International Journal of Pure and Applied Mathematics
Volume 95 No. 1 2014, 113-121
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v95i1.13
ijpam.eu

POISSON APPROXIMATION FOR THE NUMBER OF INDUCED COPIES OF A FIXED GRAPH IN A RANDOM REGULAR GRAPH

Mana Donganont

School of Mathematics
Faculty of Science
University of Phayao
Phayao, 56000, THAILAND

Abstract: Let $\mathbb{G}_{n,d}$ be a random d-regular graph with n vertices. Given a fixed graph H. W denotes the number of induced copies of H in $\mathbb{G}_{n,d}$. In this paper, we use Stein-Chen method and Local approach to show that W can approximate by the Poisson distribution and give the bound of this approximation.

Key Words: induced subgraph, a copy of graphs, Poisson distribution, Random regular graph, strictly balanced, Stein's method and local approach

1. Introduction

A random graph $\mathbb{G}_{n,p}$, written by Erdös and Rényi [5] in 1960, is the graph with 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, $0 < p < 1$.

In this work, the author interests in the part of random graph theory in which the degrees of vertices are restricted. Such work focus 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 $\mathbb{G}_{n,d}$ 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

Received: May 30, 2014

refer the readers to Wormald's survey [3] for more information (both historical and technical) about this model. As usual, $\mathbb{G}_{n,d}$ 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. Call $\rho(H) = \dfrac{e_H}{v_H}$ the density of H, and $m(H)$ the density of the densest subgraph of H, that is, $m(H) = max\{\rho(H')|H' \subseteq H\}$ and for every $H' \subseteq H$, $\rho(H') \leq \rho(H)$, then H is called balanced, and H is called strictly balanced if for every proper subgraph H', $\rho(H') < \rho(H)$. If a subgraph $H^{'}$ of a graph G is isomorphic to graph H, then $H^{'}$ is called a copy of H in G.

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

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

and

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

Then W is the number of copies of H in $\mathbb{G}_{n,d}$.

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

Clearly, the containing of at least one induced copies of H in $\mathbb{G}_{n,d}$ should be closely related to the containing of at least one copies of H in $\mathbb{G}_{n,d}$.

In 2008, Lan XIAO, Guiyin YAN, Yuwen WU and Wei REN[8] showed that the number of induced copies H in random regular graph $\mathbb{G}_{n,d}$ can be approximated by Poisson distribution with mean $\lambda = \frac{c^e}{aut(H)}$. The result as following,

Theorem 1.2 (8). *Let H be a strictly balanced graph with v vertices and $e > v$ edges.Y_H denote the number of induce copies of H in $\mathbb{G}_{n,d}$, and $Aut(H)$ is the number of automorphisms of H. If $dn^{-1+\frac{1}{m(H)}} \to c$ for some positive constant c, then Y_H converges to Poi_λ , the Poisson distribution with mean $\lambda = \frac{c^e}{Aut(H)}$.*

In 2010 ,Mana Dongoanont and Angkana Boonyued [7] show that the number of copies of H in $\mathbb{G}_{n,d}$ can approximate by the Poisson distribution and give the bound of this approximation. The result as following,

Theorem 1.3 (7). *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 $\mathbb{G}_{n,d}$. 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*

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

In this paper, we use Stein-Chen method and Local approach to show that the number of induced copies H in random regular graph $\mathbb{G}_{n,d}$ can approximate by the Poisson distribution and give the bound of this approximation. The following theorem is our main result,

Theorem 1.4. *Let H be a strictly balanced graph with v vertices and $e > v$ edges. W denotes the number of induced copies of H in $\mathbb{G}_{n,d}$, and $Aut(H)$ is the number of automorphisms of H. If $d = n^\delta$ where $\delta < 1$ such that $(1-\delta)(2e_H - e_F) \geq 2v_H$, then W converges to Poisson distribution and there exists a constant $C_H > 0$ such that*

$$\sup_{A \subseteq \mathbb{R}} |P(W \in A) - Poi_\lambda(A)| \leq \frac{C_H}{n^{(1-\delta)(2e_H - e_F) - 2v_H}},$$

where, $e_F > 0$ is the number of edges of any subgraph $F \subseteq H$ such that F isomorphic to $H^{'} \cap H^{''}$, where $H^{'}$ and $H^{''}$ are copies of H sharing at least one edge.

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 possibly different values in different places.

2. Stein-Chen method and Local Approach

In 1972, Stein[6] gave a wonderful technique to find a bound in the normal approximation to a distribution of a sum of dependent random variables. His technique was depended 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 $\Gamma = \{1, .., n\}$, 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 Poi_λ,

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

where Poi_λ is the Poisson distribution with parameter λ.

