Phase Transition on The Degree Sequence of a Random Graph Process with Vertex Copying and Deletion

Xianyuan WU

Capital Normal University

wuxy@mail.cnu.edu.cn

The 7th workshop on Markov Processes and Related Topics

Beijing Normal University, July 19-23, 2010

Outline

- 1, Scale-Free Real-World Networks

Outline

- 1, Scale-Free Real-World Networks
- 2, Other Real-World Networks

Outline

- 1, Scale-Free Real-World Networks
- 2, Other Real-World Networks
- 3, Models lead to Power Law Degree Sequence

Outline

- 1, Scale-Free Real-World Networks
- 2, Other Real-World Networks
- 3, Models lead to Power Law Degree Sequence
- 4, A Model Lead to Critical Phenomenon

Outline

- 1, Scale-Free Real-World Networks
- 2, Other Real-World Networks
- 3, Models lead to Power Law Degree Sequence
- 4, A Model Lead to Critical Phenomenon
- 5, Our Model and Main results

Outline

- 1, Scale-Free Real-World Networks
- 2, Other Real-World Networks
- 3, Models lead to Power Law Degree Sequence
- 4, A Model Lead to Critical Phenomenon
- 5, Our Model and Main results
- 6, On the proof

1, Scale-Free Real-World Networks

- For the real-world network of World Wide Web/Internet, experimental studies by Albert, Barabási & Jeong (1999) etc. demonstrated that the proportion of vertices of a given degree follows an approximate inverse power law, i.e.,

$$\frac{\text{the number of vertices of degree } k}{\text{the total unmber of vertices}} \approx Ck^{-\alpha}$$

for some constants C and α.

1, Scale-Free Real-World Networks

- For the real-world network of World Wide Web/Internet, experimental studies by Albert, Barabási & Jeong (1999) etc. demonstrated that the proportion of vertices of a given degree follows an approximate inverse power law, i.e.,

$$\frac{\text{the number of vertices of degree } k}{\text{the total unmber of vertices}} \approx Ck^{-\alpha}$$

for some constants C and α.

- The degree distribution of many real-world networks (including the Internet) are power law (It is also called scale-free).— heavy-tailed.

- For the classical random graph model $G_{n,p}$ introduced by Erdös & Rényi (1959), the proportion of vertices of a given degree follows an approximate Poisson law, i.e.,

$$\frac{\text{the number of vertices of degree } k}{\text{the total unmber of vertices}} \approx \frac{\lambda^k}{k!} e^{-\lambda},$$

where $\lambda = np$.

- For the classical random graph model $G_{n,p}$ introduced by Erdös & Rényi (1959), the proportion of vertices of a given degree follows an approximate Poisson law, i.e.,

$$\frac{\text{the number of vertices of degree } k}{\text{the total unmber of vertices}} \approx \frac{\lambda^k}{k!}e^{-\lambda},$$

where $\lambda = np$.

- The degree distribution of classical random graph model $G_{n,p}$ is light-tailed.

- For the classical random graph model $G_{n,p}$ introduced by Erdös & Rényi (1959), the proportion of vertices of a given degree follows an approximate Poisson law, i.e.,

$$\frac{\text{the number of vertices of degree } k}{\text{the total unmber of vertices}} \approx \frac{\lambda^k}{k!}e^{-\lambda},$$

  where $\lambda = np$.

- The degree distribution of classical random graph model $G_{n,p}$ is light-tailed.

- Scale-free degree distribution is one of the most important features of real-world networks.

2, Other real-world networks.

- Guassian distribution can be observed in the acquaintance network of Mormons Bernard *et al*. (1988).

2, Other real-world networks.

- Guassian distribution can be observed in the acquaintance network of Mormons Bernard *et al*. (1988).
- Exponential distribution can be observed in the powergrid of southern California Watts & Strogatz (1998).

2, Other real-world networks.

- Guassian distribution can be observed in the acquaintance network of Mormons Bernard *et al*. (1988).
- Exponential distribution can be observed in the powergrid of southern California Watts & Strogatz (1998).
- The degree distribution of the network of world airports Amaral *et al*. (2000) interpolates between Gaussian and exponential distributions.

