Mathematical Statistics
Stockholm University

# Generating simple random graphs with prescribed degree distribution

Tom Britton, Maria Deijfen and Anders Martin-Löf

## Research Report 2005:2

ISSN 1650-0377

## Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

## Internet:
http://www.math.su.se/matstat

# Generating simple random graphs with prescribed degree distribution

Tom Britton[*], Maria Deijfen[†] and Anders Martin-Löf[‡]

February 2005

**Abstract**

Let $F$ be a probability distribution with support on the non-negative integers. A number of methods for generating a simple undirected graph with degree distribution $F$ are described and, under various assumptions on $F$, they are all shown to give the correct distribution in the limit of large graph size.

*Keywords:* Simple graphs, random graphs, degree distribution, generating algorithms

AMS 2000 Subject Classification: 05C80, 05C07, 90B15, 68R10.

## 1 Introduction and overview

A graph consists of two sets of objects: a set of vertices, $\mathcal{V}$, and a set of edges, $\mathcal{E}$. Each edge corresponds to a pair of vertices and the graph is said to be *undirected* if these pairs are unordered so that no directions are associated with the edges. Graphs with no duplicate edges and no loops – that is, with at most one edge between each pair of vertices and with no edges between a vertex and itself – are called *simple*. Furthermore, a graph is referred to as random if some kind of randomness is involved in its construction. In this paper we will consider graphs that are random in that the edges are

[*]Postal address: Mathematical Statistics, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: tom.britton@math.su.se.

[†]Vrije Universiteit Amsterdam. E-mail: mia@math.su.se

[‡]Postal address: Mathematical Statistics, Stockholm University, SE-106 91 Stockholm, Sweden. E-mail: andersml@math.su.se.

generated by random mechanisms. The question at issue is, given a set of vertices and a probability distribution $F$ on the non-negative integers, how do we proceed to construct a simple undirected graph where the degree of a randomly chosen vertex has distribution $F$?

The simplest random graph model is the Erdős-Rényi graph, which was introduced in the late 50's by Paul Erdős and Alfréd Rényi (1959,1960). Given a set of $n$ vertices, an Erdős-Rényi graph is constructed by independently adding each one of the $\binom{n}{2}$ possible edges of the graph with some probability $p$. The distribution of the vertex degree is binomial with parameters $n-1$ and $p$ and, if $p$ is scaled by $1/n$ so that the expected degree is held constant as $n \to \infty$, we get a Poisson distribution in the limit. Erdős-Rényi graphs have been widely studied and thorough descriptions of the field can be found in the books by Bollobás (2001) and Janson *et al* (1999).

An area that has been subject to intensive research within statistical physics during the last few years is the use of graphs as models for various types of complex networks; see e.g. Dorogovtsev and Mendes (2003) and Newman (2003) and the references therein. Examples of network structures that have been studied are social networks, power grids, the structure of the internet and various types of collaboration networks. In the graph representations, the vertices correspond to relevant units (people, generators and transformers, routers, movie actors etc.) and the edges correspond to connections between these units (social relations, high-voltage transmission lines, physical connections between routers, appearance in a common movie etc.). Typically, the networks are very large, making it impossible to delineate them in detail. Hence we have to rely on mathematical models. A natural approach in this context is to let the edges representing the connections in the network be generated by a random procedure designed so that the resulting graph captures the features of the real-life network in question as well as possible. Since, as mentioned, the networks are usually large, it is particularly urgent that the asymptotic properties of the graph model agree with empirical observations.

An essential characteristic of a graph is the vertex degree and, in a random graph, this is a random quantity. For instance, as mentioned above, in an Erdős-Rényi graph the degree of a vertex is asymptotically Poisson distributed. The Erdős-Rényi graph has a very simple and appealing mathematical structure and a lot of work has been done on the model. However, empirical studies have shown that the degree distribution in many real-life networks differs significantly from a Poisson distribution; see e.g. Liljeros *et al* (2001) (human sexual relationships), Faloutsos *et al* (1999) (physical structure of the internet) and Barabási and Albert (1999) (movie actor collaboration network). Complex networks typically have a more heavy-tailed degree distribution, often specified by some kind of power law, meaning that the number of vertices with degree $k$ is proportional to $k^{-\tau}$ for some exponent $\tau > 1$. Barabási and Albert (1999) introduced the term *scale-free networks* for this type of graphs and pointed out that there are important features of such graphs that are missed if they are approximated by Erdős-Rényi graphs.

In view of the above, it is important to be able to generate random graphs with other degree distributions than Poisson. The aim of this paper is to contribute to this problem by describing a number of algorithms that, given a probability distribution $F$ (which will later be subject to various restrictions), produces simple undirected graphs whose vertex degree is asymptotically distributed according to $F$. Here, clearly it is required that $\text{Supp}(F) \subseteq \mathbb{N}$, where $\text{Supp}(F)$ denotes the support of $F$. Let $D_1, \ldots, D_n$ be an i.i.d. sequence with distribution $F$. Arratia and Liggett (2005) show that, if $F$ has finite mean and $\text{Supp}(F) \neq \{2k; k \geq 0\}$, then, as $n \to \infty$, with probability 1 /2 there is indeed a simple random graph where the degrees of the vertices are $D_1, \ldots, D_n$ (the fact that the probability equals 1/2 is connected to the fact that the total degree sum must be even for a simple graph to exist). In this paper, the focus is on actually producing such a graph. To be more precise, given a set of $n$ vertices and a random mechanism to generate edges between them, let $p_k^{(n)}$ denote the probability of a randomly chosen vertex having degree $k$ and write $F = \{p_k; k \geq 0\}$. Our task is then to design an edge mechanism such that

(i) $\lim_{n\to\infty} p_k^{(n)} = p_k$;

(ii) the resulting graph is simple and undirected.

In all applications mentioned above, the networks are simple and undirected. Other applications might involve more complex networks, for instance the link structure of the world-wide web constitutes a directed graph and bipartite graphs – that is, graphs with two types of vertices and edges running only between unlike types – are common within sociology. However, simple undirected networks is indeed an important class in applications and the purpose of this work is to give a rigorous treatment of the asymptotic behavior of the vertex degree in a number of possible algorithms for generating such graphs. To our knowledge, such a study does not exist previously in the literature.

The rest of the paper is organized as follows. In Section 2 we review the well-known configuration model and describe how it can be used to generate simple graphs with an arbitrary prescribed degree distribution. Section 3 treats a model inspired by Chung and Lu (2002:1,2) that generates simple graphs with mixed Poisson degree distributions. The model is obtained by associating random weights $\{W_i\}$ with the vertices and draw an edge between two vertices $v_i$ and $v_j$ with a probability $p_{ij}$ determined by $p_{ij}/(1-p_{ij}) = W_i W_j/n$. In Section 4 we propose a method that is based on the introduction of directed edges according to a suitably chosen distribution. This method produces graphs with a degree distribution whose generating function contains a Poisson factor. Finally, in Section 5 the methods are discussed and evaluated.

