RANDOM GRAPHS

Joel Spencer *

ICTP '12

Graph Theory Preliminaries A graph G, formally speaking, is a pair $(V(G), E(G))$ where the elements $v \in V(G)$ are called vertices and the elements of $E(G)$, called edges, are two element subsets $\{v, w\}$ of $V(G)$. When $\{v, w\} \in E(G)$ we say v, w are adjacent. (In standard graph theory terminology our graphs are undirected and have no loops and no multiple edges.) Pictorially, we often display the $v \in V(G)$ as points and draw an arc between v and w when they are adjacent. We call $V(G)$ the vertex set of G and $E(G)$ the edge set of G. (When G is understood we shall write simply V and E respectively. We also often write $v \in G$ or $\{v, w\} \in G$ instead of the formally correct $v \in V(G)$ and $\{v, w\} \in E(G)$ respectively.) A set $S \subseteq V$ is called a *clique* if all pairs $x, y \in S$ are adjacent. The clique number, denoted by $\omega(G)$, is the largest size of a clique in G. The complete graph on k vertices, denoted by K_k, consists of a vertex set of size k with all pairs x, y adjacent.

1 Lecture 1: Random Graphs

1.1 What is a Random Graph

Let n be a positive integer, $0 \leq p \leq 1$. The random graph $G(n, p)$ is a probability space over the set of graphs on the vertex set $\{1, \ldots, n\}$ determined by

$$\Pr[\{i, j\} \in G] = p \tag{1}$$

with these events mutually independent.

Random Graphs is an active area of research which combines probability theory and graph theory. The subject began in 1960 with the monumental paper *On the Evolution of Random Graphs* by Paul Erdős and Alfred Rényi. The book *Random Graphs* by Béla Bollobás is the standard source for the field. The book *The Probabilistic Method* by Noga Alon and this author contains much of the material in these notes, and more.

There is a compelling dynamic model for random graphs. For all pairs i, j let $x_{i,j}$ be selected uniformly from $[0, 1]$, the choices mutually independent. Imagine p going from 0 to 1. Originally, all potential edges are "off". The edge from i to j (which we may imagine as a neon light) is turned on when p reaches $x_{i,j}$ and then stays on. At $p = 1$ all edges

*Courant Institute of Mathematical Sciences (New York). E-mail: spencer@cims.nyu.edu

are "on". At time p the graph of all "on" edges has distribution $G(n,p)$. As p increases $G(n,p)$ *evolves* from empty to full.

In their original paper Erdős and Rényi let $G(n,e)$ be the random graph with n vertices and precisely e edges. Again there is a dynamic model: Begin with no edges and add edges randomly one by one until the graph becomes full. Generally $G(n,e)$ will have very similar properties as $G(n,p)$ with $p \sim \frac{e}{\binom{n}{2}}$. We will work on the probability model exclusively.

1.2 Threshold Functions

The term "the random graph" is, strictly speaking, a misnomer. $G(n,p)$ is a probability space over graphs. Given any graph theoretic property A there will be a probability that $G(n,p)$ satisfies A, which we write $\Pr[G(n,p) \models A]$. When A is monotone $\Pr[G(n,p) \models A]$ is a monotone function of p. As an instructive example, let A be the event "G is triangle free". Let X be the number of triangles contained in $G(n,p)$. Linearity of expectation gives

$$E[X] = \binom{n}{3} p^3 \tag{2}$$

This suggests the parametrization $p = c/n$. Then

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} \binom{n}{3} p^3 = c^3/6 \tag{3}$$

We shall see that the distribution of X is asymptotically Poisson. In particular

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = \lim_{n\to\infty} \Pr[X=0] = e^{-c^3/6} \tag{4}$$

Note that

$$\lim_{c\to 0} e^{-c^3/6} = 1 \tag{5}$$

$$\lim_{c\to\infty} e^{-c^3/6} = 0 \tag{6}$$

When $p = 10^{-6}/n$, $G(n,p)$ is very unlikely to have triangles and when $p = 10^6/n$, $G(n,p)$ is very likely to have triangles. In the dynamic view the first triangles almost always appear at $p = \Theta(1/n)$. If we take a function such as $p(n) = n^{-.9}$ with $p(n) \gg n^{-1}$ then $G(n,p)$ will almost always have triangles. Occasionally we will abuse notation and say, for example, that $G(n, n^{-.9})$ contains a triangle - this meaning that the probability that it contains a triangle approaches 1 as n approaches infinity. Similarly, when $p(n) \ll n^{-1}$, for example, $p(n) = 1/(n \ln n)$, then $G(n,p)$ will almost always not contain a triangle and we abuse notation and say that $G(n, 1/(n \ln n))$ is trianglefree. It was a central observation of Erdős and Rényi that many natural graph theoretic properties become true in a very narrow range of p. They made the following key definition.

Definition 1. *$r(n)$ is called a* threshold function *for a graph theoretic property A if*

1. *When $p(n) \ll r(n), \lim_{n\to\infty} \Pr[G(n,p) \models A] = 0$*
2. *When $p(n) \gg r(n), \lim_{n\to\infty} \Pr[G(n,p) \models A] = 1$*

or visa versa.

In our example, $1/n$ is a threshold function for A. Note that the threshold function, when one exists, is not unique. We could equally have said that $10/n$ is a threshold function for A.

Lets approach the problem of $G(n, c/n)$ being trianglefree once more. For every set S of three vertices let B_S be the event that S is a triangle. Then $\Pr[B_S] = p^3$. Then "trianglefreeness" is precisely the conjunction $\wedge\overline{B}_S$ over all S. If the B_S were mutually independent then we *would* have

$$\Pr[\wedge\overline{B}_S] = \prod[\overline{B}_S] = (1-p^3)^{\binom{n}{3}} \sim e^{-\binom{n}{3}p^3} \to e^{-c^3/6} \tag{7}$$

The reality is that the B_S are not mutually independent though when $|S \cap T| \leq 1$, B_S and B_T are mutually independent. This is quite a typical situation in the study of random graphs in which we must deal with events that are "almost", but not precisely, mutual independent.

1.3 Variance

Here we introduce the Variance in a form that is particularly suited to the study of random graphs. The expressions Δ and Δ^* defined in this section will appear often in these notes.

Let X be a nonnegative integral valued random variable and suppose we want to bound $\Pr[X = 0]$ given the value $\mu = E[X]$. If $\mu < 1$ we may use the inequality

$$\Pr[X > 0] \leq E[X] \tag{8}$$

so that if $E[X] \to 0$ then $X = 0$ almost always. (Here we are imagining an infinite sequence of X dependent on some parameter n going to infinity. This is the standard situation with the random graph $G(n, p(n))$.) But now suppose $E[X] \to \infty$. It does *not* necessarily follow that $X > 0$ almost always. For example, let X be the number of deaths due to nuclear war in the twelve months after reading this paragraph. Calculation of $E[X]$ can make for lively debate but few would deny that it is quite large. Yet we may believe - or hope - that $Pr[X \neq 0]$ is very close to zero. We can sometimes deduce $X > 0$ almost always if we have further information about $Var[X]$.

