Erdös-Renyi Random Graphs

Let $V = \{1, 2, \ldots, n\}$. For $1 \leq x < y \leq n$ let $\eta_{x,y}$ be independent $= 1$ with probability $p$ and $0$ otherwise. Let $\eta_{y,x} = \eta_{x,y}$. If $\eta_{x,y} = 1$ there is an edge from $x$ to $y$. Here we will be primarily concerned with situation $p = \lambda/n$ and in particular with showing that when $\lambda < 1$ most components are small, while for $\lambda > 1$ there is a giant component with $\sim g(\lambda)n$ vertices. The intuition behind this result is that a site has a $\text{Binomial}(n-1, \lambda/n)$ number of neighbors, which has mean $\approx \lambda$. Suppose we start with $I_0 = 1$, and for $n \geq 1$ let $I_n$ be the set of vertices not in $\cup_{m=0}^{n-1} I_m$ that are connected to some site in $I_{n-1}$ then when $n$ is not too large the number of points in $I_n$, $Z_n = |I_n|$ is branching process in which each individual in generation $n$ has an average of $\lambda$ children. The first section defines a branching process and gives its basic properties.

1 Branching Processes

Let $\xi_i^n$, $i, n \geq 0$, be i.i.d. nonnegative integer-valued random variables. Define a sequence $Z_n$, $n \geq 0$ by $Z_0 = 1$ and

$$Z_{n+1} = \begin{cases} \xi_1^{n+1} + \cdots + \xi_k^{n+1} & \text{if } Z_n > 0 \\ 0 & \text{if } Z_n = 0 \end{cases}$$

$Z_n$ is called a Galton-Watson process. The idea behind the definitions is that $Z_n$ is the number of people in the $n$th generation, and each member of the $n$th generation gives birth independently to an identically distributed number of children. $p_k = P(\xi_i^n = k)$ is called the offspring distribution.

**Lemma 1.1.** Let $F_n = \sigma(\xi_i^m : i \geq 1, 1 \leq m \leq n)$ and $\mu = E\xi_i^m \in (0, \infty)$. Then $Z_n/\mu^n$ is a martingale w.r.t. $F_n$.

**Proof.** Clearly, $Z_n \in F_n$. Recall, see Exercise 1.1 in Chapter 4 of Durrett (2004), that if $X = Y$ on $B \in F$ then $E(X|F) = E(Y|F)$ on $B$. On $\{Z_n = k\}$,

$$E(Z_{n+1}|F_n) = E(\xi_1^{n+1} + \cdots + \xi_k^{n+1}|F_n) = k\mu = \mu Z_n$$

Dividing both sides by $\mu^{n+1}$ now gives the desired result. \qed

**Theorem 1.2.** If $\mu < 1$ then $Z_n = 0$ for all $n$ sufficiently large.

**Proof.** $E(Z_n/\mu^n) = E(Z_0) = 1$, so $E(Z_n) = \mu^n$. Now $Z_n \geq 1$ on $\{Z_n > 0\}$ so

$$P(Z_n > 0) \leq E(Z_n; Z_n > 0) = E(Z_n) = \mu^n \to 0$$

exponentially fast if $\mu < 1$. \qed
The last answer should be intuitive. If each individual on the average gives birth to less than one child, the species will die out. The next result shows that after we exclude the trivial case in which each individual has exactly one child, the same result holds when $\mu = 1$.

**Theorem 1.3.** If $\mu = 1$ and $P(\xi_i^m = 1) < 1$ then $Z_n = 0$ for all $n$ sufficiently large.

*Proof.* When $\mu = 1$, $Z_n$ is itself a nonnegative martingale, so the martingale convergence theorem, (2.11) in Chapter 4 of Durrett (2004) converges to an a.s. finite limit $Z_\infty$. Since $Z_n$ is integer valued, we must have $Z_n = Z_\infty$ for large $n$. If $P(\xi_i^m = 1) < 1$ and $k > 0$ then $P(Z_n = k$ for all $n \geq N) = 0$ for any $N$, so we must have $Z_\infty \equiv 0$. 

**Theorem 1.4.** If $\mu > 1$ then $P(Z_n > 0$ for all $n) > 0$. 

*Proof.* For $s \in [0,1]$, let $\phi(s) = \sum_{k \geq 0} p_k s^k$ where $p_k = P(\xi_i^m = k)$. $\phi$ is the *generating function* for the offspring distribution $p_k$. Differentiating gives for $s < 1$

$$\phi'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \geq 0$$

$$\phi''(s) = \sum_{k=2} k(k-1)p_k s^{k-2} \geq 0$$

So $\phi$ is increasing and convex, and $\lim_{s \uparrow 1} \phi'(s) = \sum_{k=1}^{\infty} k p_k = \mu$. Our interest in $\phi$ stems from the following facts.

(a) If $\theta_m = P(Z_m = 0)$ then $\theta_m = \sum_{k=0}^{\infty} p_k \theta_{m-1}^k$.