## 2 The configuration model

The configuration model was originally defined by Bender and Canfield (1978) and has later been analyzed by Molloy and Reed (1995, 1998) and Newman *et al* (2001) among many others. The model describes a way to construct an undirected graph on $n$ vertices, labelled $v_1, \ldots, v_n$, having degree distribution $F = \{p_j; j \geq 0\}$. It is defined as follows: For each vertex $v_i$, generate a degree $d_i$ independently from a random variable $D$ with distribution $F$ and attach $d_i$ "stubs" to $v_i$. Then join the stubs of all vertices pairwise completely at random to form edges between the vertices. To be more precise, first pick two stubs randomly among all stubs in the graph and join them. Then pick two stubs at random from the remaining $\sum_1^n d_i - 2$ stubs and join them, etc. In general, in step $j$, two stubs chosen randomly among the $\sum_1^n d_i - 2(j-1)$ stubs that are still not connected after step $j-1$ are joined. A graph generated in this way is said to have random properties according to the configuration model.

A few problems might occur in the construction of a graph according to this algorithm. The first obvious problem is if the sum of all degrees, $\sum_i d_i$, happens to be an odd number. In this case it is not possible to join the stubs pairwise, since there will always be one remaining stub left over. However, unless $n$ is odd and the distribution of $F$ is concentrated to the odd numbers, this is not a serious problem. If the resulting degree sum is odd we can either regenerate the degrees until their sum is even or simply remove one stub chosen at random. More serious problems arise when the aim is to generate a *simple* undirected graph, that is, a graph without loops and multiple edges. In the configuration model, it is clearly possible for a stub of a given vertex $v_i$ with $d_i \geq 2$ to be matched with another one of the stubs of $v_i$, resulting in an edge from vertex $v_i$ to itself, that is, a loop. Similarly, two stubs of $v_i$ could by chance be joined with two stubs of the same other vertex, with the effect that a multiple edge is created.

So what should we do if we insist on the resulting graph being simple? Two obvious suggestions are (1): to remove loops and merge multiple edges into single edges in the generated graph to obtain a simple graph as final product, or (2): to redo the algorithm until a simple graph occurs by chance. These two methods are now analyzed in more detail under the names "Erased configuration model" and "Repeated configuration model" respectively. Both methods make the degree distribution somewhat different from the intended one, but below we show that they both of them have the right degree distribution asymptotically under certain moment conditions on the degree distribution.

## 2.1 The erased configuration model

Let $F_n = \{p_j^{(n)}; j \geq 0\}$ denote the degree distribution in the erased configuration model with stub distribution $F = \{p_j; j \geq 0\}$, that is, $p_j^{(n)}$ is the probability that a randomly selected vertex has degree $j$ in the erased configuration model on $n$ vertices.

**Theorem 2.1** *If $F$ has finite mean, then $F_n \to F$ as $n \to \infty$.*

*Proof:* We have to show that $p_j^{(n)}$ converges to $p_j$ for each $j$. To this end, first note that $\limsup_n p_j^{(n)} \leq p_j$ for all $j$, since, for each fixed $j$, we have

$$\begin{aligned}
\limsup_n p_j^{(n)} &= 1 - \liminf_n \sum_{k \neq j} p_k^{(n)} \\
&\leq 1 - \sum_{k \neq j} \liminf_n p_k^{(n)} \\
&\leq 1 - \sum_{k \neq j} p_k \\
&= p_j.
\end{aligned}$$

The theorem is proved if we can also show that $\liminf_n p_j^{(n)} \geq p_j$ for all $j$. To do this, fix $i \in \{1, \ldots, n\}$, let $D_i$ denote the degree of the vertex $v_i$ before erasing loops and multiple edges and write $E_i$ for the number of stubs attached to $v_i$ that are rubbed out in the erasing procedure. By definition of the configuration model, we have $D_i \sim F$. Writing $P^{(n)}$ for the probability law of the erased configuration model on $n$ vertices, it follows that, for all $j$, we have

$$\begin{aligned}
p_j^{(n)} &= p_j P^{(n)}(E_i = 0 | D_i = j) + \sum_{k > j} p_k P^{(n)}(E_i = k - j | D_i = k) \\
&\geq p_j P^{(n)}(E_i = 0 | D_i = j). \qquad (1)
\end{aligned}$$

We will show that

$$P^{(n)}(E_i = 0 | D_i = j) \to 1 \text{ as } n \to \infty. \tag{2}$$

By (1), this implies that $\liminf_n p_j^{(n)} \geq p_j$ for all $j$, as desired.

To prove (2), write $A_j$ for the event that a given stub belonging to a vertex with $j$ stubs in total avoids being removed in the erasing procedure. Below we show that

$$P^{(n)}(A_j) \to 1 \quad \text{for all } j \text{ as } n \to \infty. \tag{3}$$

This establishes (2): After having merged one of the $j$ stubs of the vertex $v_i$ to a stub belonging to some other vertex, saving it from being erased, the probability that a fixed one of the other $j-1$ stubs are erased equals $P^{(n)}(A_{j-1})$, since the fact that one stub from the other vertices is no longer available for merging is asymptotically negligible. This can then be repeated until there are no remaining stubs of $v_i$ and (2) follows by noting that $P^{(n)}(E_i = 0 | D_i = 0) = 1$.

For the proof to be complete it remains to prove (3). To do this, first remember that a stub can be erased for two reasons: because it forms a loop and because it is part of a multiple edge. For the sake of completeness we also include the case when a randomly selected stub is removed if the total number of stubs is odd. Now, consider a fixed stub belonging to a vertex $v$ with $j$ stubs in all. Write $A_j^{loop}$ and $A_j^{mult}$ for the events that the stub is not part of a loop and a multiple edge respectively and let $A_j^{odd}$ be the event that the stub is not removed as the randomly selected "odd" stub. To estimate the probabilities of these events, we condition on that the total number of stubs equals $m$. If $m$ is odd, the probability that the stub is removed as the "odd" stub is $1/m$ and, if $m$ is even, the probability is 0. Hence

$$P^{(n)}\left(A_j^{odd}\right) \geq 1 - \frac{1}{m}. \tag{4}$$

For the stub to form a loop it has to be joined to one of the other $j-1$ stubs of the vertex $v$ and, since clearly the stub is matched to each one of the other $m-1$ stubs in the graph with the same probability $1/(m-1)$, this happens with probability $(j-1)/(m-1)$, that is,

