The Bounds on Poisson Approximation of the Number of Copies of a Fixed Graph in a Random d-Regular Graph

Mana Dongoanont and Angkana Boonyued

Department of Mathematics
Faculty of Science Khonkean University
Khonkean, Thailand
doodeejung-me@hotmail.com
sangka@kku.ac.th

Abstract

Let $W$ be the number of copies of a fixed graph $H$ in a random d-regular graph $G_{n,d}$. In this paper, we give the bound on Poisson approximation of $W$ by using the Stein-Chen method.

Keywords: Random d-regular graph, Stein’s method and Local Approach

1 Introduction

In 1960 Erdős and Rényi wrote a fundamental paper[5], which has become the starting point of the theory of random graph.

Let $G(n, p)$ be a random graph on $n$ labeled vertices $1, 2, ..., n$ and the edges added randomly such that each of the $\binom{n}{2}$ possible edges exists with probability $p$, $0 < p < 1$.

This paper is a survey of that part of random graph theory in which the degrees of vertices are restricted. Such work concentrates on regular graphs as the most interesting examples, and the results on regular graphs often extend easily to more general degree sequences. Let $1 \leq d \leq n - 1$ be two positive integer, a random graph $G_{n,d}$ is obtained by sampling uniformly at random over the set of all simple $d$-regular graph is on a fixed set of $n$ vertices. We refer the readers to Wormald’s survey [3] for more information (both historical and technical) about this model. As usual, $G_{n,d}$ denotes the random $d$-regular graph with $n$ labeled vertices and for each vertex has the same degree $d$. For each $H = (V(H), E(H))$ be a graph, we use the notation $v_H = |V(H)|$ and $e_H = |E(H)|$ for the number of vertices and edges, respectively. If a subgraph
$H'$ of a graph $F$ is isomorphic to graph $H$, then $H'$ is called a copy of $H$ in $F$. Let

$$\Gamma = \{i =: \{i_1, i_2, \ldots, i_{v_H}\} \mid 1 \leq i_1 < \ldots < i_{v_H} \leq n\}$$

be the set of all possible combinations of $v_H$ vertices. For each $i \in \Gamma$, we define the indicator random variable

$$X_i = \begin{cases} 
1 & \text{if there is a copy of $H$ in $G_{n,d}$ that spans the vertices } i = (i_1, \ldots, i_{v_H}), \\
0 & \text{otherwise},
\end{cases}$$

and

$$W = \sum_{i \in \Gamma} X_i.$$

Then $W$ is the number of copies of $H$ in $G_{n,d}$.

In 2007 Jeong Han Kim, Benny Sodakov and Van Vu [2] proved that the distribution function of $W$ can be approximate by Poisson distribution as the following result.

**Theorem 1.1.** Let $H$ be a strictly balanced graph with $v$ vertices and $e \geq v$ edges. Let $\text{Aut}(H)$ be the number of automorphisms of $H$ and let $W$ be the number of copies of $H$ in a random $d$-regular graph $G_{n,d}$. If $(d - 1)n^{-1+\frac{1}{v_H}} \to c$ for some positive constant $c$, then $W$ converges to $\text{Poi}_\lambda$, the Poisson distribution with mean $\lambda = \frac{ce}{\text{Aut}(H)}$.

In this paper, we give the bound of this approximation by using Stein-Chen method. The following theorem is our main result.

**Theorem 1.2.** Let $H$ be a fixed graph with $v_H$ vertices and $e_H \geq v_H$ edges and let $W$ be the number of copies of $H$ in $G_{n,d}$. If $d = n^\delta$ where $\delta < 1$ such that $(1 - \delta)e_H \geq 2v_H$ then there exists a constant $C_H > 0$ such that

$$\sup_{A \subseteq \mathbb{R}} |P(W \in A) - \text{Poi}_\lambda(A)| \leq \frac{C_H}{n(1-\delta)e_H-2v_H}.$$

This paper is organized as follows. In section 2, we introduce the Stein-Chen method for Poisson approximation and local approach which use in the proof of main result in section 3. Throughout this paper, $C_H$ stands for an absolute constant depend on $H$ and which possible different values in different places.
2 Stein-Chen method and Local Approach

In 1972, Stein[6] 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 [4] applied Stein’s idea to the Poisson case.

For the rest of the section, unless explicitly stated otherwise, we use the following notation. Γ is the index set, which is finite. In most cases Γ = {1, .., n}, let $W = \sum_{i \in \Gamma} X_i$ where $X_i$ are indicator variables and $\lambda = \mathbb{E}(W)$.

Our goal will be to bound the total variation distance between distribution of $W$ and $\text{Poi}_\lambda$,

$$\sup_{A \subset \mathbb{Z}_+} |\mu(A) - \text{Poi}_\lambda(A)|,$$

where $\text{Poi}_\lambda$ is the Poisson distribution with parameter $\lambda$.

