中華民國第57屆中小學科學展覽會

作品說明書

高級中等學校組　數學科

050413

接二「連」三

學校名稱：國立桃園高級中學

作者：	指導老師：
高二 林卿恩 高二 熊品羚	蔣元召

關鍵詞：連通集、圖論演算法、機率論

摘要

在我們的研究中，先探討給定一個無向圖(undirected graph)並能使圖形必定 connected 的線段數E的條件，再以此得到圖形 connected 的機率範圍。我們接著引用已知整數數列 A001187(參見參考資料[1])類似的方法來解決隨機圖(Random Graph)的根本問題：給定一個圖$G(n,E)$，假定任兩個點連接的機率是常數p，求出該圖能形成 connected 的機率。接著我們將前述結論做推廣，假設該圖有兩子圖(subgraph)，其中各子圖中的邊，相連的機率分別為p及p_1，求出該圖能形成 connected 的機率。在這個過程中，在限制$G(n,E)$中最大連通子集的頂點數目下，我們也對於圖形不連通的機率做了許多研究。

壹、研究動機

面對浩瀚的生物疾病之謎，我們如何從中歸納哪些是關鍵的效標？在一群毫無規律的群體中，我們如何找出各種特徵和病情的關連性，又抑或是同種疾病的共同症狀？假設有 1000 份的肝病病例，我們可以從中了解到性別、居住地點、不良嗜好、身高、體重和各種其他相關資訊，並且可以整理成一些圖或表格。在這之中假如兩種資訊有共同特徵，例如居住在A處的皆為女性，這種情形我們看成是數學概念中的 Connected，而強烈與否則視為 Connected 的機率；反之，若兩種類別並沒有重要的關連性，我們視為 Disconnected。利用課本所學的排列組合這個觀點去分析整筆資料，若算出 Disconnected 的機率很高，此時代表有某些資訊發生了「孤立」的情形，使其無法與其他資訊達成關聯。以上方法我們相信能有效找出疾病和特徵的關連性，於是開啟了這次的研究，期待找出對未來醫學有所幫助的結果。

貳、研究目的

給定n個有編號的點，假設任兩點之間只有 1 條無向連線段或沒有連線。

一、為了方便計算機率，先討論下列問題：

(一) 所有連線情形的方法數；

(二) 當連線段的個數E小於何數時，圖形必為 disconnected；

(三) 當連線段的個數E大於何數時，圖形必為 connected。

二、假設任兩點連線的機率為p，求得下列通式：

(一) connected 圖形的個數A_n與圖形 connected 的機率$P(n,p)$；

(二) 求出連線段個數為E的所有 connected 圖形個數$K(n,E)$；

(三) 將原假設稍作更改：有兩點的連線機率為 p_1，其餘連線機率皆為 p。求圖形 connected 的機率 $F(n, p, p_1)$。

三、為瞭解不連通圖形的機率，我們限制最大連通子圖中的頂點數，求出了下列通式：

(一) 假設任兩點連線的機率為 p，求出圖形滿足「最大連通子圖中的頂點數小於或等於 Q」的機率 $D(n, p, Q)$；

(二) 假設有兩點連線的機率為 p_1，其餘任兩點連線的機率為 p，求出求圖形滿足「最大連通子圖中的頂點數小於或等於 Q」的機率 $B(n, p, p_1, Q)$；

四、假設 n 個點中有 m 個頂點($0 \le m \le n$)，其任兩點連線的機率為 p_1；剩餘的 $(n-m)$ 點中，任兩點的連線機率為 p，求出圖形 connected 的機率 $H(n, m, p, p_1)$。

五、(一) 假設 n 個點中有 m 個頂點($0 \le m \le n$)，其任兩點連線的機率為 p_1；剩餘的 $(n-m)$ 點中，任兩點的連線機率為 p，求出求出圖形滿足「最大連通子圖中的頂點數小於或等於 Q」的機率 $C(n, m, p, p_1, Q)$。

(二) 假設已知每條邊不同的連線機率，求 connected 的機率 $J(V_n, P_{V_n})$

參、研究設備及器材

紙、筆、電腦。

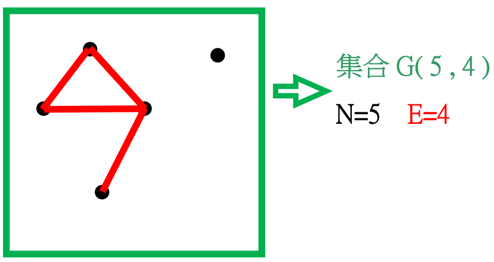
肆、研究過程或方法

一、文獻探討與符號定義

本研究範圍涉及在圖論(Graph Theory)中對圖形的基本定義以及一個與整數分割有關係的數列，如下：

(一) 圖形的基本定義(參見參考資料[2])

1. 給定一集合 $G(n, E)$ 由下列元素組成：

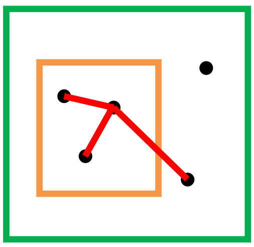


▲(圖一)有一個點沒有任何的邊與其相連

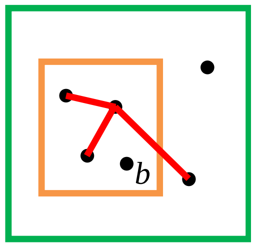
n 個點(node)，E 個邊(edge)連接任意兩個點。邊(edge)指的是沒有方向性的線段，並且限制 $0 \le E \le \dfrac{n(n-1)}{2}$，亦即任兩點若相連，則被唯一的邊所聯繫。我們也稱集合 $G(n, E)$ 為一個圖(graphy)。我們的研究中可以允許某些點沒有任何一條邊與其

相連，如(圖一)，有一個點沒有任何的邊與其相連。

2. 設 $G'(n', E')$ 為集合 $G(n, E)$ 中一個含有 n' 個點和 E' 條邊的子集合。如果在 $G'(n', E')$ 中，任意兩點 a, b 均可藉由若干個邊，從 a 點沿著邊移動到 b 點，則稱該子集 $G'(n', E')$ 為一個連通子集 (connected subset)，或簡稱連通(connected)。否則，則稱該子集 $G'(n', E')$ 為非連通 (disconnected)，如(圖二)。



原集合　　G(5 , 3)
子集合　　G’ (3 , 2)
子集合 G’ (3 , 2)是 connected，因為任兩點都可以藉由紅色的邊，由一點移動到另一點。



原集合　　G(6 , 3)
子集合　　G’ (4 , 2)
子集合 G’ (4 , 2)是 disconnected，因為點 b 無法移動到任何一個子集 G’ 的點。

▲(圖二) Connected ＆ Disconnected 說明

3. n 個點各有其編號(Labeled Graph，意即所有點視為不同)且點和點之間的連線沒有方向性(non-oriented line)，意即任兩點只有 connected 和 disconnected 兩種情況。

(二) 數列 A001187—Number of connected labeled graphs with n nodes (參見參考資料[1])

編號為 A001187 的整數數列其遞迴關係式為 $\begin{cases} A_0 = A_1 = 1 \\ A_n = 2^{C_2^n} - \sum_{k=1}^{n-1} C_{k-1}^{n-1} \cdot 2^{C_2^{n-k}} \cdot A_k \end{cases}$，我們在研究中得到與這個數列一模一樣的結果，並且以此為基礎，解決後續的問題。

(三) 數學式說明方式

我們在研究中大量使用遞迴關係式來得到所有結果，在說明式子計算想法時，會在式子中畫底線並給予註記或編號，以方便說明。

二、前置研究

(一)給定 n 個點，所有連線情形的方法數為 $2^{C_2^n}$ 種。

(二)給定 n 個點，邊的個數 E 小於 $(n-1)$ 時，該圖必為 disconnected。

【證明】

1. 不失一般性，我們可以假設 $P_1, \cdots, P_k$ 為 k 個單點(沒有任何邊與其相連)，及 $(n-k)$ 個有邊點(至少一條邊與該點相連)。

2. 定義 $G(n, E)$ 的子集合 $G'(P_i, S_i)$ $=\{P_1, \cdots, P_i\}$，$S_i =$ 子集合 $G'(P_i, S_i)$ 裡面邊的個數，其中 $1 \le i \le n$。顯然，$P_n = G(n, E)$，$S_n = E$，並且依據假設，有

$$\begin{cases} S_i = 0 \text{ ，對於 } 0 \le i \le k \\ S_i \ge 1 \text{ ，對於 } k+1 \le i \le n \end{cases}$$ 。

3. 對於 $k+1 \le i \le n$ ，我們以 a_i 表示點 P_i 與子集 $G'(P_{i-1}, S_{i-1})$ 中，任取若干點連線，所得的邊數。

4. 因為共有 k 個<u>單點</u>，導致於所有<u>有邊點</u> P_i ($k+1 \le i \le n$)與單點的連線數，至少需要 k 條才能使 $G(n, E)$ 構成 connected set，亦即當 $i = k+1$ 時， $a_i \ge k$ 。我們期望利用最少邊數，使得 $G(n, E)$ 能達成 connected set， $a_{i+1} - a_i = 1$, $k+1 \le i \le n$ 。於是，

$$\begin{cases} S_{k+1} = S_k + a_{k+1} \\ S_{k+2} \ge S_{k+1} + a_{k+2} + 1 \\ \quad\vdots \\ S_n \ge S_{n-1} + a_n + 1 \end{cases}$$

（註：點 P_{k+2} 和 $G'(P_{k+1}, S_{k+1})$ 至少需要連1條邊才構成connected ，依此類推。）

$$\Rightarrow S_n \ge S_k + (a_{k+1} + \cdots + a_n) + (n-k-1) \ge 0 + k + (n-k-1) = n-1$$

$\therefore E \ge n-1$ 時才可能connected ，亦即 $E < n-1$ 時必為disconnected 。

(三) 給定 n 個點，當邊的個數 $E > C_2^{n-1}$ 時，圖形必為 connected。即所有 $G(n, E)$ 的子集合 G_i 皆為連通，其中 $0 \le i \le n$ 且 $\mathrm{G}_i = \mathrm{G}_i(n_i, E_i)$ 為一個有 n_i 個點、 E_i 個邊的子集合。

【證明】

要證明上述命題，證明過程中須搭配證明【引理 1】～【引理 3】，並將命題主要證明分成兩部份分別放在步驟 2 與步驟 5。

1. 【引理 1】 $\boxed{\text{若固定 } n \text{ 值，} n_1 + n_2 = n,\ n_1 \in \mathbb{N}, n_2 \in \mathbb{N} \text{ ，則 } \frac{1}{4}n^2 \ge n_1 n_2 \ge n-1 \text{ 。}}$

 (1)固定 n 值， $n_1 + n_2 = n$, $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}$ ，

$$n_1 n_2 = n_1(n - n_1) = -n_1^2 + n(n_1) = -(n_1 - \frac{1}{2}n)^2 + \frac{1}{4}n^2 \text{ 。}$$