*Proof of (a).* If $Z_1 = k$, an event with probability $p_k$, then $Z_m = 0$ if and only if all $k$ families die out in the remaining $m - 1$ units of time, an independent event with probability $\theta_{m-1}^k$. Summing over the disjoint possibilities for each $k$ gives the desired result.

(b) If $\phi'(1) = \mu > 1$ there is a unique $\rho < 1$ so that $\phi(\rho) = \rho$.

*Proof of (b).* $\phi(0) \geq 0$, $\phi(1) = 1$, and $\phi'(1) > 1$, so $\phi(1 - \epsilon) < 1 - \epsilon$ for small $\epsilon$. The last two observations imply the existence of a fixed point. To see it is unique, observe that $\mu > 1$ implies $p_k > 0$ for some $k > 1$, so $\phi''(\theta) > 0$ for $\theta > 0$. Since $\phi$ is strictly convex, it follows that if $\rho < 1$ is the smallest fixed point, then $\phi(x) < x$ for $x \in (\rho,1)$.

(c) As $m \uparrow \infty$, $\theta_m \uparrow \rho$.

*Proof of (c).* $\theta_0 = 0$, $\phi(\rho) = \rho$, and $\phi$ is increasing, so induction implies $\theta_m$ is increasing and $\theta_n \leq \rho$. Let $\theta_\infty = \lim \theta_m$. Taking limits in $\theta_m = \phi(\theta_{m-1})$, we see $\theta_\infty = \phi(\theta_\infty)$. Since $\theta_\infty \leq \rho$, it follows that $\theta_\infty = \rho$.

Combining (a)–(c) shows $P(Z_n = 0$ for some $n) = \lim_{n \to \infty} \theta_n = \rho < 1$. 


Example. Consider the Poisson distribution with mean \( \lambda \), i.e.,

\[
P(\xi = k) = e^{-\lambda} \frac{\lambda^k}{k!}
\]

In this case \( \phi(s) = \sum_{k=0}^{\infty} e^{-\lambda} s^k \lambda^k / k! = \exp(\lambda(s - 1)) \) so the fixed point equation is

\[
\rho = \exp(\lambda(\rho - 1)) \tag{1.1}
\]

Theorem 1.4 shows that when \( \mu > 1 \), the limit of \( Z_n/\mu^n \) has a chance of being nonzero. The best result on this question is due to Kesten and Stigum:

**Theorem 1.5.** Suppose \( \mu > 1 \). \( W = \lim Z_n/\mu^n \) is not \( \equiv 0 \) if and only if \( \sum p_k k \log k < \infty \).

For a proof, see Athreya and Ney (1972), p. 24–29. We will now show

**Theorem 1.6.** If \( \mu = E(\xi^m_i) > 1 \) and \( \sum k^2 p_k < \infty \) then \( W = \lim Z_n/\mu^n \) is not \( \equiv 0 \).

**Proof.** Let \( \sigma^2 = \text{var}(\xi^m_i) \). Let \( X_n = Z_n/\mu^n \).

\[
E(X_n^2|\mathcal{F}_{n-1}) = X_{n-1}^2 + 2X_{n-1}E(X_n - X_{n-1}|\mathcal{F}_{n-1}) + E((X_n - X_{n-1})^2|\mathcal{F}_{n-1})
\]

\[
= X_{n-1}^2 + E((X_n - X_{n-1})^2|\mathcal{F}_{n-1}) \tag{1.2}
\]

since \( X_n \) is a martingale. To compute the second term, we observe

\[
E((X_n - X_{n-1})^2|\mathcal{F}_{n-1}) = E((Z_n/\mu^n - Z_{n-1}/\mu^{n-1})^2|\mathcal{F}_{n-1})
\]

\[
= \mu^{-2n} E((Z_n - \mu Z_{n-1})^2|\mathcal{F}_{n-1}) \tag{1.3}
\]

On \( \{Z_{n-1} = k\} \),

\[
E((Z_n - \mu Z_{n-1})^2|\mathcal{F}_{n-1}) = E\left(\left( \sum_{i=1}^{k} \xi^m_i - \mu k \right)^2 \bigg| \mathcal{F}_{n-1} \right) = k\sigma^2 = Z_{n-1}\sigma^2
\]

Combining the last three equations gives

\[
EX_n^2 = EX_{n-1}^2 + E(Z_{n-1}\sigma^2/\mu^{2n}) = EX_{n-1}^2 + \sigma^2/\mu^{n+1}
\]

since \( E(Z_{n-1}/\mu^{n-1}) = EZ_0 = 1 \). Now \( EX_0^2 = 1 \), so \( EX_1^2 = 1 + \sigma^2/\mu^2 \), and induction gives

\[
EX_n^2 = 1 + \sigma^2 \sum_{k=2}^{n+1} \mu^{-k} \tag{1.4}
\]

This shows \( \sup EX_n^2 < \infty \), so by the \( L^2 \) convergence theorem for martingales, (4.5) in Chapter 4 of Durrett (2004), \( X_n \to W \) in \( L^2 \), and hence \( EX_n \to EW \). \( EX_n = 1 \) for all \( n \), so \( EW = 1 \) and \( W \) is not \( \equiv 0 \).
Our next result shows that when $W$ is not equivalent to 0 it is positive on the set where the branching process does not die out.