2, Other real-world networks.

- Guassian distribution can be observed in the acquaintance network of Mormons Bernard *et al*. (1988).
- Exponential distribution can be observed in the powergrid of southern California Watts & Strogatz (1998).
- The degree distribution of the network of world airports Amaral *et al*. (2000) interpolates between Gaussian and exponential distributions.
- The degree distribution of the citation network in high energy physics Lehmann, Lautrup & Jackson (2003) interpolates between exponential and power law distributions.

An example: a model which exhibits more than one D.S.

- For a general model of collaboration networks in Zhou *et al.* (2005) indicate that:

  while a relevant parameter α increases from 0 to 1.5, four kinds of degree distributions appear as:

  1, exponential,

  2, arsy-varsy,

  3, semi-power law and

  4, power law

  in turn.

3, Models lead to Power Law D. S.

- Why Power Law?: What is the underlying causes for the emergence of power law degree distributions?

3, Models lead to Power Law D. S.

- Why Power Law?: What is the underlying causes for the emergence of power law degree distributions?

- Actually, people use the Random Graph processes $\{G_t = (V_t, E_t) : t \geq 1\}$ to model the evolving real-world network.

- For given $t \geq 1$, one may select a vertex from V_t uniformly at random, and let ξ_t be the degree of the selected vertex, then the distribution of ξ_t is called the degree distribution of G_t.

- For given $t \geq 1$, one may select a vertex from V_t uniformly at random, and let ξ_t be the degree of the selected vertex, then the distribution of ξ_t is called the degree distribution of G_t.
- Let $D_k(t)$ be the number of vertices with degree $k \geq 0$ in G_t and let $\overline{D}_k(t)$ be the expectation of $D_k(t)$. Suppose that $v_t = |V_t|$ grow linearly, then

$$\mathbb{P}(\xi_t = k) = \mathbb{E}(\frac{D_k(t)}{v_t}) \propto \frac{\overline{D}_k(t)}{t}.$$

- For given $t \geq 1$, one may select a vertex from V_t uniformly at random, and let ξ_t be the degree of the selected vertex, then the distribution of ξ_t is called the degree distribution of G_t.
- Let $D_k(t)$ be the number of vertices with degree $k \geq 0$ in G_t and let $\overline{D}_k(t)$ be the expectation of $D_k(t)$. Suppose that $v_t = |V_t|$ grow linearly, then

$$\mathbb{P}(\xi_t = k) = \mathbb{E}(\frac{D_k(t)}{v_t}) \propto \frac{\overline{D}_k(t)}{t}.$$

- Our goal is to study the limit $\lim_{t\to\infty} \dfrac{\overline{D}_k(t)}{t}$ under various setting of G_t.

Some random graph models were introduced to explain "Why Power law?":

- 'LCD model' of Bollobás & O. Riordan (2004);

Some random graph models were introduced to explain "Why Power law?":

- 'LCD model' of Bollobás & O. Riordan (2004);
- the generalization of 'LCD' model due to Buckley & Osthus (2004);

Some random graph models were introduced to explain "Why Power law?":

- 'LCD model' of Bollobás & O. Riordan (2004);
- the generalization of 'LCD' model due to Buckley & Osthus (2004);
- 'copying' models of Kumar *et al* (2000);

Some random graph models were introduced to explain "Why Power law?":

- 'LCD model' of Bollobás & O. Riordan (2004);
- the generalization of 'LCD' model due to Buckley & Osthus (2004);
- 'copying' models of Kumar *et al* (2000);
- the very general models defined by Copper & Frieze (2003);

Some random graph models were introduced to explain “Why Power law?":

- ‘LCD model’ of Bollobás & O. Riordan (2004);
- the generalization of ‘LCD’ model due to Buckley & Osthus (2004);
- ‘copying’ models of Kumar *et al* (2000);
- the very general models defined by Copper & Frieze (2003);
- the other model with random deletions defined by Copper, Frieze & Vera (2004).