 (2)當 $n_1 = \frac{1}{2}n$ 時， $n_1 n_2$ 為 $\max : \frac{1}{4}n^2$ ；當 $n_1 = 1$, 或 $n_1 = n\text{-}1, n_2 = 1$ 時， $n_1 n_2$ 為 $\min : n-1$

$$\therefore \frac{1}{4}n^2 \ge n_1 n_2 \ge n-1$$ ，【引理 1】證畢。

2. 先考慮將 $G(n, E)$ 分成 2 群且 $E > C_2^{n-1}$ 時， $G(n, E)$ 必為 connected。

 (1)設 $G_1(n_1, E_1)$ 與 $G_2(n_2, E_2)$ 為 $G(n, E)$ 中的兩個子集合滿足

$$G = G_1 \cup G_2, G_1 \cap G_2 = \varnothing$$ ，

則有 $E_1 \le C_2^{n_1}$, $E_2 \le C_2^{n_2}$ 。

$$(2)\ E = E_1 + E_2 \le C_2^{n_1} + C_2^{n_2} = \frac{n_1\left(n_1-1\right)}{2} + \frac{n_2\left(n_2-1\right)}{2} = \frac{\left(n_1+n_2\right)^2 - n - 2n_1n_2}{2} = C_2^n - n_1n_2$$

$$\Rightarrow E = E_1 + E_2 \le C_2^n - n_1n_2$$

3. 【引理 2】 設 $\sum_{i=1}^{k} n_i = n$ ，$n_i \in \mathbb{N}$ 。當 $\begin{cases} n_k = n-k+1 \\ n_i = 1 \\ 1 \le i \le k-1 \end{cases}$ 時，$\sum_{i=1}^{k} n_i^{\,2}$ 有最大值。

現改變 n_k 和 n_i 的值(其中 $1 \le i \le k-1$)，得一新值 n_i' ，考慮

$$\begin{cases} n_k' = n-k+1-\sum_{i=1}^{k-1} m_i \\ n_i' = 1+m_i (1 \le i \le k-1) \end{cases} ,\quad \begin{cases} m_i \in \mathbb{N} \cup \{0\} \\ \sum_{i=1}^{k-1} m_i \le n-k \end{cases} \quad (\because n_i \in \mathbb{N})$$

$$\Rightarrow \sum_{i=1}^{k} (n_i')^2 = \sum_{i=1}^{k-1} (1+m_i)^2 + (n-k+1-\sum_{i=1}^{k-1} m_i)^2$$

$$= (k-1) + 2\sum_{i=1}^{k-1} m_i + \sum_{i=1}^{k-1} \left(m_i\right)^2 + (n-k+1)^2 - 2(n-k+1)\sum_{i=1}^{k-1} m_i + \left(\sum_{i=1}^{k-1} m_i\right)^2$$

$$= (k-1) + (n-k+1)^2 - 2(n-k)\sum_{i=1}^{k-1} m_i + \left(\sum_{i=1}^{k-1} m_i\right)^2 + \sum_{i=1}^{k-1} m_i^2$$

$$= \underline{\left[(k-1)+(n-k+1)^2\right]} + \left(\sum_{i=1}^{k-1} m_i\right)\left[\sum_{i=1}^{k-1} m_i - 2(n-k)\right] + \sum_{i=1}^{k-1} m_i^2$$

$$= \underline{\left[\sum_{i=1}^{k} n_i^{\,2}\right]} + \left(\sum_{i=1}^{k-1} m_i\right)\left[\sum_{i=1}^{k-1} m_i - 2(n-k)\right] + \sum_{i=1}^{k-1} m_i^{\,2}$$

(原預設之最大值)

$$\because m_i \ge 0 \ ,\ \therefore (\sum_{i=1}^{k-1} m_i)^2 \ge \sum_{i=1}^{k-1} m_i^{\,2} \quad \Rightarrow (n-2)^2 \ge (n-k)^2 \ge (\sum_{i=1}^{k-1} m_i)^2 \ge \sum_{i=1}^{k-1} m_i^{\,2}$$

$(\forall k \in \mathbb{N}, \forall k \ge 2)$

$$\therefore \sum_{i=1}^{k} n_i'^2 - \sum_{i=1}^{k} n_i^{\,2} = \underline{(\sum_{i=1}^{k-1} m_i)^2} + \sum_{i=1}^{k-1} m_i^{\,2} - 2\underline{\underline{(\sum_{i=1}^{k-1} m_i)}}(n-k)$$ 將上式代入

$$\le \underline{(n-k)^2} + \sum_{i=1}^{k-1} m_i^{\,2} - 2\underline{\underline{(n-k)}}(n-k) \quad \le (\sum_{i=1}^{k-1} m_i)^2 - (n-k)^2 \quad \le (n-k)^2 - (n-k)^2 = 0$$

$$\therefore \sum_{i=1}^{k} n_i'^2 - \sum_{i=1}^{k} n_i^{\,2} \le 0 \quad \therefore \sum_{i=1}^{k} n_i^{\,2} \ge \sum_{i=1}^{k} n_i'^2$$ 。

得知當k群中的其中一群有$(n-k+1)$個點，其餘$k-1$群皆只有 1 個點時，

$\sum_{i=1}^{k} n_i^{\,2}$ 為最大值，【引理 2】證畢。

4.【引理 3】n個點分割的群數k越大，則形成的 disconnected sets 中的邊數E最大值越小。

設n個點被切割成k個 connected sets $G_1, G_2, \cdots, G_k$，其中

$G_i \cap G_j = \varnothing, 1 \le i < j \le k$ 且兩兩 disconnected 的情況下，定義

$\{e_i \mid$第 i 群 *connected set* G_i 中的邊數$, 1 \le i \le k\}$為出現的邊數所成的集合，

$$\text{設 } e(n,k) = \sum_{i=1}^{k} e_i \le \sum_{i=1}^{k} C_2^{n_i} = \frac{\left[\sum_{i=1}^{k} n_i^2\right]-n}{2} \le \frac{\left[(n-k+1)^2+(k-1)\right]-n}{2}$$

--<由【引理 2】>

$$=\frac{n^2-n}{2}-n(k-1)+\frac{(k-1)^2+(k-1)}{2}=C_2^n+C_2^k-n(k-1)\text{ 。}$$

再定義$E(n,\mathrm{k})=\max\left[e(n,\mathrm{k})\right]$，則$E(n,k)=C_2^n+C_2^k-n(k-1)$ 。

設 $n \ge k_i > k_j > \cdots \ge 2$ ，則

$$E(n,k_i)-E(n,k_j)=C_2^{k_i}-C_2^{k_j}-n(k_i-k_j)=\frac{k_i^{\,2}-k_j^{\,2}-(k_i-k_j)}{2}-n(k_i-k_j)$$

$$=\frac{(k_i-k_j)\left[(k_i-n)+(k_j-n)-1\right]}{2}<0$$

$\therefore E(n,k_j) > E(n,k_i)$ ， $1 < k_i, k_j \le n$ 。

得知群數k越大，disconnected的邊數最大值越小，【引理 3】證畢。

5. 將$G(n,E)$分成k群($k \ge 2$)且$E > C_2^{n-1}$ 時，$G(n,E)$必 connected

將n個點分成k群，即$G(n,E)$分成$G_1, \cdots, G_k$，其中任兩群不相連且滿足

$$G=\bigcup_{i=1}^{k} G_i \quad ,1 \le k \le n$$

$$G_i(n_i, E_i) \quad ,n=\sum_{i=1}^{k} n_i \quad ,E=\sum_{i=1}^{k} n_i$$

已知$E(n,2) \ge E(n,k)$--<由【引理 3】>

$$\therefore C_2^n+C_2^k-n(k-1) \le C_2^n+C_2^2-n(2-1)=C_2^n+1-n$$

$$=\frac{n(n-1)}{2}-n+1=\frac{n^2-3n+2}{2}=C_2^{n-1}$$

設 $G_1,\cdots,G_k$ 互為 disconnected 時，由上述結果可知其最大邊數為 C_2^{n-1} ，也就是說，當 $E>C_2^{n-1}$(已知E,n)時，整張圖形必為 connected。本命題證畢。

伍、研究結果

一、給 n 個點，求出 connected 圖形的個數 A_n

我們利用下列的遞迴關係式來求得 A_n ，結果與已知的數列 A001187 一致，說明如下。

$$\begin{cases} A_0=A_1=1 \\ A_n=\underbrace{2^{C_2^n}}_{①}-\underbrace{\sum_{k=2}^{n}}_{②}\underbrace{\sum_{\substack{\sum_{i=1}^{k}n_i=n\\(n_i\in\mathbb{N})}}}_{③}\underbrace{\left[\prod_{j=1}^{n}\frac{j}{\left(\sum_{i=1}^{k}\delta_{n_{i,j}}\right)!}\prod_{i=1}^{k}\frac{A_{n_i}}{(n_i)!}\right]}_{④黑色部份⑤藍色部份} \end{cases}$$

【說明】我們將上式分成五個圖示部分(畫底線部分)來說明。

①n 個頂點時，全部圖的數目；

②頂點數為 n 之不連通子圖數目介於 $2\sim n$ 個之間；③總頂點數為 n ；

④將 n 個具有編號的頂點分為 $n_1,\cdots,n_k$ ，共 k 堆的分堆方法數；

⑤每個分堆方法，要再乘上各自的連通方法數，因為每個連通子圖的連通情況皆互相獨立，所以用乘的。

$$\Rightarrow\sum_{k=1}^{n}\sum_{\sum_{i=1}^{k}n_i=n}\left[\prod_{j=1}^{n}\frac{j}{\left(\sum_{i=1}^{k}\delta_{n_{i,j}}\right)!}\prod_{i=1}^{k}\frac{A_{n_i}}{(n_i)!}\right]=2^{C_2^n}\text{ 且 }\delta_{n_{i,j}}=\begin{cases}0,\text{when } i\neq j\\1,\text{when } i=j\end{cases}。$$

若令 m 為一個包含一個固定頂點 v_i 的頂點子圖之中的頂點個數，且限制這個 m 點的子圖是連通的，再使用上面的結果，可得到以下推導。

$$\begin{cases} A_0=A_1=1 \\ A_n=2^{C_2^n}-\sum_{m=1}^{n-1}C_{m-1}^{n-1}\left(\sum_{k=1}^{n-m}\sum_{\sum_{i=1}^{k}n_i=n-m}\left[\prod_{j=1}^{n-m}\frac{j}{\left(\sum_{i=1}^{k}\delta_{n_{i,j}}\right)!}\prod_{i=1}^{k}\frac{A_{n_i}}{(n_i)!}\right]\right) \end{cases}\Rightarrow\begin{cases} A_0=A_1=1 \\ A_n=2^{C_2^n}-\sum_{m=1}^{n-1}C_{m-1}^{n-1}\cdot\underbrace{2^{C_2^{n-m}}}_{⑥}\cdot\underbrace{A_m}_{⑦} \end{cases}$$

上式最後兩個部分說明如下：

⑥剩下來的$(n-m)$個頂點的所有圖的總數；

⑦包含v_i的具有m個頂點的連通子集的連通方法數。

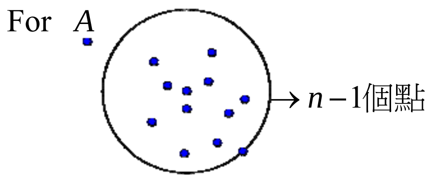
最後將m換成k，我們可以得到：$\begin{cases} A_0 = A_1 = 1 \\ A_n = 2^{C_2^n} - \sum_{k=1}^{n-1} C_{k-1}^{n-1} \cdot 2^{C_2^{n-k}} \cdot A_k \end{cases}$，此數列即為已知編號為 A001187 的整數數列。

二、給n個點，每條E連線機率為p，求出圖形 connected 的機率$P(n, p)$

為簡化符號，在本節中我們設$P(k) = P(k, p)$，其中$0 \le k \le n$。已知任兩點間連通機率p，設不通為$q = 1 - p$。所以$\begin{cases} P(1) = 1 \\ P(2) = 1 - q \end{cases}$，當$n \ge 2$時，$P(n) = 1 - \sum_{k=1}^{n-1} C_{k-1}^{n-1} P(k) q^{k(n-k)}$