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

$$I_A(j) - Poi_\lambda(A) = \lambda f_A(j + 1) - j f_A(j) \tag{1}$$

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 $I_A : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be defined by

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

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

$$P(W \in A) - Poi_\lambda(A) = \mathbb{E}(\lambda f_A(W + 1)) - W f_A(W)). \tag{2}$$

From this we get

$$\sup_{A\subset\mathbb{Z}_+} |P(W \in A) - Poi_\lambda(A)| = \sup_{A\subset\mathbb{Z}_+} |\mathbb{E}(\lambda f_A(W + 1) - W f_A(W))|. \tag{3}$$

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 (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* Γ_i^s *and* Γ_i^w , *so that, informally,*

$$\Gamma_i^s = \{j \in \Gamma \setminus \{i\}; X_j \text{“strongly” dependent on } X_i\}.$$

Let $Z_i = \sum_{\Gamma_i^s} X_j$ *and* $W_i = \sum_{\Gamma_i^w} X_j$. *Then*

$$\begin{aligned} &d_{TV}(\mathcal{L}(W), Poi_\lambda) \\ &\qquad \le k_2(\lambda) \sum_{i\in\Gamma} (p_i \mathbb{E}(X_i + Z_i) + \mathbb{E}(X_i Z_i)) + k_1(\lambda) \sum_{i\in\Gamma} \mathbb{E}|p_i - \mathbb{E}(X_i|W_i)|, \end{aligned}$$

where $k_1(\lambda) := (1 \wedge \sqrt{\frac{2}{e\lambda}})$ *and* $k_2(\lambda) := (\frac{1 - e^{-\lambda}}{\lambda})$ *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 $e_H > v_H$. Let

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

and

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

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

$$\Gamma_i^w = \{j \in \Gamma \setminus \{i\} \mid i \cap j = \phi\}, \Gamma_i^s = \{j \in \Gamma \setminus \{i\} \mid i \cap j \ne \phi\},$$

and define $Z_i = \sum_{j\in\Gamma_i^s} X_j$ and $W_i = \sum_{j\in\Gamma_i^w} X_j$.

Lan XIAO, Guiyin YAN, Yuwen WU and Wei REN[8] show that the expectation of the number of induced copies of a graph H in $\mathbb{G}_{n,d}$ is asymptotically the same as the expectation of the number of induced copies of H in $\mathbb{G}(n,p)$, where $p=\dfrac{d}{n}$.

Then

$$\mathbb{E}(X_i)=p_i=P(X_i=1)=\frac{v_H!}{Aut(H)}(\frac{d}{n})^{e_H}(1-\frac{d}{n})^{\binom{v_H}{2}-e_H} \tag{4}$$

for all $i\in\Gamma$ and

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

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

By definition of Z_i, we have

$$\mathbb{E}(Z_i)=\mathbb{E}(\sum_{j\in\Gamma_i^s}X_j)=\sum_{j\in\Gamma_i^s}\mathbb{E}(X_j)=p_i\Big[\binom{n}{v_H}-\binom{n-v_H}{v_H}-1\Big]. \tag{5}$$

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

Let $H^{'}$ and $H^{''}$ be induced copies of H in $\mathbb{G}_{n,d}$ such that spans by the vertices i and j respectively.

Case 1. For any two copies $H^{'}$,$H^{''}$ with at most one vertex in common

$$\mathbb{E}(X_iX_j)=P(X_i=1,X_j=1)=\left[\frac{v_H!}{Aut(H)}\right]^2(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H}. \tag{6}$$

Case2 : For any two copies $H^{'}$,$H^{''}$ sharing at least one edge given any subgraph $F\subseteq H$ with $e_F>0$ such that F isomorphic to $H^{'}\cap H^{''}$. Then we have

$$\begin{aligned}\mathbb{E}(X_iX_j)&=P(X_i=1,X_j=1)\\&=\left[\frac{v_H!}{Aut(H)}\right]^2(\frac{d}{n})^{2e_H-e_F}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H+e_F-\binom{v_F}{2}}.\end{aligned} \tag{7}$$

Case3 : Given t vertices, there are n^{2v-t} pairs of edge-disjoint copies $H^{'}$,$H^{''}$ in K_n sharing the t vertices, where $t\geq 2$. Hence,

$$\begin{aligned}\mathbb{E}(X_iX_j)&=P(X_i=1,X_j=1)\\&=\left[\frac{v_H!}{Aut(H)}\right]^2(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H-\binom{t}{2}}.\end{aligned} \tag{8}$$

From (6),(7)and (8), we have

$$\mathbb{E}(X_iX_j) = \left(\frac{v_H!}{Aut(H)}\right)^2 \Bigg[(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H} + (\frac{d}{n})^{2e_H-e_F}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H+e_F-\binom{v_F}{2}} + (\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H-\binom{t}{2}}\Bigg].$$

Thus

$$\begin{aligned}\mathbb{E}(X_iZ_i) &= \mathbb{E}(X_i \sum_{i\in\Gamma_i^s} X_j)\\ &= \sum_{i\in\Gamma_i^s} \mathbb{E}(X_iX_j)\\ &\le C_H\binom{n}{v_H}\Bigg[(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H}\\ &\quad +(\frac{d}{n})^{2e_H-e_F}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H+e_F-\binom{v_F}{2}}\\ &+(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H-\binom{t}{2}}\Bigg], \qquad (9)\end{aligned}$$

where C_H stands for an absolute constant depend on H.