Theorem 1.1.

$$\Pr[X = 0] \leq \frac{Var[X]}{E[X]^2} \tag{9}$$

Proof. Set $\lambda = \mu/\sigma$ in Chebyschev's Inequality. Then

$$\Pr[X = 0] \leq \Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2} = \frac{\sigma^2}{\mu^2} \tag{10}$$

□

We generally apply this result in asymptotic terms.

Corollary 1.2. *If* $Var[X] = o(E[X]^2)$ *then* $X > 0$ *a.a.*

The proof of Theorem 1.3 actually gives that for any $\epsilon > 0$

$$\Pr[|X - E[X]| \geq \epsilon E[X]] \leq \frac{Var[X]}{\epsilon^2 E[X]^2} \tag{11}$$

and thus in asymptotic terms we actually have the following stronger assertion:

Corollary 1.3. *If* $Var[X] = o(E[X]^2)$ *then* $X \sim E[X]$ *a.a.*

Suppose again $X = X_1 + \ldots + X_m$ where X_i is the indicator random variable for event A_i. For indices i, j write $i \sim j$ if $i \neq j$ and the events A_i, A_j are not independent. We set (the sum over ordered pairs)

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j] \tag{12}$$

Note that when $i \sim j$

$$Cov[X_i, X_j] = E[X_i X_j] - E[X_i]E[X_j] \leq E[X_i X_j] = \Pr[A_i \wedge A_j] \tag{13}$$

and that when $i \neq j$ and not $i \sim j$ then $Cov[X_i, X_j] = 0$. Thus

$$Var[X] \leq E[X] + \Delta \tag{14}$$

Corollary 1.4. *If* $E[X] \to \infty$ *and* $\Delta = o(E[X]^2)$ *then* $X > 0$ *almost always. Furthermore* $X \sim E[X]$ *almost always.*

Let us say $X_1, \ldots, X_m$ are *symmetric* if for every $i \neq j$ there is an automorphism of the underlying probability space that sends event A_i to event A_j. Examples will appear in the next section. In this instance we write

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j] = \sum_i \Pr[A_i] \sum_{j \sim i} \Pr[A_j | A_i] \tag{15}$$

and note that the inner summation is independent of i. We set

$$\Delta^* = \sum_{j \sim i} \Pr[A_j | A_i] \tag{16}$$

where i is any fixed index. Then

$$\Delta = \sum_i \Pr[A_i] \Delta^* = \Delta^* \sum_i \Pr[A_i] = \Delta^* E[X] \tag{17}$$

Corollary 1.5. *If* $E[X] \to \infty$ *and* $\Delta^* = o(E[X])$ *then* $X > 0$ *almost always. Furthermore* $X \sim E[X]$ *almost always.*

The condition of Corollary 1.4 has the intuitive sense that conditioning on any specific A_i holding does not substantially increase the expected number $E[X]$ of events holding.

1.4 Appearance of Small Subgraphs

What is the threshold function for the appearance of a given graph H. This problem was solved in the original papers of Erdős and Rényi. We begin with an instructive special case.

Theorem 1.6. *The property $\omega(G) \geq 4$ has threshold function $n^{-2/3}$.*

Proof. For every 4-set S of vertices in $G(n,p)$ let A_S be the event "S is a clique" and X_S its indicator random variable. Then

$$E[X_S] = \Pr[A_S] = p^6 \tag{18}$$

as six different edges must all lie in $G(n,p)$. Set

$$X = \sum_{|S|=4} X_S \tag{19}$$

so that X is the number of 4-cliques in G and $\omega(G) \geq 4$ if and only if $X > 0$. Linearity of Expectation gives

$$E[X] = \sum_{|S|=4} E[X_S] = \binom{n}{4} p^6 \sim \frac{n^4 p^6}{24} \tag{20}$$

When $p(n) \ll n^{-2/3}$, $E[X] = o(1)$ and so $X = 0$ almost surely.

Now suppose $p(n) \gg n^{-2/3}$ so that $E[X] \to \infty$ and consider the Δ^* of Corollary 1.5. (All 4-sets "look the same" so that the X_S are symmetric.) Here $S \sim T$ if and only if $S \neq T$ and S, T have common edges - i.e., if and only if $|S \cap T| = 2$ or 3. Fix S. There are $O(n^2)$ sets T with $|S \cap T| = 2$ and for each of these $\Pr[A_T|A_S] = p^5$. There are $O(n)$ sets T with $|S \cap T| = 3$ and for each of these $\Pr[A_T|A_S] = p^3$. Thus

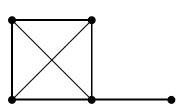
$$\Delta^* = O(n^2 p^5) + O(np^3) = o(n^4 p^6) = o(E[X])$$

since $p \gg n^{-2/3}$. Corollary 1.5 therefore applies and $X > 0$, i.e., there *does* exist a clique of size 4, almost always. □

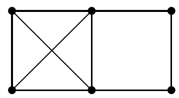
The proof of Theorem 1.6 appears to require a fortuitous calculation of Δ^*. The following definitions will allow for a description of when these calculations work out.

Definition 2. *Let H be a graph with v vertices and e edges. We call $\rho(H) = e/v$ the* density *of H. We call H* balanced *if every subgraph H' has $\rho(H') \leq \rho(H)$. We call H* strictly balanced *if every proper subgraph H' has $\rho(H') < \rho(H)$.*

`Examples.` K_4 and, in general, K_k are strictly balanced. The graph



is not balanced as it has density $7/5$ while the subgraph K_4 has density $3/2$. The graph



is balanced but not strictly balanced as it and its subgraph K_4 have density 3/2.

Theorem 1.7. *Let H be a balanced graph with v vertices and e edges. Let $A(G)$ be the event that H is a subgraph (not necessarily induced) of G. Then $p = n^{-v/e}$ is the threshold function for A.*

Proof. We follow the argument of Theorem 1.6. For each v-set S let A_S be the event that $G|_S$ contains H as a subgraph. Then

$$p^e \leq \Pr[A_S] \leq v!p^e \tag{21}$$

(Any particular placement of H has probability p^e of occuring and there are at most $v!$ possible placements. The precise calculation of $\Pr[A_S]$ is, in general, complicated due to the overlapping of potential copies of H.) Let X_S be the indicator random variable for A_S and

$$X = \sum_{|S|=v} X_S \tag{22}$$

so that A holds if and only if $X > 0$. Linearity of Expectation gives

$$E[X] = \sum_{|S|=v} E[X_S] = \binom{n}{v} \Pr[A_S] = \Theta(n^v p^e) \tag{23}$$

If $p \ll n^{-v/e}$ then $E[X] = o(1)$ so $X = 0$ almost always.