$= 1 - (C_0^{n-1} P(1) q^{1(n-1)} + C_1^{n-1} P(2) q^{2(n-2)} + C_2^{n-1} P(3) q^{3(n-3)} + ... + C_{n-2}^{n-1} P(n-1) q^{n-1(1)})$

將某A點孤立，使剩餘$n-1$個點自成一群，如(圖四)，$C_0^{n-1} P(1) q^{1(n-1)}$指從$n-1$個點中選 0 個點，並使$A$點 connected(即$P(1)$)，且與剩餘$n-1$個 disconnected(即$q^{1(n-1)}$)。



▲(圖四)孤立某點A

$C_1^{n-1} P(2) q^{2(n-2)}$指從$n-1$個點中選1點，並使$A$點和選中的 1 點 connected (即$P(2)$)，且與剩餘$n-2$個點 disconnected(即$q^{2(n-2)}$)，此後以此類推。最後 connected 機率即為「1 減掉所有 disconnected 情形機率的和」。

舉例說明：假設有 4 個點，且兩點連通機率為 0.2，藉由公式可得 connected 機率為 0.082496。

三、連線段個數為E的所有 connected 圖形個數$K(n, E)$

設函數$K(n,E)$為「一個具有n個有編號之頂點的無向圖，且限定邊數為E，且連通」，滿足上述條件之圖的個數我們得到以下遞迴：

$$\begin{cases} K(n, n-1) = n^{n-2} \quad \cdots ① \\ K(n,E) = \underbrace{C_E^{C_2^n}}_{②} - \underbrace{\sum_{k=1}^{n-1}}_{③} \underbrace{\sum_{e=0}^{E}}_{④} \underbrace{C_{k-1}^{n-1}}_{⑤} \underbrace{C_{E-e}^{C_2^{n-k}}}_{⑥} \underbrace{K(k,e)}_{⑦} \end{cases} \text{且} E \in \left[n-1, C_2^n\right]$$

【說明】①由於在此條件下的連通圖必為一個具有n個頂點的樹，所以就直接使用 Caley 公式來計算這個條件下的圖之數目；

②所有的情形總數；

③固定其中一頂點v_i，讓它和其他$k-1$個頂點相連通，使這k個頂點形成一個連通子集，其中在$k \le n-1$條件下整張圖不連通；

④將其中的e條邊分給k個頂點；

⑤從v_i以外的$n-1$個頂點之中選出$k-1$個頂點的方法數；

⑥剩下來的$n-k$個頂點，且限制$E-e$條邊的所有圖的總數；

⑦限制k個頂點，e條邊的連通方法數

也可以改寫成遞迴式：
$$\begin{cases} K(n,n-1)=n^{n-2} \\ K(n,E)=\underbrace{C_E^{C_2^n}}_{①*}-\underbrace{\sum_{k=2}^{n}}_{②*}\underbrace{\sum_{\substack{\sum_{i=1}^{k} n_i=n \\ \sum_{i=1}^{k} E_i=E \\ E_i\in\left[n_i-1,C_2^{n_i}\right]}}}_{③*}\underbrace{\prod_{j=1}^{n}\frac{j}{\left(\sum_{i=1}^{k}\delta n_{i,j}\right)!}\prod_{i=1}^{k}\frac{K(n_i,E_i)}{(n_i)!}}_{\text{④*黑色部分 ⑤*藍色部分}} \end{cases}$$

【說明】①* n個頂點時，限制E條邊的圖的數目；

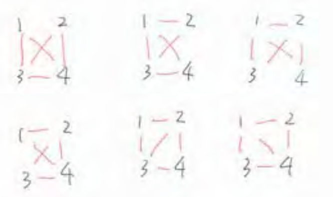
②* 頂點數為n之不連通子圖數目介於$2\sim n$個之間；

③* 總頂點數必為n，且各圖之邊數必須合理；

④* 將n個具有編號的頂點分為$n_1,\cdots,n_k$，k堆的分堆方法數；

⑤* 每個分堆方法，要再乘上各自的連通方法數，因為每個連通子圖的連通情況皆互相獨立，所以用乘的。

舉例說明：假設有 4 個點，5 條邊，藉由上述公式可以得到 connected 的圖形數為 6，如(圖五)。



▲(圖五)

四、機率$F(n,p,p_1)$

已知n個點中，有一邊會相連的機率為p_1，其餘的邊會相連的機率皆為p。設函數$F(n,p,p_1)$為圖形連通的機率。我們得到以下遞迴式：

$$\begin{cases} F(o,p,p_1)=F(1,p,p_1)=1 \\ F(n,p,p_1)=\underbrace{1}_{①}-\underbrace{\sum_{k=1}^{n-1}}_{②}\left[\underbrace{C_{k-2}^{n-2}F(k,p,p_1)}_{③}+\underbrace{\frac{1-p_1}{1-p}C_{k-1}^{n-2}P(k,p)}_{④}\right]\underbrace{(1-p)^{k(n-k)}}_{⑤} \end{cases}$$

【說明】①所有事件之機率和；

②固定其中一頂點v_i，讓它和其他$k-1$個頂點相連通，使這k個頂點形成一個連通子集，條件$k \le n-1$下整張圖不連通；

③將機率為p_1之邊納入了k個頂點的連通子集之中的機率大小；

④只納入了被固定的那一點v_i的機率大小；

⑤使相連通的k個頂點和剩下的$n-k$個頂點完全不相連的機率大小。

舉例說明：假設有 4 個點，p 為 0.2，p_1 為 0.8，藉由公式可得 connected 機率為 0.186944。

五、機率$D(n, p, Q)$

已知n個點中，所有的邊會相連的機率為p。設函數$D(n, p, Q)$為事件「最大的連通子圖的頂點數小於等於Q」發生的機率。我們可以得到以下式子：

$$D(n, p, Q)=\begin{cases} 1 & , when\ n \le Q \\ \underbrace{\sum_{k=1}^{Q}}_{①} \underbrace{C_{k-1}^{n-1}}_{②} \underbrace{(1-p)^{k(n-k)}}_{③} \underbrace{P(k, p)}_{④} A(n-k, p, Q) & , when\ n > Q \end{cases}$$

【說明】①限定$k \le Q$，使得最大連通子圖之頂點數小於等於Q；

②從v_i以外的$n-1$個頂點之中選出$k-1$個頂點的方法數；

③使相連通的k個頂點和剩下的$n-k$個頂點完全不相連的機率大小；

④k個頂點相連通的機率大小。

舉例說明：假設有 4 個點，兩點連通機率為 0.2，且最大的連通子圖的頂點數為 3，由算式可得 disconnected 機率為 0.917504，也就是$1-P(4, 0.2)$。

六、機率$B(n, p, p_1, Q)$

已知n個點中，有一邊會相連的機率為p_1，其餘的邊會相連的機率皆為p。設函數$B(n, p, p_1, Q)$為事件「最大的連通子圖的頂點數小於等於Q」會發生的機率。我們可以得到以下式子：

$$B(n, p, p_1, Q)=\begin{cases} 1 & , when\ n \le Q \\ \underbrace{\sum_{k=1}^{Q}}_{①} \left(\underbrace{C_{k-2}^{n-2} F(k, p, p_1)}_{②} + \underbrace{\frac{1-p_1}{1-p} C_{k-1}^{n-2} P(k, p)}_{③} \right) \underbrace{(1-p)^{k(n-k)}}_{④} B(n-k, p, p_1, Q) & , when\ n > Q \end{cases}$$

【說明】①從 m 個特殊的頂點中選出 i 個頂點，放入具有 k 個頂點的連通子集之中，且固定特殊頂點中的其中一個頂點 vi 是被選取的；

②將機率為 p_1 之邊納入了 k 個頂點的連通子集之中的機率大小；

③只納入了被固定的那一點 v_i 的機率大小；

④使相連通的 k 個頂點和剩下的 $n-k$ 個頂點完全不相連的機率大小；

七、機率 $H\left(n, m, p, p_1\right)$

已知 n 個點中有 m 個頂點($0 \le m \le n$)，其任兩點連線的機率為 p_1 ；剩餘的 $\left(n-m\right)$ 點中，任兩點的連線機率為 p ，設函數 $H(n,m,p,p_1)$ 為圖形連通的機率。我們得到以下遞迴：

$$\begin{cases} H(n, 0, p, p_1) = H(n, 1, p, p_1) = P(n, p) \\ H(n,m,p,p_1) = \underbrace{1}_{①} - \underbrace{\sum_{k=1}^{n-1}}_{②}\sum_{i=1}^{m} \underbrace{C_{k-i}^{n-m}C_{i-1}^{m-1}}_{③} \underbrace{\left(\frac{1-p_1}{1-p}\right)^{i(m-i)}\left(1-p\right)^{k(n-k)}}_{④} \underbrace{H\left(k, i, p, p_1\right)}_{⑤} & ,\forall m > 1, m \le n \end{cases}$$

【說明】①所有事件之機率和；

②固定其中一頂點 v_i ，讓它和其他 $k-1$ 個頂點相連通，使這 k 個頂點形成一個連通子集，在 $k \le n-1$ 這個條件下整張圖不連通；

③從 m 個特殊的頂點中選出 i 個頂點，放入具有 k 個頂點的連通子集之中，且固定特殊頂點中的其中一個頂點 v_i 是被選取的；

④使相連通的 k 個頂點和剩下的 $n-k$ 個頂點完全不相連的機率大小；

⑤具有 i 個特殊的頂點，共 k 個頂點連通的方法數。

八、機率 $C\left(n, m, p, p_1, Q\right)$

已知 n 個點中有 m 個頂點($0 \le m \le n$)，其任兩點連線的機率為 p_1 ；剩餘的 $\left(n-m\right)$ 點中，任兩點的連線機率為 p ，設函數 $C\left(n, m, p, p_1, Q\right)$ 為事件「最大的連通子圖的頂點數小於等於 Q 」會發生的機率，我們推得以下式子：

$$C\left(n, m, p, p_1, Q\right) = \begin{cases} 1 & ,\text{when } n \le Q \\ \underbrace{\sum_{k=1}^{Q}}_{①}\sum_{i=1}^{m} \underbrace{C_{k-i}^{n-m}C_{i-1}^{m-1}}_{②} \underbrace{\left(\frac{1-p_1}{1-p}\right)^{i(m-i)}\left(1-p\right)^{k(n-k)}}_{③} \underbrace{H(k,i,p,p_i)}_{④} C\left(n-k, m-i, p, p_1, Q\right) & ,\text{when } n > Q \end{cases}$$

【說明】①限定 $k \le Q$ ，使得最大連通子圖之頂點數小於等於 Q ；

②從 m 個特殊的頂點中選出 i 個頂點，放入具有 k 個頂點的連通子集之中，且固定特殊頂點中的其中一個頂點 v_i 是被選取的；

③使相連通的 k 個頂點和剩下的 $n-k$ 個頂點完全不相連的機率大小；

④具有 i 個特殊的頂點，共 k 個頂點連通的方法數。

舉例說明：1. 假設有 4 個點，p 為 0.2，p_1 為 0.8，且最大的連通子圖的頂點數為 3，由算式可得 $B(4\,,0.2\,,0.8\,,3)$ 為 0.813056；2. 假設有 4 個頂點，其中有 3 個點 $p_1 = 0.8$，其餘為 $p = 0.2$ ，且最大的連通子圖的頂點數為 3，則 $C(4\,,3\,,0.2\,,0.8\,,3)$ 為 0.555776。