**Theorem 1.7.** If $P(W = 0) < 1$ then \( \{ W > 0 \} = \{ Z_n > 0 \text{ for all } n \} \).

**Proof.** Let $\rho = P(W = 0)$. In order for $Z_n/\mu^n$ to converge to 0 this must also hold for the branching process started by each of the children in the first generation. Breaking things down according to the number of children in the first generation

$$\rho = \sum_{k=0}^{\infty} p_k \rho^k = \phi(\rho)$$

so $\rho < 1$ is a fixed point of the generating function and hence $\rho = P(Z_n = 0$ for some $n$).

Clearly, $\{ W > 0 \} \subset \{ Z_n > 0 \text{ for all } n \}$. Since the two sets have the same probability $P(\{ Z_n > 0 \text{ for all } n \} - \{ W > 0 \}) = 0$, which is the desired result. \qed

Our final question: What happens in a supercritical branching process when it dies out?

**Theorem 1.8.** A supercritical branching process conditioned to become extinct is a subcritical branching process. If the original offspring distribution is Poisson($\lambda$) with $\lambda > 1$ then the conditioned one is Poisson($\lambda \rho$) where $\rho$ is the extinction probability.

**Proof.** Let $T_0 = \inf \{ n : Z_n = 0 \}$ and consider $\check{Z}_n = (Z_n | T_0 < \infty)$. To check the Markov property for $\check{Z}_n$ note that the Markov property for $Z_n$ implies:

$$P(Z_{n+1} = z_{n+1}, T_0 < \infty | Z_n = z_n, \ldots, Z_0 = z_0) = P(Z_{n+1} = z_{n+1}, T_0 < \infty | Z_n = z_n, \ldots, Z_0 = z_0)$$

To compute the transition probability for $\check{Z}_n$, observe that if $\rho$ is the extinction probability then $P_x(T_0 < \infty) = \rho^x$. Let $p(x, y)$ be the transition probability for $Z_n$. we note that the Markov property implies

$$\check{p}(x, y) = \frac{P_x(Z_1 = y, T_0 < \infty)}{P_x(T_0 < \infty)} = \frac{P_x(Z_1 = y)p_y(T_0 < \infty)}{P_x(T_0 < \infty)} = \frac{p(x, y)\rho^y \rho^x}{\rho^x}$$

Taking $x = 1$ and computing the generating function

$$\sum_{y=0}^{\infty} \check{p}(1, y)s^y = \rho^{-1}\sum_{y=0}^{\infty} p(1, y)(sp)^y = \rho^{-1}\phi(sp)$$

where $p_y = p(1, y)$ is the offspring distribution. In words, we take the graph of $\phi$ up to the fix point and rescale to make the domain and range $[0, 1]$.

$\check{p}_y = \check{p}(1, y)$ is the distribution of the size of the family of an individual, conditioned on the branching process dying out. If we start with $x$ individuals then in $Z_n$ each gives rise to
an independent family. In \( Z_n \) each family must die out, so \( Z_n \) is a branching process with offspring distribution \( \bar{p}(1, y) \). To prove this formally observe that

\[
p(x, y) = \sum_{j_1, \ldots, j_x \geq 0, j_1 + \cdots + j_x = y} p_{j_1} \cdots p_{j_x}
\]

Writing \( \sum' \) as shorthand for the sum in the last display

\[
\frac{p(x, y) \rho^y}{\rho^x} = \sum' \frac{p_{j_1} \rho^{j_1}}{\rho} \cdots \frac{p_{j_x} \rho^{j_x}}{\rho} = \sum' \bar{p}_{j_1} \cdots \bar{p}_{j_x}
\]

In the case of the Poisson distribution \( \phi(s) = \exp(\lambda(s - 1)) \) so if \( \lambda > 1 \), using the fixed point equation (1.1)

\[
\frac{\phi(s \rho)}{\rho} = \frac{\exp(\lambda(s \rho - 1))}{\exp(\lambda(\rho - 1))} = \exp(\lambda \rho(s - 1))
\]

which completes the proof. \( \square \)

2 Epidemics

To investigate the cluster containing 1. we let \( S_0 = \{2, 3, \ldots, n\}, I_0 = \{1\}, \) and \( R_0 = \emptyset \). The letters are motivated by the epidemic interpretation of the growing cluster. \( S_t \) are the susceptibles, \( I_t \) are infected, and \( R_t \) are removed. In graph terms, we have already examined the connection of all sites in \( R_t \), \( I_t \) are the sites to be investigated on this turn, and \( S_t \) are unexplored.

\[
R_{t+1} = R_t \cup I_t \\
I_{t+1} = \{y : \eta_{x,y} = 1 \text{ for some } x \in I_t\} \\
S_{t+1} = S_t - I_{t+1}
\]

(2.1)

The cluster containing 1, \( C_1 = \cup_{t=0}^{\infty} I_t \).