Now assume $p \gg n^{-v/e}$ so that $E[X] \to \infty$ and consider the Δ^* of Corollary 1.5. (All v-sets look the same so the X_S are symmetric.) Here $S \sim T$ if and only if $S \neq T$ and S, T have common edges - i.e., if and only if $|S \cap T| = i$ with $2 \leq i \leq v-1$. Let S be fixed. We split

$$\Delta^* = \sum_{T \sim S} \Pr[A_T|A_S] = \sum_{i=2}^{v-1} \sum_{|T \cap S|=i} \Pr[A_T|A_S] \tag{24}$$

For each i there are $O(n^{v-i})$ choices of T. Fix S, T and consider $\Pr[A_T|A_S]$. There are $O(1)$ possible copies of H on T. Each has - since, critically, H is balanced - at most $\frac{ie}{v}$ edges with both vertices in S and thus at least $e - \frac{ie}{v}$ other edges. Hence

$$\Pr[A_T|A_S] = O(p^{e-\frac{ie}{v}}) \tag{25}$$

and

$$\Delta^* = \sum_{i=2}^{v-1} O(n^{v-i} p^{e-\frac{ie}{v}}) = \sum_{i=2}^{v-1} O((n^v p^e)^{1-\frac{i}{v}}) \tag{26}$$

so that

$$\Delta^* == \sum_{i=2}^{v-1} o(n^v p^e) = o(E[X]) \tag{27}$$

as $n^v p^e \to \infty$. Hence Corollary 1.5 applies. □

Theorem 1.8. *In the notation of Theorem 1.7 if H is* not *balanced then $p = n^{-v/e}$ is* not *the threshold function for A.*

Proof. Let H_1 be a subgraph of H with v_1 vertices, e_1 edges and $e_1/v_1 > e/v$. Let α satisfy $v/e < \alpha < v_1/e_1$ and set $p = n^{-\alpha}$. The expected number of copies of H_1 is then $o(1)$ so almost always $G(n,p)$ contains no copy of H_1. But if it contains no copy of H_1 then it surely can contain no copy of H. □

The threshold function for the property of containing a copy of H, for general H, was examined in the original papers of Erdős and Rényi. Let H_1 be that subgraph with maximal density $\rho(H_1) = e_1/v_1$. (When H is balanced we may take $H_1 = H$.) They showed that $p = n^{-v_1/e_1}$ is the threshold function. This will follow fairly quickly from the methods of theorem 1.7.

We finish this section with two strengthenings of Theorem 1.7.

Theorem 1.9. *Let H be strictly balanced with v vertices, e edges and a automorphisms. Let X be the number of copies of H in $G(n,p)$. Assume $p \gg n^{-v/e}$. Then almost always*

$$X \sim \frac{n^v p^e}{a} \tag{28}$$

Proof. Label the vertices of H by $1, \ldots, v$. For each ordered $x_1, \ldots, x_v$ let $A_{x_1,\ldots,x_v}$ be the event that $x_1, \ldots, x_v$ provides a copy of H in that order. Specifically we define

$$A_{x_1,\ldots,x_v} : \{i,j\} \in E(H) \Rightarrow \{x_i, x_j\} \in E(G) \tag{29}$$

We let $I_{x_1,\ldots,x_v}$ be the corresponding indicator random variable. We define an equivalence class on v-tuples by setting $(x_1, \ldots, x_v) \equiv (y_1, \ldots, y_v)$ if there is an automorphism σ of $V(H)$ so that $y_{\sigma(i)} = x_i$ for $1 \leq i \leq v$.Then

$$X = \sum I_{x_1,\ldots,x_v} \tag{30}$$

gives the number of copies of H in G where the sum is taken over one entry from each equivalence class. As there are $(n)_v/a$ terms

$$E[X] = \frac{(n)_v}{a} E[I_{x_1,\ldots,x_v}] = \frac{(n)_v p^e}{a} \sim \frac{n^v p^e}{a} \tag{31}$$

Our assumption $p \gg n^{-v/e}$ implies $E[X] \to \infty$. It suffices therefore to show $\Delta^* = o(E[X])$. Fixing $x_1, \ldots, x_v$,

$$\Delta^* = \sum_{(y_1,\ldots,y_v)\sim(x_1,\ldots,x_v)} \Pr[A_{(y_1,\ldots,y_v)} | A_{(x_1,\ldots,x_v)}] \tag{32}$$

There are $v!/a = O(1)$ terms with $\{y_1, \ldots, y_v\} = \{x_1, \ldots, x_v\}$ and for each the conditional probability is at most one (actually, at most p), thus contributing $O(1) = o(E[X])$ to Δ^*. When $\{y_1, \ldots, y_v\} \cap \{x_1, \ldots, x_v\}$ has i elements, $2 \leq i \leq v-1$ the argument of Theorem 1.7 gives that the contribution to Δ^* is $o(E[X])$. Altogether $\Delta^* = o(E[X])$ and we apply Corollary 1.5 □

Theorem 1.10. *Let H be* any *fixed graph. For every subgraph H' of H (including H itself) let $X_{H'}$ denote the number of copies of H' in $G(n,p)$. Assume p is such that $E[X_{H'}] \to \infty$ for every H'. Then*

$$X_H \sim E[X_H] \tag{33}$$

almost always.

Proof. Let H have v vertices and e edges. As in Theorem 4.4 it suffices to show $\Delta^* = o(E[X])$. We split Δ^* into a finite number of terms. For each H' with w vertice and f edges we have those $(y_1, \ldots, y_v)$ that overlap with the fixed $(x_1, \ldots, x_v)$ in a copy of H'. These terms contribute, up to constants,

$$n^{v-w}p^{e-f} = \Theta\left(\frac{E[X_H]}{E[X_{H'}]}\right) = o(E[X_H])$$

to Δ^*. Hence Corollary 1.5 does apply. □

1.5 Connectivity

In this section we give a relatively simple example of what we call the Poisson Paradigm: the rough notion that if there are many rare and nearly independent events then the number of events that hold has approximately a Poisson distribution. This will yield one of the most beautiful of the Erdős- Rényi results, a quite precise description of the threshold behavior for connectivity. A vertex $v \in G$ is *isolated* if it is adjacent to no $w \in V$. In $G(n,p)$ let X be the number of isolated vertices.

Theorem 1.11. *Let $p = p(n)$ satisfy $n(1-p)^{n-1} = \mu$. Then*

$$\lim_{n\to\infty} \Pr[X = 0] = e^{-\mu} \tag{34}$$

Proof. We let X_i be the indicator random variable for vertex i being isolated so that $X = X_1 + \ldots + X_n$. Then $E[X_i] = (1-p)^{n-1}$ so by linearity of expectation $E[X] = \mu$. Now consider the r-th factorial moment $E[(X)_r]$ ($(X)_r := \prod i = 0^{r-1}(X-i)$) for any fixed r. By the symmetry $E[(X)_r] = (n)_r E[X_1 \cdots X_r]$. For vertices $1, \ldots, r$ to all be isolated the $r(n-1) - \binom{r}{2}$ pairs $\{i, x\}$ overlapping $1, \ldots, r$ must all not be edges. Thus

$$E[(X)_r] = (n)_r (1-p)^{r(n-1)-\binom{r}{2}} \sim n^r (1-p)^{r(n-1)} \sim \mu^r \tag{35}$$

(That is, the dependence among the X_i was asymptotically negligible.) All the moments of X approach those of $P(\mu)$. This implies (a nonobvious result in probability theory) that X approaches $P(\mu)$ in distribution. □

Now we give the Erdős-Rényi famous "double exponential" result.