九、機率 $J(V_n, P_{V_n})$

已知每條邊不同的連線機率，且定義 $V_n := \{v_i \mid 1 \le i \le n\,, i \in \mathbb{N}\}$ ， $N_n := \{a \big| a \in [1,n], a \in \mathbb{N}\}$ ，

$$P_{V_n} := \left\{ p_{ij} \left| \begin{cases} p_{ij} \in [0,1] \\ p_{ii} = 0 \\ p_{ij} = p_{ji} \end{cases} ,\forall\, 1 \le i\,,\, j \le n\,, i \in \mathbb{N}\,,\, j \in \mathbb{N} \right. \right\} ,\quad V^k := \left\{ V_k' \left| \begin{cases} k \in N_{n-1} \\ V_k' \subset V_n\,, \left\| \bigcap_{k=1}^{n} V_k' \right\| = 1 \\ |V_k = k| \end{cases} \right. \right\}$$

推得以下遞迴式:

$$\begin{cases} J\left(\{v_i, v_j\}, \{p_{ij}\}\right) = p_{ij}, \forall\, 1 \le i\,,\, j \le n\,, i \in \mathbb{N}\,,\, j \in \mathbb{N} \\ J(V_n, \underbrace{P_{V_n}}_{①}) \underbrace{=}_{②} 1 - \underbrace{\sum_{V_m \in V^k} J(V_m, P_{V_m})}_{③} \underbrace{\prod_{\begin{cases} v_i \in V_m \\ v_j \in V_n - V_m \end{cases}} \left(1 - J\left(\{v_i, v_j\}, \{p_{ij}\}\right)\right)}_{④} \end{cases}$$

【說明】①所有事件之機率和；

②取 k 個點，且 $1 \le k \le n$ ，其集合都有包含一個固定點；

③使 k 個點為 connected 的機率；

④其 connected 的 k 個點和其他點 disconnected 的機率。

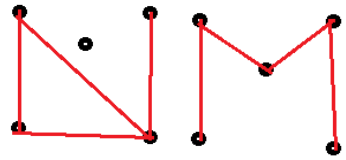
陸、討論

一、在前置研究中進一步分析發現，若$1\le n\le 4$時，$n-1\ge C_2^{n-1}$恆成立。倘若我們考慮邊的數滿足$n-1>E>C_2^{n-1}$………條件(A)，其結果並不與前置研究的結果(二)與(三)產生矛盾。原因是滿足條件(A)的邊數E並不存在，我們將可能的E列表如下：

	C_2^{n-1}	n-1	滿足條件(A)的 E 數
n=2	0	1	不存在
n=3	1	2	不存在
n=4	3	3	不存在

所以今後我們可以只考慮$n\ge 5$才適用本研究的結果。

二、在前置研究(二)與(三)的解決過程裡，我們原本希望得到一個判斷圖形是否為 connected 的等價條件；但實際上發現，僅有邊和點的個數，無法唯一決定圖形的拓樸結構(topological structure)，例如(圖六)，兩個圖形都是$G(5,4)$，但左圖是 disconnected，右圖卻是 connected。因此，此研究僅找到決定圖形 connected 的*必要條件*。



▲(圖六)

拓樸結構 Connected ＆ Disconnected 比較

三、問題(一)，已知一張連通圖有n個頂點，E條邊(此圖連通機率為p_1)，若多加入了一個與原圖相連的頂點且增加e條邊(此圖連通機率為p_2)，則連通的機率的增加率$r_1(n,E,1,e)$會是多少？

(連通邊機率×不連通邊機率×連通圖個數)

↑

$$r_1(n,E,1,e)=\frac{p_2-p_1}{p_1}=\frac{(p^{E+e}(1-p)^{C_2^{n+1}-(E+e)}K(n+1,E+e))-p^E(1-p)^{C_2^n-E}K(n,E)}{p^E(1-p)^{C_2^n-E}K(n,E)}$$

$$=\frac{p^e(1-p)^{n-e}K(n+1,E+e)-K(n,E)}{K(n,E)},\begin{cases}n-1\le E\le C_2^n\\1\le e\le n\end{cases},E,e\in\mathbb{N}$$

例子:當　$E=5,n=4,e=1,p=0.9$時

$$\frac{p^e(1-p)^{n-e}K(n+1,E+e)-K(n,E)}{K(n,E)}$$

$$=\frac{0.9\times 0.1^3K(5,6)-K(4,5)}{K(4,5)}$$

$$=\frac{0.9\times 0.1^3\times 205-6}{6}$$

$$=-0.96925$$

問題(二)，已知一張連通圖有 n 個頂點，E 條邊(此圖連通機率為 p_3)，
若多加入了一個與原圖不相連的頂點(此圖連通機率為 p_4)，則連通的機率的增加率 r_2 是多少？

$$r_2=\frac{p_4-p_3}{p_3}=\frac{0-p^E(1-p)^{C_2^n-E}K(n,E)}{p^E(1-p)^{C_2^n-E}K(n,E)}=-1\ ,$$

而且 $n-1\le E\le C_2^n$ 。

四、我們運用引理四及以下過程步驟估計 $K(n,E)$ 圖形的趨勢和範圍。

(一)【引理 4】

設 $a,b\in\mathbb{Z}$ ，

$$\begin{cases}\dfrac{a!}{(b-1)!\left(a-b+1\right)!}\le\dfrac{a!}{b!\left(a-b\right)!}\\[2ex]\dfrac{a!}{(b+1)!\left(a-b-1\right)!}\le\dfrac{a!}{b!\left(a-b\right)!}\end{cases}\Rightarrow\text{此時}C_b^a\text{為最大值}\Rightarrow\frac{a-1}{2}\le b\le\frac{a+1}{2}$$

當 $a\to\infty$ ，C_b^a 的最大值 $\approx C_{\frac{a}{2}}^a$ 。

(二) $\because\begin{cases}E<n-1\text{時，則圖形必為disconnected，}\\ K\left(n,E\right)=C_E^{C_2^n}-\displaystyle\sum_{k=1}^{n-1}\sum_{e=0}^{E}C_{k-1}^{n-1}C_{E-e}^{C_2^{n-k}}K(k,e)\ge 0\\ C_b^a\text{ 的最小值為 }C_{\frac{a}{2}}^a\text{且}a,b\in\mathbb{Z}--\text{【引理4】}\end{cases}$ ，

$$\therefore\ C_{n-1}^{C_2^n}\le K(n,E)\le C_E^{C_2^n}\le C_{C_2^n/2}^{C_2^n}\Rightarrow\ln(C_{n-1}^{C_2^n})\le\ln(K(n,E))\le\ln(C_E^{C_2^n})\le\ln(C_{C_2^n/2}^{C_2^n})$$

利用 Strilng formula$\left\{\ln(N!)\approx\left(N\right)\ln(N)-N\right\}$我們可以得到：

$$C_{C_2^n/2}^{C_2^n}\approx\frac{2^{C_2^n}}{C_2^n}\le 2^{C_2^n}\le e^{n^2}\Rightarrow\ln(C_{C_2^n/2}^{C_2^n})\approx\ln(2^{C_2^n})\le n^2\ 。$$

綜合以上我們可以得到：

$$\left|\frac{\ln(K(n,E))}{\ln(C_E^{C_2^n})}\right|\le\left|\frac{\ln(K(n,E))}{\ln(C_{C_2^n/2}^{C_2^n})}\right|\le\left|\frac{\ln(K(n,E))}{\ln(2^{C_2^n})}\right|\le\left|\frac{\ln(K(n,E))}{n^2}\right|\le 1$$

$$\Rightarrow\mathrm{O}(\ln(K(n,E)))\le\mathrm{O}(\ln(K(n,\frac{1}{2}C_2^n)))=\mathrm{O}(n^2)--(1)$$

當 $E \approx n-1$，

$$\because \begin{cases} E < n-1\text{時，則圖形必為disconnected} \\ K(n,n-1) = n^{n-2} \end{cases},$$

$$\therefore \left|\frac{\ln(K(n,n-1))}{n\ln(n)}\right| \le \left|\frac{\ln(K(n,n-1))}{(n-2)\ln(n)}\right| = 1$$

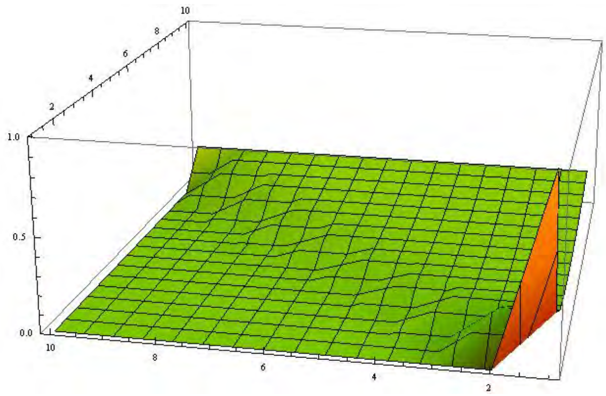
$$\Rightarrow O(n\ln(n)) \le O(\ln(K(n,n-1))) \le O(\ln(K(n,E))) --(2)$$

綜合(1)(2)我們可以得到：

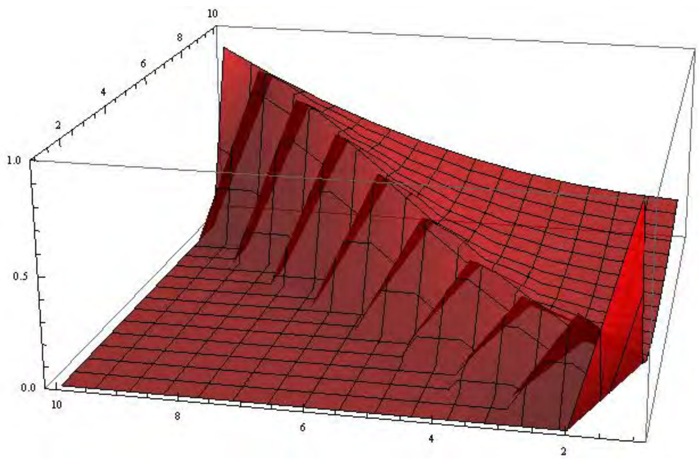
$O(n\ln(n)) \le O(\ln(K(n,E))) \le O(n^2)$

由 n 得知連通時 E 的範圍，進而可估計 $K(n,E)$ 圖形的趨勢和範圍。

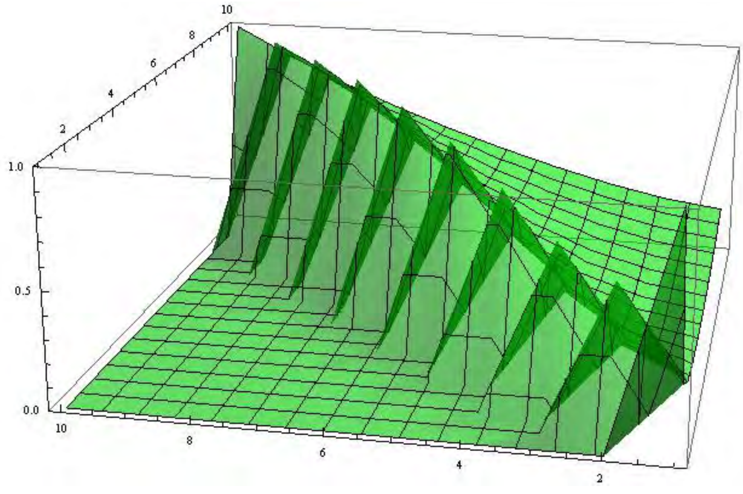
五、我們藉由 Excel VBA 寫出 $H(n, m, p, p_1)$ 的計算程式，在固定 p 和 p_1 的條件中，求出 n 、 m 分別為 1~10 時不同的 $H(n, m, p, p_1)$，並用數學軟體 Mathematica 畫出這些值在三維空間中的圖形，如下(圖七)、(圖八)、(圖九)、(圖十)。



