta}]}d\theta)\\
&= \frac{C\sigma_n^{r\beta}n^2\mu}{2\varepsilon n\sigma_n^{r+3}} + \frac{1}{2\varepsilon\sigma_n^\beta}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(|X_iZ_{ij}|P\Big(z-\varepsilon-\frac{1}{\sigma_n^\beta}\le W_{ij}\le z+\varepsilon+\frac{1}{\sigma_n^\beta}\Big))\\
&\le \frac{Cn\mu}{\varepsilon\sigma_n^{r(1-\beta)+3}} + \frac{3(d+2)n^2\mu}{4\varepsilon n\sigma_n^{\beta+2}}\Big[2\delta + \frac{2\varepsilon}{\sqrt{2\pi}} + \frac{4}{\sqrt{2\pi}\sigma_n^\beta}\Big]\\
&\quad + \frac{Cn^2\mu}{\varepsilon n\sigma_n^{\beta+2}}E|Z_i+V_{ij}|^{\frac{1}{1-\beta}}\sigma_n^{\beta\frac{1}{1-\beta}}\\
&\le \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}} + \frac{3(d+2)}{2\varepsilon\sigma_n^\beta}\Big[2\delta + \frac{2\varepsilon}{\sqrt{2\pi}} + \frac{4}{\sqrt{2\pi}\sigma_n^\beta}\Big] + \frac{C}{\varepsilon\sigma_n^\beta}\Big[\frac{\sigma_n^{\frac{\beta}{1-\beta}}}{\sigma_n^{\frac{1}{1-\beta}}}\Big]\\
&\le \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}} + \frac{3(d+2)}{2\varepsilon\sigma_n^\beta}\Big[2\delta + \frac{2\varepsilon}{\sqrt{2\pi}} + \frac{4}{\sqrt{2\pi}\sigma_n^\beta}\Big] + \frac{C}{\varepsilon\sigma_n^{\beta+1}}.
\end{aligned}
$$

From (3.34) we can choose r be such that

$$
B_2(\varepsilon) \le \frac{3(d+2)\delta}{\varepsilon\sigma_n^\beta} + \frac{C}{\sigma_n^\beta} \le \frac{\delta}{6} + \frac{C(d)}{\sigma_n^\beta}.
$$

Step 4. We will show that for $0<\beta<1$,

$$
B_3(\varepsilon) = \le \frac{\delta}{6} + \frac{C(d)}{\sigma_n^\beta}.
$$

From (3.30) and Proposition 3.3 we have

$$\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E|X_i Z_{ij}| \leq \frac{3(d+2)n^2\mu}{2n\sigma_n^2} \leq 3(d+2). \tag{3.38}$$

By Proposition 3.8(3), (3.10), (3.15) and (3.27) we have

$$\begin{aligned}
E|I_{z,\varepsilon}(W_n) - I_{z,\varepsilon}(W_{ij})| &\leq \frac{1}{2\varepsilon}E(|Z_i + V_{ij}|\int_0^1 I_{[z-\varepsilon\leq W_n - \theta(Z_i+V_{ij})\leq z+\varepsilon]}d\theta)\\
&\leq \frac{1}{2\varepsilon}E(|Z_i + V_{ij}|\int_0^1 I_{[z-\varepsilon\leq W_n - \theta(Z_i+V_{ij})\leq z+\varepsilon, |Z_i+V_{ij}|\leq \frac{1}{\sigma_n^\beta}]}d\theta)\\
&\quad + \frac{1}{2\varepsilon}E(|Z_i + V_{ij}|\int_0^1 I_{[z-\varepsilon\leq W_n - \theta(Z_i+V_{ij})\leq z+\varepsilon, |Z_i+V_{ij}|> \frac{1}{\sigma_n^\beta}]}d\theta)\\
&\leq \frac{1}{2\varepsilon\sigma_n^\beta}E(\int_0^1 I_{[z-\varepsilon-\frac{1}{\sigma_n^\beta}\leq W_n\leq z+\varepsilon+\frac{1}{\sigma_n^\beta}]}d\theta) + \frac{\sigma_n^{r\beta}}{2\varepsilon}E|Z_i+V_{ij}|^{r+1}\\
&= \frac{1}{2\varepsilon\sigma_n^\beta}P\Big(z-\varepsilon-\frac{1}{\sigma_n^\beta}\leq W_n\leq z+\varepsilon+\frac{1}{\sigma_n^\beta}\Big) + \frac{C\sigma_n^{r\beta}}{\varepsilon\sigma_n^{r+1}}\\
&\leq \frac{1}{2\varepsilon\sigma_n^\beta}\Big[2\delta + \frac{2\varepsilon}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta}\Big] + \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}}.
\end{aligned}$$

From this fact, (3.34) and (3.38) we can choose r be such that

$$\begin{aligned}
B_3(\varepsilon) &\leq \frac{3(d+2)}{2\varepsilon\sigma_n^\beta}\Big[2\delta + \frac{2\varepsilon}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta}\Big] + \frac{C}{\varepsilon\sigma_n^{r(1-\beta)+1}}\\
&\leq \frac{3(d+2)\delta}{\varepsilon\sigma_n^\beta} + \frac{C}{\sigma_n^\beta}\\
&\leq \frac{\delta}{6} + \frac{C(d)}{\sigma_n^\beta}.
\end{aligned}$$

Therefore by (3.25), step 1 - step 4 and (3.34) we have

$$\begin{aligned}
\delta &\leq \frac{C(d)}{\sigma_n} + \frac{C(d)}{\sigma_n^\beta} + \frac{\delta}{2}\\
&\leq \frac{C(d)}{\sigma_n^\beta} + \frac{\delta}{2}.
\end{aligned}$$

Therefore, there exists a constant $C(d) > 0$ such that for $0 < \beta < 1$

$$\sup_{z\in\mathbb{R}}\Big|P(W_n\leq z) - \Phi(z)\Big| \leq \frac{C(d)}{\sigma_n^\beta}.$$

□

Remark.

We can use the same argument of Theorem 3.1 in proving Theorem 3.2.

CHAPTER IV

POISSON APPROXIMATION OF THE NUMBER OF VERTICES OF A FIXED DEGREE IN A RANDOM GRAPH

In this chapter, we give bounds in Poisson approximation of number of vertices of a fixed degree in a random graph with n vertices.

Let $\mathbb{G}(n,p)$ be a random graph on n labeled vertices $\{1,2,\ldots,n\}$ where possible edge $\{i,j\}$ is present randomly and independently with the probability p, $0<p<1$. For each $i \in \{1,2,\ldots,n\}$, we define the indicator random variable X_i, as follows:

$$X_i = \begin{cases} 1 & \text{if vertex } i \text{ has degree } d \text{ in } \mathbb{G}(n,p), \\ 0 & \text{otherwise.} \end{cases}$$

Then $W := \sum_{i=1}^{n} X_i$ is the number of vertices of degree d in $\mathbb{G}(n,p)$.

In 1992, Barbour, Holst and Janson[7] proved that the distribution of W can be approximated by Poisson distribution with parameter

$$\lambda := E(W) = nP(X_i = 1) = n\binom{n-1}{d} p^d q^{n-1-d} \tag{4.1}$$

where $q = 1-p$ and $d = 1,2,\ldots,\binom{n}{2}$, and the uniform bound is given by the following.