Theorem 1.12. *Let*

$$p = p(n) = \frac{\log n}{n} + \frac{c}{n} + o(\frac{1}{n}) \tag{36}$$

Then

$$\lim_{n\to\infty} \Pr[G(n,p) \text{ is connected}] = e^{-e^{-c}} \tag{37}$$

Proof. For such p, $n(1-p)^{n-1} \sim \mu = e^{-c}$ and by the above argument the probability that X has no isolated vertices approaches $e^{-\mu}$. If G has no isolated vertices but is not connected there is a component of k vertices for some $2 \leq k \leq \frac{n}{2}$. Letting B be this event

$$\Pr[B] \leq \sum_{k=2}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-1)-\binom{k}{2}} \tag{38}$$

The first factor is the choice of a component set $S \subset V(G)$. The second factor is a choice of tree on S. The third factor is the probability that those tree pairs are in $E(G)$. The final factor is that there be no edge from S to $V(G)-S$. Some calculation (which we omit but note that $k=2$ provides the main term) gives that $\Pr[B] = o(1)$ so that $X \neq 0$ and connectivity have the same limiting probability. □

1.6 The Janson Inequalities

In many instances we would like to bound the probability that none of a set of bad events $B_i, i \in I$ occur. If the events are mutually independent then

$$\Pr[\wedge_{i \in I} \overline{B_i}] = \prod_{i \in I} \Pr[\overline{B_i}] \tag{39}$$

When the B_i are "mostly" independent the Janson Inequalities allow us, sometimes, to say that these two quantities are "nearly" equal.

Let Ω be a finite universal set and let R be a random subset of Ω given by

$$\Pr[r \in R] = p_r, \tag{40}$$

these events mutually independent over $r \in \Omega$. (In application to $G(n,p)$, Ω is the set of pairs $\{i,j\}$, $i,j \in V(G)$ and all $p_r = p$ so that R is the edge set of $G(n,p)$.) Let $A_i, i \in I$, be subsets of Ω, I a finite index set. Let B_i be the *event* $A_i \subseteq R$. (That is, each point $r \in \Omega$ " flips a coin" to determine if it is in R. B_i is the event that the coins for all $r \in A_i$ came up "heads".) Let X_i be the indicator random variable for B_i and $X = \sum_{i \in I} X_i$ the number of $A_i \subseteq R$. The event $\wedge_{i \in I} \overline{B_i}$ and $X = 0$ are then identical. For $i, j \in I$ we write $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Note that when $i \neq j$ and not $i \sim j$ then B_i, B_j are independent events since they involve separate coin flips. Furthermore, and this plays a crucial role in the proofs, if $i \notin J \subset I$ and not $i \sim j$ for all $j \in J$ then B_i is mutually independent of $\{B_j | j \in J\}$, i.e., independent of any Boolean function of those B_j. This is because the coin flips on A_i and on $\cup_{j \in J} A_j$ are independent. We define

$$\Delta = \sum_{i \sim j} \Pr[B_i \wedge B_j] \tag{41}$$

Here the sum is over ordered pairs so that $\Delta/2$ gives the same sum over unordered pairs. (This will be the same Δ as defined earlier.) We set

$$M = \prod_{i \in I} \Pr[\overline{B_i}], \tag{42}$$

the value of $\Pr[\wedge_{i \in I} \overline{B_i}]$ if the B_i were independent.

Theorem 1.13 (The Janson Inequality). *Let $B_i, i \in I$, Δ, M be as above and assume all $\Pr[B_i] \le \epsilon$. Then*

$$M \le \Pr[\wedge_{i\in I}\overline{B_i}] \le Me^{\frac{1}{1-\epsilon}\frac{\Delta}{2}} \tag{43}$$

Now set

$$\mu = E[X] = \sum_{i\in I} \Pr[B_i] \tag{44}$$

For each $i \in I$

$$\Pr[\overline{B_i}] = 1 - \Pr[B_i] \le e^{-\Pr[B_i]} \tag{45}$$

so, multiplying over $i \in I$,

$$M \le e^{-\mu} \tag{46}$$

It is often more convenient to replace the upper bound of Theorem 1.13 with

$$\Pr[\wedge_{i\in I}\overline{B_i}] \le e^{-\mu+\frac{1}{1-\epsilon}\frac{\Delta}{2}} \tag{47}$$

As an example, set $p = cn^{-2/3}$ and consider the probability that $G(n,p)$ contains no K_4. The B_i then range over the $\binom{n}{4}$ potential K_4 - each being a 6-element subset of Ω. Here, as is often the case, $\epsilon = o(1)$, $\Delta = o(1)$ (as calculated previously) and μ approaches a constant, here $k = c^6/24$. In these instances $\Pr[\wedge_{i\in I}\overline{B_i}] \to e^{-k}$. Thus we have the fine structure of the threshold function of $\omega(G) = 4$.

As Δ becomes large the Janson Inequality becomes less precise. Indeed, when $\Delta \ge 2\mu(1-\epsilon)$ it gives an upper bound for the probability which is larger than one. At that point (and even somewhat before) the following result kicks in.

Theorem 1.14. *(Generalized Janson Inequality) Under the assumptions of Theorem 1.13 and the further assumption that $\Delta \ge \mu(1-\epsilon)$*

$$\Pr[\wedge_{i\in I}\overline{B_i}] \le e^{-\frac{\mu^2(1-\epsilon)}{2\Delta}} \tag{48}$$

The Generalized Janson Inequality (when it applies) often gives a much stronger result than Chebyschev's Inequality as used earlier. We can bound $Var[X] \le \mu + \Delta$ so that

$$\Pr[\wedge_{i\in I}\overline{B_i}] = \Pr[X = 0] \le \frac{Var[X]}{E[X]^2} \le \frac{\mu + \Delta}{\mu^2} \tag{49}$$

Suppose $\epsilon = o(1)$, $\mu \to \infty$, $\mu \ll \Delta$, and $\gamma = \frac{\mu^2}{\Delta} \to \infty$. Chebyschev's upper bound on $\Pr[X = 0]$ is then roughly γ^{-1} while Janson's upper bound is roughly $e^{-\gamma}$.

1.7 The Proofs

The original proofs of Janson are based on estimates of the Laplace transform of an appropriate random variable. The proof presented here follows that of Boppana and Spencer [1989]. We shall use the inequalities

$$\Pr[B_i | \wedge_{j\in J} \overline{B_j}] \le \Pr[B_i] \tag{50}$$

valid for all index sets $J \subset I, i \notin J$ and

$$\Pr[B_i | B_k \wedge \bigwedge_{j \in J} \overline{B_j}] \leq \Pr[B_i | B_k] \tag{51}$$

valid for all index sets $J \subset I, i, k \notin J$. The first follows from general Correlation Inequalities. The second is equivalent to the first since conditioning on B_k is the same as assuming $p_r = \Pr[r \in R] = 1$ for all $r \in A_k$. We note that Janson's Inequality actually applies to any set of events B_i and relation $\sim$ for which (50,51) apply.