Some random graph models were introduced to explain "Why Power law?":

- 'LCD model' of Bollobás & O. Riordan (2004);
- the generalization of 'LCD' model due to Buckley & Osthus (2004);
- 'copying' models of Kumar *et al* (2000);
- the very general models defined by Copper & Frieze (2003);
- the other model with random deletions defined by Copper, Frieze & Vera (2004).
- "hard copying" model of Ning, Wu & Cai (2008). etc.

4, A Model Leads to Critical Phenomenon

- Researches reveal that PREFERENTIAL ATTACHMENT (also called “BA mechanism") and COPYING are the most important mechanisms which lead to power law degree sequences

4, A Model Leads to Critical Phenomenon

- Researches reveal that PREFERENTIAL ATTACHMENT (also called “BA mechanism") and COPYING are the most important mechanisms which lead to power law degree sequences

- Our Problem:

  Does it exist some dynamically evolving random graph process which brings forth various degree distributions by continuous changing of its parameters only?

4, A Model Leads to Critical Phenomenon

- Researches reveal that PREFERENTIAL ATTACHMENT (also called “BA mechanism") and COPYING are the most important mechanisms which lead to power law degree sequences

- Our Problem:

  Does it exist some dynamically evolving random graph process which brings forth various degree distributions by continuous changing of its parameters only?

- Our goal:

  Answer the above problem in a mathematically rigorous manner.

The First Result (A simplified version!)

Model 1 [Wu, Dong, Liu and Cai (2009),JAP]:

$\{G_t = (V_t, E_t), t \geq 1\}$, Write $e_t = |E_t|$, $v_t = |V_t|$.

- Let $G_1 = \{x_1\}$

The First Result (A simplified version!)

Model 1 [Wu, Dong, Liu and Cai (2009),JAP]:

$\{G_t = (V_t, E_t), t \geq 1\}$, Write $e_t = |E_t|$, $v_t = |V_t|$.

- Let $G_1 = \{x_1\}$
- At Time-Step $t \geq 2$, to define G_t from G_{t-1}, one of the two following substeps is executed.

The First Result (A simplified version!)

Model 1 [Wu, Dong, Liu and Cai (2009),JAP]:

$\{G_t = (V_t, E_t), t \geq 1\}$, Write $e_t = |E_t|$, $v_t = |V_t|$.

- Let $G_1 = \{x_1\}$
- At Time-Step $t \geq 2$, to define G_t from G_{t-1}, one of the two following substeps is executed.
- With probability $\alpha > 0$ we add a vertex x_t to G_{t-1}. We then add m random edges incident with x_t. When a edge is added, the random neighbour w of x_t is chosen in the manner of preferential attachment, namely,

$$\mathbb{P}(w = v) = \frac{d_v(t-1)}{2e_{t-1}},$$

where $d_v(t-1)$ denotes the degree of vertex v in G_{t-1}.

- With probability $1-\alpha \geq 0$ we delete $\min\{m, e_{t-1}\}$ randomly chosen edges from E_{t-1}.

  Remark 1: In this setting, $\{e_t : t \geq 1\}$ is Markovian and

$$\mathbb{E}(e_t) \approx (2\alpha - 1)mt.$$

  - Note that, to the best of my knowledge, almost all studied models have linear edge and vertex growth. To some extent, this makes the argument standard (or, relatively easy!).

- With probability $1 - \alpha \geq 0$ we delete $\min\{m, e_{t-1}\}$ randomly chosen edges from E_{t-1}.

  Remark 1: In this setting, $\{e_t : t \geq 1\}$ is Markovian and

  $$\mathbb{E}(e_t) \approx (2\alpha - 1)mt.$$

  - Note that, to the best of my knowledge, almost all studied models have linear edge and vertex growth. To some extent, this makes the argument standard (or, relatively easy!).