Theorem 4.1. *Let W be the number of vertices of degree d, $d \geq 1$, in a random graph $\mathbb{G}(n,p)$ and $A \subseteq \{0,1,\ldots,n\}$. Then*

$$\sup_A |P(W \in A) - Poi_\lambda(A)| \leq \mu\Big(1 + R_1 + R_2\Big) \tag{4.2}$$

where Poi_λ is a Poisson distribution with parameter λ, i.e., $Poi_\lambda(A) = \sum_{k\in A} \frac{e^{-\lambda}\lambda^k}{k!}$, *and*

$$\mu = P(X_i = 1),$$
$$R_1 = \Big[\frac{(n-1-d)}{(n-1)(1-p)} + \frac{d}{(n-1)p}\Big]E(d-deg(i))^+,$$
$$R_2 = \frac{(n-1-d)}{(n-1)(1-p)}\Big[1 + \frac{(n-d-2)p}{(d+1)(1-p)}\Big]E(deg(i)-d)^+$$

where $deg(i)$ *is degree of a vertex* i. *In particular, a bound in* (4.2) *converges to* 0 *as* $n \to \infty$ *if either*

1. $np \to 0$ *and* $d \geq 2$;
2. np *is bounded away from* 0 *and* $(np)^{-\frac{1}{2}}|d-np| \to \infty$.

In this chapter, we used the result from Barbour, Holst, Janson([7],1992) and Santiwipanont, Teerapabolarn([52],2006) to give non-uniform and uniform bounds of this approximation for a fixed $d = 0, 1, 2, \ldots, \binom{n}{2}$ by using Stein-Chen method. The followings are our main results.

Theorem 4.2. *Let* W *be the number of vertices of degree* d, $d \geq 1$, *in a random graph* $\mathbb{G}(n,p)$ *and* $A \subseteq \{0, 1, \ldots, n\}$. *Then*

1. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq C(\lambda, A)\mu(1 + R_1 + R_2)$,
2. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq (1 - e^{-\lambda})\mu(1 + R_1 + R_2)$

where $C(\lambda, A)$ *is a constant defined by*

$$C(\lambda, A) = \min\Big\{1, \lambda, \frac{\Delta(\lambda)}{M_A + 1}\Big\},$$

$$\Delta(\lambda) = \begin{cases} e^\lambda + \lambda - 1 & \textit{if} \quad \lambda^{-1}(e^\lambda - 1) \leq M_A, \\ 2(e^\lambda - 1) & \textit{if} \quad \lambda^{-1}(e^\lambda - 1) > M_A, \end{cases}$$

$$M_A = \begin{cases} \max\{w | C_w \subseteq A\} & \textit{if} \quad 0 \in A, \\ \min\{w | w \in A\} & \textit{if} \quad 0 \notin A, \end{cases} \qquad \textit{and}$$

$$C_w = \{0, 1, \ldots, w\}.$$

Furthermore, we know from ([46],[57]) *that*

$$C(\lambda,\{0,1,\ldots,w_0\}) \le (1-e^{-\lambda})\min\Big\{1,\frac{e^{\lambda}}{w_0+1}\Big\},$$

where $w_0 = 0,1,\ldots,n$ *and*

$$C(\lambda,\{w_0\}) \le \min\Big\{1,\frac{\lambda}{w_0}\Big\}$$

where $w_0 = 1,2,\ldots,n$.

Corollary 4.3. *Let* W *be the number of vertices of degree* d, $d \ge 1$, *in a random graph* $\mathbb{G}(n,p)$ *and* $p = \frac{1}{n^{\gamma}}$ *for any* $\gamma \in \mathbb{R}^+$. *Then for* $A \subseteq \{0,1,\ldots,n\}$

1. *if* $\gamma > 1$ *then*

$$1.1 \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le \frac{C(\lambda,A,d)}{n^{(\gamma-1)(d-1)}},$$

$$1.2 \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le \frac{6d^2}{d!q^{d+3}}\frac{(1-e^{-\lambda})}{n^{(\gamma-1)(d-1)}},$$

where $C(\lambda,A,d) = \frac{6d^2}{d!q^{d+3}}C(\lambda,A)$,

2. *if* $0 < \gamma < 1$ *then*

$$2.1 \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le \frac{C(\lambda,A,d)}{n^{d(1-\gamma)}},$$

$$2.2 \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le \frac{6d^2(2d+2)!}{d!q^{3+d}}\frac{(1-e^{-\lambda})}{n^{d(1-\gamma)}},$$

where $C(\lambda,A,d) = \frac{6d^2(2d+2)!}{d!q^{3+d}}C(\lambda,A)$.

Theorem 4.4. *Let* W *be the number of isolated vertices, i.e.,* $d=0$, *in a random graph* $\mathbb{G}(n,p)$. *Then, for* $A \subseteq \{0,1,2,\ldots,n\}$,

$$1. \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le C(\lambda,A)[(n-2)p+1](1-p)^{n-2} \tag{4.3}$$

$$2. \quad \Big|P(W\in A) - Poi_{\lambda}(A)\Big| \le (1-e^{-\lambda})[(n-2)p+1](1-p)^{n-2} \tag{4.4}$$

where Poi_{λ} *is a Poisson distribtuion with parameter* $\lambda = nq^{n-1}$.

Using the fact that $(1-p) < \frac{1}{e^p}$, we see that the bounds in Theorem 4.4 converge to 0 when $np \to \infty$, that is $p = \frac{1}{n^{\gamma}}$ for $0 < \gamma < 1$.

Corollary 4.5. *Let W be the number of isolated vertices, i.e., $d = 0$, in a random graph $\mathbb{G}(n,p)$ and $p = \frac{1}{n^\gamma}$ for any $0 < \gamma < 1$. Then, for $A \subseteq \{0, 1, 2, \ldots, n\}$,*

1. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq \frac{3C(\lambda, A)}{q^2 n^{1-\gamma}}$
2. $\Big|P(W \in A) - Poi_\lambda(A)\Big| \leq \frac{3(1-e^{-\lambda})}{q^2 n^{1-\gamma}}$.

This chapter is organized as follows. In section 4.1, we introduce Stein-Chen and coupling methods which are used in our work. In section 4.2 we give the proof of Theorem 4.2 while the proof of Corollary 4.3 is given in section 4.3. The proof of Theorem 4.4 and Corollary 4.5 are given in section 4.4 and section 4.5, respectively.

4.1 Stein-Chen and coupling methods

In 1972, Stein[54] gave a new technique to find a bound in the normal approximation to a distribution of a sum of dependent random variables. His technique was relied instead on the elementary differential equation, and in 1975, Chen[14] applied Stein's idea to the Poisson case. The central idea of the Stein-Chen method is the difference equation

$$I_A(j) - Poi_\lambda(A) = \lambda g_{\lambda,A}(j+1) - j g_{\lambda,A}(j), \qquad j \in \mathbb{N} \cup \{0\} \tag{4.5}$$

where $\lambda > 0$ and $A \subseteq \mathbb{N} \cup \{0\}$ and $I_A : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be defined by

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

We always call equation (4.5) that Stein's equation for Poisson distribution function and it is well-known that the solution $g_{\lambda,A}$ of (4.5) is of the form,

$$g_{\lambda,A}(w) = \begin{cases} (w-1)!\lambda^{-w}e^{\lambda}[\mathcal{P}_\lambda(I_{A\cap C_{w-1}}) - \mathcal{P}_\lambda(I_A)\mathcal{P}_\lambda(I_{C_{w-1}})] & \text{if } \ w \geq 1, \\ 0 & \text{if } \ w = 0 \end{cases}$$

where

$$\mathcal{P}_\lambda(I_A) = e^{-\lambda}\sum_{l=0}^{\infty} I_A(l)\frac{\lambda^l}{l!},$$

and

$$C_{w-1} = \{0, 1, \ldots, w-1\}$$

([7], p.7, 1992).

By substituting j and λ in (4.5) by any integer-valued random variable W and $\lambda = E(W)$, we have

$$P(W \in A) - Poi_\lambda(A) = E(\lambda g_{\lambda,A}(W+1)) - E(W g_{\lambda,A}(W)). \tag{4.6}$$

So far W could be $\sum_{i=1}^{n} X_i$ and $\lambda = E(W) = \sum_{i=1}^{n} p_i$ where $p_i = E(X_i) = P(X_i = 1)$.