To define a comparison branching process we introduce a new independent set of variables \( \zeta_{x,y}^t, x, t \geq 1, y \leq n \) that are independent, \( = 1 \) with probability \( c/n \), and 0 otherwise. Let \( Z_0 = 1 \) and

\[
Z_{t+1} = \sum_{x \in I_t, y \in S_t} \eta_{x,y} + \sum_{x \in I_t} \sum_{y=1}^{n} \zeta_{x,y}^t + \sum_{x=|I_t|+1}^{Z_t} \sum_{y=1}^{n} \zeta_{x,y}^t
\]

(2.2)

The third term represents children of individuals in the branching process that are not in \( S_t \). The second term, which we will denote by \( B_t \), is the set of extra births in the branching process due to the fact that \( |S_t| < n \). As for the first term,

\[
C_{t+1} = \sum_{x \in I_t, y \in S_t} \eta_{x,y} - |I_{t+1}| \geq 0
\]
is the number of collisions, i.e., the number of extra edges. It is immediate from the con-
struction that $Z_t$ is a branching process with offspring distribution Binomial($n, \lambda/n$) and $Z_t \geq |I_t|$

To bound the number of removed individuals we let $Y_0 = 0$ and for $t \geq 0$

$$Y_{t+1} = \sum_{s=0}^{t} Z_s = Y_t + Z_t \geq |R_{t+1}|$$

This is enough to take care of the case $\lambda < 1$. $EZ_t = \lambda^t$, so the mean cluster size

$$E|C_1| = E \left( \sum_{t=0}^{\infty} |I_t| \right) \leq \sum_{t=0}^{\infty} \lambda^t = \frac{1}{1-\lambda} < \infty$$ \hspace{1cm} (2.3)

To study the case $\lambda > 1$ (and to show that the last result is asymptotically sharp) we need to bound the difference between $Z_t$ and $|I_t|$. Let $\mathcal{F}_t$ be the $\sigma$-field generated by $S_s, I_s, R_s$ for $s \leq t$ and $\zeta_{s,y}$ with $s \leq t$.

$$E(B_{t+1} | \mathcal{F}_t) = \frac{\lambda}{n} |I_t| (|I_t| + |R_t|) \leq \frac{\lambda}{n} Z_t (Z_t + Y_t) = \frac{\lambda}{n} Z_t Y_{t+1}$$

To bound the collision term we observe that

$$C_{t+1} \leq |\{(x,x',y) : x, x' \in I_t, y \in S_t, x < x', \eta_{x,y} = \eta_{x',y} = 1\}|$$

so using $|S_t| \leq n$ we have

$$E(C_{t+1} | \mathcal{F}_t) \leq \left( \frac{\lambda}{n} \right)^2 |I_t|^2 |S_t| \leq \frac{\lambda^2}{n} |Z_t|^2$$

Adding the last two bounds, using $Z_t \leq Y_{t+1}$ and replacing $t+1$ by $t$ we have

$$E(B_t + C_t | \mathcal{F}_{t-1}) \leq \frac{\lambda + \lambda^2}{n} Y_t^2$$

**Lemma 2.1.** If $\lambda > 1$ then there is a constant $C$ so that

$$E(B_t + C_t) \leq \frac{C}{n} \lambda^{2t}$$

**Proof.** $Y_t^2 = (\sum_{s=0}^{t-1} Z_s)^2 = \sum_{0 \leq r, s < t} Z_r Z_s$. The mean and variance of the offspring distribution are $\mu = \lambda$ and $\sigma^2 = n(\lambda/n)(1 - \lambda/n) \leq \lambda$, so (1.4) gives

$$EZ_s^2/\mu^{2s} = 1 + \sigma^2 \sum_{k=2}^{s+1} \mu^{-k} \leq C$$
If \( r < s \) then
\[
E(Z_r Z_s | \mathcal{F}_r) = Z_r E(Z_s | \mathcal{F}_r) = \mu^{s-r} Z_r^2 \leq C \mu^{r+s}
\]
Using the last two results we have
\[
E \sum_{0 \leq r, s < t} Z_r Z_s \leq C \mu^{2t-2} \sum_{k=0}^{\infty} (k+1) \mu^{-k}
\]
so \( E(Y_t^2) \leq C \lambda^{2t} \) and the desired result follows. \( \Box \)

To estimate the third term in (2.2), we call the \( C_s + B_s \) individuals added at time \( s \) immigrants, and let \( A_{s,t} \) be the children at time \( t \geq s \) of immigrants at \( s \). Clearly
\[
E(A_{s,t}) = \lambda^{t-s} E(B_s + C_s)
\]
Since \( A_{t,t} = B_t + C_t \), using (2.2) and Lemma 2.1 we have
\[
E(Z_t - |I_t|) \leq E \left( \sum_{s=0}^{t} A_{s,t} \right) \leq E(Y_{t+1}^2) \leq \frac{C}{n} \lambda^{2t} \sum_{k=0}^{\infty} \lambda^{-k}
\]
and we have established