$$P^{(n)}\left(A_j^{loop}\right) = 1 - \frac{j-1}{m-1}. \tag{5}$$

To compute the probability that the stub is not part of a multiple edge, assume that it does not make up a loop and condition on the degree $k$ of the vertex $v'$ of the stub to which it is joined. Also, number the remaining $j-1$ stubs of the vertex $v$ in some arbitrary way from 1 to $j-1$ and let $B_j$ be the event that there are no loops among these stubs. Trivially,

$$P^{(n)}\left(A_j^{mult}\right) \geq P^{(n)}\left(A_j^{mult} \cap B_j\right).$$

For the event $A_j^{mult} \cap B_j$ to happen, none of the remaining $j-1$ stubs of the vertex $v$ can connect to another stub of $v$ or to a stub originating from $v'$. Considering the stubs $1, \ldots, j-1$ one in a turn, we see that the probability that stub number 1 avoids being connected to a stub of $v$ or $v'$ is $(m-j-k)/(m-3)$ (the denominator comes from that two stubs are already used in the fixed connection between $v$ and $v'$ and the stub cannot join to itself). Then, given that stub number 1 is connected to some other vertex, the probability that stub number 2 is so as well is $(m-j-k-1)/(m-5)$, and so on. Furthermore, it is only possible for $A_j^{mult} \cap B_j$ to happen if $j-1 \leq m-k-j$, since otherwise loops among the $j-1$ stubs on vertex $v$ or multiple edges between $v$ and $v'$ cannot be avoided. Hence we have

$$P^{(n)}\left(A_j^{mult} \cap B_j\right) = \begin{cases} \frac{m-k-j}{m-3}\frac{m-k-j-1}{m-5} \cdots \frac{m-k-(j-1)}{m-(2j-1)} & \text{if } j-1 \leq m-k-j; \\ 0 & \text{otherwise,} \end{cases}$$

implying that

$$P^{(n)}\left(A_j^{mult}\right) \geq \left(\frac{m-k-j+1}{m-3}\right)_+^{j-1}, \tag{6}$$

where $r_+ = \max\{r, 0\}$. Combining (4), (5) and (6), it follows that

$$\begin{aligned} P^{(n)}(A_j) &= P^{(n)}\left(A_j^{odd} \cap A_j^{loop} \cap A_j^{mult}\right) \\ &\geq \left(1 - \frac{1}{m}\right)\left(1 - \frac{j-1}{m-1}\right)\left(\frac{m-k-j+1}{m-3}\right)_+^{j-1}. \end{aligned}$$

Removing the conditioning on $m$ and $k$ and denoting the corresponding random variable $M_n$ and $K_n$ respectively, we get

$$P^{(n)}(A_j) \geq \mathrm{E}\left[\left(1 - \frac{j}{M_n - 1}\right)\left(\frac{M_n - K_n - j + 1}{M_n - 3}\right)_+^{j-1}\right]. \tag{7}$$

To complete the proof, we use dominated convergence to show that the right hand side of (7) converges to 1 as $n \to \infty$, establishing (3). Recall that $M_n$ is the total number of stubs in the configuration and $K_n$ is the number of stubs connected to the vertex of a randomly selected stub. The total number of stubs is obtained as a sum of $n$ i.i.d. random variables $\{D_l\}$ with distribution $F$ and mean $\mu_F$, which is finite by assumption. Hence, the law of large numbers implies that $M_n/n \to \mu_F$ almost surely as $n \to \infty$. The distribution of $K_n$ is specified by $P(K_n = i) = iN_i^{(n)}/\sum_r rN_n(r)$, where $N_i^{(n)}$ is the number of $D_l$:s that equal $i$. Since $N_i^{(n)}/n \to p_i$ as $n \to \infty$ and $\mu_F < \infty$, it follows that $K_n$ converges in distribution to a proper random variable $K$ with distribution $P(K = i) = ip_i/\sum_r rp_r = ip_i/\mu_F$ and hence $K_n/n \to 0$ almost surely as $n \to \infty$. Combining these two observations, we get that

$$\left(1 - \frac{j}{M_n - 1}\right)\left(\frac{M_n - K_n - j + 1}{M_n - 3}\right)_+^{j-1} \to 1 \quad \text{a.s. as } n \to \infty.$$

The first factor on the left hand side above lies between 0 and 1 and for the second factor, since $j$ and $K_n$ are both strictly positive, we have

$$0 \le \left(\frac{M_n - K_n - j + 1}{M_n - 3}\right)_+^{j-1} \le \left(\frac{M_n - 1}{M_n - 3}\right)_+^{j-1} \le 3^{j-1}$$

for $n \ge 4$. By dominated convergence, it follows that the right hand side of (7) converges to 1 and the proof is complete. □

## 2.2 The repeated configuration model

Recall that the repeated configuration model consists of performing the configuration model until it produces a simple graph. Using results from McKay (1985), Molloy and Reed (1995) have showed that the probability of obtaining a simple graph converges to a strictly positive constant $c$ under the assumption that the degree distribution has finite second moment. This implies that the configuration model will then produce a simple graph after a geometrically distributed number of tries. Of course, such a graph might not be typical for the configuration model. In particular one might suspect that the number of edges is somewhat smaller than normal, since there by chance were no multiple edges or loops. However, below we use the result of Molloy and Reed (1995) to show that the resulting degree distribution converges to the right one provided that it has finite second moment. In fact, we show the stronger result that the empirical degree distribution converges to the intended distribution.

To formulate the result, let $F'_n = \{p_j^{(n)}; j \ge 0\}$ denote the degree distribution of the repeated configuration model on $n$ vertices with stub distribution $F = \{p_j; j \ge 0\}$ and write $N_j^{(n)}$ for the number of vertices having degree $j$ in the resulting graph.

**Proposition 2.1** *Assume that $F$ has finite second moment. Then*

(a) $F'_n \to F$ *as* $n \to \infty$*;*

(b) $N_j^{(n)}/n \to p_j$ *in probability, that is, the empirical distribution converges in probability to* $F$*.*

*Proof:* We first show part (b). Let $D_1, \ldots, D_n$ be i.i.d. random variables with distribution $F$ and let $\tilde{p}_j^{(n)} = |\{D_i; D_i = j, i = 1, \ldots, n\}|/n$ denote the empirical distribution of these $n$ variables; here $|\cdot|$ denotes set cardinality. Also, write $S_n$ for the event that the configuration model on $n$ vertices produces a simple graph. The empirical distribution of the repeated configuration model is the same as the distribution of the vector with elements $\tilde{p}_j^{(n)}$ conditioned on $S_n$ and we hence have to show that

$$P(|\tilde{p}_j^{(n)} - p_j| > \epsilon \mid S_n) \to 0 \quad \text{as } n \to \infty \text{ for any } \epsilon > 0 \text{ and any } j. \tag{8}$$

Trivially, we have

$$\begin{aligned} P\left(|\tilde{p}_j^{(n)} - p_j| > \epsilon \mid S_n\right) &= \frac{P\left(|\tilde{p}_j^{(n)} - p_j| > \epsilon\ ,\ S_n\right)}{P(S_n)} \\ &\leq \frac{P\left(|\tilde{p}_j^{(n)} - p_j| > \epsilon\right)}{P(S_n)}. \end{aligned}$$

The numerator here converges to 0 by the law of large numbers and, by the cited result of Molloy and Reed (1995), the assumption that $F$ has finite second moment implies that $P(S_n) \to c > 0$. Hence (8) follows.