Barbour, Holst and Janson([7],1992) used Stein-Chen method and construct coupling random variable W_i to find the bound in Poisson approximation. He assumed that for each i we can construct a random variable W_i, on the same probability space as W, such that the distribution $\mathscr{L}(W_i)$ of W_i equals the conditional distribution $\mathscr{L}(W - X_i | X_i = 1)$. Hence, for each $i \in \{1, 2, \ldots, n\}$,

$$\begin{aligned} E(X_i g_{\lambda,A}(W)) &= E(E(X_i g_{\lambda,A}(W)|X_i)) \\ &= E(X_i g_{\lambda,A}(W)|X_i = 0)P(X_i = 0) + E(X_i g_{\lambda,A}(W)|X_i = 1)P(X_i = 1) \\ &= E(g_{\lambda,A}(W)|X_i = 1)P(X_i = 1) \\ &= p_i E(g_{\lambda,A}(W_i + 1)). \end{aligned} \tag{4.7}$$

Then by (4.6) and (4.7), we have

$$\begin{aligned} \Big|P(W \in A) - Poi_\lambda(A)\Big| &= \Big|E(\lambda g_{\lambda,A}(W+1)) - E(W g_{\lambda,A}(W))\Big| \\ &= \Big|\lambda E(g_{\lambda,A}(W+1)) - \sum_{i=1}^{n} E(X_i g_{\lambda,A}(W))\Big| \\ &= \Big|\sum_{i=1}^{n} p_i E(g_{\lambda,A}(W+1)) - \sum_{i=1}^{n} p_i E(g_{\lambda,A}(W_i+1))\Big| \\ &\le \sum_{i=1}^{n} p_i E(|g_{\lambda,A}(W+1) - g_{\lambda,A}(W_i+1)|) \end{aligned}$$

$$\leq \sum_{i=1}^{n} p_i E(|\sup_w [g_{\lambda,A}(w+1) - g_{\lambda,A}(w)][(W+1) - (W_i+1)]|)$$

$$\leq \sup_w [g_{\lambda,A}(w+1) - g_{\lambda,A}(w)] \sum_{i=1}^{n} p_i E(|W - W_i|).$$

From the estimates above, we arrive at our fundamental result.

Theorem 4.6. *If W and W_i are defined as above, then*

$$|P(W \in A) - Poi_\lambda(A)| \leq ||\Delta g(\lambda, A)|| \sum_{i=1}^{n} p_i E(|W - W_i|) \tag{4.8}$$

where $||\Delta g(\lambda, A)|| := \sup_w [g_{\lambda,A}(w+1) - g_{\lambda,A}(w)]$.

In order to justify the Poisson approximation we therefore only have to

1. bound $||\Delta g(\lambda, A)||$ and
2. find couplings (W, W_i) which made $E(|W - W_i|)$ small.

Many authors would like to determine a bound of $||\Delta g(\lambda, A)||$. For $A \subseteq \mathbb{N} \cup \{0\}$, Chen ([14],1975) proved that

$$||\Delta g(\lambda, A)|| \leq \min\{1, \lambda^{-1}\}$$

and Janson([33],1994) showed that

$$||\Delta g(\lambda, A)|| \leq \lambda^{-1}(1 - e^{-\lambda}). \tag{4.9}$$

In case of non-uniform bound, Neammanee([46],2003) showed that

$$||\Delta g(\lambda, \{w_0\})|| \leq \min\left\{\frac{1}{w_0}, \lambda^{-1}\right\} \tag{4.10}$$

and Teerapabolarn and Neammanee([57],2005) gave a bound of $||\Delta g(\lambda, A)||$ where $A = \{0, 1, \ldots, w_0\}$ in the terms of

$$||\Delta g(\lambda, \{0, 1, \ldots, w_0\})|| \leq \lambda^{-1}(1 - e^{-\lambda}) \min\left\{1, \frac{e^\lambda}{w_0 + 1}\right\}. \tag{4.11}$$

In general case for any subset A of $\{0, 1, \ldots, n\}$, Santiwipanont and Teerapabolarn ([52],2006) gave a bound in the form of

$$||\Delta g(\lambda, A)|| \leq \lambda^{-1} \min\left\{1, \lambda, \frac{\Delta(\lambda)}{M_A + 1}\right\} \tag{4.12}$$

where

$$\Delta(\lambda) = \begin{cases} e^{\lambda} + \lambda - 1 & \text{if} \quad \lambda^{-1}(e^{\lambda} - 1) \le M_A, \\ 2(e^{\lambda} - 1) & \text{if} \quad \lambda^{-1}(e^{\lambda} - 1) > M_A, \end{cases}$$

and

$$M_A = \begin{cases} \max\{w | C_w \subseteq A\} & \text{if} \quad 0 \in A, \\ \min\{w | w \in A\} & \text{if} \quad 0 \notin A. \end{cases}$$

The difficult part in applying Theorem 4.6 is to find W_i which make $E(|W - W_i|)$ small enough. This is not the solution in general. For the case of $X_1, X_2, \ldots, X_n$ are independent, we let $W_i = W - X_i$. Then $E(|W - W_i|) = p_i$ and, from (4.8), we have

$$|P(W \in A) - Poi_{\lambda}(A)| \le ||\Delta g(\lambda, A)|| \sum_{i=1}^{n} p_i^2.$$

The problem of the construction of W_i is difficult in the case of dependent indicator summands. In this case W_i is vary and depends on X_i, see examples in [52] (p.17-24).

In section 4.2, we will use Theorem 4.6 to prove our main result by constructing the random variable W_i which make $E|W - W_i|$ small.

4.2 Proof of Theorem 4.2

By (4.8) and (4.12), it suffices to bound $E(|W - W_i|)$ for any $i \in \{1, 2, \ldots, n\}$ where $W_i \sim (W - X_i)|\ X_i = 1$. Barbour, Holst and Janson([7],1992) constructed a random variable W_i as follows.

Let $G = \{E_{lj} : l, j \in \{1, 2, \ldots, n\}\}$ be a sampled graph in $\mathbb{G}(n, p)$ and $deg(i)$ be degree of a vertex i. To determine W_i, construct a new graph $G' = \{E'_{lj} : l, j \in \{1, 2, \ldots, n\}\}$ where $E'_{lj} = E_{lj}$ for all $l, j \neq i$. In case $l = i$ or $j = i$ we define E'_{lj} as follows:

1. If $deg(i) = d$ then we define $E'_{ij} = E_{ij}$ for $j \neq i$.

2. If $deg(i) < d$ then we will choose $d - deg(i)$ vertices from all vertex j such that $E_{ij} = 0$. For all j which are chosen we define $E'_{ij} = 1$ and defind $E'_{ij} = E_{ij}$ for all other j.

3. If $deg(i) > d$ then we will choose $deg(i) - d$ vertices from each vertex j such that $E_{ij} = 1$. For all j which are chosen we define $E'_{ij} = 0$ and defind $E'_{ij} = E_{ij}$ for all other j.