**Theorem 2.2.** If \( \lambda > 1 \) and \( n \geq n_0 \), \( E(Z_t - |I_t|) \leq \frac{C}{n} \lambda^{2t} \).

When \( t = a \log n/ \log \lambda \) we have
\[
E(Z_t - |I_t|) \leq C n^{2a-1}
\]
If \( a < 1/2 \) then the right-hand side tends to 0 and \( Z_t = |I_t| \) with a probability that tends to 1 as \( n \to \infty \). \( EZ_t = \lambda^t = n^a \) so if \( a < 1 \), \( E(Z_t - |I_t|) = o(EZ_t) \).

### 3 The random walk viewpoint

Although the connection with branching processes is intuitive, it is more convenient technically to expose the cluster one site at a time to get a random walk. Since we have an emotional attachment to using \( S_t \) for a random walk, we will change the previous notation and let \( R_0 = \emptyset \), \( U_0 = \{2, 3, \ldots, n\} \), and \( A_0 = \{1\} \). \( R_t \) is the set of removed sites, \( U_t \) are the unexplored sites and \( A_t \) is the set of active sites. At time \( \tau = \inf \{t : A_t = \emptyset\} \) the process stops. If \( A_t \neq \emptyset \), pick \( i_t \) from \( A_t \) according to some rule that is measurable with respect to \( A_t = \sigma(A_0, \ldots A_t) \) and let
\[
R_{t+1} = R_t \cup \{i_t\} \\
A_{t+1} = A_t - \{i_t\} \cup \{y \in U_t : \eta_{i_t,y} = 1\} \\
U_{t+1} = U_t - \{y \in U_t : \eta_{i_t,y} = 1\}
\] (3.1)
This time $|R_t| = t$ for $t \leq \tau$, so the cluster size is $\tau$.

**Upper bound for $\lambda < 1$.** To define a comparison random walk, we introduce a new independent set of variables $\zeta^t_y$, $t \geq 1$, $y \leq n$ that are independent, $= 1$ with probability $\lambda/n$, and $0$ otherwise. Let $S_0 = 1$ and for $t \geq 0$

$$S_{t+1} = S_t - 1 + \sum_{y \in U_t} \eta^t_{y} + \sum_{y = |U_t| + 1}^{n} \zeta^t_y \text{ if } A_t \neq \emptyset$$

$$S_{t+1} = \sum_{y=1}^{n} \zeta^t_y \text{ if } A_t = \emptyset$$

$S_t$ is a random walk with $S_t \geq |A_t|$ if $t \leq \tau$, so if $T = \inf\{t : S_t = 0\}$ then $\tau \leq T$.

The increments $X_t$ of the random walk are $-1 + \text{Binomial}(n, \lambda/n)$. If $\lambda < 1$ stopping the martingale $S_t - (\lambda - 1)t$ at the bounded stopping time $T \wedge t$ gives

$$E S_{T \wedge t} - (\lambda - 1)E(T \wedge t) = ES_0 = 1$$

from which it follows that $E(T \wedge t) \leq 1/(1-\lambda)$. Letting $t \to \infty$ we have $ET \leq 1/(1-\lambda)$. Having verified that $ET < \infty$ we can now use Wald’s equation, see e.g., (1.6) in Chapter 3 of Durrett (2004), to conclude

$$ET = 1/(1-\lambda) \quad (3.2)$$

We can get a much better result by using the moment generating function:

**Theorem 3.1.** Suppose $\lambda < 1$ and let $\alpha = \lambda - 1 - \log(\lambda) > 0$. If $a > 1/\alpha$ then

$$P(\max_{1 \leq x \leq n} |C_x| \geq a \log n) \to 0$$

**Remark.** This bound is very accurate. Corollary 5.11 of Bollobas (2001) shows that the largest component is asymptotically

$$\frac{1}{\alpha} \left( \log n - \frac{5}{2} \log \log n \right)$$

**Proof.** We begin by computing the moment generating function:

$$\phi(\theta) = E \exp(\theta X_i)) = e^{-\theta}(1 - \lambda/n + (\lambda/n)e^{\theta})^n$$

$$\leq \exp(-\theta + \lambda(e^{\theta} - 1)) = \psi(\theta) \quad (3.3)$$

since $1 + x \leq e^x$. Note that the right-hand side is the moment generating function of $-1 + \text{Poisson mean } \lambda$. $\psi'(0) = EX_i = 1 - \lambda$ so if $\lambda < 1$, $\psi(\theta) < 1$ when $\theta > 0$ is small. The derivative

$$(-\theta + \lambda(e^{\theta} - 1))' = -1 + \lambda e^{\theta} = 0$$

when $\theta_1 = -\log \lambda$. At this point $\psi(\theta_1) = \exp(\log(\lambda) + 1 - \lambda) \equiv e^{-\alpha} < 1$. $M_t = \exp(\theta_t S_t)/\phi(\theta_t)$ is a nonnegative martingale, so using the optional stopping theorem for nonnegative supermartingales, see e.g., (7.6) in Chapter 4 of Durrett (2004)

$$1/\lambda = e^{\theta_1} \geq E(\phi(\theta_1)^{-T}) \geq E(\psi(\theta_1)^{-T}) = E(e^{\alpha T})$$
so using Chebyshev’s inequality

\[ P(T \geq k) \leq e^{-ka}/\lambda \]  