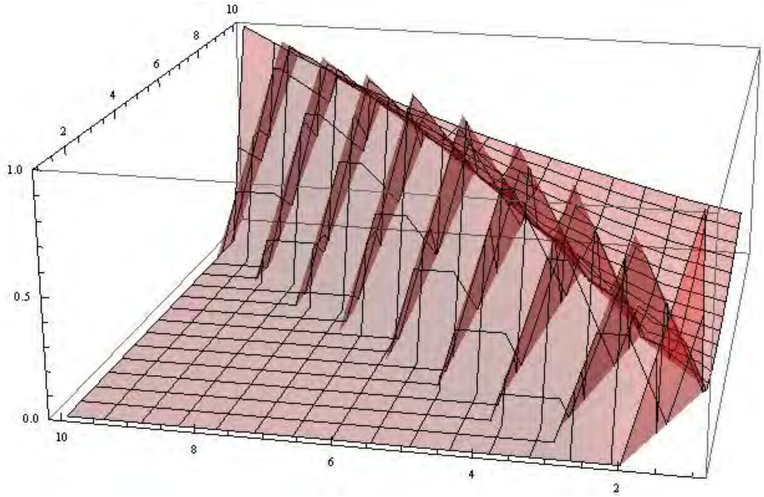
▲(圖七) $p = p_1 = 0.2$ 時



▲(圖八) $p = 0.2$ ， $p_1 = 0.4$ 時

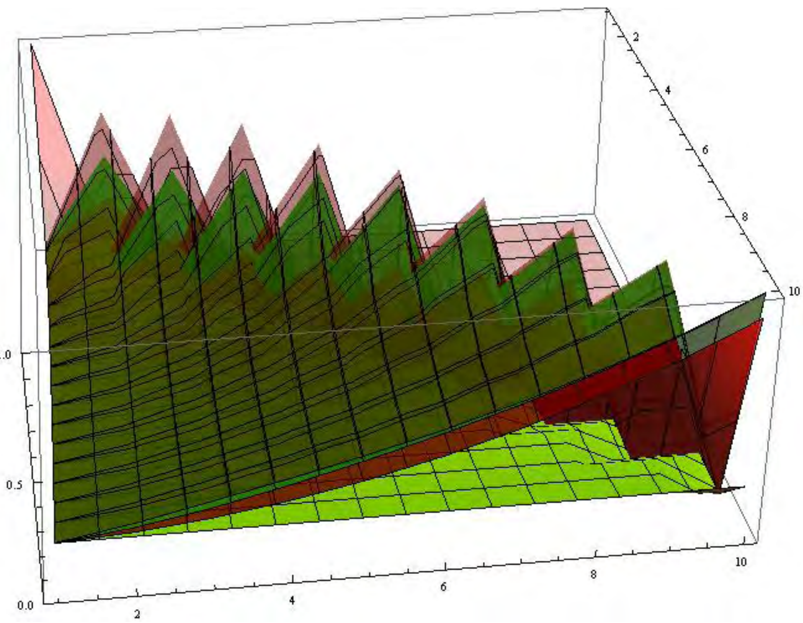


▲(圖九) $p=0.2$ ， $p_1=0.6$時

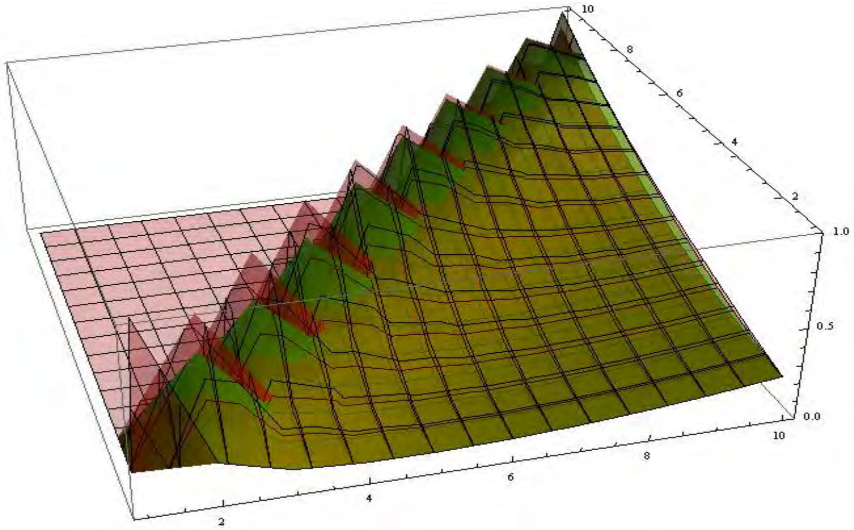


▲(圖十) $p=0.2$ ， $p_1=0.8$時

再將上面的圖疊合，並從不同角度看。得到(圖十一)、(圖十二)。



▲(圖十一)疊合圖角度 1



▲(圖十二)疊合圖角度 2

柒、結論

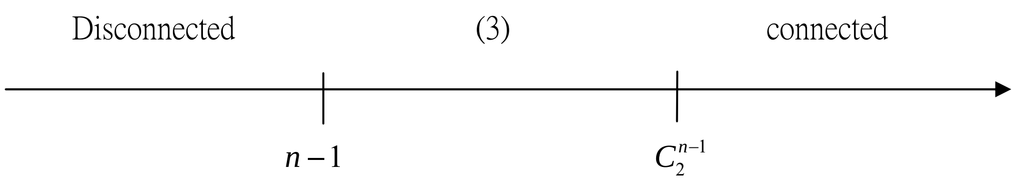
一、前置研究中在不考慮機率並且只考慮點數n以及連線邊數E的前提下，得到以下結果：

(一) 1.若$E<n-1$時，則圖形必為 disconnected；

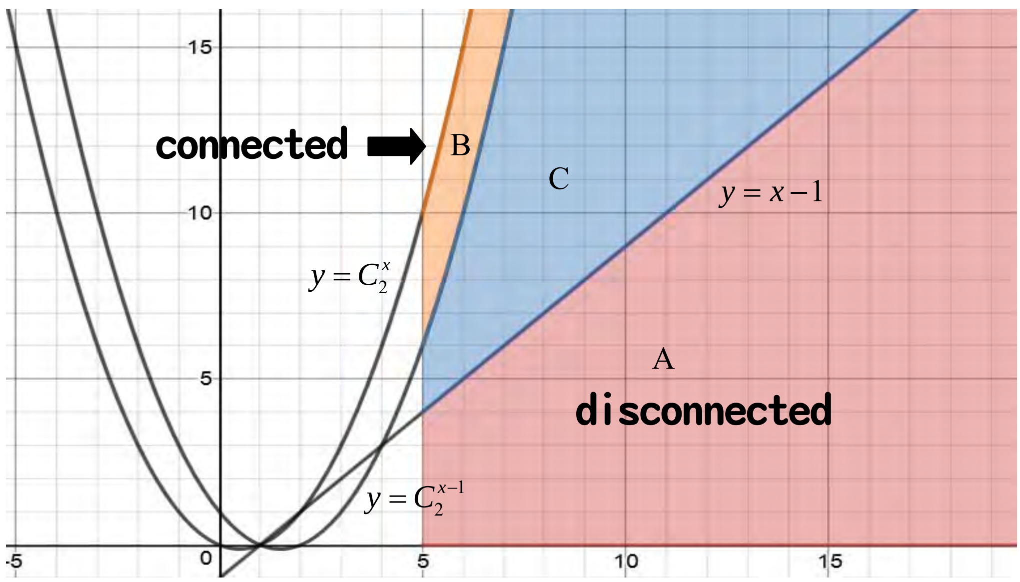
2.若$E>C_2^{n-1}$時，則圖形必為 connected；

3.若集合G有 Hamiltonia path，圖形必為 connected(不考慮 Hamiltonia circuit)

總結以上三點，我們可以以E為一軸標出三個範圍：



(二)若假設$\left\{f(x)=x-1,\ g(x)=C_2^x,\ h(x)=C_2^{x-1}\right\}$，我們有下列圖形(圖十三)：



我們以點座標(n, E)表示圖形點的個數和邊的個數。當點(n, E)落在橘色區域時，圖形$G(n, E)$為 disconnected，當點(n, E)落在紅色區域時，圖形$G(n, E)$為 connected。當點落在藍色區域時，目前無法判定圖形是否為連通集。我們未來將積極尋找相對應的條件，使其成為是否連通的*充分條件*，再搭配本論文結果，提供一個圖形連通的等價判斷準則。

(三)承上述第 3 點的分析，我們可以回答下面的積分幾何問題：

給定固定點數N_1，固定邊數E_1 (N_1, E_1可以很大)，任意選取一個$G(n,E)$，其中$5 \le n \le N_1$且$4 \le E \le E_1$，其圖形必為 connected 的機率必定大於等於

$$\frac{(4+E_{1})N_1 \cdot \frac{1}{2}}{(N_1-5)(E_1)} \begin{array}{l} \rightarrow \text{connected area B} \\ \rightarrow \text{總面積=(A+B+C)} \end{array} ;$$

其圖形必為 disconnected 的機率必大於等於

$$\frac{Area\ A}{(N_1-5)E_1} \begin{array}{l} \\ \rightarrow \text{總面積=(A+B+C)} \end{array}$$ ，其中 Area A 可用微積分求出。

二、藉由研究結果二到三的遞迴式我們發現：

在n個點中，假設已知任兩點的連線機率最多有兩種：p及p_1，則圖形連通的機率問題可藉由排容原理搭配遞迴式分解成我們研究結果中$P(n, p)$與$F(n, p, p_1)$的級數和。

三、總結研究結果一至八：若以箭頭【A ⟸ B】表示【結果 B 為結果 A 的推廣】，我們有：

結果五

結果二 ⟸ 結果四 ⟸ 結果六 ⟸ 結果七 ⟸ 結果八

，亦即結果八是最終的推廣型態。例如，不難驗證在結果八中，當$Q = n, m = 0, p = p_1$時，其公式退化為結果一。其中結果一和結果三的推導是為了計算各式機率時，所需使用到計算各種子圖個數的遞迴公式。

四、本研究結果有許多有趣的應用，例如：由<u>結論七</u>可知，任意給定m, n, p, p_1，我們可以計算該圖連通的機率$H(n, m, p, p_1)$。可以反問，給定m, n, p並限制連通的機率H需要大於某一給定數字r，$0 \le r \le 1$，則可利用遞迴公式不斷增加p_1值，直到H大於r。換言

之，在實際應用上，我們可以找出最小的可能 p_1 值，使得該圖整體連通機率大過給定某數字， 於是此最小 p_1 值便是控制該圖聯通與否的關鍵值。另外也可以更動其他變量，例如給定 H , m , p , p_1 , r ，可以求出可能的 n 值範圍等等，其應用範圍之廣。除此之外，我們的結果也和 Erdos-Renyi Random Graphs 問題[2]緊密結合，並且可以應用至數學[3]或其他科學領域[4]，特別是許多網路模型[5, 6]，計算社交網站人群的關聯性等。總的來說，以純粹數學的角度作為理論基礎的結論一，其本質是來判斷一個圖是否為 connected 充分條件，其應用層面顯而可見；例如最近討論有關 nontransitive dice[7]與 Game theory 相關的文章裡面提到的方法，其圖形就必須是 connected。

捌、參考資料及其他

[1](網頁資料)The On-Line Encyclopedia Of Integer Sequence：http://oeis.org/A001187.