Let $W_i = W' - 1$, where W' is the number of vertices of degree d obtained from G'. Then Barbour, Holst and Janson showed in [7](pp.99) that the distribution $\mathscr{L}(W_i + 1) = \mathscr{L}(W|X_i = 1)$ which implied that for any $k \in \mathbb{N}$

$$\begin{aligned} P(W_i + 1 = k) &= P(W = k|X_i = 1) \\ P(W_i = k - 1) &= P(W - X_i = k - 1|X_i = 1) \\ P(W_i = k) &= P(W - X_i = k|X_i = 1) \end{aligned}$$

that is $W_i \sim (W - X_i)|\ X_i = 1$. By Barbour, Holst and Janson([7],1992), (p.100), showed that

$$E(|W - W_i|) \le \mu\Big(1 + R_1 + R_2\Big) \tag{4.13}$$

where $$R_1 = \Big[\frac{(n-1-d)}{(n-1)(1-p)} + \frac{d}{(n-1)p}\Big]E(d - deg(i))^+, \qquad \text{and}$$
$$R_2 = \frac{(n-1-d)}{(n-1)(1-p)}\Big[1 + \frac{(n-d-2)p}{(d+1)(1-p)}\Big]E(deg(i) - d)^+.$$

Then by (4.8), (4.9), (4.12) and (4.13) we have

$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le C(\lambda, A)\mu(1 + R_1 + R_2) \quad \text{and}$$
$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le (1 - e^{-\lambda})\mu(1 + R_1 + R_2)$$

where $C(\lambda, A)$ is a constant which defined by

$$C(\lambda, A) = \min\Big\{1, \lambda, \frac{\Delta(\lambda)}{M_A + 1}\Big\}.$$

Furthermore, we know from (4.10) and (4.11) that

$$C(\lambda, \{0, 1, \ldots, w_0\}) \le (1 - e^{-\lambda})\min\Big\{1, \frac{e^\lambda}{w_0 + 1}\Big\},$$

where $w_0 = 0, 1, \ldots, n$ and

$$C(\lambda, \{w_0\}) \le \min\Big\{1, \frac{\lambda}{w_0}\Big\}$$

where $w_0 = 1, 2, \ldots, n$.

This completes the proof of Theorem 4.2. $\square$

Remark. We observe that the uniform bound of Theorem 4.2 is better than the bound by Barbour, Holst and Janson(1992), since $1 - e^{-\lambda} < 1$ for $\lambda = n\binom{n-1}{d}p^d q^{n-1-d}$.

4.3 Proof of Corollary 4.3

For the asymtotic results, we note that, from $1 - p \leq \frac{1}{e^p}$,

$$\mu = P(X_i = 1) = \binom{n-1}{d}p^d(1-p)^{n-1-d} \leq \frac{(np)^d}{d!q^{1+d}e^{np}}. \tag{4.14}$$

By [7], pp.100, we have

$$E(d - deg(i))^+ \leq d \quad \text{and} \quad E(deg(i) - d)^+ = \sum_{\substack{j=1 \\ j \neq i}}^{n} EI_{[E'_{ij} < E_{ij}]} = (n-1)p \leq np$$

where $(d - deg(i))^+ = \max\{d - deg(i), 0\}$ and $(deg(i) - d)^+ = \max\{deg(i) - d, 0\}$. Then from (4.13) and (4.14), we have

$$\begin{aligned}
E(|W - W_i|) &\leq \frac{(np)^d}{d!q^{1+d}e^{np}}\Big\{1 + \Big[\frac{(n-1-d)}{(n-1)(1-p)} + \frac{d}{(n-1)p}\Big]E(d - deg(i))^+ \\
&\quad + \frac{(n-1-d)}{(n-1)(1-p)}\Big[1 + \frac{(n-d-2)p}{(d+1)(1-p)}\Big]E(deg(i) - d)^+\Big\} \\
&\leq \frac{(np)^d}{d!q^{1+d}e^{np}}\Big\{\frac{d}{q} + \Big[\frac{1}{q} + \Big(1 + \frac{1}{n-1}\Big)\frac{d}{np}\Big]d + \Big[\frac{1}{q} + \frac{np}{(d+1)q^2}\Big]np\Big\} \\
&\leq \frac{(np)^d}{d!q^{1+d}e^{np}}\Big\{\Big[\frac{2}{q} + \frac{2d}{np}\Big]d + \Big[\frac{np}{q} + \frac{(np)^2}{(d+1)q^2}\Big]\Big\}.
\end{aligned} \tag{4.15}$$

We suppose that $p = \frac{1}{n^\gamma}$ for any $\gamma \in \mathbb{R}^+$.

1. If $\gamma > 1$ then we observe that

$$\begin{aligned}
E(|W - W_i|) &\leq \frac{(np)^d}{d!q^{1+d}e^{np}}\Big\{\Big[\frac{2}{q} + \frac{2d}{np}\Big]d + \Big[\frac{np}{q} + \frac{(np)^2}{(d+1)q^2}\Big]\Big\} \\
&\leq \frac{(np)^{d+2}}{d!q^{1+d}e^{np}}\Big\{\Big[\frac{2d}{q(np)^2} + \frac{2d^2}{(np)^3}\Big] + \Big[\frac{1}{qnp} + \frac{1}{dq^2}\Big]\Big\} \\
&\leq \frac{(np)^{d+2}}{d!q^{1+d}e^{np}}\Big\{\frac{6d^2}{q^2(np)^3}\Big\} \\
&\leq \frac{6d^2(np)^{d-1}}{d!q^{3+d}} \\
&= \frac{6d^2}{d!q^{3+d}}\frac{1}{n^{(\gamma-1)(d-1)}}.
\end{aligned} \tag{4.16}$$

From (4.8), (4.9), (4.12) and (4.16) we have

$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le \frac{C(\lambda, A, d)}{n^{(\gamma-1)(d-1)}} \qquad \text{and}$$
$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le \frac{6d^2}{d!q^{d+3}}\frac{(1-e^{-\lambda})}{n^{(\gamma-1)(d-1)}},$$

where $C(\lambda, A, d) = \dfrac{6d^2}{d!q^{d+3}}C(\lambda, A)$.

2. If $\gamma < 1$. From (4.15), we observe that

$$\begin{aligned}
E(|W - W_i|) &\le \frac{(np)^d}{d!q^{1+d}e^{np}}\Big\{\Big[\frac{2}{q} + \frac{2d}{np}\Big]d + \Big[\frac{np}{q} + \frac{(np)^2}{(d+1)q^2}\Big]\Big\} \\
&\le \frac{(np)^d}{d!q^{1+d}e^{np}}\Big[\frac{2d}{q} + \frac{2d^2(np)^2}{q^2} + \frac{(np)^2}{q} + \frac{(np)^2}{dq^2}\Big] \\
&\le \frac{(np)^d}{d!q^{1+d}e^{np}}\Big[\frac{2d^2(np)^2}{q^2} + \frac{2d^2(np)^2}{q^2} + \frac{d^2(np)^2}{q^2} + \frac{d^2(np)^2}{q^2}\Big] \\
&= \frac{6d^2(np)^{d+2}}{d!q^{3+d}e^{np}} \\
&\le \frac{6d^2(2d+2)!(np)^{d+2}}{d!q^{3+d}(np)^{2d+2}} \\
&= \frac{6d^2(2d+2)!}{d!q^{3+d}}\frac{1}{n^{d(1-\gamma)}}.
\end{aligned} \tag{4.17}$$

From (4.8), (4.9), (4.12) and (4.17) we have

$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le \frac{C(\lambda, A, d)}{n^{d(1-\gamma)}} \qquad \text{and}$$
$$\Big|P(W \in A) - Poi_\lambda(A)\Big| \le \frac{6d^2(2d+2)!}{d!q^{3+d}}\frac{(1-e^{-\lambda})}{n^{d(1-\gamma)}},$$

where $C(\lambda, A, d) = \dfrac{6d^2(2d+2)!}{d!q^{3+d}}C(\lambda, A)$.

This complete the proof of Corollary 4.3. □

4.4 Proof of Theorem 4.4

A vertex i is an isolated vertex in $\mathbb{G}(n,p)$ if the number of edges incident to it is 0. Then in this case, W is the number of isolated vertices in $\mathbb{G}(n,p)$ and $W = \sum_{i=1}^{n} X_i$, where

$$X_i = \begin{cases} 1 & \text{if vertex } i \text{ is an isolated vertex in } \mathbb{G}(n,p), \\ 0 & \text{otherwise} \end{cases}$$

and $P(X_i = 1) = (1-p)^{n-1}$.