- The main results for Model 1 follow as

- **Theorem**: Let $\alpha_c = \dfrac{2}{3}$, then it is a critical point for the degree sequence of the model satisfying:

- **Theorem**: Let $\alpha_c = \dfrac{2}{3}$, then it is a critical point for the degree sequence of the model satisfying:
  1. if $\alpha > \alpha_c$, then there exists constant $C_1 = C_1(m, \alpha)$ such that,

$$\lim_{t\to\infty} \frac{\overline{D}_k(t)}{t} = C_1 k^{-1-\beta} + O(k^{-2-\beta});$$

- **Theorem**: Let $\alpha_c = \dfrac{2}{3}$, then it is a critical point for the degree sequence of the model satisfying:
  1. if $\alpha > \alpha_c$, then there exists constant $C_1 = C_1(m, \alpha)$ such that,

  $$\lim_{t\to\infty} \frac{\overline{D}_k(t)}{t} = C_1 k^{-1-\beta} + O(k^{-2-\beta});$$

  2. if $\frac{4}{7} < \alpha < \alpha_c$, then there exists constant $C_2 = C_2(m, \alpha)$ such that

  $$\lim_{t\to\infty} \frac{\overline{D}_k(t)}{t} = C_2 \gamma^k k^{-1+\beta} + O(\gamma^k k^{-2+\beta});$$

3 if $\alpha = \alpha_c$, then there exists constant $C_c = C_c(m, \alpha)$ such that,

$$\lim_{t \to \infty} \frac{\overline{D}_k(t)}{t} = C_c u_c(k).$$

- Where

$$u_c(k) = \int_0^1 t^{k-1} e^{-\frac{1}{1-t}} dt$$

and

$$\beta = \frac{4\alpha - 2}{3\alpha - 2}, \quad \gamma = \frac{\alpha}{2(1-\alpha)}.$$

- Remark 2: Three years later, I get the following

$$u_c(k) = e^{-O(\sqrt{k})}$$

  - I will thank Prof. Yong-Hua Mao in this point, it's him who suggest me to consider the function $e^{-\sqrt{k}}$.

- Remark 2: Three years later, I get the following

$$u_c(k) = e^{-O(\sqrt{k})}$$

  - I will thank Prof. Yong-Hua Mao in this point, it's him who suggest me to consider the function $e^{-\sqrt{k}}$.

- So, Model 1 exhibits a perfect critical phenomenon on its degree distribution!

- Remark 2: Three years later, I get the following

$$u_c(k) = e^{-O(\sqrt{k})}$$

  - I will thank Prof. Yong-Hua Mao in this point, it's him who suggest me to consider the function $e^{-\sqrt{k}}$.
- So, Model 1 exhibits a perfect critical phenomenon on its degree distribution!
- Recall that Model 1 only consider edge deletion, we note that the argument fails when vertex deletion is considered additionally.

5. The Model (2) and Main results

Fix a constant $\mu > 0$. Consider the random graph process $\{G_t = (V_t, E_t) : t \geq 1\}$ defined as follows:

- *Time-Step 1.* Let G_1 consist of an isolated vertex x_1.

5. The Model (2) and Main results

Fix a constant $\mu > 0$. Consider the random graph process $\{G_t = (V_t, E_t) : t \geq 1\}$ defined as follows:

- *Time-Step 1.* Let G_1 consist of an isolated vertex x_1.
- *Time-Step $t \geq 2$.* In the case of $v_{t-1} = 0$, let G_t consist of an isolated vertex x_t. Otherwise, we do the following:

5. The Model (2) and Main results

Fix a constant $\mu > 0$. Consider the random graph process $\{G_t = (V_t, E_t) : t \geq 1\}$ defined as follows:

- *Time-Step 1.* Let G_1 consist of an isolated vertex x_1.
- *Time-Step $t \geq 2$.* In the case of $v_{t-1} = 0$, let G_t consist of an isolated vertex x_t. Otherwise, we do the following:
  1. With probability $\alpha > 0$, we add a vertex x_t to G_{t-1}. Then, for each vertex $x \in V_{t-1}$, independently, we add the edge between x and x_t with probability $\min\{\mu/v_{t-1}, 1\}$.