[2] D. West, Introduction to Graph Theory(2nd ed.), 2001, Prentice Hall.

[3] L. Lovasz , Random Walks on Graphs: A Survey, Combinatorics, Paul Erdos is Eighty is Eighty (Volume 2), Keszthely (Hungary), 1993, pp. 1 – 46.

[4] Reza Zafarani, Mohammad Ali Abbasi, , Huan Liu, Social Media Mining: An Introduction, 2014, Cambridge University.

[5](網頁資料)實踐大學影音平台：network models-Random Graphs (媒體中心>資訊科技>社群媒體探勘>http://media.usc.edu.tw/media/314)

[6] 楊振翔、劉維中、潘建興 (統計科學研究所)，知識天地：大型網絡之分析，中央研究院週報，第 1382 期, 2012.

[7] Month Horizons: What Nontransitive Dice Exist among Us? MAA press, April 2017, pp. 14-17.

【評語】050413

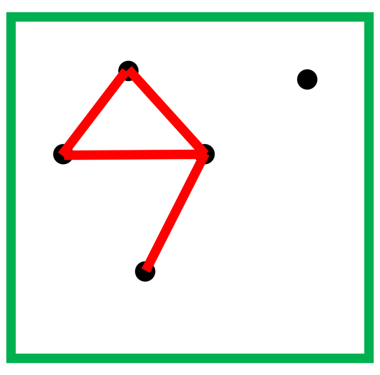
本作品探討隨機圖的課題：在給定一個圖 G(n,E)，假定任意兩點連接的機率是 P，求該圖連通的機率。之後，並將此結果做一推廣，假設該圖有兩子圖，其中各子圖的邊，其相連的機率分別為 P 及 P_1，求出該圖能形成連通的機率。本研究架構尚為完整，結果有一程度的興趣。但作品之研究主題，宜加強說明。另文獻探討亦可再加強，以說明與現有文獻之相關性。

作品海報

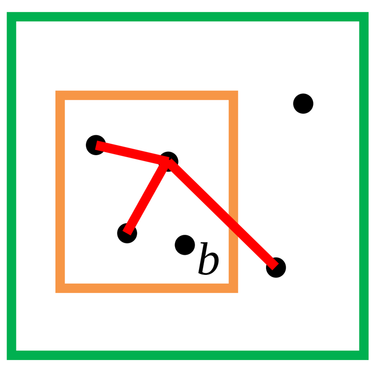
壹、研究大綱

在我們的研究中，先探討給定一個無向圖(undirected graph)並能使圖形必定 connected 的線段數 E 的條件，再以此得到圖形 connected 的機率範圍。我們接著引用已知整數數列 A001187(參見參考資料[1])類似的方法來解決隨機圖(Random Graph)的根本問題：給定一個圖 $G(n,E)$，假定任兩個點連接的機率是常數 p，求出該圖能形成 connected 的機率。接著我們將先前結論做推廣，假設該圖有兩子圖(subgraph)，其中個子圖中的邊，相連的機率分別為 p 及 p_1，求出該圖能形成 connected 的機率。在這個過程中，在限制 $G(n,E)$ 中最大連通子集的頂點數目下，我們也對於圖形不連通的機率做了許多研究。

貳、研究定義



▲(圖一)



▲(圖二)

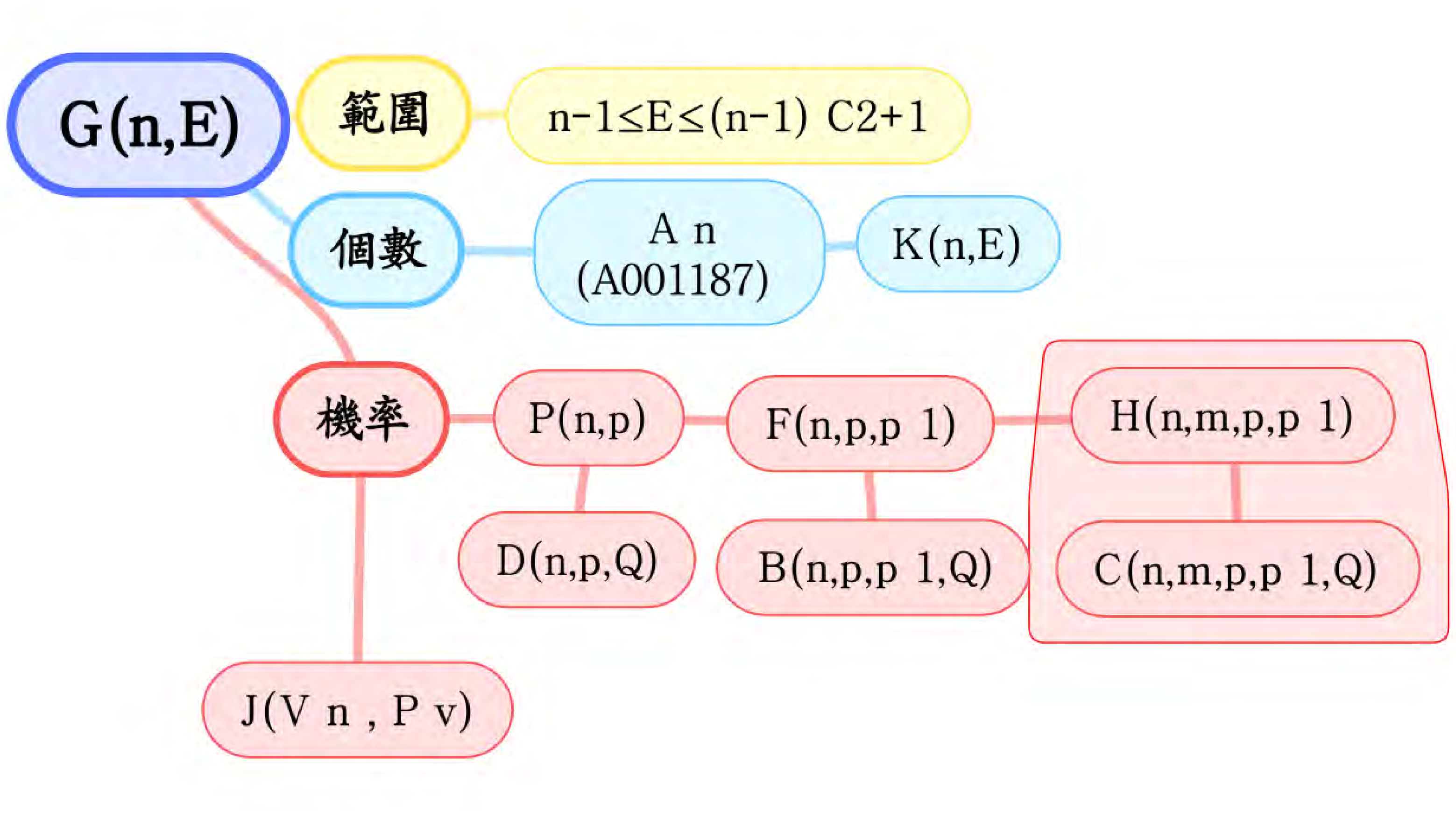
一、給定一集合 $G(n,E)$ 由下列元素組成：n 個點(node)，E 個邊(edge)連接任意兩個點。邊指的是沒有方向性的線段，並且限制 $0 \le E \le C_2^n$，亦即任兩點若相連，則被唯一的邊所聯繫。我們也稱集合 $G(n,E)$ 為一個圖(graphy)。我們的研究中可以允許某些點沒有任何一條邊與其相連，如(圖一):有一個點沒有任何的邊與其相連。

二、設 $G'(n',E')$ 為集合 $G(n,E)$ 中一個含有 n' 個點和 E' 條邊的子集合。如果在 $G'(n',E')$ 中，任意兩點 a,b 均可藉由若干個邊，從 a 點沿著邊移動到 b 點，則稱該子集 $G'(n',E')$ 為一個連通子集 (connected subset)，或簡稱連通(connected)。否則，則稱該子集 $G'(n',E')$ 為非連通 (disconnected)，如(圖二):Connected & Disconnected 說明。

三、n 個點各有其編號(Labeled Graph，意即所有點視為不同)且點和點之間的連線沒有方向性(non-oriented line)，意即任兩點只有 connected 和 disconnected 兩種情況。

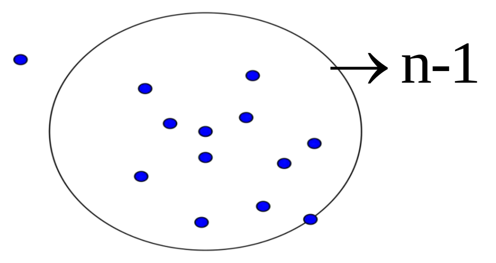
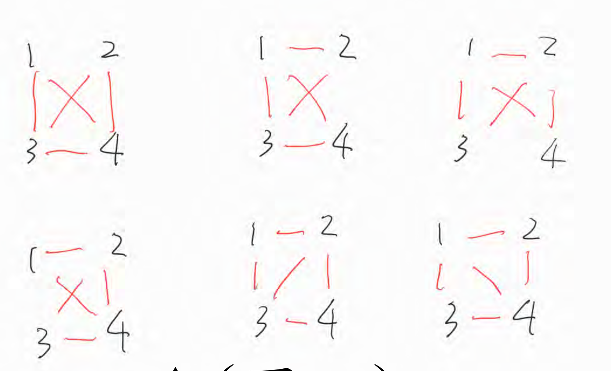
參、研究過程

一、研究架構及前置研究



問題一:所有連線情形的方法數	$2^{C_2^n}$
問題二:小於何數時，圖形必為disconnected	$E < n-1$
問題三:大於何數時，圖形必為 connected	$E > C_2^{n-1}$

二、研究過程

分類	題目	公式	例子
假任兩點連線的機率為 p	1. 給 n 個點，求 connected 圖形的個數 A_n	$\begin{cases} A_0 = A_1 = 1 \\ A_n = \underline{2^{C_2^n}} - \sum\limits_{k=2}^{n} \underline{\sum\limits_{\substack{\sum_{i=1}^{k} n_i = n \\ (n_i \in \mathbb{N})}}} \underline{\left[\prod\limits_{j=1}^{n} \frac{j}{\left(\sum\limits_{i=1}^{k} \delta_{n_{i,j}}\right)!} \prod\limits_{i=1}^{k} \frac{A_{n_i}}{(n_i)!} \right]} \end{cases}$ $\begin{cases} A_0 = A_1 = 1 \\ A_n = 2^{C_2^n} - \sum\limits_{k=1}^{n-1} C_{k-1}^{n-1} \cdot 2^{C_2^{n-k}} \cdot A_k \end{cases}$，A001187 的整數數列	 ▲(圖四)  ▲(圖五)
	2.給 n 個點，每條 E 連線機率為 p，求出圖形 connected 的機率 $P(n,p)$	$\begin{cases} P(n,p) = 1 - \sum\limits_{k=1}^{n-1} C_{k-1}^{n-1} P(k,p) q^{k(n-k)}, n \ge 3 \\ P(0,p) = P(1,p) = P(n,1) = 1 \\ P(2,p) = p = 1-q \end{cases}$	假設有 4 個點，且兩點連通機率為 0.2，connected 機率為 0.082496
	3. 連線段個數為 E 的所有 connected 圖形個數 $K(n,E)$	$\begin{cases} K(n,n-1) = n^{n-2} \\ K(n,E) = \underline{C_E^{C_2^n}} - \sum\limits_{k=1}^{n-1} \sum\limits_{e=0}^{E} \underline{C_{k-1}^{n-1}} \underline{C_{E-e}^{C_2^{n-k}}} \underline{K(k,e)} \end{cases}$ 且 $E \in [n-1, C_2^n]$	假設有 4 個點，5 條邊，connected 的圖形數為 6，如(圖五)