Proof. (Thm. 1.13) The lower bound follows immediately. Order the index set $I = \{1, \ldots, m\}$ for convenience. For $1 \leq i \leq m$

$$Pr[B_i | \wedge_{1 \leq j < i} \overline{B_j}] \leq \Pr[B_i] \tag{52}$$

so

$$Pr[\overline{B_i} | \wedge_{1 \leq j < i} \overline{B_j}] \geq Pr[\overline{B_i}] \tag{53}$$

and

$$\Pr[\wedge_{i \in I} \overline{B_i}] = \prod_{i=1}^{m} Pr[\overline{B_i} | \wedge_{1 \leq j < i} \overline{B_j}] \geq \prod_{i=1}^{m} Pr[\overline{B_i}] \tag{54}$$

Now the upper bound. For a given i renumber, for convenience, so that $i \sim j$ for $1 \leq j \leq d$ and not for $d+1 \leq j < i$. We use the inequality $\Pr[A|B \wedge C] \geq \Pr[A \wedge B|C]$, valid for any A, B, C. With $A = B_i$, $B = \overline{B_1} \wedge \ldots \wedge \overline{B_d}$, $C = \overline{B_{d+1}} \wedge \ldots \wedge \overline{B_{i-1}}$

$$Pr[B_i | \wedge_{1 \leq j < i} \overline{B_j}] = \Pr[A|B \wedge C] \geq Pr[A \wedge B|C] = Pr[A|C] Pr[B|A \wedge C] \tag{55}$$

From the mutual independence $Pr[A|C] = Pr[A]$. We bound

$$\Pr[B|A \wedge C] \geq 1 - \sum_{j=1}^{d} Pr[B_j | B_i \wedge C] \geq 1 - \sum_{j=1}^{d} Pr[B_j | B_i] \tag{56}$$

from the Correlation Inequality. Thus

$$Pr[B_i | \wedge_{1 \leq j < i} \overline{B_j}] \geq \Pr[B_i] - \sum_{j=1}^{d} Pr[B_j \wedge B_i] \tag{57}$$

Reversing

$$Pr[\overline{B_i} | \wedge_{1 \leq j < i} \overline{B_j}] \leq \Pr[\overline{B_i}] + \sum_{j=1}^{d} Pr[B_j \wedge B_i] \tag{58}$$

$$\leq \Pr[\overline{B_i}] \left(1 + \frac{1}{1-\epsilon} \sum_{j=1}^{d} Pr[B_j \wedge B_i] \right) \tag{59}$$

since $\Pr[\overline{B_i}] \geq 1 - \epsilon$. Employing the inequality $1 + x \leq e^x$,

$$Pr[\overline{B_i} | \wedge_{1 \leq j < i} \overline{B_j}] \leq \Pr[\overline{B_i}] e^{\frac{1}{1-\epsilon} \sum_{j=1}^{d} Pr[B_j \wedge B_i]} \tag{60}$$

For each $1 \le i \le m$ we plug this inequality into

$$\Pr[\wedge_{i\in I}\overline{B_i}] = \prod_{i=1}^{m} Pr[\overline{B_i} | \wedge_{1\le j<i} \overline{B_j}] \tag{61}$$

The terms $\Pr[\overline{B_i}]$ multiply to M. The exponents add: for each $i, j \in I$ with $j < i$ and $j \sim i$ the term $Pr[B_j \wedge B_i]$ appears once so they add to $\Delta/2$. □

Proof. (Theorem 1.14) As discussed earlier, the proof of Theorem 1.13 gives

$$\Pr[\wedge_{i\in I}\overline{B_i}] \le e^{-\mu + \frac{1}{1-\epsilon}\frac{\Delta}{2}} \tag{62}$$

which we rewrite as

$$-\ln[\Pr[\wedge_{i\in I}\overline{B_i}]] \ge \sum_{i\in I} \Pr[B_i] - \frac{1}{2(1-\epsilon)} \sum_{i\sim j} \Pr[B_i \wedge B_j] \tag{63}$$

For any set of indices $S \subset I$ the same inequality applied only to the $B_i, i \in S$ gives

$$-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]] \ge \sum_{i\in S} \Pr[B_i] - \frac{1}{2(1-\epsilon)} \sum_{i,j\in S, i\sim j} \Pr[B_i \wedge B_j] \tag{64}$$

Let now S be a random subset of I given by

$$\Pr[i \in S] = p \tag{65}$$

with p a constant to be determined, the events mutually independent. (Here we are using probabilistic methods to prove a probability theorem!)

$$E\left[-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]]\right] \ge E\left[\sum_{i\in S} \Pr[B_i]\right] - \frac{1}{2(1-\epsilon)} E\left[\sum_{i,j\in S, i\sim j} \Pr[B_i \wedge B_j]\right] \tag{66}$$

Each term $\Pr[B_i]$ then appears with probability p and each term $\Pr[B_i \wedge B_j]$ with probability p^2 so that

$$E\left[-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]]\right] \ge p\mu - \frac{1}{1-\epsilon} p^2 \frac{\Delta}{2} \tag{67}$$

We set

$$p = \frac{\mu(1-\epsilon)}{\Delta} \tag{68}$$

so as to maximize this quantity. The added assumption of Theorem 1.14 assures us that the probability p is at most one. Then

$$E\left[-\ln[\Pr[\wedge_{i\in S}\overline{B_i}]]\right] \ge \frac{\mu^2(1-\epsilon)}{2\Delta} \tag{69}$$

Therefore there is a specific $S \subset I$ for which

$$-\ln[\Pr[\wedge_{i\in S}\overline{B_i}] \ge \frac{\mu^2(1-\epsilon)}{2\Delta} \tag{70}$$

That is,

$$\Pr[\wedge_{i\in S}\overline{B_i}] \le e^{-\frac{\mu^2(1-\epsilon)}{2\Delta}} \tag{71}$$

But

$$\Pr[\wedge_{i\in I}\overline{B_i}] \le \Pr[\wedge_{i\in S}\overline{B_i}] \tag{72}$$

completing the proof. □

1.8 Appearance of Small Subgraphs Revisited

Generalizing the fine threshold behavior for the appearance of K_4 we find the fine threshold behavior for the appearance of any strictly balanced graph H.

Theorem 1.15. *Let H be a* strictly *balanced graph with v vertices, e edges and a automorphisms. Let $c > 0$ be arbitrary. Let A be the property that G contains no copy of H. Then with $p = cn^{-v/e}$,*

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = exp[-c^e/a] \tag{73}$$

Proof. Let $A_\alpha, 1 \le \alpha \le \binom{n}{v}v!/a$, range over the edge sets of possible copies of H and let B_α be the event $G(n,p) \supseteq A_\alpha$. We apply Janson's Inequality. As

$$\lim_{n\to\infty} \mu = lim_{n\to\infty}\binom{n}{v}v!p^e/a = c^e/a \tag{74}$$

we find

$$\lim_{n\to\infty} M = exp[-c^e/a] \tag{75}$$

Now we examine (similar to Theorem 1.7)

$$\Delta = \sum_{\alpha\sim\beta} \Pr[B_\alpha \wedge B_\beta] \tag{76}$$

We split the sum according to the number of *vertices* in the intersection of copies α and β. Suppose they intersect in j vertices. If $j = 0$ or $j = 1$ then $A_\alpha \cap A_\beta = \emptyset$ so that $\alpha \sim \beta$ cannot occur. For $2 \le j \le v$ let f_j be the maximal $|A_\alpha \cap A_\beta|$ where $\alpha \sim \beta$ and α, β intersect in j vertices. As $\alpha \ne \beta$, $f_v < e$. When $2 \le j \le v-1$ the critical observation is that $A_\alpha \cap A_\beta$ is a subgraph of H and hence, as H is strictly balanced,

$$\frac{f_j}{j} < \frac{e}{v} \tag{77}$$

There are $O(n^{2v-j})$ choices of α, β intersecting in j points since α, β are determined, except for order, by $2v - j$ points. For each such α, β

$$\Pr[B_\alpha \wedge B_\beta] = p^{|A_\alpha \cup A_\beta|} = p^{2e-|A_\alpha\cap A_\beta|} \le p^{2e-f_j} \tag{78}$$