For $i, j \in \{1, 2, \ldots n\}$, we define

$$X_j^{(i)} = \begin{cases} 1 & \text{if vertex } j \text{ is an isolated vertex in } \mathbb{G}(n,p) - \{i\}, \\ 0 & \text{otherwise.} \end{cases}$$

Let W_i be the number of isolated vertices in a random graph $\mathbb{G}(n,p) - \{i\}$ obtained from $\mathbb{G}(n,p)$ by removing the vertex i and all the edges incident to it. Then the distribution $\mathscr{L}(W_i)$ of W_i equals the conditional distribution $\mathscr{L}(W - X_i | X_i = 1)$, that is for $k \in \{0, 1, \ldots, n-1\}$,

$$\begin{aligned} P(W - X_i = k | X_i = 1) &= \frac{P(W - X_i = k, X_i = 1)}{P(X_i = 1)} \\ &= \frac{P(W = k+1, X_i = 1)}{P(X_i = 1)} \\ &= \frac{\binom{n-1}{k} q^{(n-2)k} q^{n-1}}{q^{n-1}} \\ &= \binom{n-1}{k} q^{(n-2)k} \\ &= P(W_i = k). \end{aligned}$$

We observe that in case of $X_i = 1$,

$$W_i = W - 1 \tag{4.18}$$

and in case of $X_i = 0$,

$W_i = W$ + (the number of vertices of degree 1 in $\mathbb{G}(n,p)$ that adjacent to vertex i), i.e.

$$W_i = W + \sum_{\substack{j=1 \\ j \neq i}}^{n} E_{ij} X_j^{(i)} \tag{4.19}$$

where

$$E_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent in } \mathbb{G}(n,p), \\ 0 & \text{otherwise.} \end{cases}$$

From (4.18) and (4.19), we have

$$(W - W_i)^+ \leq X_i \text{ and } (W_i - W)^+ \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} E_{ij} X_j^{(i)}. \tag{4.20}$$

We know that

$$|W - W_i| = (W - W_i)^+ + (W - W_i)^-,$$

where $(W - W_i)^+ = \max\{W - W_i, 0\}$ and $(W - W_i)^- = -\min\{W - W_i, 0\}$.
Since $-\min\{W - W_i, 0\} = \max\{W_i - W, 0\} = (W_i - W)^+$, we have

$$E(|W - W_i|) = E(W - W_i)^+ + E(W_i - W)^+. \tag{4.21}$$

From the fact that

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} E(E_{ij} X_j^{(i)}) = (n-1)P(E_{ij} = 1, X_j^{(i)} = 1) = (n-1)p(1-p)^{n-2},$$

(4.21) and $(1-p) < \dfrac{1}{e^p}$ we have,

$$\begin{aligned} E(|W - W_i|) &\leq E(X_i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} E(E_{ij} X_j^{(i)}) \\ &= (1-p)^{n-1} + (n-1)p(1-p)^{n-2} \\ &= [(n-2)p + 1](1-p)^{n-2}. \end{aligned} \tag{4.22}$$

Therefore, by (4.8), (4.12) and (4.22), we have

$$\Big| P(W \in A) - Poi_\lambda(A) \Big| \leq C(\lambda, A)[(n-2)p + 1](1-p)^{n-2}$$

where

$$C(\lambda, A) = \min\Big\{1, \lambda, \frac{\Delta(\lambda)}{M_A + 1}\Big\}.$$

In case of uniform bound by (4.8), (4.9) and (4.22), we have

$$\Big| P(W \in A) - Poi_\lambda(A) \Big| \leq (1 - e^{-\lambda})[(n-2)p + 1](1-p)^{n-2}$$

This completes the proof of Theorem 4.4. $\square$

4.5 Proof of Corollary 4.5

By using the fact that $(1-p) \leq \dfrac{1}{e^p}$ and (4.22) we have

$$
\begin{aligned}
E(|W-W_i|) &\leq [(n-2)p+1](1-p)^{n-2} \\
&\leq \frac{np+1}{q^2 e^{np}} \\
&\leq \frac{1}{q^2}\Big[\frac{2np}{(np)^2}+\frac{1}{np}\Big] \\
&= \frac{3}{q^2 n^{1-\delta}}
\end{aligned} \tag{4.23}
$$

where $q = 1-p$ and $p = \dfrac{1}{n^{\gamma}}$, $0 < \gamma < 1$.

Then by (4.8), (4.9), (4.12) and (4.23) we obtains Corollary 4.5.

Remark.

In case of non-isolated vertices, (i.e. $w = 0$). Teerapabolarn, Neammanee and Chongcharoen ([58],2004) gave the approximation in the form of

$$
\Big|P(W=0)-e^{-\lambda}\Big| \leq (\lambda + e^{-\lambda}-1)\Big[\frac{(n-2)p+1}{n(1-p)}\Big]
$$

where $\lambda = n(1-p)^{n-1}$, i.e.,

$$
\Big|P(W=0)-e^{-\lambda}\Big| \leq \frac{1}{\lambda}(\lambda + e^{-\lambda}-1)[(n-2)p+1](1-p)^{n-2}. \tag{4.24}
$$

We note from Theorem 4.4 that $C(\lambda, \{0\}) = (1-e^{-\lambda})$. By the fact that $e^{\lambda} \geq 1+\lambda$ we can show that

$$
\frac{1}{\lambda}(\lambda + e^{-\lambda}-1) \leq C(\lambda, \{0\}).
$$

Thus a bound in (4.24) is better than a bound from (4.4).

CHAPTER V

FUTURE RESEARCH

In this chapter, we describe about some future research in normal approximation of the number of isolated trees in a random graph.

A tree is, by definition, a connected graph containing no cycles and a tree in $\mathbb{G}(n,p)$ is isolated if there is no edge in $\mathbb{G}(n,p)$ with one vertex in the tree and the other outside of the tree.

Let $\Lambda := \Lambda(n,k) = \left\{\bar{i} := \{i_1, i_2, \ldots, i_k\} \middle| 1 \le i_1 < i_2 < \cdots < i_k \le n\right\}$ be the set of all possible combinations of k vertices, $k > 1$. For each $\bar{i} \in \Lambda$, we define

$$Y_{\bar{i}} = \begin{cases} 1 & ;\ \text{if there is an isolated tree in } \mathbb{G}(n,p) \text{ that spans the vertices} \\ & \quad \bar{i} = \{i_1, i_2, \ldots, i_k\}, \\ 0 & ;\ \text{otherwise,} \end{cases}$$

Let S be the number of isolated trees of a fixed order k, $k > 1$, in $\mathbb{G}(n,p)$. Then $S = \sum_{\bar{i} \in \Lambda} Y_{\bar{i}}$.

In 1986, Stein[55] proved that the distribution of S can be approximated by Poisson distribution with parameter

$$\lambda = E(S) = \binom{n}{k} P(Y_{\bar{i}} = 1) = \binom{n}{k} k^{k-2} p^{k-1} q^{\binom{k}{2} - (k-1)} q^{(n-k)k}$$

and the uniform bound is given by

$$|P(S \in A) - Poi_\lambda(A)| \le \frac{C}{\sqrt{k}} (1 + c_n) e^{1 - c_n} (c_n e^{1 - c_n})^{k-1}$$

for all $A \subseteq \mathbb{N} \cup \{0\}$, $n \in \mathbb{N}$, and $k \le n$, where Poi_λ is a Poisson random variable with parameter λ and $c_n = -n \log(1-p)$.

K.Neammanee ([47], 2005) gave a pointwise approximation of S by Poisson distribution as follows:

Theorem 5.1. ([47], pp.90) *Suppose* $2k < n$ *and* $w_0 \neq 0$. *Then*

1. $\left|P(S = w_0) - \dfrac{e^{-\lambda}\lambda^{w_0}}{w_0!}\right| \leq \lambda \min\left\{\dfrac{1}{w_0}, \dfrac{1}{\lambda}\right\} \min\left\{2, \dfrac{\lambda k^2}{n}\left(1 + c_n e^{\frac{k^2}{n}(c_n - 1)}\right)\right\}$, *and*

2. $\left|P(S = 0) - e^{-\lambda}\right| \leq \min\left\{1, \lambda\right\} \min\left\{2, \lambda\right\}$.