(3.4)

Letting \( C_x \) denote the cluster containing \( x \) and taking \( k = (1 + \epsilon)(\log n)/\alpha \)

\[ P(|C_x| \geq (1 + \epsilon)(\log n)/\alpha) \leq n^{-(1+\epsilon)/\lambda} \]

from which the desired result follows.

\[ \square \]

**Lower Bound for \( \lambda > 1 \).** To get a lower bound on the growth of the cluster let \( \hat{U}_t^\delta \) consists of the \((1 - \delta)n\) smallest members of \( \hat{U}_t \). As long as \( \hat{A}_t \neq \emptyset \) and \(|\hat{A}_t| + t \leq n\delta\), we can define

\[
\begin{align*}
\hat{R}_{t+1} &= \hat{R}_t \cup \{j_t\} \\
\hat{A}_{t+1} &= \hat{A}_t - \{j_t\} \cup \{y \in \hat{U}_t^\delta : \eta_{j_t,y} = 1\} \\
\hat{U}_{t+1} &= \hat{U}_t - \{y \in \hat{U}_t^\delta : \eta_{j_t,y} = 1\}
\end{align*}
\]

(3.5)

where \( j_t = \min \hat{A}_t \). It is easy to see that if we take \( i_t = j_t \) in (3.1) then \(|A_t| \geq |\hat{A}_t|\). To define a comparison random walk, we let \( W_0 = 1 \) and for

\[ t \leq T_W = \inf\{s : W_s = 0, \text{ or } W_s + s \geq n\delta\} \]

define

\[
W_{t+1} = W_t - 1 + \begin{cases} 
\sum_{y \in U_t^\delta} \eta_{i_t,y} & \text{if } t < T_W \\
\sum_{y=1}^{n(1-\delta)} \zeta_{y} & \text{if } t \geq T_W 
\end{cases}
\]

It is easy to see that for \( t \leq T_W, |\hat{A}_t| = W_t \) so \( \tau \geq T_W \).

We will use the comparison random walk to prove

**Theorem 3.2.** Suppose \( \lambda > 1 \). There is a constant \( \beta \) so that with probability \( \to 1 \), there is only one component of the random graph with more than \( \beta \log n \) vertices. The size of this component \( \sim (1 - \rho(\lambda))n \) where \( \rho(\lambda) \) is the extinction probability for the Poisson(\( \lambda \)) branching process.

**Remark.** This time our constant is the end result of several choices and is not so good. Corollary 5.11 of Bollobas (2001) shows that when \( \lambda > 0 \) the largest component is asymptotically \((1/\alpha) \log n\) where \( \alpha = \lambda - 1 - \log(\lambda) > 0 \).

**Proof.** The increments of \( W \) have the distribution \(-1 + \text{Binomial}((1-\delta)n, \lambda/n)\). By (3.3) the moment generating function of an increment

\[ \phi_\delta(\theta) \leq \exp(-\theta + \lambda(1-\delta)(e^\theta - 1)) \equiv \psi_\delta(\theta) \]  

(3.6)
Choose $\delta > 0$ so that $\lambda(1 - \delta) > 1$. $\psi'_\delta(0) > 0$ so $\psi'_\delta(-\theta) < 1$ when $\theta > 0$ is small. Since $\psi'_\delta(-\theta)$ is convex and tends to $\infty$ as $\theta \to \infty$ there is a unique positive solution of $\psi'_\delta(-\theta) = 1$. $\phi_\delta(-\theta) \leq 1$ so $M_t = \exp(-\theta W_t)$ is a supermartingale. Let $T_0 = \inf\{t : W_t = 0\}$. Stopping at time $T_0 \wedge t$ we have $e^{-\theta t} \geq P_1(T_0 \leq t)$ where the subscript 1 indicates that $W_0 = 1$. Letting $t \to \infty$

$$P_1(T_0 < \infty) \leq e^{-\theta t} \tag{3.7}$$

To compare with the previous approach using branching processes, note that $\rho = e^{-\theta_0} < 1$ has

$$\theta_0 + \lambda(e^{-\theta_0} - 1) = 0$$

which is the fixed point equation for the branching process.

From (3.7) it immediate that if $P_m$ denotes the probability when $W_0 = m$

$$P_m(T < \infty) \leq e^{-\theta m} \tag{3.8}$$

We want to make sure that all of the components $C_x, 1 \leq x \leq n$ behave as we expect, so we take $m_\delta = (2/\theta) \log n$ to make the right-hand side $n^{-2}$. To control the behavior of $W_t$ with this degree of certainty we use the following large deviations result.