To show (a), note that, since $0 \leq N_j^{(n)}/n \leq 1$, by dominated convergence, the result in (b) implies that $\mathrm{E}[N_j^{(n)}/n] \to p_j$. But $\mathrm{E}[N_j^{(n)}] = \sum_{i=1}^n p_j^{(n)} = np_j^{(n)}$, and the desired result follows. □

# 3 The generalized random graph

In an Erdős-Rényi graph on $n$ vertices, the edges are defined by independent Bernoulli random variables $\{X_{ij}\}$ with $P(X_{ij} = 1) = p$, the event $X_{ij} = 1$ signifying the presence of an undirected edge between $v_i$ and $v_j$. By definition, $X_{ji} = X_{ij}$ for $i < j$ and $X_{ii} = 0$ for all $i$. In this section, we consider a model where the probability $p_{ij}$ of an edge between two vertices $v_i$ and $v_j$ is allowed to depend on $i$ and $j$. Special cases of this have been considered in Chung and Lu (2002:1,2). We will show that, if the probabilities $\{p_{ij}\}$ are picked randomly in a suitable way, we get a graph with a degree distribution that is easy to characterize in the limit when $n \to \infty$; see Theorem 3.1. Also, the degrees of the vertices are approximatively independent. The model will be referred to as the generalized random graph.

First, we develop the model in more detail. To this end, let $X = \{X_{ij}\}_{i<j}$ be the array of edge indicators and write $P(X_{ij} = 1) = p_{ij} = 1 - q_{ij}$. Since the indicators are independent, the probability density of $X$ is given by

$$P(X = x) = \prod_{i<j} p_{ij}^{x_{ij}} q_{ij}^{1-x_{ij}}.$$

Introducing the odds ratios $r_{ij} = p_{ij}/q_{ij}$ and noting that $p_{ij} = r_{ij}/(1 + r_{ij})$ and $q_{ij} = 1/(1 + r_{ij})$, this can be written

$$P(X = x) = \prod_{i<j}(1 + r_{ij})^{-1} \prod_{i<j} r_{ij}^{x_{ij}}.$$

Moreover, if we specialize to the situation where $r_{ij} = u_i u_j$ for some parameters $\{u_i\}_{i=1}^n$ with $u_i \geq 0$ and define $G(u) := \prod_{i<j}(1 + u_i u_j)$, we get

$$\begin{aligned} P_u(X = x) &= G^{-1}(u) \prod_{i<j} (u_i u_j)^{x_{ij}} \\ &= G^{-1}(u) \prod_i u_i^{d_i(x)}, \end{aligned} \tag{9}$$

where $d_i(x)$ is the degree of the vertex $v_i$ in the configuration $x$, that is, $d_i(x) := \sum_j x_{ij}$. This is a "canonical" distribution in the sense of statistical mechanics with sufficient statistics $\{d_i(X)\}$ and from (9) we see that the conditional distribution of $X$ given that $\{d_i(X) = d_i\}$ is uniform, that is, all graphs with a given degree sequence $\{d_i\}$ have the same probability. This is indeed a nice property of the model, motivating the use of the parametrization $r_{ij} = u_i u_j$ instead of the one defined by $p_{ij} = u_i u_j$ used in Chung and Lu (2002:1,2).

To obtain a formula for the joint generating function of the degree vector $\{d_i(X)\}$, note that, by (9), we have

$$\begin{aligned} \mathrm{E}_u\Big[\prod_i t_i^{d_i(X)}\Big] &= \sum_x P_u(X = x) \prod_i t_i^{d_i(x)} \\ &= G^{-1}(u) \sum_x \prod_i (t_i u_i)^{d_i(x)}. \end{aligned}$$

Since $\sum_x P_u(X = x) = 1$, it follows from (9) that $\sum_x \prod_i u_i^{d_i(x)} = G(u)$, and hence we get

$$\begin{aligned} \mathrm{E}_u\Big[\prod_i t_i^{d_i(X)}\Big] &= G^{-1}(u)G(tu) \\ &= \prod_{i<j} \frac{1+t_iu_it_ju_j}{1+u_iu_j}. \end{aligned} \tag{10}$$

Now consider the situation where the parameters $\{u_i\}$ are suitably scaled random variables, more precisely, we set $u_i = W_i/\sqrt{n}$, where $\{W_i\}$ are i.i.d. random variables with finite mean $\mu_W$. Write $\{D_i\}$ for the degrees of the vertices in this setting, that is, $D_i = d_i(X) = d_i(X(W))$. The following theorem specifies the liming distribution of the $D_i$:s.

**Theorem 3.1** *Consider a generalized random graph on $n$ vertices with edge probabilities defined by $p_{ij}/q_{ij} = W_iW_j/n$, where $\{W_i\}$ are i.i.d. random variables with mean $\mu_W$ and finite moment of order $1+\varepsilon$ for some $\varepsilon > 0$. We have:*

(a) *The limiting distribution of a degree variable $D_k$ as $n \to \infty$ is mixed Poisson with parameter $W_k\mu_W$, that is, in the limit we have*

$$P(D_k = l) = \mathrm{E}\left[\frac{(W_k\mu_W)^l}{l!}e^{-W_k\mu_W}\right].$$

(b) *For any $m$, the variables $D_1, \dots D_m$ are asymptotically independent.*

*Proof of Theorem 3.1:* By taking $t_k = t$ and $t_i = 1$ for $i \neq k$ in (10), it follows that

$$\mathrm{E}\Big[t^{D_k}\Big] = \mathrm{E}\Bigg[\prod_{i=1}^{n}\frac{1+W_iW_kt/n}{1+W_iW_k/n}\Bigg].$$