If we use the idea in chapter 3 to give a uniform bound between $P(W \leq z)$ and $\Phi(z)$, the following is our expected result.

Theorem 5.2. *Let* $k \geq 2$ *and* $p = \dfrac{1}{n^\gamma}$ *for a fixed* $\gamma \in \left[1, 1 + \dfrac{1}{k-1}\right)$, *then there exists a constant* C *such that for* $0 < \beta < 1$,

$$|P(W \leq z) - \Phi(z)| \leq \frac{C}{\sigma_n^\beta}.$$

To proof Theorem 5.2, we define the random variables which the same idea as in chapter 3.

For each $\bar{i} \in \Lambda$, let

$$\Lambda_1^{\bar{i}} = \left\{\bar{j} \in \Lambda |\ \bar{j} \cap \bar{i} = \varnothing\right\} \qquad \text{and} \qquad \Lambda_2^{\bar{i}} = \left\{\bar{j} \in \Lambda |\ \bar{j} \cap \bar{i} \neq \varnothing\right\}$$

and for $\bar{j}, \bar{l} \in \Lambda_1^{\bar{i}}$, let

$$Y_{\bar{i}}^{(\bar{j})} = \begin{cases} 1 & ;\ \text{if there is an isolated tree in } \mathbb{G}(n,p) - \bar{j} \text{ which spaned by a} \\ & \quad \text{vertices } \bar{i} = \{i_1, i_2, \ldots, i_k\}, \\ 0 & ;\ \text{otherwise,} \end{cases}$$

and

$$Y_{\bar{i}}^{(\bar{j},\bar{l})} = \begin{cases} 1 & ;\ \text{if there is an isolated tree in } \mathbb{G}(n,p) - (\bar{j} \cup \bar{l}) \text{ which spaned by a} \\ & \quad \text{vertices } \bar{i} = \{i_1, i_2, \ldots, i_k\}, \\ 0 & ;\ \text{otherwise,} \end{cases}$$

where a random graph $\mathbb{G}(n,p) - \bar{j}$ and $\mathbb{G}(n,p) - (\bar{j} \cup \bar{l})$ obtained from $\mathbb{G}(n,p)$ by removing the vertex in $\bar{j}$ and vertex in $\bar{j} \cup \bar{l}$, respectively. For $\bar{i} \in \Lambda$, let

$$X_{\bar{i}} = \frac{Y_{\bar{i}} - EY_{\bar{i}}}{\sigma},$$

$$Z_{\bar{i}\bar{j}} = \begin{cases} \frac{1}{\sigma} Y_{\bar{j}} & ;\ \bar{j} \in \Lambda_2^{\bar{i}} \\ \frac{1}{\sigma}(Y_{\bar{j}} - Y_{\bar{j}}^{(\bar{i})}), & ;\ \bar{j} \in \Lambda_1^{\bar{i}}, \end{cases}$$

$$Z_{\bar{i}} = \sum_{\bar{j} \in \Lambda} Z_{\bar{i}\bar{j}} = \frac{1}{\sigma}\Big\{ \sum_{\bar{j} \in \Lambda_2^{\bar{i}}} Y_{\bar{j}} + \sum_{\bar{j} \in \Lambda_1^{\bar{i}}} (Y_{\bar{j}} - Y_{\bar{j}}^{(\bar{i})}) \Big\},$$

$$W_{\bar{i}} = W - Z_{\bar{i}},$$

$$V_{\bar{i}\bar{j}} = \begin{cases} 0 & ; \bar{i} = \bar{j} \\ \frac{1}{\sigma} Y_{\bar{j}}^{(\bar{i})} & ;\ \bar{i} \neq \bar{j} \text{ and } \bar{j} \in \Lambda_1^{\bar{i}} \\ \frac{1}{\sigma}\Big\{ \sum_{\bar{l} \in \Lambda_1^{\bar{i}} \cap \Lambda_2^{\bar{j}}} Y_l^{(\bar{i})} + \sum_{\bar{l} \in \Lambda_1^{\bar{i}} \cap \Lambda_1^{\bar{j}}} (Y_l^{(\bar{i})} - Y_l^{(\bar{i},\bar{j})}) \Big\} & ; \bar{i} \neq \bar{j} \text{ and } \bar{j} \in \Lambda_2^{\bar{i}}, \end{cases}$$

$$\text{and} \quad W_{\bar{i}\bar{j}} = \sum_{\substack{\bar{l} \in \Lambda \\ \bar{l} \cap (\bar{i} \cup \bar{j}) = \varnothing}} \frac{1}{\sigma}\Big\{ Y_{\bar{l}}^{(\bar{i}\cup\bar{j})} - E(Y_l^{(\bar{i}\cup\bar{j})}) \Big\} - E(V_{\bar{i}\bar{j}}) - E(Z_{\bar{i}}) = W_{\bar{i}} - V_{\bar{i}\bar{j}}$$

where $\sigma^2 = \mathrm{Var} S$.

Note that $E(X_{\bar{i}}) = 0$, $W = \sum_{\bar{i} \in \Lambda} X_{\bar{i}}$ and $VarW = 1$. By Cayley's Theorem(see, for example, Graver and Watkins [31], p. 322, 1977) there are k^{k-2} different trees on k specified vertices. For a given isolated tree on these k vertices it is necessary and sufficient that the $k-1$ connections of the specified tree be made, but none of the $\binom{k}{2} - (k-1)$ other connections among these k vertices, and that none of the $(n-k)k$ possible connections of these k vertices to vertices outside this set be made. Then we have the expectation of $Y_{\bar{i}}$ for $\bar{i} \in \Lambda$

$$\mu = P(Y_{\bar{i}} = 1) = k^{k-2} p^{k-1} q^{\binom{k}{2} - (k-1)} q^{(n-k)k} \leq k^{k-2} p^{k-1} \tag{5.1}$$

where $q = 1 - p$ and $k \geq 2$.

Barbour, Karoński and Ruciński ([9], 1989) showed that

$$\sigma^2 \geq \Big[1 - \frac{1}{\sqrt{2\pi(k-1)}} \Big] E(S) \tag{5.2}$$

for large n and $k>1$.

To prove the main theorem, we need the following properties of $X_{\bar{i}}, Z_{\bar{i}}, Z_{\bar{i}\bar{j}}$ and $V_{\bar{i}\bar{j}}$.

Proposition 5.3. *For every $\bar{i},\bar{j}\in\Lambda$ and for any $r,r_1,r_2\in\mathbb{N}$.*

1. $E(|Z_{\bar{i}}^r|)\le r2^rk^{r(2k-1)}\dfrac{n^{k-1}p^{k-1}}{\sigma^r}$.
2. $E(|X_{\bar{i}}^{r_1}Z_{\bar{i}}^{r_2}|)\le\left[2^{r_1+r_2}r_2(1+k^{r_2(2k-1)})\right]\dfrac{\mu}{\sigma^{r_1+r_2}}$.
3. $E(|V_{\bar{i}\bar{j}}^r|)\le\left[2^rrk^{r(2k-1)}+k^{k-2}\right]\dfrac{n^{k-1}p^{k-1}}{\sigma^r}$.
4. $E(|X_{\bar{i}}Z_{\bar{i}\bar{j}}|)\le\begin{cases}\dfrac{k^{k-2}\mu p^k}{\sigma^2} & ;\ \bar{j}\in\Lambda_1^{\bar{i}}\\ \dfrac{\mu^2}{\sigma^2} & ;\ \bar{j}\in\Lambda_2^{\bar{i}} \text{ and } \bar{i}\ne\bar{j}\\ \dfrac{\mu}{\sigma^2} & ;\ \bar{j}\in\Lambda_2^{\bar{i}} \text{ and } \bar{i}=\bar{j}.\end{cases}$