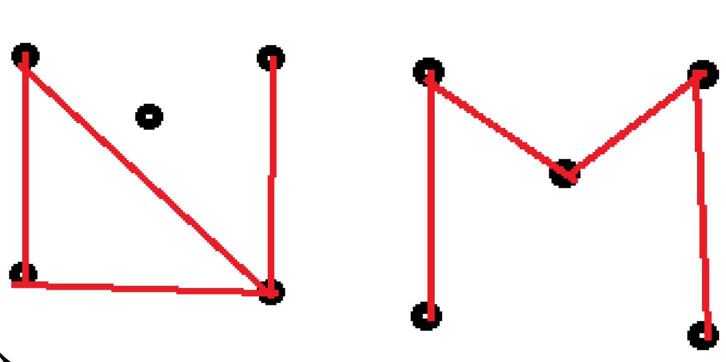
		$\begin{cases} K(n,n-1)=n^{n-2} \\ K(n,E)= C_E^{C_2^n} - \sum_{k=2}^{n} \sum_{\begin{cases}\sum_{i=1}^{k} n_i=n \\ \sum_{i=1}^{k} E_i=E \\ E_i\in\left[n_i-1,C_2^{n_i}\right]\end{cases}} \prod_{j=1}^{n}\frac{j}{\left(\sum_{i=1}^{k}\delta n_{i,j}\right)!}\prod_{i=1}^{k}\frac{K(n_i,E_i)}{(n_i)!} \end{cases}$	
	4. 機率 $F(n,p,p_1)$	$\begin{cases} F(0,p,p_1)=F(1,p,p_1)=1 \\ F(n,p,p_1)= 1 - \sum_{k=1}^{n-1}\left[C_{k-2}^{n-2}F(k,p,p_1)+\frac{1-p_1}{1-p}C_{k-1}^{n-2}P(k,p)\right](1-p)^{k(n-k)} \end{cases}$	假設有 4 個點， p 為 0.2，E 為 0.8，connected 機率為 0.186944
限制Q	5. 機率 $D(n,p,Q)$	$D(n,p,Q)=\begin{cases}1 \\ \sum_{k=1}^{Q}C_{k-1}^{n-1}(1-p)^{k(n-k)}P(k,p)A(n-k,p,Q) \quad ,when\ n>Q\end{cases}$	假設有 4 個點，兩點連通機率為 0.2，且最大的連通子圖的頂點數為 3，disconnected 機率為 0.917504
	6.機率 $B(n,p,p_1,Q)$	$B(n,p,p_1,Q)=\begin{cases}1 \\ \sum_{k=1}^{Q}\left(C_{k-2}^{n-2}F(k,p,p_1)+\frac{1-p_1}{1-p}C_{k-1}^{n-2}P(k,p)\right)(1-p)^{k(n-k)}B(n-k,p,p_1,Q) \\ ,when\ n>Q\end{cases}$	假設有 4 個點， p 為 0.2， p_1 為 0.8，且最大的連通子圖的頂點數為 3，$B(4,0.2,0.8,3)$ 為 0.813056
	7. 機率 $H(n,m,p,p_1)$	$\begin{cases} H(n,0,p,p_1)=H(n,1,p,p_1)=P(n,p) \\ H(n,m,p,p_1)= 1 - \sum_{k=1}^{n-1}\sum_{i=1}^{m}C_{k-i}^{n-m}C_{i-1}^{m-1}\left(\frac{1-p_1}{1-p}\right)^{i(m-i)}(1-p)^{k(n-k)}H(k,i,p,p_1) \\ ,\forall m>1,m\le n \end{cases}$	
	8. 機率 $C(n,m,p,p_1,Q)$	$C(n,m,p,p_1,Q)=\begin{cases}1 \\ \sum_{k=1}^{Q}\sum_{i=1}^{m}C_{k-i}^{n-m}C_{i-1}^{m-1}\left(\frac{1-p_1}{1-p}\right)^{i(m-i)}(1-p)^{k(n-k)}H(k,i,p,p_i)C(n-k,m-i,p,p_1,Q) \\ ,\text{when } n>Q\end{cases}$	假設有 4 個頂點，其中有 3 個點 $p_1=0.8$ ，其餘為 $p=0.2$ ，且最大的連通子圖的頂點數為 3，則 $C(4,3,0.2,0.8,3)$ 為 0.555776

肆、延伸

一、在前置研究中進一步分析發現，若邊的範圍為[1,4] 時，$n-1\ge C_2^{n-1}$ 恆成立。倘若我們考慮邊的數滿足 ………條件(A)，其結果並不與前置研究的結果(二)與(三)產生矛盾。原因是滿足條件(A)的邊數並不存在，我們將可能的邊數列表如右：

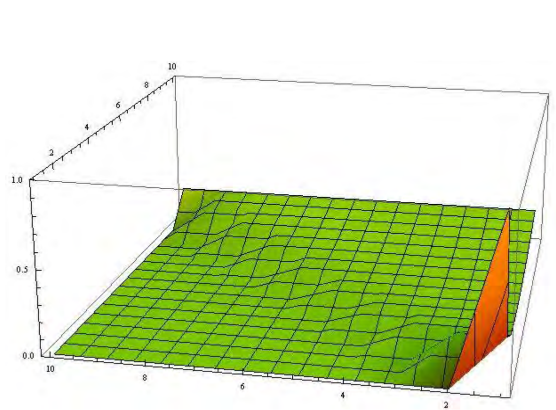
	n(n-1)/2	n-1	符合條件(A)的邊數
n=2	0	1	不存在
n=3	1	2	不存在
n=4	3	3	不存在

二、在前置研究(二)與(三)的解決過程裡，我們原本希望得到一個判斷圖形是否為connected的等價條件；但實際上發現，僅有邊和點的個數，無法唯一決定圖形的拓樸結構(topological structure)，例如(圖六)，兩個圖形都是 ，但左圖是disconnected，右圖卻是connected。因此，此研究僅找到決定圖形connected的*必要條件*。

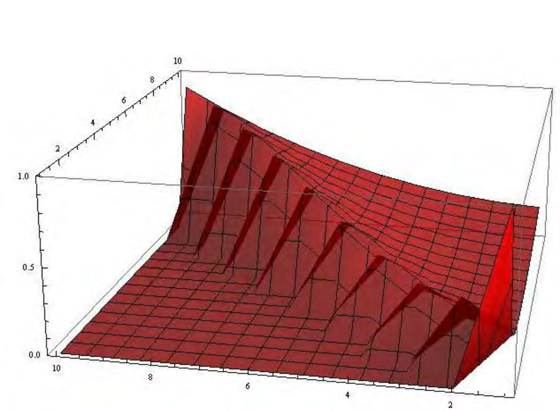


▲(圖六)

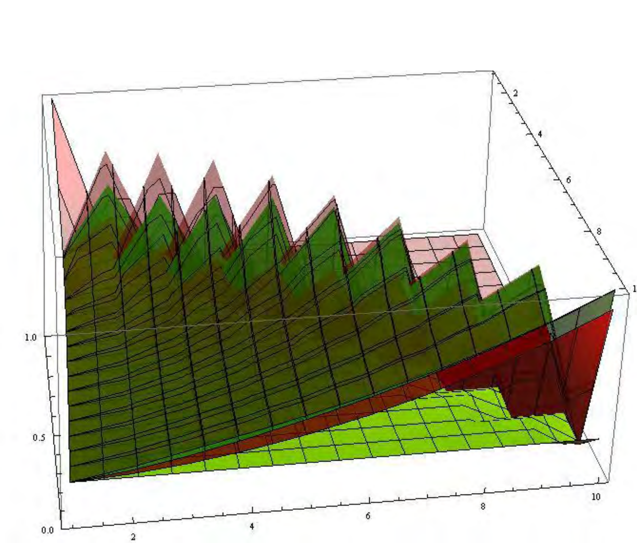
三、我們藉由Excel VBA寫出 $H(n,m,p,p_1)$ 的計算程式，在固定兩不同機率的條件中，求1~10時不同的值，並用數學軟體Mathematica畫出這些值在三維空間中的圖形，如左下。疊合圖如右下:

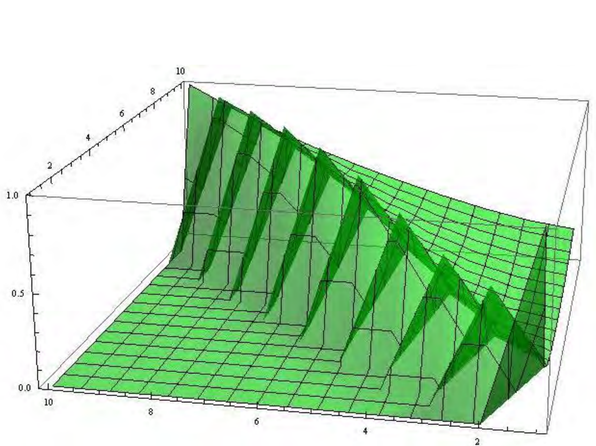


$p=p_1=0.2$

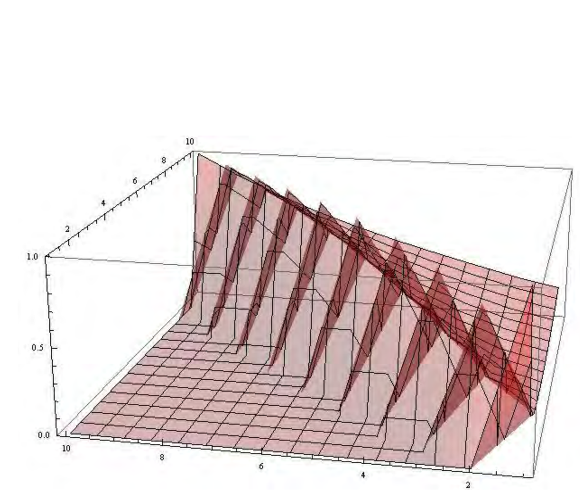


$p=0.2, p_1=0.4$

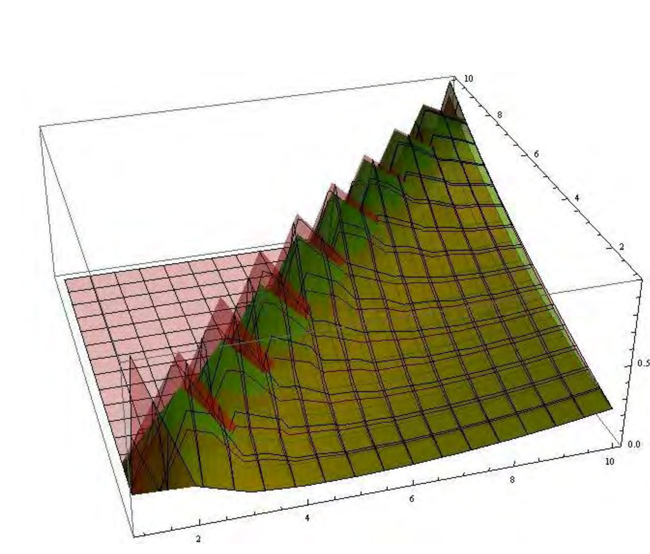




$p=0.2, p_1=0.6$



$p=0.2, p_1=0.8$



四、得到最終目的機率 $J(V_n,P_{V_n})$

已知每條邊不同的連線機率，且定義 $V_n:=\left\{v_i \mid 1\le i\le n, i\in\mathbb{N}\right\}$ ， $N_n:=\left\{a \mid a\in[1,n], a\in\mathbb{N}\right\}$ ， $P_{V_n}:=\left\{p_{ij} \,\middle|\, \begin{cases} p_{ij}\in[0,1] \\ p_{ii}=0 \\ p_{ij}=p_{ji} \end{cases}, \forall 1\le i, j\le n, i\in\mathbb{N}, j\in\mathbb{N}\right\}$ ，