Using the Taylor expansion $\log(1+x) = x + O(x^2)$, we see that

$$\prod_i \frac{1+W_iW_kt/n}{1+W_iW_k/n} = \exp\left\{\frac{W_k\sum_i W_i}{n}(t-1) + R_n\right\},$$

where $R_n = O(W_k^2\sum_i W_i^2/n^2)$. To estimate $R_n$, note that $W_i^2 \leq \max_l\{W_l\}W_i$. The law of large numbers implies that $\sum_i W_i/n \to \mu_W$ and, since the $W_l$:s have finite $1+\varepsilon$-moment, we have that $\max_{1\leq l\leq n}\{W_l\}/n \to 0$. It follows that $R_n$ converges almost surely to 0 as $n \to \infty$. Hence

$$\mathrm{E}\Big[t^{D_k}\Big] \to \mathrm{E}\Big[e^{W_k\mu_W(t-1)}\Big] \quad \text{as } n \to \infty,$$

and part (a) follows. To establish (b), note that, by taking $t_i = 1$ for $i > m$ in (10) and proceeding as in proving (a), it can be seen that

$$\mathrm{E}\left[\prod_{i=1}^{m} t_i^{D_i}\right] \to \prod_{i=1}^{m} \mathrm{E}\left[e^{W_i \mu_W (t_i - 1)}\right] \quad \text{as } n \to \infty.$$

Hence the joint generating function of $(D_1, \ldots, D_m)$ asymptotically factorizes into a product of mixed Poisson generating functions, as desired. □

Now recall that our task is to generate a simple random graph with a given degree distribution $F$. The above theorem tells us that, if $F$ is a mixed Poisson distribution with parameter distribution $Q$ with finite moment of order $1+\varepsilon$, then this can be done by using the generalized random graph model with i.i.d. weights $\{W_i\}$ distributed according to $Q/\sqrt{\mu_Q}$, where $\mu_Q$ denotes the mean of $Q$.

As mentioned in the introduction, the degree distribution in many real-life networks is heavy-tailed, the probability of a vertex having degree $k$ being proportional to $k^{-\tau}$ for some exponent $\tau > 1$. According to the following simple proposition, heavy-tailed mixed Poisson distributions with this type of power law behavior can be accomplished by choosing a heavy-tailed parameter distribution. Here '$\sim$' means that the quotient between the right hand side and the left hand side tends to a finite constant as $k \to \infty$.

**Proposition 3.1** *Let $F = \{p_k; k \geq 0\}$ be a mixed Poisson distribution with parameter distribution $Q$. If $Q$ has a density $q(x)$ that satisfies $q(x) \sim x^{-\tau}$ for some exponent $\tau > 1$, then $p_k \sim k^{-\tau}$.*

*Proof:* Pick $k$ large so that $k - \tau > 0$. Then we have

$$\begin{aligned} p_k &= \int_0^\infty \frac{x^k}{k!} e^{-x} q(x) dx \\ &\sim \frac{1}{k!} \int_{\xi_Q}^\infty x^{k-\tau} e^{-x} dx \\ &= \frac{1}{k!} \left( \int_0^\infty x^{k-\tau} e^{-x} dx - \int_0^{\xi_Q} x^{k-\tau} e^{-x} dx \right), \end{aligned} \tag{11}$$

where $\xi_Q := \inf \mathrm{Supp}(Q)$. Trivially $\int_0^{\xi_Q} x^{k-\tau} e^{-x} dx \leq \xi_Q^{k-\tau}$. To deal with the first integral in (11), remember the definition of the Gamma function, $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$, and recall Stirling's formula, which states that $\Gamma(r) \sim r^{r-1/2} e^{-r}$. This yields

$$\int_0^\infty x^{k-\tau}e^{-x}dx = \Gamma(k-\tau+1) \sim (k-\tau+1)^{k-\tau+1/2}e^{-(k-\tau+1)}.$$

Also note that, since $\Gamma(k+1) = k!$, it follows from Stirling's formula that $k! \sim (k+1)^{k+1/2}e^{-(k+1)}$. Substituting these estimates in (11) gives

$$p_k \sim \frac{(k-\tau+1)^{k-\tau+1/2}e^{-(k-\tau+1)} - \xi_Q^{k-\tau}}{(k+1)^{k+1/2}e^{-(k+1)}} \sim k^{-\tau},$$

as desired. □

We finish this section by showing that the empirical degree distribution in the generalized random graph converges to the asymptotic mixed Poisson degree distribution in the graph. To this end, write $N_k^{(n)}$ for the number of vertices having degree $k$ in the generalized random graph on $n$ vertices with edge probabilities defined by $p_{ij}/q_{ij} = W_iW_j/n$, where $\{W_i\}$ are i.i.d. random variables with mean $\mu_W$. Also, let $F = \{p_k; k \geq 0\}$ be a mixed Poisson distribution with parameter described by $W\mu_W$.

**Proposition 3.2** *If $\mu_W < \infty$, then, for all $k$, we have that $N_k^{(n)}/n \to p_k$ in probability as $n \to \infty$.*

*Proof:* Write $P^{(n)}$ for the probability law of the generalized random graph on $n$ vertices and let $\mathbf{1}_{\{\cdot\}}$ denote the indicator function. Clearly $N_k^{(n)} = \sum_{i=1}^n \mathbf{1}_{\{D_i=k\}}$ and hence, using symmetry, it follows that

$$\mathrm{E}\Big[N_k^{(n)}\Big] = \sum_{i=1}^n P^{(n)}(D_i = k) = nP^{(n)}(D_1 = k).$$

By Theorem 3.1 (a), we have $P^{(n)}(D_1 = k) \to p_k$ as $n \to \infty$, meaning that $\mathrm{E}[N_k^{(n)}/n] \to p_k$. The desired result is now obtained from Chebyshev's inequality if we can show that $\mathrm{Var}(N_k^{(n)}/n) \to 0$. To do this, note that

$$\mathrm{E}\Big[\Big(N_k^{(n)}\Big)^2\Big] = \mathrm{E}\Big[\sum_i \mathbf{1}^2_{\{D_i=k\}} + \sum_{i\neq j}\mathbf{1}_{\{D_i=k\}}\mathbf{1}_{\{D_j=k\}}\Big] = nP^{(n)}(D_1 = k) + n(n-1)P^{(n)}(D_1 = k, D_2 = k),$$

where the last equality follows from symmetry. By Theorem 3.1 (b), the variables $D_1$ and $D_2$ are asymptotically independent and hence we have $P^{(n)}(D_1 = k, D_2 = k) \to p_k^2$ as $n \to \infty$. Using the formula $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$, it follows that $\mathrm{Var}(N_k^{(n)}/n) \to 0$ as $n \to \infty$ and we are done. $\square$

# 4 Directed graphs with removed directions

In this section we propose a construction method where directed edges are introduced according to some distribution $G$. The directions of the edges are then disregarded and multiple edges are fused together, producing a graph whose asymptotic degree distribution $F$ is the convolution of the distribution $G$ and a Poisson distribution with parameter $\mu_G$, where $\mu_G$ denotes the mean of $G$; see Proposition 4.1. Roughly the construction is as follows:

1. Each vertex is independently assigned a random number of outgoing edges according to some distribution $G$. For each vertex, the vertices that the outgoing edges will connect to are chosen randomly without replacement from the rest of the graph.