Proposition 5.4. *Let $k>1$ and $p=\dfrac{1}{n^\gamma}$ where $\gamma\ge 1$. For $n\ge 2$, $r_1,r_2,r_3\in\mathbb{N}$ and $i,j\in\{1,2,\dots,n\}$, there exists a positive constant C such that*

$$E(|X_i^{r_1}Z_i^{r_2}V_{ij}^{r_3}|)\le\frac{C\mu}{\sigma^{r_1+r_2+r_3}}.$$

Proposition 5.5. *Let $p=\dfrac{1}{n^\gamma}$ for $\gamma\ge1$. Then for $r_1,r_2,r_3\in\mathbb{N}$ and for every $\bar{i},\bar{j}\in\Lambda$, $\bar{i}\ne\bar{j}$,*

$$E(|X_{\bar{i}}^{r_1}Z_{\bar{i}\bar{j}}^{r_2}V_{\bar{i}\bar{j}}^{r_3}|)=\begin{cases}\dfrac{C(k,r_1)\mu^{r_1}p^k}{\sigma^{r_1+r_2+r_3}} & ;\quad \bar{j}\in\Lambda_1^{\bar{i}}\\ 0 & ;\quad \bar{j}\in\Lambda_2^{\bar{i}}.\end{cases}$$

where $C(k,r_1)=2^{r_1-1}k^{k-2}2^{k^2+k-1}$.

The idea of the proof of Proposition 5.3 - Proposition 5.5 follows directly from Proposition 3.4 and Proposition 3.5.

By the same argument of Martin([44], 2003) we can show that (3.25) holds and by using the same technique of chapter 3 we have

$$\begin{aligned}A_1(\varepsilon)&\le\frac{Cn^{\frac{k-1}{2}}}{\sigma},\\ A_2(\varepsilon)+A_3(\varepsilon)&\le\frac{C}{\sigma},\\ B_1(\varepsilon)+B_2(\varepsilon)+B_3(\varepsilon)&\le\frac{\delta}{2}+\frac{C}{\sigma^\beta}.\end{aligned}$$

From (5.1) and (5.2) we can see that $\sigma \sim Cn^{\frac{k}{2}}p^{\frac{(k-1)}{2}}$, then the bound of $A_1(\varepsilon)$ is

$$\frac{Cn^{\frac{k-1}{2}}}{\sigma} \sim \frac{Cn^{\frac{k-1}{2}}}{n^{\frac{k}{2}-\frac{\gamma(k-1)}{2}}} = \frac{C}{n^{\frac{1}{2}-\frac{\gamma(k-1)}{2}}} \not\to 0 \quad \text{as } n \to \infty$$

when $k \geq 2$ and $\gamma \geq 1$.

Hence, to complete Theorem 5.2 it still to improve the bound on $A_1(\varepsilon)$ only.

$\square$

REFERENCES

[1] Abello J., Buchsbaum A.L., and Westbrook J.R. (1998). A functional approach to external graph algorithms, In Algprithms-ESA'98 (Venice). *Lecture Notes in Comput. Sci.*, **1461**:332-343.

[2] Anderson R.M. and May R.M. (1995). Susceptible-infectious-recovered epidemic models with dynamic partnerships. *J. Math. Biology*, **33**:661-675.

[3] Ball F., Mollison and Scalia-Tomba G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Probab.*, **7**:46-89.

[4] Barabási A., Albert R. and Jeong H. (1999). Scale-free characteristics of random networks: The topology of the world wide web. *Physica A*, **272**:173-187.

[5] Barbour A.D., Eagleson G.K. (1985). Multiple comparisons and sums of dissociated random variables. *Adv. Appl. Probab.*, **17**:147-162.

[6] Barbour A.D., Eagleson G.K. (1986). Random association of symmetric arrays. *Stochastic Anal. Appl.*, **4**:239-281.

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

[8] Barbour A.D. (1982). Poisson convergence and random graphs. *Math. Proc. Cambridge Philos. Soc.*, **92**:349-359.

[9] Barbour A.D., Karoński M., Ruciński A. (1989). A central limit theorem for decomposable random variables with application to random graphs. *J. Comb. Th. Ser. B*, **47**:125-145 .

[10] Barbour A.D. and Hall P. (1985). On bounds to the rate of convergence in the central limit theorem. *Bull. London Math. Soc.*, **17**:151-156.

[11] Bhalla U.S. and Iyengar R. (1999). Emergent properties of networks of biological signalling pathways. *Science*, **283**:381-387.

[12] Bower J.M. and Bolouri H. (2001). *Computational Modeling of genetic and biochemical networks*, The MIT Press, Cambridge, MA.

[13] Broder A., Kumar S.R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tompkins A. and Wiener J. (2000). Graph structure in the web. *Comput.Netw.*, **33**:209-321.

[14] Chen L.H.Y. (1975). Poisson approximation for dependent trials, *Annals of probability*, **3**:534-545.

[15] 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.

[16] Chen L.H.Y. and Shao Q.M. (2005). Stein's method for normal approximation. *An Introduction to Stein's Method, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Singapore University Press and World Scientific*, **4**:1-59.

[17] Cohen J.E., Briand E. and Newman C.M. (1990). Community food webs, *Data and Theory.* Springer-Verlag, Berlin.

[18] Cohen R., Erez K., Ben Avraham D. and Havlin S. (2000). Resilience of the Internet to random breakdowns. *Phys. Rev. Lett.*, **85**:4626-4628 .

[19] Derrida B. and Pomeau Y. (1986). Random networks of automata: A simple annealed approximation. *Europhys. Lett.*, **1**:45-49.

[20] Durrett, R. (1991). *Probability: Theory and example.* U.S.A.: Brooks/Cole Publishing Company.

[21] Erdős P. (1947). Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, **53**:292-294.

[22] Erdös P. and Rény A. (1959). On random graphs. I. *Publ. Math. Debrecen*, **6**:290-297 .

[23] Erdös P. and Rény A. (1960). On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*, **5**:17-61.

[24] Erdös P. and Rény A. (1961). On the strength of connectedness of a random graph. *Acca Math. Acad. Sci. Hungar.*, **12**:261-267.

[25] Erdős P. and Rényi A. (1964). On random matrices. *Publ. Math. Inst. Hung. Acad. Sci.*, **8**:455-461.

[26] Erdős P. and Rényi A. (1966). On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hung.*, **17**:359-368.

[27] Erdős P. and Rényi A. (1968). On random matrices II. *Stud. Sci. Math. Hung.*, **3**:459-464.

[28] Faloutsos M., Faloutsos P. and Faloutsos C. (1999). On power-law relationships of the internet topology. *Comput. Comm. Rev.*, **29**:251-262 .

[29] Granovetter M. (1978). Threshold models of collective behavior. *Am.J.Soc.*, **83**:1420-1443.

[30] Glance N.S. and Huberman B.A. (1993). The outbreak of cooperation. *J.Math.Soc.*, **17**:281-302.

[31] Graver J. E. and Watkins M. E. (1977). *Combinatorics with Emphasis on the Theory of Graphs.* Springer-Verlag, New York.

[32] Janson S. (1987), Poisson convergence and Poisson processes with applications to random graphs, *Stochastic Process. Appl.*, **26**: 1-30.

[33] Janson S. (1994). Coupling and Poisson approximation. *Acta Applicandae Mathematicae*, **34**:7-15.

[34] Karoński M. *Balanced subgraphs of large random graphs*, Adam Mickiewicz, Univ. Press, Poznan.

[35] Karoński M. and Ruciński A. (1987). Poisson convergence and semi-induced properties random graphs. *Math.Proc. Cambridge Philos. Soc.*, **101**:291-300.