**Lemma 3.3.** Let $Z = X_1 + \cdots + X_t$ where the $X_i$ are independent Binomial$((1 - \delta)n, \lambda/n)$. Let $\mu = EZ = t\lambda(1 - \delta)$. Let $\gamma(x) = x \log x - x + 1$ which is $> 0$ when $x \neq 1$. If $x < 1 < y$

then

$$P(Z \leq x \mu) \leq e^{-\gamma(x)\mu} \quad P(Z \geq y \mu) \leq e^{-\gamma(y)\mu}$$

**Proof.**

$$E \exp(\theta Z) \leq \exp(\mu(e^\theta - 1)) \text{ by (3.6). If } \theta > 0 \text{ Markov’s inequality implies}$$

$$P(Z \geq y \mu) \leq \exp((-\theta y + e^\theta - 1)\mu)$$

Since $y > 1$, $-\theta y + e^\theta - 1 < 0$ for small $\theta > 0$. Differentiating we see that the bound is optimized by taking $\theta = \log b$. If $\theta < 0$ Markov’s inequality implies

$$P(Z \leq x \mu) \leq \exp((-\theta x + e^\theta - 1)\mu)$$

Since $x < 1$, $-\theta x + e^\theta - 1 < 0$ for small $\theta < 0$. Differentiating we see that the bound is optimized by taking $\theta = \log x$. 

Let $\epsilon = (\lambda - 1)/2$. Applying Lemma 3.3 to $S_t - S_0 + t$ with $x = (1 + \epsilon)/\lambda$ we see that there is an $\eta_1 > 0$ so that

$$P(S_t - S_0 \leq \epsilon t) \leq \exp(-\eta_1 t) \tag{3.9}$$

Taking $y = 2$ in Lemma 3.3 we have an $\eta_2 > 0$ so that

$$P(S_t - S_0 + t \geq 2\lambda t) \leq \exp(-\eta_2 t) \tag{3.10}$$
Recall $S_0 = 1$. When $S_t + t \leq 2\lambda t$, we have $|U_s| \geq n - 2\lambda t$ for all $s \leq t$, so the number of births lost in $A_s$ for $s \leq t \wedge \tau$

\[ \leq \sum_{s=0}^{t-1} \sum_{y=|U_s|+1}^{n} \zeta^t \leq \text{Binomial}(2\lambda t^2, \lambda/n) \]  

From this it follows that

\[ P(|A_t| > 0, S_t - |A_t| \geq 2) \leq \left( \frac{2\lambda t^2}{2} \right) \left( \frac{\lambda}{n} \right)^2 \]  

(3.12)

We are going to use the results in the previous paragraph at time $r = \beta \log n$ where $\beta$ is chosen so that $\beta \epsilon \geq 3/\theta$, $\beta \eta_1 \geq 2$ and $\beta \eta_2 \geq 2$. Combining (3.9), (3.10), and (3.12) we have

\[ P(0 < |A_r| \leq (3/\theta) \log n - 2) \leq C \frac{(\log n)^2}{n^2} \]

If $n$ is large $(3/\theta) \log n - 2 \geq (2/\theta) \log n = m_\theta$, so (3.8) implies that if $|A_r| > 0$ it is unlikely that the lower bounding random walk will ever hit 0. Let $\epsilon_\delta = (\lambda(1 - \delta) - 1)/2$. Using Lemma 3.3 twice we have

\[ P_1(W(n^{2/3}) - W(0) \leq \epsilon_\delta n^{2/3}) \leq \exp(-\eta_3 n^{2/3}) \]
\[ P_1(W(n^{2/3}) - W(0) + n^{2/3} \geq 2\lambda n^{2/3}) \leq \exp(-\eta_4 n^{2/3}) \]  

(3.13)

Here we take $W(0) = |A_r| \leq 2\lambda \beta \log n$. Since $W_t + t$ is nondecreasing this shows that with probability $1 - O(n^{-2})$, $W_s + s \leq \delta n$ for all $s \leq n^{2/3}$, and the coupling between $W_s$ and $|A(s + r)|$ remains valid for $0 \leq s \leq n^{2/3}$.

The first bound in (3.13) implies that if a cluster reaches size $r = \beta \log n$ then the set of active sites at time $r + n^{2/3}$ is $\geq \epsilon_\delta n^{2/3}$ with high probability. Thus if we have two such clusters of size $\geq \beta \log n$ then either (a) they will intersect by time $r + n^{2/3}$ of (b) at time $r + n^{2/3}$ they have disjoint sets $I$ and $J$ of active sites of size $\geq \epsilon_\delta n^{2/3}$. The probability of no edge connecting $I$ and $J$ is

\[ = \left( 1 - \frac{\lambda}{n} \right)^{\epsilon_\delta^{4/3} n^{2/3}} \leq \exp(-\eta_5 n^{1/3}) \]

where $\eta_5 = \lambda \epsilon_\delta^2$. This proves the first assertion in Theorem 3.2.