Our starting point is the Stein equation for Poisson distribution, which gives,

$$I_A(j) - \text{Poi}_\lambda(A) = \lambda f_A(j + 1) - j f_A(j)$$

(1)

where $\lambda > 0$ and $j \in \mathbb{N} \cup \{0\}$, $f_A$ is the unique solution for Stein’s equation, $A \subseteq \mathbb{N} \cup \{0\}$ and $I_A : \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$ be defined by

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

By substituting $j$ and $\lambda$ in (1) by any integer-valued random variable $W$ and $\lambda = \mathbb{E}(W)$, we have

$$P(W \in A) - \text{Poi}_\lambda(A) = \mathbb{E}(\lambda f_A(W + 1)) - W f_A(W))$$

(2)

From this we get

$$\sup_{A \subset \mathbb{Z}_+} |P(W \in A) - \text{Poi}_\lambda(A)| = \sup_{A \subset \mathbb{Z}_+} |\mathbb{E}(\lambda f_A(W + 1) - W f_A(W))|.$$  

(3)

To bound the right hand side in (3), a local and coupling approach have been suggested. The first one was used by Chen(1975) and in convenient when the dependence structure of the indicator variables is local (meaning that each indicator is independent of “most” of the other). The idea of combining the Stein’s equation with local approach was stated in 2005 by A.D. Barbour and Louis H.Y. Chen[1]. It follows that
Theorem 2.1. (The local approach). Let $W = \sum_{i \in \Gamma} X_i$, $\{X_i; i \in \Gamma\}$ are indicator variables. For each $i \in \Gamma$, divide $\Gamma \setminus \{i\}$ into two subsets $\Gamma^s_i$ and $\Gamma^w_i$, so that, informally,

\[ \Gamma^s_i = \{ j \in \Gamma \setminus \{i\}; X_j \text{"strongly" dependent on } X_i \}. \]

Let $Z_i = \sum_{j \in \Gamma^s_i} X_j$ and $W_i = \sum_{j \in \Gamma^w_i} X_j$. Then

\[ d_{TV}(\mathcal{L}(W), \text{Poi}_\lambda) \leq k_2(\lambda) \sum_{i \in \Gamma} (p_i \mathbb{E}(X_i + Z_i) + \mathbb{E}(X_i Z_i)) + k_1(\lambda) \sum_{i \in \Gamma} \mathbb{E}|p_i - \mathbb{E}(X_i|W_i)|, \]

where $k_1(\lambda) := (1 \wedge \sqrt{\frac{2}{e\lambda}})$ and $k_2(\lambda) := (1 - e^{-\lambda})$ and $p_i = \mathbb{E}(X_i)$ for each $i \in \Gamma$.

In next section, we will use Theorem 2.1 to prove our main result.

3 Proof of Main Result

In this section we prove the our main theorem. Let $H$ be a fixed graph with $v_H$ vertices and $e_H$ edges such that $e_H \geq v_H$. Let

\[ \Gamma = \{i =: \{i_1, i_2, \ldots, i_{v_H}\} | 1 \leq i_1 < \ldots < i_{v_H} \leq n\}, \]

and

\[ X_i = \begin{cases} 
1 & \text{if there is a copy of } H \text{ in } \mathbb{G}_{n,d} \text{ that spans the vertices } i = (i_1, \ldots, i_{v_H}), \\
0 & \text{otherwise.}
\end{cases} \]

We divide $\Gamma \setminus \{i\}$ into two subsets as follow

\[ \Gamma^w_i = \{ j \in \Gamma \setminus \{i\} | i \cap j = \phi \}, \Gamma^s_i = \{ j \in \Gamma \setminus \{i\} | i \cap j \neq \phi \}, \]

and define $Z_i = \sum_{j \in \Gamma^s_i} X_j$ and $W_i = \sum_{j \in \Gamma^w_i} X_j$.

J.H. Kim, B. Sodakov and V. Vu [2] show that the expectation of the number of copies of a graph $H$ in $\mathbb{G}_{n,d}$ is asymptotically the same as the expectation of the number of copies of $H$ in $\mathbb{G}(n,p)$, where $p = \frac{d}{n}$. Then

\[ \mathbb{E}(X_i) = p_i = P(X_i = 1) = \frac{\nu_H!}{\text{Aut}(H)} \left( \frac{d}{n} \right)^{e_H} \left( 1 - \frac{d}{n} \right)^{v_H - e_H}. \]
Poisson approximation in random d-regular

for all $i \in \Gamma$ and

$$
\lambda = \mathbb{E}(W) = \binom{n}{v_H} P(X_i = 1) = \binom{n}{v_H} \frac{v_H!}{\text{Aut}(H)} \left( \frac{d}{n} \right)^{e_H} (1 - \frac{d}{n})^{\frac{ad - e_H}{2}},
$$

where $\frac{v_H!}{\text{Aut}(H)}$ is a number copies of $H$ which spans the vertex $i$.

By definition of $Z_i$, we have

$$
\mathbb{E}(Z_i) = \mathbb{E} \left( \sum_{j \in \Gamma_i^s} X_j \right) = \sum_{j \in \Gamma_i^s} \mathbb{E}(X_j) = p_i \left[ \binom{n}{v_H} - \left( \frac{n - v_H}{v_H} \right) - 1 \right].
$$

(5)

Now, we consider $\mathbb{E}(X_iX_j)$ for $j \in \Gamma_i^s$.