5. The Model (2) and Main results

Fix a constant $\mu > 0$. Consider the random graph process $\{G_t = (V_t, E_t) : t \geq 1\}$ defined as follows:

- *Time-Step 1.* Let G_1 consist of an isolated vertex x_1.
- *Time-Step $t \geq 2$.* In the case of $v_{t-1} = 0$, let G_t consist of an isolated vertex x_t. Otherwise, we do the following:
  1. With probability $\alpha > 0$, we add a vertex x_t to G_{t-1}. Then, for each vertex $x \in V_{t-1}$, independently, we add the edge between x and x_t with probability $\min\{\mu/v_{t-1}, 1\}$.
  2. With probability $0 \leq \beta \leq 1 - \alpha$, we generate vertex x_t by copying an existing vertex x from V_{t-1} uniformly at random.

5. The Model (2) and Main results

Fix a constant $\mu > 0$. Consider the random graph process $\{G_t = (V_t, E_t) : t \geq 1\}$ defined as follows:

- *Time-Step 1.* Let G_1 consist of an isolated vertex x_1.
- *Time-Step $t \geq 2$.* In the case of $v_{t-1} = 0$, let G_t consist of an isolated vertex x_t. Otherwise, we do the following:
  1. With probability $\alpha > 0$, we add a vertex x_t to G_{t-1}. Then, for each vertex $x \in V_{t-1}$, independently, we add the edge between x and x_t with probability $\min\{\mu/v_{t-1}, 1\}$.
  2. With probability $0 \leq \beta \leq 1 - \alpha$, we generate vertex x_t by copying an existing vertex x from V_{t-1} uniformly at random.
  3. With probability $\gamma := 1 - \alpha - \beta$, we delete a randomly chosen vertex x from V_{t-1}.

5. Why we study Model 2?

- To model the www network more properly: to this end, we consider vertex deletion!
  - In general, vertex deletion makes the problem more complicated. As discussed in Copper, Frieze & Vera (2004), vertex deletion will make edge estimation “surprisingly difficult": if a high degree vertex is deleted, the edge number may change by a large amount in one step, and the use of standard concentration inequalities is prohibited.

5. Why we study Model 2?

- To model the www network more properly: to this end, we consider vertex deletion!
  - In general, vertex deletion makes the problem more complicated. As discussed in Copper, Frieze & Vera (2004), vertex deletion will make edge estimation “surprisingly difficult": if a high degree vertex is deleted, the edge number may change by a large amount in one step, and the use of standard concentration inequalities is prohibited.
- To get a power law with exponent lies in (1,2]:
  - Almost all other studied models possess a power law with an exponent larger than 2.
  - In this point, I shall thank Prof. Fu-Zhou Gong, it’s him who draw my attention to this aspect.

- To this end, we consider vertex copying, which makes the edge number of G_t grow super-linearly. A pre-research in Ning, Wu & Cai (2008) suggests that vertex copying may lead to such a power law exponent.

- To this end, we consider vertex copying, which makes the edge number of G_t grow super-linearly. A pre-research in Ning, Wu & Cai (2008) suggests that vertex copying may lead to such a power law exponent.
- Copying comes from the basic idea that a new web page is often made by copying an old one and then changing some of the links. One will see latter that, in the copying step of our model, the degree increasing rate of a vertex is just proportional to the degree of the vertex, and this coincides with that of models with preferential attachment mechanism.

- To this end, we consider vertex copying, which makes the edge number of G_t grow super-linearly. A pre-research in Ning, Wu & Cai (2008) suggests that vertex copying may lead to such a power law exponent.
- Copying comes from the basic idea that a new web page is often made by copying an old one and then changing some of the links. One will see latter that, in the copying step of our model, the degree increasing rate of a vertex is just proportional to the degree of the vertex, and this coincides with that of models with preferential attachment mechanism.
- Other than preferential attachment mechanism, copying plays a key role in our argument. Actually, due to copying, any estimate for edge number is avoided, this ENABLE us to consider vertex deletion!