To prove the second one it suffices to show that

\[ |\{ x : |C_x| \leq \beta \log n \}|/n \rightarrow \rho(\lambda) \]  

(3.14)

The first step is to show $P(|C_x| \leq \beta \log n) \rightarrow \rho(\lambda)$. Let $T_0 = \inf\{ t : S_t = 0 \}$. Because of the comparison between $S_t$ and $|A_t|$, the probability in question is $\geq P(T_0 \leq \beta \log n) \rightarrow \rho(\lambda)$. For the other direction we note that (3.10) and (3.11) show that $P(T_0 > \beta \log n, S_t \neq |A_t|) \rightarrow 0$. 
To complete the proof of (3.14) we will show that the random variables \( Y_x = 1 \) if \(|C_x| \leq \beta \log n\) and 0 otherwise are asymptotically uncorrelated. To do this it suffices to consider \( Y_1 \) and \( Y_2 \). Let \( R_t, U_t \) and \( A_t \) be the process introduced earlier for the cluster starting at 1, where for concreteness we suppose \( i_t = \min A_t \). To generate a random variable \( Y_2' \) so that \((Y_1, Y_2')\) has the same distribution as \((Y_1, Y_2)\), introduce independent copies of the basic indicator random variables \( \eta_{t,y}' \); let \( R_0' = \emptyset, A_0' = \{2\} \) and \( U_0' = \{1, 2, \ldots, n\} - \{2\}. \) If \( A_t' \neq \emptyset \), pick \( i_t' = \min A_t' \). If \( i_t' \not\in R_{\beta \log n} \) let

\[
\begin{align*}
R_{t+1}' &= R_t' \cup \{i_t'\} \\
A_{t+1}' &= A_t' - \{i_t'\} \cup \{y \in U_t' : \eta_{t,y}' = 1\} \\
U_{t+1}' &= U_t' - \{y \in U_t' : \eta_{t,y}' = 1\}
\end{align*}
\]

(3.15)

However if \( i_t' \in R_{\beta \log n} \), an event we call a collision, we use \( \eta_{t,y}' \) instead of \( \eta_{t,y}' \). In words if while growing cluster 2 we choose a site that was used in the growth of cluster 1, we use the original random variables \( \eta_{t,y} \). Otherwise we use independent random variables.

It should be clear from the construction that

\[
P(Y_1 = 1, Y_2 = 1) - P(Y_1 = 1)P(Y_2 = 1) \leq P(R_{\beta \log n} \cap R_{\beta \log n}' \neq \emptyset)
\]

Let \( G = \{|A_{\beta \log n} \cup R_{\beta \log n}'| \geq 2\lambda \beta \log n\} \). (3.10) and the choice of \( \beta \) imply that \( P(G^c) \leq n^{-2} \). Since \(|R_{\beta \log n}| = \beta \log n\), the probability of a collision when \( G \) occurs is at most \( 2\lambda(\beta \log n)^2/n \). Using our bound on the covariance

\[
\var \left( \sum_{x=1}^{n} Y_x \right) \leq n + \left( \frac{n}{2} \right) \frac{2\lambda(\beta \log n)^2}{n} \leq Cn \log n
\]

so it follows from Chebyshev’s inequality that

\[
P\left( \sum_{x=1}^{n} (Y_x - E Y_x) \geq n^{2/3} \right) \leq \frac{Cn \log n}{n^{4/3}} \to 0
\]

This proves (3.14) and completes the proof of Theorem 3.2.

Having proved the existence of the giant component, we can use the branching process results from Section 2 to compute the average distance between two points.

**Theorem 3.4.** Suppose \( \lambda > 1 \) and pick two points \( x \) and \( y \) at random from the giant cluster. Then \( d(x, y)/\log n \to 1/\log \lambda \) in probability.

**Proof.** The growth of the cluster containing \( x \) is dominated by the branching process. The size at time \((1-\epsilon) \log n/\log \lambda \) is \( \sim n^{1-\epsilon} \), so most of the points are at least that distance \((1-\epsilon) \log n/\log \lambda \). As Theorem 3.2 shows, membership in the giant component is asymptotically equivalent to the cluster size being larger than \( \beta \log n \). Let \( Y_{\infty} \) be the total progeny of the branching process. When \( \lambda > 1 \), \( P(\beta \log n < \infty) \to 0 \) as \( n \to \infty \), so points in the infinite
cluster are with high probability associated with branching processes that don’t die out. The size of one of these processes at time \((1+\epsilon) \log n/(2 \log \lambda)\) is \(\sim n^{(1+\epsilon)/2} \delta\) with high probability. If we consider two of them at this time then either (a) they have already intersected or (b) they will have \(n^{1+\epsilon} \delta^2\) chances to intersect at the next time, and the probability they will fail is

\[
\leq \left(1 - \frac{\lambda}{n}\right)^{n^{1+\epsilon} \delta^2} \leq \exp(-n^\epsilon \delta^2) \to 0
\]