Case 1 $E(H') \cap E(H'') = \phi$.

$$
\mathbb{E}(X_iX_j) = P(X_i = 1, X_j = 1) = \left[ \frac{v_H!}{\text{Aut}(H)} \right]^2 \left( \frac{d}{n} \right)^{e_H} (1 - \frac{d}{n})^{\frac{ad - e_H}{2}}.
$$

(6)

Case 2 $E(H') \cap E(H'') \neq \phi$.

Let $F$ be a subgraph of $H$ such that $F$ isomorphic to $H' \cap H''$. Then we have

$$
\mathbb{E}(X_iX_j) = P(X_i = 1, X_j = 1) = \left[ \frac{v_H!}{\text{Aut}(H)} \right]^2 \left( \frac{d}{n} \right)^{2e_H-e_F} (1 - \frac{d}{n})^{\frac{ad - (2e_H - e_F)}{2}}.
$$

(7)

From (6) and (7), we have

$$
\mathbb{E}(X_iX_j) = \left( \frac{v_H!}{\text{Aut}(H)} \right)^2 \left( \frac{d}{n} \right)^{e_H} (1 - \frac{d}{n})^{\frac{ad - e_H}{2}} \left( \frac{d}{n} \right)^{2e_H-e_F} (1 - \frac{d}{n})^{\frac{ad - (2e_H - e_F)}{2}}.
$$

Thus,

$$
\mathbb{E}(X_iZ_i) = \mathbb{E}(X_i \sum_{j \in \Gamma_i^s} X_j)
$$

$$
= \sum_{i \in \Gamma_i^s} \mathbb{E}(X_iX_j)
$$

$$
\leq C_H \left( \frac{n}{v_H} \right) \left[ \left( \frac{d}{n} \right)^{e_H} (1 - \frac{d}{n})^{\frac{ad - e_H}{2}} \left( \frac{d}{n} \right)^{2e_H-e_F} (1 - \frac{d}{n})^{\frac{ad - (2e_H - e_F)}{2}} \right],
$$

(8)
where $C_H$ stands for an absolute constant depend on $H$.

By Theorem 2.1, (4), (5) and (8), we get

$$\sup_{A \subset \mathbb{Z}^+} |P(W \in A) - Po\lambda(A)|$$

$$\leq k_2(\lambda) \sum_{i \in \Gamma} (p_i \mathbb{E}(X_i + Z_i) + \mathbb{E}(X_i Z_i))$$

$$= k_2(\lambda) \sum_{i \in \Gamma} [p_i \mathbb{E}(X_i) + p_i \mathbb{E}(Z_i) + \mathbb{E}(X_i Z_i)]$$

$$\leq k_2(\lambda) C_H \left( \frac{n}{v_H} \right) \left\{ \left[ \left( \frac{d}{n} \right)^{e_H (1 - \frac{d}{n} \frac{n^d - e_H}{2})} \right]^2 + \left( \frac{n}{v_H} \right) \left( \left( \frac{d}{n} \right)^{e_H (1 - \frac{d}{n} \frac{n^d - e_H}{2})} \right)^2 \right\}$$

$$+ \left( \frac{n}{v_H} \right) \left[ \left( \frac{d}{n} \right)^{e_H (1 - \frac{d}{n} \frac{n^d - e_H}{2})} \right]^2 + \left( \frac{d}{n} \right)^{2e_H - e_F (1 - \frac{d}{n} \frac{n^d - e_H - e_F}{2})} \right\}.$$

By the fact that $e_F \leq e_H$ and $(1 - \frac{d}{n} \frac{n^d - e_H}{2})$ converges to some positive constant when $d = n^\delta$ for $\delta < 1$ such that $(1 - \delta)e_H \geq 2v_H$, we have for $d = n^\delta$

$$\sup_{A \subset \mathbb{Z}^+} |P(W \in A) - Po\lambda(A)|$$

$$\leq C_H \left( \frac{n^{v_H}}{n^{2(1-\delta)e_H}} + \frac{n^{v_H}}{n^{2(1-\delta)e_H - 2v_H}} + \frac{n^{v_H}}{n^{2(1-\delta)e_H - 2v_H}} + \frac{n^{v_H}}{n^{2(1-\delta)(2e_H - e_F) - 2v_H}} \right)$$

$$= C_H \left[ \frac{1}{n^{2(1-\delta)e_H - v_H}} + \frac{1}{n^{2(1-\delta)e_H - 2v_H}} + \frac{1}{n^{2(1-\delta)e_H - 2v_H}} + \frac{1}{n^{2(1-\delta)(2e_H - e_F) - 2v_H}} \right]$$

$$\leq C_H \left[ \frac{1}{n^{(1-\delta)e_H - 2v_H}} \right].$$

This complete the proof.

**References**


Received: August, 2009