[36] Kauffman S.A. (1984). Emergent properties in random complex automata. *Physica D*, **10**:145-456.

[37] Kauffman S.A. (1993). *The origins of order: Self-organization and selection in evolution*, Oxford University Press, New york.

[38] Keeling M.J. (1999). The effects of local spatial structure on epidemiological invasions. *Proc. Royal Soc. Lond. B*, **266**:859-867.

[39] Kordecki W. (1987). Normal approximation and isolated vertices in random graphs. *Random graphs' 97*, Proceedings, Pozanń, eds. M. Karoński, J. Jaworski and A. Ruciński, John Wiley&Sons, Chichester, 131-139.

[40] Kosterev D.N., Taylor C.W. and Mittelstadt W.A. (1999). A model validation for the August 10, 1996 WSCC system outage. *IEEE Trans. Power Systems*, **14**:967-979.

[41] Kretschmar M. and Morris M. (1996). Measures of concurrency in networks and the spread of infectious disease. *Math. Biosciences*, **133**:165-195.

[42] Laha, R.G. and Rohatgi, V.K. (1979). *Probability Theory.* New York: John Wiley & Son.

[43] Lynch J.F. (2002). Critical points for random boolean networks. *Physica D*, **172**:49-64.

[44] Martin R. (2003). Normal Aproximation by Stein's method. *Proc.of the seventh young statisticians meeting*, 71-97.

[45] Neammanee K. and Suntadkarn A. (2007). A bound on normal approximation of number of vertices of a fixed degree in a random graph. *International Journal of Pure and Applied Mathematics* **40**:241-255.

[46] Neammanee K. (2003). Pointwise approximation of Poisson binomial by Poisson distribution. *Stochastic Modelling and Applications*, **6**:20-26

[47] Neammanee K. (2005). A pointwise approximation of isolated trees in a random graph,*Acta. Math.* 2189-99.

[48] Newman M.E.J. (2001). The structure of scientific collaboration networks.*Prod. Nat. Acad. Sci.*,**98**:404-409.

[49] Newman M.E.J., Strogatz S.H. and Watts D.J. (2001). Random graphs with arbitary degree distributions and their applications. *Phys. Rev. E*,**64**:026118.

[50] Nowicki K. (1987). Asymtotic Poisson distributions with applications to statistical analysis of graphs, *Appl. Probab.*, in press

[51] Petrov, V.V. (1995). *Limit theorem of probability: Sequence of independent random variables.* Oxford studies in Probability 4, Clarendon Press, Oxford.

[52] Santiwipanont T. and Teerapabolarn K. (2006). Two formulas of non-uniform bounds on Poisson approximation for dependent indicators. *Thai journal of mathematics.*

[53] Sattenspiel L. and Simon C.P. (1988). The spread and persistence of infectious diseases in structured populations. *Math. Biosci.*, **90**:367-383.

[54] Stein C.M. (1972). A bound for the error in normal approximation to the distribution of a sum of dependent random variables. *Proc.Sixth Berkeley Sympos, Math. Statist. Probab.* **3**:583-602.

[55] Stein C.M. (1986). Approximate computation of expectations, Institute of Mathematical Statistics *Lecture Notes-Monograph Series,* Institute of Mathematical Statistics, Hayward, CA,**7**.

[56] Strogatz S.H. (2001). Exploring complex networks. *Nature*, **410**:268-276

[57] Teerapabolarn K. and Neammanee K. (2005). A non-uniform bound on Poisson approximation for dependent trials. *Stochastic Modelling and Applications*, **8**:17-31.

[58] Teerapabolarn K., Neammanee K. and Chongcharoen S. (2004). Approximation of the probability of non-isolated vertices in random graph. *Annual Meeting in Applied Statistics 2004*, National Institute of Department Administration, 9-18.

[59] Valente T.W. (1996). Social network thresholds in the diffusion of innovations. *Social Netw.*, **18**:69-89.

[60] Williams R.J. and Martinez N.D. (2000). Simple rules yield complex food webs. **Nature**, **404**:180-183.

APPENDIX

APPENDIX A

The solution of Stein's equation $f'(w) - wf(w) = I_z(w) - \Phi(z)$ is of the form

$$f_z(w) = \begin{cases} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(z)] & \text{if } w \le z \\ \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(z)[1-\Phi(w)] & \text{if } w > z. \end{cases}$$

Moreover the first derivative of f_z does not exists at $w = z$.

Proof. From Stein ([55], pp.22) we have

$$f_z(w) = \begin{cases} \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(w)[1-\Phi(z)] & \text{if } w \le z \\ \sqrt{2\pi}e^{\frac{w^2}{2}}\Phi(z)[1-\Phi(w)] & \text{if } w > z. \end{cases}$$

We use this form and L'Hopital's rule to consider the left and the right derivatives of f_z as follows.

$$\begin{aligned} \lim_{h\to 0^-} \frac{f_z(z+h) - f_z(z)}{h} &= \lim_{h\to 0^-} \frac{\sqrt{2\pi}e^{\frac{(z+h)^2}{2}}\Phi(z+h)[1-\Phi(z)] - \sqrt{2\pi}e^{\frac{z^2}{2}}\Phi(z)[1-\Phi(z)]}{h} \\ &= \sqrt{2\pi}[1-\Phi(z)] \lim_{h\to 0^-} \frac{e^{\frac{(z+h)^2}{2}}\Phi(z+h) - e^{\frac{z^2}{2}}\Phi(z)}{h} \\ &= \sqrt{2\pi}[1-\Phi(z)] \lim_{h\to 0^-} \Big[\frac{1}{\sqrt{2\pi}} + \Phi(z+h)(z+h)e^{\frac{(z+h)^2}{2}}\Big] \\ &= \sqrt{2\pi}[1-\Phi(z)]\Big[\frac{1}{\sqrt{2\pi}} + z\Phi(z)e^{\frac{z^2}{2}}\Big] \\ &= [1-\Phi(z)]\Big[1 + \sqrt{2\pi}z\Phi(z)e^{\frac{z^2}{2}}\Big] \end{aligned}$$

and

$$\begin{aligned} \lim_{h\to 0^+} \frac{f_z(z+h) - f_z(z)}{h} &= \lim_{h\to 0^+} \frac{\sqrt{2\pi}e^{\frac{(z+h)^2}{2}}\Phi(z)[1-\Phi(z+h)] - \sqrt{2\pi}e^{\frac{z^2}{2}}\Phi(z)[1-\Phi(z)]}{h} \\ &= \sqrt{2\pi}\Phi(z) \lim_{h\to 0^+} \frac{e^{\frac{(z+h)^2}{2}}[1-\Phi(z+h)] - e^{\frac{z^2}{2}}[1-\Phi(z)]}{h} \\ &= \sqrt{2\pi}\Phi(z) \lim_{h\to 0^+} \Big[e^{\frac{(z+h)^2}{2}} - \frac{1}{\sqrt{2\pi}} + [1-\Phi(z+h)](z+h)e^{\frac{(z+h)^2}{2}}\Big] \\ &= \sqrt{2\pi}\Phi(z)\Big[e^{\frac{z^2}{2}} - \frac{1}{\sqrt{2\pi}} + [1-\Phi(z)]ze^{\frac{z^2}{2}}\Big]. \end{aligned}$$

Thus $\lim_{h\to 0^-} \frac{f_z(z+h) - f_z(z)}{h} \neq \lim_{h\to 0^+} \frac{f_z(z+h) - f_z(z)}{h}$, this implies that $f_z'(z)$ does not exists.

$\square$

VITA

Name	:Miss Angkana Suntadkarn
Date of Birth	:December 2, 1976
Place of Birth	:Chumporn, Thailand
Previous Studies	:B.Sc.(Mathematics), Khon Kean University, 1999,
	M.Sc.(Mathematics), Chulalongkorn University, 2001.
Work Experience	:Khon Kean University.