2. The directions of the edges are disregarded and any multiple edges between pairs of vertices are reduced to one single edge.

The degree of a vertex $v_i$ in the resulting graph is the sum of the outgoing edges from $v_i$ that were added in step 1 and the incoming edges to $v_i$ from vertices that are not hit by an outgoing edge from $v_i$. We will show that the asymptotic distributions of these quantities are $G$ and Poisson $\mu_G$ respectively, and that their sum is asymptotically distributed according to the convolution of these distributions. Questions that arise and that will be partly answered here are: Which distributions $F$ can be obtained in this way, by convolving a distribution $G$ with a Poisson distribution with the same mean as $G$? How should the distribution $G$ be chosen for an intended degree distribution $F$?

Let us first describe the construction in more detail. To begin with, fix $n$ and let $\mathcal{V}$ be a set of $n$ labelled vertices $v_1, \ldots, v_n$. Pick a probability distribution $G$ with finite mean $\mu_G$ and $\mathrm{Supp}(G) \subset \mathbb{N}$, and let $\{g_k\}$ denote the probabilities associated with $G$. Also, write $G_n$ for the distribution defined by the probabilities

$$g_k^{(n)} := \begin{cases} g_k & \text{for } k = 0, 1, \ldots, n-2; \\ \sum_{k=n-1}^{\infty} g_k & \text{for } k = n-1; \\ 0 & \text{for } k \geq n, \end{cases}$$

that is, $G_n$ is a truncated version of $G$ with support on $\{0, 1, \ldots, n-1\}$. Now, to each vertex $v_i$, associate independently a random variable $Y_i$ with distribution $G_n$, and add to the graph $Y_i$ directed edges pointing out from $v_i$. The vertices to be hit by the edges starting at $v_i$ are chosen randomly without replacement from $\mathcal{V} \setminus \{v_i\}$, independently for all vertices. This defines a directed random graph $\mathcal{G}_{dir}(n, G) = \{\mathcal{V}, \mathcal{E}_{dir}\}$. To obtain a simple undirected graph we disregard the directions of the edges and erase multiple edges, that is, an undirected edge between the vertices $v_i$ and $v_j$ is included in $\mathcal{E}$ as soon as at least one of the directed edges $(v_i, v_j)$ and $(v_j, v_i)$ is present in $\mathcal{E}_{dir}$.

Let $D_i$ denote the degree of the vertex $v_i$ in the resulting undirected graph $\mathcal{G}(n, G) = \{\mathcal{V}, \mathcal{E}\}$. To find an expression for $D_i$, write $\mathcal{V}_i^{out}$ for the set of vertices that are hit by edges pointing out from $v_i$ in $\mathcal{E}_{dir}$ and write $\mathcal{V}_i^{in}$ for the set of vertices that sends outgoing edges to $v_i$ in $\mathcal{E}_{dir}$, that is,

$$\mathcal{V}_i^{out} = \{v_j;\ (v_i, v_j) \in \mathcal{E}_{dir}\}$$

and

$$\mathcal{V}_i^{in} = \{v_j;\ (v_j, v_i) \in \mathcal{E}_{dir}\}.$$

Define $Z_i = |\mathcal{V}_i^{in} \cap \neg \mathcal{V}_i^{out}|$ so that, in words, $Z_i$ indicates the number of edges in $\mathcal{E}_{dir}$ pointing at $v_i$ and starting at vertices that are not hit by outgoing edges from $v_i$. Some thought reveals that

$$D_i = Y_i + Z_i.$$

Clearly all variables $\{Z_i\}$ have the same distribution, which we denote by $H_{(n,G)}$. Also remember that $Y_i \sim G_n$ for all $i$. Hence the degree variables $\{D_i\}$ for the graph model $\mathcal{G}(n, G)$ are identically distributed and we write $F_{(n,G)}$ for their distribution. The following theorem is the aforementioned characterization of the asymptotic degree distribution as the convolution of $G$ and a Poisson distribution with the same mean as $G$.

**Theorem 4.1** *As $n \to \infty$, we have*

(a) $G_n \to G$*;*

(b) $H_{(n,G)} \to \mathrm{Po}(\mu_G)$*, where* $\mathrm{Po}(\mu_G)$ *is a Poisson distribution with mean* $\mu_G$*;*

(c) $F_{(n,G)} \to G * \mathrm{Po}(\mu_G)$*.*

*Proof:* The claim in (a) is immediate from the definition of $G_n$. To prove (b), fix a vertex $v_k$ and, for $i \neq k$, let $X_{ik}$ be a 0-1 variable, indicating whether there is a directed edge from $v_i$ to $v_k$ in $\mathcal{E}_{dir}$ or not, that is,

$$X_{ik} = \begin{cases} 1 & \text{if } v_k \in \mathcal{V}_i^{out}; \\ 0 & \text{otherwise.} \end{cases}$$

Since there are $Y_i$ outgoing edges from $v_i$ and the vertices to be hit by these edges are chosen randomly without replacement from the $n-1$ vertices in $\mathcal{V}\backslash\{v_i\}$, we have

$$P(X_{ik} = 1) = \frac{Y_i}{n-1}.$$

Also, since the targets of the outgoing edges are chosen independently for all vertices, the variables $\{X_{ik}\}_i$ are independent. Now consider the variable $Z_k \sim H_{(n,G)}$ and note that

$$Z_k = \sum_{i \neq k; v_i \in \neg \mathcal{V}_k^{out}} X_{ik}. \tag{12}$$

By conditioning on $\{Y_i\}$ and using (12), we obtain

$$\mathrm{E}\left[t^{Z_k}\right] = \mathrm{E}\left[\prod_{i \neq k; v_i \in \neg \mathcal{V}_k^{out}} \left(1 + \frac{Y_i}{n-1}(t-1)\right)\right],$$

and, since $\{Y_i\}$ are i.i.d. with distribution $G_n$ and $|\neg \mathcal{V}_k^{out} \backslash \{v_k\}| = n - 1 - Y_k$, it follows that

$$\begin{aligned} \mathrm{E}\left[t^{Z_k}\right] &= \mathrm{E}\left[\left(1 + \frac{\mu_{G_n}}{n-1}(t-1)\right)^{n-1-Y_k}\right] \\ &= \left(1 + \frac{\mu_{G_n}}{n-1}(t-1)\right)^{n-1} \mathrm{E}\left[\left(\frac{n-1}{n-1+\mu_{G_n}(t-1)}\right)^{Y_k}\right]. \end{aligned}$$

Here $\mu_{G_n}$ denotes the mean of the distribution $G_n$. As $n \to \infty$, we have

$$\left(1 + \frac{\mu_{G_n}}{n-1}(t-1)\right)^{n-1} \longrightarrow e^{\mu_G(t-1)} \tag{13}$$

and

$$\mathrm{E}\left[\left(\frac{n-1}{n-1+\mu_{G_n}(t-1)}\right)^{Y_k}\right] \longrightarrow 1.$$

The limit in (13) is recognized as the moment generating function of a Poisson variable with parameter $\mu_G$ and hence part (b) is established.