$V^k:=\left\{V_k \,\middle|\, \begin{cases} k\in N_{n-1} \\ V_k\subset V_n, \left\|\bigcap_{k=1}^{n} V_k\right\|=1 \\ |V_k|=k \end{cases}\right\}$ ，推得以下遞迴式：$\begin{cases} J\left(\{v_i,v_j\},\{p_{ij}\}\right)=p_{ij}, \forall 1\le i, j\le n, i\in\mathbb{N}, j\in\mathbb{N} \\ J(V_n,P_{V_n})=1-\sum\limits_{V_m\in V^k} J(V_m,P_{V_m}) \prod\limits_{\begin{cases} v_i\in V_m \\ v_j\in V_n-V_m \end{cases}} \left(1-J\left(\{v_i,v_j\},\{p_{ij}\}\right)\right) \end{cases}$

五、問題(一)，已知一張連通圖有 n 個頂點， E 條邊(此圖連通機率為 p_1)，若多加入了一個與原圖相連的頂點且增加 e 邊(此圖連通機率為 p_2)，則連通的機率的增加率 $r_1(n,E,1,e)$ 會是多少？

答: $\dfrac{p^e(1-p)^{n-e}K(n+1,E+e)-K(n,E)}{K(n,E)}, \begin{cases} n-1\le E\le C_2^n \\ 1\le e\le n \end{cases}, E,e\in\mathbb{N}$

問題(二)，已知一張連通圖有 n 個頂點， E 條邊(此圖連通機率為 p_3)，若多加入了一個與原圖不相連的頂點(此圖連通機率為 p_4)，則連通的機率的增加率 r_2 是多少？

答:-1

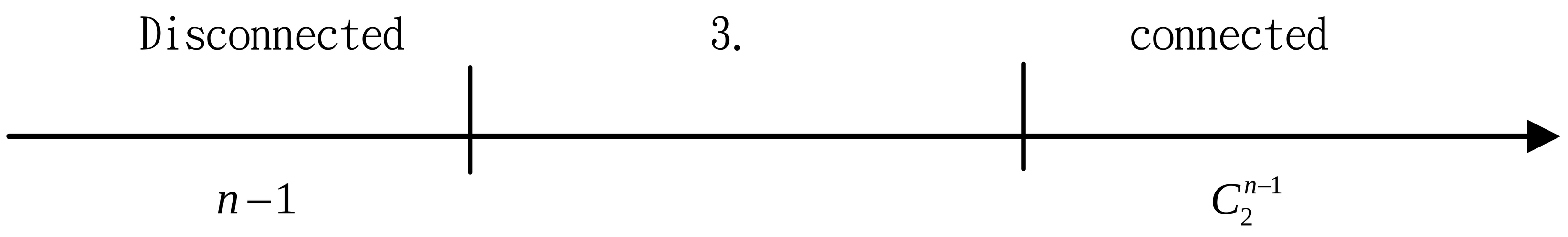
六、我們估計 $K(n,E)$ 圖形的趨勢和範圍。

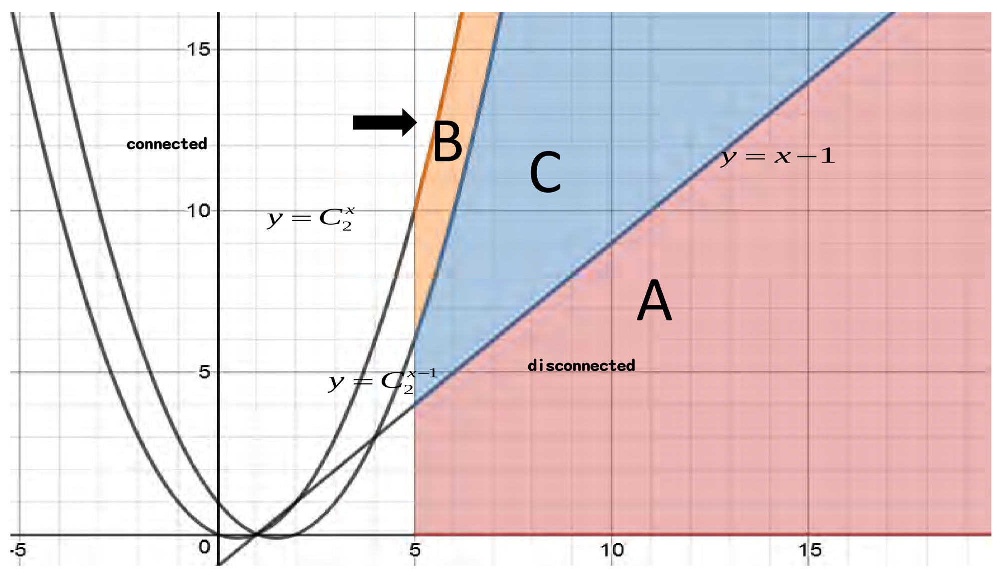
伍、結論

一、前置研究中在不考慮機率並且只考慮點數 n 以及連線邊數 E 的前提下，得到以下結果：

1. 若 $E<n-1$ 時，則圖形必為 disconnected；
2. 若 $E>C_2^{n-1}$ 時，則圖形必為 connected。
3. 若集合 G 有 Hamiltonia path，圖形必為 connected(不考慮 Hamiltonia circuit)

總結以上，可以得到以下:





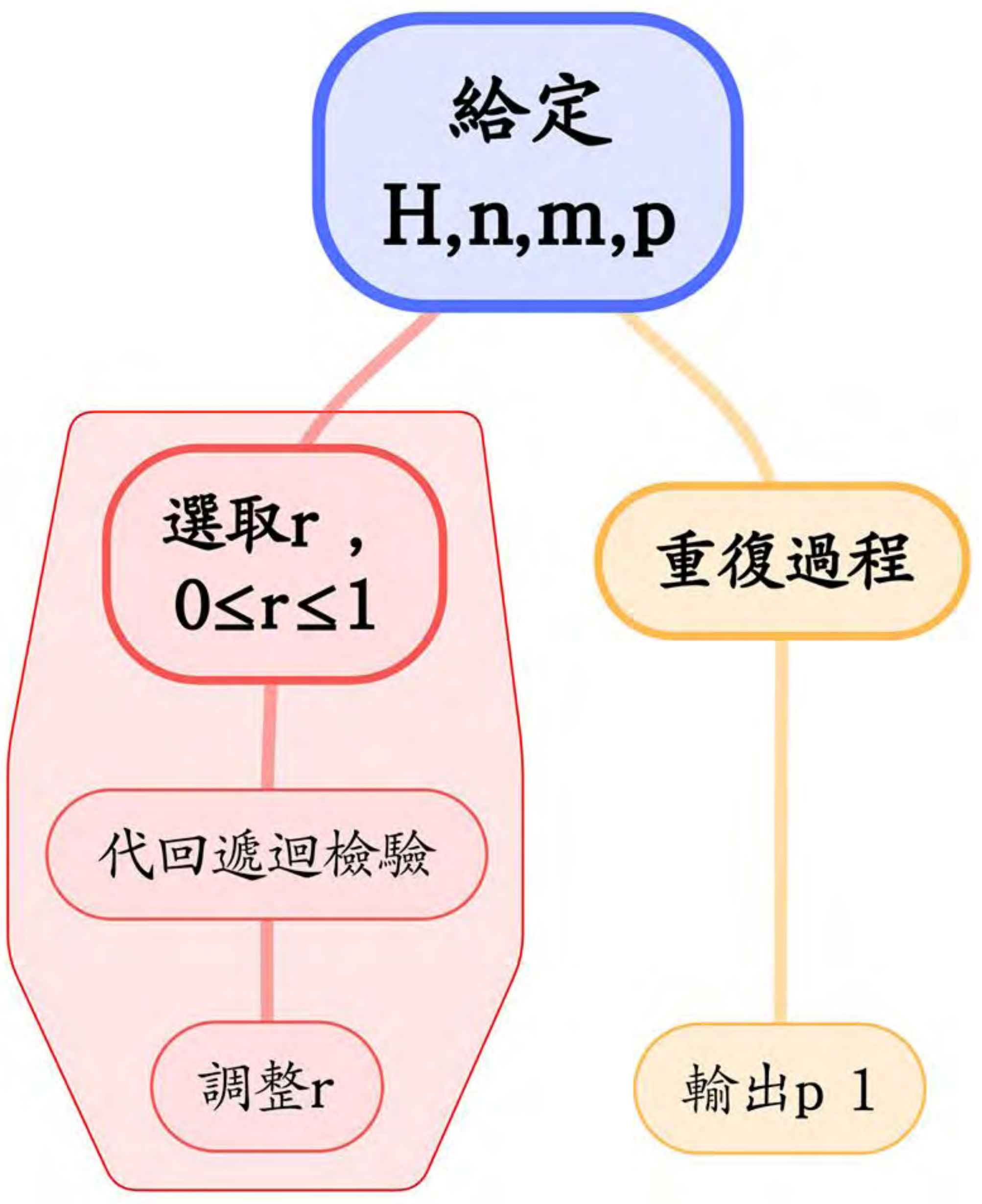
若假設 $\{f(x)=x-1, g(x)=C_2^x, h(x)=C_2^{x-1}\}$ ，我們有右圖：

我們以點座標 (n,E) 表示圖形點的個數和邊的個數。當點 (n,E) 落在橘色區域 時，圖形 $G(n,E)$ 為 disconnected，當點 (n,E) 落在紅色區域時，圖形 $G(n,E)$ 為 connected。當點落在藍色區域時，目前無法判定圖形是否為連通集。我們未來將積極尋找相對應的條件，使其成為是否連通的*充分條件*，再搭配本論文結果，提供一個圖形連通的等價判斷準則。

二、實際應用上，我們可以找出最小的可能 p_1 值，使得該圖整體連通機率大過給定某數字， 於是此最小 p_1 值便是控制該圖聯通與否的關鍵值。另外也可以更動其他變量。

例如給定 H, m, p, p_1, r，可以求出可能的 n 值範圍等等，其應用範圍之廣。除此之外，我們的結果也和 Erdos-Renyi Random Graphs 問題[2]緊密結合，並且可以應用至數學[3]或其他科學領域[4]，特別是許多網路模型[5,6]，計算社交網站人群的關聯性等。總的來說，以純粹數學的角度作為理論基礎的結論一，其本質是來判斷一個圖是否為 connected 充分條件，其應用層面顯而可見；例如最近討論有關 nontransitive dice[7]與 Game theory 相關的文章裡面提到的方法，其圖形就必須是 connected。

三、



四、