Thus

$$\Delta = \sum_{j=2}^{v} O(n^{2v-j})O(n^{-\frac{v}{e}(2e-f_j)}) \tag{79}$$

But

$$2v - j - \frac{v}{e}(2e - f_j) = \frac{vf_j}{e} - j < 0 \tag{80}$$

so each term is $o(1)$ and hence $\Delta = o(1)$. By Janson's Inequality

$$\lim_{n\to\infty} \Pr[\wedge \overline{B}_\alpha] = \lim_{n\to\infty} M = exp[-c^e/a] \tag{81}$$

completing the proof. □

The fine threshold behavior for the appearance of an arbitrary graph H has been worked out but it can get quite complicated.

1.9 Some Very Low Probabilities

Let A be the property that G does not contain K_4 and consider $\Pr[G(n,p) \models A]$ as p varies. (Results with K_4 replaced by an arbitrary H are discussed at the end of this section.) We know that $p = n^{-2/3}$ is a threshold function so that for $p \gg n^{-2/3}$ this probability is $o(1)$. Here we want to estimate that probability. Our estimates here will be quite rough, only up to a $o(1)$ additive factor in the hyperexponent, though with more care the bounds differ by "only" a constant factor in the exponent. If we were to consider all potential K_4 as giving mutually independent events then we would be led to the estimate $(1-p^6)^{\binom{n}{4}} = e^{-n^{4+o(1)}p^6}$. For p appropriately small this turns out to be correct. But for, say, $p = \frac{1}{2}$ it would give the estimate $e^{-n^{4+o(1)}}$. This must, however, be way off the mark since with probability $2^{-\binom{n}{2}} = e^{-n^{2+o(1)}}$ the graph G could be empty and hence trivially satisfy A.

Rather than giving the full generality we assume $p = n^{-\alpha}$ with $\frac{2}{3} > \alpha > 0$. The result is:

$$\Pr[G(n,p) \models A] = e^{-n^{4-6\alpha+o(1)}} \tag{82}$$

for $\frac{2}{3} > \alpha \geq \frac{2}{5}$ and

$$\Pr[G(n,p) \models A] = e^{-n^{2-\alpha+o(1)}} \tag{83}$$

for $\frac{2}{5} \geq \alpha > 0$.

The upper bound follows from the inequality

$$\Pr[G(n,p) \models A] \geq \max\left[(1-p^6)^{\binom{n}{4}}, (1-p)^{\binom{n}{2}}\right] \tag{84}$$

This is actually two inequalities. The first comes from the probability of G not containing a K_4 being at most the probability as if all the potential K_4 were independent. The second is the same bound on the probability that G doesn't contain a K_2 - i.e., that G has no edges. Calculation shows that the "turnover" point for the two inequalities occurs when $p = n^{-2/5+o(1)}$.

The upper bound follows from the Janson inequalities. For each four set α of vertices B_α is that that 4-set gives a K_4 and we want $\Pr[\wedge \overline{B_\alpha}]$. We have $\mu = \Theta(n^4p^6)$ and $-\ln M \sim \mu$ and (as shown earlier) $\Delta = \Theta(\mu\Delta^*)$ with $\Delta^* = \Theta(n^2p^5 + np^3)$. With $p = n^{-\alpha}$ and $\frac{2}{3} > \alpha > \frac{2}{5}$ we have $\Delta^* = o(1)$ so that

$$\Pr[\wedge \overline{B_\alpha}] \leq e^{-\mu(1+o(1))} = e^{-n^{4-6\alpha+o(1)}}$$

When $\frac{2}{5} > \alpha > 0$ then $\Delta^* = \Theta(n^2p^5)$ (somewhat surprisingly the np^3 never is significant in these calculations) and the extended Janson inequality gives

$$\Pr[\wedge \overline{B_\alpha}] \leq e^{-\Theta(\mu^2/\Delta)} = e^{-\Theta(\mu/\Delta^*)} = e^{-n^{2-\alpha}}$$

The general result has been found by T. Łuczak, A. Rucinski and S. Janson (1990). Let H be any fixed graph and let A be the property of not containing a copy of H. For any subgraph H' of H the correlation inequality gives

$$\Pr[G(n,p) \models A] \leq e^{-E[X_{H'}]} \tag{85}$$

where $X_{H'}$ is the number of copies of H' in G. Now let $p = n^{-\alpha}$ where we restrict to those α for which p is past the threshold function for the appearance of H. Then

$$\Pr[G(n,p) \models A] = e^{n^{o(1)}} \min_{H'} e^{-E[X_{H'}]} \tag{86}$$

2 Lecture II: The Phase Transition

2.1 Branching Processes

Paul Erdős and Alfred Rényi, in their original 1960 paper, discovered that the random graph $G(n,p)$ undergoes a remarkable change at $p = 1/n$. Speaking roughly, let first $p = c/n$ with $c < 1$. Then $G(n,p)$ will consist of small components, the largest of which is of size $\Theta(\ln n)$. But now suppose $p = c/n$ with $c > 1$. In that short amount of "time" many of the components will have joined together to form a "giant component" of size $\Theta(n)$. The remaining vertices are still in small components, the largest of which has size $\Theta(\ln n)$. They dubbed this phenomenon the *Double Jump*. We prefer the descriptive term Phase Transition because of the connections to percolation.

To better understand the Phase Transition we make a lengthy detour into the subject of Branching Processes. Imagine that we are in a unisexual universe and we start with a single organism. Imagine that this organism has a number of children given by a given random variable Z. (For us, Z will be Poisson with mean c.) These children then themselves have children, the number again being determined by Z. These grandchildren then have children, etc. As $Z = 0$ will have nonzero probability there will be some chance that the line dies out entirely. We want to study the total number of organisms in this process, with particular eye to whether or not the process continues forever.

Now lets be more precise. Let $Z_1, Z_2, \ldots$ be independent random variables, each with distribution Z. Define $Y_0, Y_1, \ldots$ by initial value $Y_0 = 1$ and the recursion

$$Y_i = Y_{i-1} + Z_i - 1 \tag{87}$$

and let T be the least t for which $Y_t = 0$. If no such t exists (the line continuing forever) we say $T = +\infty$. The Y_i and Z_i mirror the Branching Process as follows. We view all organisms as living or dead. Initially there is one live organism and no dead ones. At each time unit we select one of the live organisms, it has Z_i children, and then it dies. The number Y_i of live organisms at time i is then given by the recursion. The process stops when $Y_t = 0$ (extinction) but it is a convenient fiction to define the recursion for all t. Note that T is not affected by this fiction since once $Y_t = 0$, T has been defined. T (whether finite or infinite) is the total number of organisms, including the original, in this process. (A natural approach, found in many probability texts, is to have all organisms of a given generation have their children at once and study the number of children of each generation. While we may think of the organisms giving birth by generation it will not affect our model.)