To prove (c), note that

$$\mathrm{E}\left[t^{Y_k+Z_k}\right] = \mathrm{E}\left[t^{Y_k}\mathrm{E}\left[t^{Z_k}|Y_k\right]\right].$$

By a computation analogous to the one used in deriving the moment generating function of $H_{(n,G)}$ above, we get for the inner expectation that

$$\mathrm{E}\left[t^{Z_k}|Y_k\right] \quad = \quad \left(1+\frac{\mu_{G_n}}{n-1}(t-1)\right)^{n-1-Y_k}$$

and hence

$$\mathrm{E}\left[t^{Y_k+Z_k}\right] = \left(1+\frac{\mu_{G_n}}{n-1}(t-1)\right)^{n-1}\mathrm{E}\left[\left(\frac{t(n-1)}{n-1+\mu_{G_n}(t-1)}\right)^{Y_k}\right].$$

As $n \to \infty$, the second factor converges to the generating function of $G$ and, as pointed out above, the limit of the first factor is the moment generating function of a Poisson distribution with mean $\mu_G$. Hence (c) is established.□

Our task here is to construct a graph with given asymptotic degree distribution $F$. Having proved that the asymptotic distribution of the vertex degree in the graph model $\mathcal{G}(n,G)$ is $G{*}\mathrm{Po}(\mu_G)$, the obvious question is which distributions $F$ that can be characterized in this way. We will not give a full answer to this question – it is presumably difficult – but rather give a few examples of distributions $F$ that can indeed be obtained and also specify how the distribution $G$ should be constructed in these cases.

### Poisson distribution

The simplest example is to let $F$ be a Poisson distribution with parameter $\mu_F$. Clearly this is accomplished by choosing $G$ to be a Poisson distribution with parameter $\mu_F/2$.

### Mixed Poisson distribution

A mixed Poisson distribution can be obtained as a limiting degree distribution in the $\mathcal{G}(n,G)$-model given a certain condition (14) on the law of the parameter. To see this, first note that, since we require that $F = G{*}\mathrm{Po}(\mu_G)$, the moment generating functions of $F$ and $G$, denoted by $\psi_F(t)$ and $\psi_G(t)$ respectively, must satisfy

$$\psi_F(t) = \psi_G(t)e^{\mu_G(t-1)}.$$

Now assume that the desired degree distribution $F$ is mixed Poisson with parameter law $Q$ where $\mathrm{Supp}(Q) \subseteq \mathbb{R}^+$ and $\mu_Q < \infty$. Then

$$\psi_F(t) = \int_0^\infty e^{x(t-1)} dQ(x)$$

and, since $\mu_G = \mu_F/2 = \mu_Q/2$, we must have

$$\psi_G(t) = \int_0^\infty e^{(x-\mu_Q/2)(t-1)} dQ(x).$$

To ensure that this is the generating function of a probability distribution, let $\xi_Q = \inf \mathrm{Supp}(Q)$ and assume that

$$\xi_Q - \mu_Q/2 > 0. \tag{14}$$

Then we can write

$$\psi_G(t) = \int_0^\infty e^{y(t-1)} d\widetilde{Q}(y),$$

where $\widetilde{Q}(y) = Q(y + \mu_Q/2)$, that is, $\widetilde{Q}$ is the distribution $Q$ translated $\mu_Q/2$ units to the left. This means that $G$ is a mixed Poisson distribution with parameter distribution $\widetilde{Q}$. Hence, a mixed Poisson distribution with parameter distribution $Q$ that satisfies $\xi_Q - \mu_Q/2 > 0$, is obtained as a limiting degree distribution in $\mathcal{G}(n, G)$ by choosing $G$ to be mixed Poisson with parameter distribution $\widetilde{Q}$.

**Compound Poisson distribution**

The compound Poisson distribution specifies the law of a sum of a Poisson number of i.i.d. random variables and we will now see that such a distribution is possible to attain as degree distribution in $\mathcal{G}(n, G)$ in the limit of large $n$ provided that the summand variables satisfy a certain condition (16).

Let $F$ be a compound Poisson distribution with Poisson parameter $\lambda$ and discrete summand distribution $R$ with finite mean $\mu_R$ and generating function $\psi_R(t)$. Then $\mu_F = \lambda\mu_R$ and

$$\psi_F(t) = e^{\lambda(\psi_R(t)-1)}.$$

If we want the limiting degree distribution in $\mathcal{G}(n, G)$ to be $F$, we must have

$$\begin{aligned} \psi_G(t) &= \psi_F(t)e^{-\mu_G(t-1)} \\ &= e^{\lambda(\psi_R(t)-1)-\mu_G(t-1)}. \end{aligned} \tag{15}$$

Here, since $\mu_G = \mu_F/2 = \lambda\mu_R/2$ and $\psi_R(t) = \sum_0^\infty r_k t^k$, where $\{r_k\}$ denotes the probabilities associated with $R$, the exponent in (15) can be written as

$$\lambda(\psi_R(t)-1) - \mu_G(t-1) = \lambda\left(\left(r_0 + \frac{\mu_R}{2}\right) + \left(r_1 - \frac{\mu_R}{2}\right)t + \sum_{k=2}^{\infty} r_k t^k - 1\right).$$

Now assume that

$$r_1 > \mu_R/2, \tag{16}$$

and introduce a new distribution $R'$ by defining

$$r'_k = \begin{cases} r_0 + \mu_R/2 & \text{for } k = 0; \\ r_1 - \mu_R/2 & \text{for } k = 1; \\ r_k & \text{for } k \geq 2, \end{cases}$$

that is, $R'$ is obtained by transferring the mass $\mu_R/2$ from the point 1 to the point 0 in the distribution $R$. Note that a consequence of (16) is that $R$ must be chosen so that $\mu_R < 2$. With $R'$ defined in this way we have

$$\psi_G(t) = e^{\lambda(\psi_{R'}(t)-1)},$$

that is, $G$ is a compound Poisson distribution with Poisson parameter $\lambda$ and summand distribution $R'$. Hence a compound Poisson distribution $F$ with summand distribution $R$ that satisfies (16) is obtained as limiting degree distribution in $\mathcal{G}(n, G)$ by choosing $G$ to be a compound Poisson distribution with summands distributed according to $R'$.