By Theorem 2.1, (4), (5) and (9), we get

$$\begin{aligned}&\sup_{A\subset\mathbb{Z}_+} |P(W\in A) - Poi_\lambda(A)|\\ \le& k_2(\lambda)\sum_{i\in\Gamma}(p_i\mathbb{E}(X_i+Z_i)+\mathbb{E}(X_iZ_i))\\ =& k_2(\lambda)\sum_{i\in\Gamma}[p_i\mathbb{E}(X_i)+p_i\mathbb{E}(Z_i)+\mathbb{E}(X_iZ_i)]\\ \le& k_2(\lambda)C_H\binom{n}{v_H}\Bigg[\left((\frac{d}{n})^{e_H}(1-\frac{d}{n})^{\binom{v_H}{2}-e_H}\right)^2 + \binom{n}{v_H}\left((\frac{d}{n})^{e_H}(1-\frac{d}{n})^{\binom{v_H}{2}-e_H}\right)^2\Bigg]\\ &+ k_2(\lambda)C_H\binom{n}{v_H}\binom{n}{v_H}\Bigg[(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H}\\ &+(\frac{d}{n})^{2e_H-e_F}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H+e_F-\binom{v_F}{2}}\\ &+(\frac{d}{n})^{2e_H}(1-\frac{d}{n})^{2\binom{v_H}{2}-2e_H-\binom{t}{2}}\Bigg].\end{aligned}$$

By the fact that $e_F \le e_H$ and $(1-\frac{d}{n})^{\binom{v_H}{2}-e_H}$ converges to some positive constant when $d = n^\delta$ for $\delta < 1$ such that $(1-\delta)(2e_H - e_F) \geq 2v_H$. Hence:

$$\begin{aligned}
&\sup_{A\subset\mathbb{Z}_+} |P(W\in A) - Poi_\lambda(A)| \\
\leq& C_H(n^{v_H})\left[\frac{1}{n^{2(1-\delta)e_H}} + n^{v_H}\frac{1}{n^{2(1-\delta)e_H}} + n^{v_H}\frac{1}{n^{2(1-\delta)e_H}}\right. \\
&\left. + n^{v_H}\frac{1}{n^{(1-\delta)(2e_H-e_F)}} + n^{v_H}\frac{1}{n^{2(1-\delta)e_H}}\right] \\
=& C_H\left[(\frac{1}{n^{2(1-\delta)e_H - v_H}} + \frac{1}{n^{2(1-\delta)e_H-2v_H}} + \frac{1}{n^{2(1-\delta)e_H-2v_H}}\right. \\
&\left. + \frac{1}{n^{(1-\delta)(2e_H-e_F)-2v_H}} + \frac{1}{n^{2(1-\delta)e_H-2v_H}}\right] \\
\leq& \frac{C_H}{n^{(1-\delta)(2e_H-e_F)-2v_H}} \to 0,
\end{aligned}$$

as $n \to \infty$. The proof is complete.

Acknowledgments

We are thankful for financial support, School of Science, University of Phayao, Thailand.

References

[1] A.D. Barbour and Louis H.Y. Chen, An introduction to stein's method, 2005, Singapore university Press and world scientific Publishing Co. Pte. Ltd.

[2] J.H. Kim, B. Sodakov and V. Vu, Note Small subgraphs of random regular graphs. "Math" Debrecen, 6:290-297, 1959.

[3] N. Wormald, Models of random regular graphs, in: Surveys in Combinatorics, Proceeding Cambridge 1999, London Mathematical Society Lecture Note Ser., vol. 276, Cambridge University Press, Cambridge, 1999, pp. 239298.

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

[5] P. Erdős and A. Rényi. On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960) 1761.

[6] C.M. Stein (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.

[7] M. Donganont and A. Boonyued (2010), The bound on Poisson Approximation of the number of copies of a fixed graph in a random d-regular graph, Proc. Sixth Berkeley Sympos, Math.Statist.Probab. 3:583-602.

[8] Lan XIAO, Guiyin YAN, Yuwen WU and Wei REN (2008), Induced subgraph in random regular graph, Jrl Syst Sci and Complexity. 21:645-650.