We shall use the major result of Branching Processes that when $E[Z] = c < 1$ with probability one the process dies out ($T < \infty$) but when $E[Z] = c > 1$ then there is a nonzero probability that the process goes on forever ($T = \infty$).

When a branching process dies we call $H = (Z_1, \ldots, Z_T)$ the *history* of the process. A sequence $(z_1, \ldots, z_t)$ is a possible history if and only if the sequence y_i given by $y_0 = 1, y_i = y_{i-1} + z_i - 1$ has $y_i > 0$ for $0 \leq i < t$ and $y_t = 0$. When Z is Poisson with mean λ

$$\Pr[H = (z_1, \ldots, z_t)] = \prod_{i-1}^{t} \frac{e^{-\lambda}\lambda^{z_i}}{z_i!} = \frac{e^{-\lambda}(\lambda e^{-\lambda})^{t-1}}{\prod_{i=1}^{t} z_i!} \tag{88}$$

since $z_1 + \ldots + z_t = t - 1$.

Definition 3. *We call* $d < 1 < c$ *a* conjugate pair *if*

$$de^{-d} = ce^{-c} \tag{89}$$

The function $f(x) = xe^{-x}$ increases from 0 to e^{-1} in [0,1) and decreases back to 0 in $(1, \infty)$ so that all $c \neq 1$ have a uniqe conjugate. Let $c > 1$ and $y = \Pr[T < \infty]$ so that $y = e^{c(y-1)}$. Then $(cy)e^{-cy} = ce^{-c}$ so

$$d = cy \tag{90}$$

For future use we note that if $c = 1 + \epsilon$ then its dual d satisfies

$$d = 1 - \epsilon + O(\epsilon^2) \tag{91}$$

as $\epsilon \to 0^+$.

Theorem 2.1. *(Duality Principle) Let* $d < 1 < c$ *be conjugates. The Branching Process with mean* c*, conditional on extinction, has the same distribution as the Branching Process with mean* d*.*

Proof. It suffices to show that for every history $H = (z_1, \ldots, z_t)$

$$\frac{e^{-c}(ce^{-c})^{t-1}}{y\prod_{i=1}^{t} z_i!} = \frac{e^{-d}(de^{-d})^{t-1}}{\prod_{i=1}^{t} z_i!} \tag{92}$$

This is immediate as $ce^{-c} = de^{-d}$ and $ye^{-d} = ye^{-cy} = e^{-c}$. ☐

2.2 The Giant Component

Now let's return to random graphs. We define a procedure to find the component $C(v)$ containing a given vertex v in a given graph G. We are motivated by Karp [1990] in which this approach is applied to random digraphs. In this procedure vertices will be live, dead or neutral. Originally v is live and all other vertices are neutral, time $t = 0$ and $Y_0 = 1$. Each time unit t we take a live vertex w and check all pairs $\{w, w'\}$, w' neutral, for membership in G. If $\{w, w'\} \in G$ we make w' live, otherwise it stays neutral. After searching all neutral w' we set w dead and let Y_t equal the new number of live vertices. When there are no live vertices the process terminates and $C(v)$ is the set of dead vertices. Let Z_t be the number of w' with $\{w, w'\} \in G$ so that $Y_0 = 1$ and

$$Y_t = Y_{t-1} + Z_t - 1 \tag{93}$$

With $G = G(n, p)$ each neutral w' has independent probability p of becoming live. Here, critically, no pair $\{w, w'\}$ is ever examined twice so that the conditional probability for $\{w, w'\} \in G$ is always p. As $t - 1$ vertices are dead and Y_{t-1} are live

$$Z_t \sim B[n - (t - 1) - Y_{t-1}, p] \tag{94}$$

Let T be the least t for which $Y_t = 0$. Then $T = |C(v)|$. As in Section 2.1 we continue the recursive definition of Y_t, this time for $0 \le t \le n$.

Theorem 2.2. *For all* t

$$Y_t \sim B[n - 1, 1 - (1 - p)^t] + 1 - t \tag{95}$$

Proof. It is more convenient to deal with

$$N_t = n - t - Y_t \tag{96}$$

the number of neutral vertices at time t and show, equivalently,

$$N_t \sim B[n - 1, (1 - p)^t] \tag{97}$$

This is reasonable since each $w \neq v$ has independent probability $(1-p)^t$ of staying neutral t times. Formally, as $N_0 = n - 1$ and

$$\begin{aligned} N_t \quad = n - t - Y_t \quad &= n - t - B[n - (t - 1) - Y_{t-1}, p] - Y_{t-1} + 1 \\ &= N_{t-1} - B[N_{t-1}, p] \\ &= B[N_{t-1}, 1 - p] \end{aligned}$$

the result follows by induction. □

We set $p = c/n$. When t and Y_{t-1} are small we may approximate Z_t by $B[n, c/n]$ which is approximately Poisson with mean c. Basically small components will have size distribution as in the Branching Process. The analogy must break down for $c > 1$ as the Branching Process may have an infinite population whereas $|C(v)|$ is surely at most n.

Essentially, those v for which the Branching Process for $C(v)$ does not "die early" all join together to form the giant component.

Fix c. Let $Y_0^*, Y_1^*, \ldots, T^*, Z_1^*, Z_2^*, \ldots, H^*$ refer to the Branching Process and let the unstarred $Y_0, Y_1, \ldots, T, Z_1, Z_2, \ldots, H$ refer to the Random Graph process. For any possible history $(z_1, \ldots, z_t)$

$$\Pr[H^* = (z_1, \ldots, z_t)] = \prod_{i=1}^{t} \Pr[Z^* = z_i] \tag{98}$$

where Z^* is Poisson with mean c while

$$\Pr[H = (z_1, \ldots, z_t)] = \prod_{i=1}^{t} \Pr[Z_i = z_i] \tag{99}$$

where Z_i has Binomial Distribution $B[n-1-z_1-\ldots-z_{i-1}, c/n]$. The Poisson distribution is the limiting distribution of Binomials. When $m = m(n) \sim n$ and c, i are fixed

$$\lim_{n\to\infty} \Pr[B[m, c/n] = i] = \lim_{n\to\infty} \binom{m}{z} (\frac{c}{n})^z (1 - \frac{c}{n})^{m-z} = e^{-c}c^z/z! \tag{100}$$

hence

$$\lim_{n\to\infty} \Pr[H = (z_1, \ldots, z_t)] = \Pr[H^* = (z_1, \ldots, z_t)] \tag{101}$$

Assume $c < 1$. For any fixed t, $\lim_{n\to\infty} \Pr[T = t] = \Pr[T^* = t]$. We now bound the size of the largest component. For any t

$$\Pr[T > t] \leq \Pr[Y_t > 0] = \Pr[B[n-1, 1-(1-p)^t] \geq t] \leq \Pr[B[n, tc/n] \geq t] \tag{102}$$

as $1 - (1-p)^t \leq tp$ and $n - 1 < n$. By Large Deviation Results

$$\Pr[T > t] < e^{-\alpha t} \tag{103}$$

where $\alpha = \alpha(c) > 0$. Let $\beta = \beta(c)$ satisfy $\alpha\beta > 1$. Then

$$\Pr[T > \beta \ln n] < n^{-\alpha\beta} = o(n^{-1}) \tag{104}$$

There are n choices for initial vertex v. Thus almost always *all* components have size $O(\ln n)$.