**Power law distribution**

Finally, we consider the case when the exact form of the degree distribution is not important, but only its tail-behavior. More precisely, it might be the case that our only demand on the degree distribution is that its tail should obey some kind of power-law. The following simple proposition states that, if $G$ is a power-law distribution, then the resulting degree distribution $G * \text{Po}(\mu_G)$ in the $\mathcal{G}(n, G)$-model will be so as well. Remember that '$\sim$' means that the quotient between the right hand side and the left hand side is a constant in the limit.

**Proposition 4.1** *Write* $G = \{g_k; k \geq 0\}$ *and* $G * \text{Po}(\mu_G) = \{p_k; k \geq 0\}$. *If* $g_k \sim k^{-\tau}$ *for some exponent* $\tau > 2$, *then* $p_k \sim k^{-\tau}$.

*Proof:* We have

$$
\begin{aligned}
p_k &= \sum_{n=0}^{k} g_{k-n} \frac{\mu_G^n}{n!} e^{-\mu_G} \\
&= k^{-\tau} \sum_{n=0}^{k} g_{k-n} \frac{\mu_G^n k^\tau}{n!} e^{-\mu_G}.
\end{aligned}
$$

Since $g_k \sim k^{-\tau}$, the last sum is clearly convergent as $k \to \infty$ and the proposition follows. $\square$

# 5 Concluding comments

In the present paper, four different ways of generating simple undirected graphs with a prescribed degree distribution are described. The methods are referred to as the erased configuration model, the repeated configuration model, the generalized random graph and the directed graph with removed directions (DGRD) respectively. None of the methods is able to produce a graph that has the desired distribution exactly – that is, in a finite graph, a randomly selected vertex will not have exactly the correct degree distribution – but under certain regularity assumptions, it is shown that all four methods gives the right distribution in the limit as the number of vertices $n$ tends to infinity.

Let us summarize the assumptions on the degree distribution for the different methods: In order for the repeated configuration model to produce a simple graph in stochastically bounded time as $n \to \infty$, the second moment of the degree distribution has to be finite and for the generalized random graph model to be applicable, finite moment of order $1 + \varepsilon$ for some $\varepsilon > 0$ is required. For the other two methods, finite mean is sufficient to ensure that the degree distribution converges to the prescribed one. As for the class of achievable distributions, the erased configuration model and the repeated configuration model are both able to generate graphs with any limiting distribution, provided it has finite mean and finite second moment respectively. The generalized random graph model can only produce mixed Poisson distributions and the DGRD-model gives distributions that can be expressed as the convolution of a discrete distribution with finite mean and a Poisson distribution with the same mean – a class containing certain types of mixed Poisson and compound Poisson distributions for instance. However, if only tail properties of the desired distribution are specified, both the generalized random graph model and the DGRD-model can do the job, given of course that the distribution has finite $1 + \varepsilon$-moment and finite mean respectively.

Concerning the number of operations needed to produce the graph, it is easily seen to be of order $n$ for all methods except the generalized random graph, which requires $O(n^2)$ operations. Note however that an approximation of the generalized random graph model that uses only $O(n)$ operations can be obtained by replacing the conditional degree distribution of a given vertex, conditional on the $W_i$:s, with a Poisson distribution with the same mean. Among the methods using $O(n)$ operations the erased configuration model and the DGRD-model require less operations: The repeated configuration model generates a graph a geometrically distributed number of times whereas the other two methods only generates a graph once and then erases a few edges.

A perhaps more subjective opinion is that the generalized random graph model is probabilistically more tractable than the other methods. Its construction is straightforward, implying that it is easier to show property results for this model. For example, without much extra effort, it was shown that the empirical degree distribution converges to the prescribed distribution, a result that was only shown for the generalized random graph model and – under the additional assumption of finite second moment – for the repeated configuration model. Also, it was easily seen from the construction of the general random graph that the obtained graph is uniform in the sense that all graphs with a given degree sequence have the same probability.

Finally we mention that, apart from the degree distribution, there is of course a number of other properties of a graph that it would also be desirable to have control over. When modelling a social network for instance, the clustering and path length is essential and, in some contexts, various types of preferential attachment algorithms are appropriate. Rigorous analyzes of algorithms incorporating these aspects is to a large extent lacking and it is a task for the future to further investigate and develop methods for generating random graphs.

# References


Arratia, R. and Liggett, T.M. (2005): How likely is an i.i.d. degree sequence to be graphical? *Ann. Appl. Prob.* **15**, 652-670.

Barabási, A.-L. and Albert, R. (1999): Emergence of scaling in random networks, *Science* **286**, 509-512.

Bender, E. and Canfield, E. (1978): The asymptotic number of labelled graphs with given degree sequences, *J. Comb. Th. A* **24**, 296-307.

Bollobás, B. (2001): *Random Graphs*, Academic Press, 2nd ed.

Chung, F. and Lu, L. (2002:1): Connected components in random graphs with given degrees sequences, *Ann. Comb.* **6**, 125-145.

Chung, F. and Lu, L. (2002:2): The average distances in random graphs with given expected degrees, *Proc. Natl. Acad. Sci.* **99**, 15879-15882.

Dorogovtsev, S. and Mendes, J. (2003): *Evolution of Networks, from Biological Nets to the Internet and WWW*, Oxford University Press.

Erdős, P. and Rényi, A. (1959): On random graphs, *Publ. Math.* **6**, 290-297.

Erdős, P. and Rényi, A. (1960): On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17-61.

Faloutsos, M., Faloutsos P. and Faloutsos, C. (1999): On power-law relationships of the internet topology, *Comp. Comm. Rev.* **29**, 251-262.

Janson, S., Luczak, T. and Rucinski, A. (1999): *Random Graphs*, John Wiley.

McKay, B. (1985): Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars. Comb.* **19A**, 15-20.

Liljeros, F., Edling, C., Amaral, L., Stanley, H. and Åberg, Y. (2001): The web of human sexual contacts, *Nature* **411**, 907-908.

Molloy, M. and Reed, B. (1995): A critical point for random graphs with a given degree sequence, *Rand. Struct. Alg.* **6**, 161-179.

Molloy, M. and Reed, B. (1998): The size of the giant component of a random graphs with a given degree sequence, *Comb. Prob. Comp.* **7**, 295-305.

Newman (2003): The structure and function of complex networks, *SIAM Rev.* **45**, 167-256.

Newman, M., Strogatz, S. and Watts, D. (2001): Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E* **64**, 026118.