Main Results

- We assume

$$0 < \alpha < 1, \quad \alpha + \beta > 1/2$$

and

$$\mu \geq \frac{\min\{\beta, \gamma\}}{\alpha} - 1.$$

Main Results

- We assume

$$0 < \alpha < 1, \quad \alpha + \beta > 1/2$$

and

$$\mu \geq \frac{\min\{\beta, \gamma\}}{\alpha} - 1.$$

- Given α, β, γ, let

$$\beta_c = \frac{1-\alpha}{2}, \quad \xi = \frac{\alpha}{\beta - \gamma} \quad \text{and} \quad \zeta = \frac{\beta}{\gamma}.$$

Main Results

- We assume
$$0 < \alpha < 1, \quad \alpha + \beta > 1/2$$
and
$$\mu \geq \frac{\min\{\beta, \gamma\}}{\alpha} - 1.$$
- Given α, β, γ, let
$$\beta_c = \frac{1-\alpha}{2}, \quad \xi = \frac{\alpha}{\beta - \gamma} \quad \text{and} \quad \zeta = \frac{\beta}{\gamma}.$$
- The main results of Model 2 follow as:

Main Theorem:

- if $\beta > \beta_c$, then there exists a constant $C_1 = C_1(\alpha, \beta, \mu)$ such that, for any $\epsilon \in (0, \frac{1}{2})$,

$$\left| \frac{\overline{D}_k(t)}{t} - C_1 k^{-1-\xi} \right| = O(t^{-\epsilon}) + O(k^{-2-\xi});$$

Main Theorem:

- if $\beta > \beta_c$, then there exists a constant $C_1 = C_1(\alpha, \beta, \mu)$ such that, for any $\epsilon \in (0, \frac{1}{2})$,

$$\left| \frac{\overline{D}_k(t)}{t} - C_1 k^{-1-\xi} \right| = O(t^{-\epsilon}) + O(k^{-2-\xi});$$

- if $\beta < \beta_c$ and $\mu < \frac{\beta}{\alpha} + \frac{\beta}{\gamma - \beta}$, then there exists a constant $C_2 = C_2(\alpha, \beta, \mu)$ such that, for any $\epsilon \in (0, \frac{1}{2})$,

$$\left| \frac{\overline{D}_k(t)}{t} - C_2 \zeta^k k^{-1+\xi} \right| = O(t^{-\epsilon}) + O\left(\zeta^k k^{-2+\xi}\right);$$

- if $\beta = \beta_c$, then there exists a constant $C_c = C_c(\alpha, \mu)$ such that, for any $\epsilon \in (0, \frac{1}{2})$,

$$\left| \frac{\overline{D}_k(t)}{t} - C_c u_c(k) \right| = O(t^{-\epsilon}) + O\left(\frac{\mu^k}{k!}\right),$$

where

$$u_c(k) = \int_0^1 t^{k+\frac{\alpha\mu}{\beta}-1}(1-t)^{\frac{\alpha\mu}{\beta}} \exp\left\{\frac{-\alpha}{\beta(1-t)}\right\} dt = e^{-O(\sqrt{k})}$$

and

$$\sqrt{\alpha/\beta}\sqrt{k} \leq O(\sqrt{k}) \leq 2\sqrt{\alpha/\beta}\sqrt{k}.$$

- if $\beta = \beta_c$, then there exists a constant $C_c = C_c(\alpha, \mu)$ such that, for any $\epsilon \in (0, \frac{1}{2})$,

$$\left| \frac{\overline{D}_k(t)}{t} - C_c u_c(k) \right| = O(t^{-\epsilon}) + O\left(\frac{\mu^k}{k!}\right),$$

  where

$$u_c(k) = \int_0^1 t^{k+\frac{\alpha\mu}{\beta}-1}(1-t)^{\frac{\alpha\mu}{\beta}} \exp\left\{\frac{-\alpha}{\beta(1-t)}\right\} dt = e^{-O(\sqrt{k})}$$

  and

$$\sqrt{\alpha/\beta}\sqrt{k} \le O(\sqrt{k}) \le 2\sqrt{\alpha/\beta}\sqrt{k}.$$

- Remark 3: In this paper, the conditions on μ are all technical, and it is conjectured that our results hold for any $\mu > 0$.