Now assume $c > 1$. For any fixed t, $\lim_{n\to\infty} \Pr[T = t] = \Pr[T^* = t]$ but what corresponds to $T^* = \infty$? For $t = o(n)$ we may estimate $1 - (1-p)^t \sim pt$ and $n - 1 \sim n$ so that

$$\Pr[Y_t \leq 0] = \Pr[B[n-1, 1-(1-p)^t] \leq t-1] \sim \Pr[B[n, tc/n] \leq t] \tag{105}$$

drops exponentially in t by Large Deviation results. When $t = \alpha n$ we estimate $1 - (1-p)^t$ by $1 - e^{-c\alpha}$. The equation $1 - e^{-c\alpha} = \alpha$ has solution $\alpha = 1 - y$ where y is the extinction probability. For $\alpha < 1 - y$, $1 - e^{-c\alpha} > \alpha$ and

$$\Pr[Y_t \leq 0] \sim \Pr[B[n, 1 - e^{-c\alpha}] \leq \alpha n] \tag{106}$$

is exponentially small while for $\alpha > 1-y$, $1-e^{-c\alpha} < \alpha$ and $\Pr[Y_t \le 0] \sim 1$. Thus almost always $Y_t = 0$ for some $t \sim (1-y)n$. Basically, $T^* = \infty$ corresponds to $T \sim (1-y)n$. Let $\epsilon, \delta > 0$ be arbitrarily small. With somewhat more care to the bounds we may show that there exists t_0 so that for n sufficiently large

$$\Pr[t_0 < T < (1-\delta)n(1-y) \text{ or } \mathrm{T > (1+\delta)n(1-y)}] < \epsilon \tag{107}$$

Pick t_0 sufficiently large so that

$$y - \epsilon \le \Pr[T^* \le t_0] \le y \tag{108}$$

Then as $\lim_{n\to\infty} \Pr[T \le t_0] = \Pr[T^* \le 0]$ for n sufficiently large

$$y - 2\epsilon \le \Pr[T \le t_0] \le y + \epsilon \tag{109}$$

$$1 - y - 2\epsilon \le \Pr[(1-\delta)n(1-y) < T < (1+\delta)n(1-y)] < 1 - y + 3\epsilon \tag{110}$$

Now we expand our procedure to find graph components. We start with $G \sim G(n,p)$, select $v = v_1 \in G$ and compute $C(v_1)$ as before. Then we delete $C(v_1)$, pick $v_2 \in G - C(v_1)$ and iterate. At each stage the remaining graph has distribution $G(m,p)$ where m is the number of vertices. (Note, critically, that no pairs $\{w, w'\}$ in the remaining graph have been examined and so it retains its distribution.) Call a component $C(v)$ small if $|C(v)| \le t_0$, giant if $(1-\delta)n(1-y) < |C(v)| < (1+\delta)n(1-y)$ and otherwise failure. Pick $s = s(\epsilon)$ with $(y+\epsilon)^s < \epsilon$. (For ϵ small $s \sim K \ln \epsilon^{-1}$.) Begin this procedure with the full graph and terminate it when either a giant component or a failure component is found or when s small components are found. At each stage, as only small components have thus far been found, the number of remaining points is $m = n - O(1) \sim n$ so the conditional probabilities of small, giant and failure remain asymptotically the same. The chance of ever hitting a failure component is thus $\le s\epsilon$ and the chance of hitting all small components is $\le (y+\epsilon)^s \le \epsilon$ so that with probability at least $1-\epsilon'$, where $\epsilon' = (s+1)\epsilon$ may be made arbitrarily small, we find a series of less than s small components followed by a giant component. The remaining graph has $m \sim yn$ points and $pm \sim cy = d$, the conjugate of c as defined earlier. As $d < 1$ the previous analysis gives the maximal components. In summary: almost always $G(n, c/n)$ has a giant component of size $\sim (1-y)n$ and all other components of size $O(\ln n)$. Furthermore, the Duality Principle has a discrete analog.

Theorem 2.3. *(Discrete Duality Principle) Let $d < 1 < c$ be conjugates. The structure of $G(n, c/n)$ with its giant component removed is basically that of $G(m, d/m)$ where m, the number of vertices not in the giant component, satisfies $m \sim ny$.*

Well, this is not a pricisely stated theorem – but the concept is quite compelling.

2.3 A Static View

The small components of $G(n, c/n)$ can also be examined from a static view. For a fixed k let X be the number of tree components of size k. Then

$$E[X] = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{k}{2}-(k-1)} \tag{111}$$

Here we use the nontrivial fact, due to Cayley, that there are k^{k-2} possible trees on a given k-set. For c, k fixed

$$E[X] \sim n\frac{e^{-ck}k^{k-2}c^{k-1}}{k!} \tag{112}$$

As trees are strictly balanced a second moment method gives $X \sim E[X]$ almost always. Thus $\sim p_k n$ points lie in tree components of size k where

$$p_k = \frac{e^{-ck}(ck)^{k-1}}{k!} \tag{113}$$

It can be shown analytically that $p_k = \Pr[T = k]$ in the Branching Process with mean c. Let Y_k denote the number of cycles of size k and Y the total number of cycles. Then

$$E[Y_k] = \frac{(n)_k}{2k}(\frac{c}{n})^k \sim \frac{c^k}{2k} \tag{114}$$

for fixed k. For $c < 1$

$$E[Y] = \sum E[Y_k] \to \sum_{k=1}^{\infty} \frac{c^k}{2k} \tag{115}$$

has a finite limit whereas for $c > 1$, $E[Y] \to \infty$. Even for $c > 1$ for any fixed k the number of k-cycles has a limiting expectation and so do not asymptotically affect the number of components of a given size.

References

N. Alon and J. Spencer, *The Probabilistic Method*, John Wiley

B. Bollobás, *Random Graphs*, Academic Press.

R. B. Boppana and J. H. Spencer (1989), A useful elementary correlation inequality, J. Combinatorial Theory Ser. A, *50*, 305-307.

P. Erdős and A. Rényi (1960), On the Evolution of Random Graphs, Mat Kutató Int. Közl. *5*, 17-60

S. Janson, T. Łuczak, A. Rucinski (1990), An Exponential Bound for the Probability of Nonexistence of a Specified Subgraph in a Random Graph, *in* Random Graphs '87 (M. Karonski, J. Jaworski, A. Rucinski, eds.), John Wiley, 73-87

R.M. Karp (1990), The transitive closure of a Random Digraph, Random Structures and Algorithms *1*, 73-94

E.M. Wright (1977), The number of connected sparsely edged graphs, Journal of Graph Theory *1*, 317-330.