6, On the proof

The proof of the main Theorem is divided into two parts:

- Part 1: To prove that the limit $\lim_{t\to\infty}\frac{\overline{D}_k(t)}{t}$ exist and be the solution of the following recurrence in k: $d_{-1}=0$ and for $k\geq -1$,

$$(A_2(k+2)+B_2)d_{k+2}+(A_1(k+1)+B_1)d_{k+1}+(A_0k+B_0)d_k$$

$$=-\frac{\alpha\mu^{k+1}}{(k+1)!}e^{-\mu}=:-b_{k+1}. \qquad (*)$$

where $A_2=\frac{\gamma}{\eta}, A_1=-\frac{\beta+\gamma}{\eta}, A_0=\frac{\beta}{\eta}, B_2=0, B_1=-1-\frac{\alpha\mu-\beta+\gamma}{\eta},\ B_0=\frac{\alpha\mu}{\eta}$ and $\eta=2(\alpha+\beta)-1$.

To finish part 1, we need the following two useful estimates:

- 1, an upper estimate for Δ_t, the maximum degree of G_t:

$$\Delta_t \leq t^{\rho}, \quad q.s.$$

for some $0 < \rho < 1$, and

To finish part 1, we need the following two useful estimates:

- 1, an upper estimate for Δ_t, the maximum degree of G_t:

$$\Delta_t \leq t^{\rho}, \quad q.s.$$

  for some $0 < \rho < 1$, and

- 2, an uniform bound for the convergence rate of the so-called *law of small numbers*: For any $\epsilon > 0$,

$$\sup_{k \geq 1} \left| \mathbb{P}(Y_t = k) - \frac{\nu^k}{k!} e^{-\nu} \right| = O(t^{-1+\epsilon})$$

  where $Y_t = b(t, \nu/t)$ be the general binomial random variable. (*Thank Prof. Fu-Xi Zhang, she gives me a hint!*)

- Part 2: Solve the above difference equation (*) in three cases respectively. To finish part 2, we need the following three substeps:

- Part 2: Solve the above difference equation (*) in three cases respectively. To finish part 2, we need the following three substeps:
  1. Use the Laplace's method to solve the following homogenous equation

$$(A_2(k+2)+B_2)f_{k+2}+(A_1(k+1)+B_1)f_{k+1}+(A_0k+B_0)f_k$$

$$= 0, \quad k \geq 1.$$

- Part 2: Solve the above difference equation (*) in three cases respectively. To finish part 2, we need the following three substeps:
    1. Use the Laplace's method to solve the following homogenous equation

$$(A_2(k+2)+B_2)f_{k+2}+(A_1(k+1)+B_1)f_{k+1}+(A_0k+B_0)f_k$$

$$=0, \quad k \geq 1.$$

    2. For any given integer $m \geq 0$, we solve the following non-homogenous equation: $f_{-1}=0$ and for $k \geq -1$

$$(A_2(k+2)+B_2)f_{k+2}+(A_1(k+1)+B_1)f_{k+1}+(A_0k+B_0)f_k$$

$$=-I_{\{k=m-1\}}.$$

- 3. Finally, we construct the solution as follows to solve the general non-homogenous equation (*)

$$d_k := \sum_{m=0}^{\infty} b_m f_k^m, \quad k \geq -1$$

where $\{f_k^m : k \geq -1\}$ be the solution of the equation considered in substep 2.

- 3. Finally, we construct the solution as follows to solve the general non-homogenous equation (*)

$$d_k := \sum_{m=0}^{\infty} b_m f_k^m, \quad k \geq -1$$

where $\{f_k^m : k \geq -1\}$ be the solution of the equation considered in substep 2.

- To make the general non-homogenous equation (*) solvable and the problem meaningful, b_k must decay at least as $o(e^{-k})$. In this case we have $b_k = O(\frac{\mu^k}{k!})$.

- 3. Finally, we construct the solution as follows to solve the general non-homogenous equation (*)

$$d_k := \sum_{m=0}^{\infty} b_m f_k^m, \quad k \geq -1$$

where $\{f_k^m : k \geq -1\}$ be the solution of the equation considered in substep 2.

- To make the general non-homogenous equation (*) solvable and the problem meaningful, b_k must decay at least as $o(e^{-k})$. In this case we have $b_k = O(\frac{\mu^k}{k!})$.
- To solve such an equation successfully is a by-product of the present paper.

- Thank You Very Much!