Random Graphs

SVANTE JANSON
Uppsala University
Sweden

TOMASZ ŁUCZAK
Adam Mickiewicz University
Poland

ANDRZEJ RUCINSKI
Emory University
Atlanta, Georgia
Preface

The theory of random graphs originated in a series of papers published in the period 1959–1968 by two outstanding Hungarian mathematicians, Paul Erdős and Alfred Rényi. Over the forty years that have passed since then, the theory has developed into an independent and fast-growing branch of discrete mathematics, located at the intersection of graph theory, combinatorics and probability theory, with applications to theoretical computer science, reliability of transportation and communication networks, natural and social sciences and to discrete mathematics itself. Aside from applications, random graphs continue to serve as nontrivial, but simple enough models for other, more complex random structures, paving the road for more advanced theories.

In the early days, the literature on the subject was scattered around several probabilistic, combinatorial and general mathematics journals. In the late seventies, Béla Bollobás became the leading scientist in the field and contributed dozens of papers, which gradually made up a framework for his excellent, deep and extensive monograph *Random Graphs*, printed in 1985. The appearance of that book stimulated the research even further, shaping up a new theory.

Two other ingredients that added to this trend were the ongoing series of international conferences on random graphs and probabilistic methods in combinatorics held biennially in Poznań, Poland, since 1983, and the journal, *Random Structures and Algorithms*, launched by Wiley in 1990. Both have established a forum for the exchange of ideas and cooperation in the theory of random graphs and related fields.
It is not accidental then that tremendous progress has been made since 1985. Over the last decade several new, beautiful results have been proved and numerous fine techniques and methods have been introduced. Our goal is to present many of these new developments, including results on threshold functions (Ch. 1), small subgraphs (Ch. 3), generalized matchings (Ch. 4), phase transition (Ch. 5), limit distributions (Ch. 6), chromatic number (Ch. 7), partition and extremal properties (Ch. 8), Hamiltonian cycles in random regular graphs (Ch. 9), and zero-one laws (Ch. 10). We emphasize new techniques and tools such as the martingale, Talagrand and correlation inequalities (Ch. 2), the orthogonal decomposition (Ch. 6), the Regularity Lemma of Szemerédi (Ch. 8), the Contiguity Theorem (Ch. 9), and the analysis of variance (Ch. 9).

In a sense, our book can be viewed as an update on Bollobás's 1985 book. However, the topics selected for the book reflect the interest of its authors and do not pretend to exhaust the entire field. In fact, in order not to duplicate Bollobás's work, we do not include subjects which are covered there, on which only a little progress has been made. In particular, we have no sections on degree sequences, long paths and cycles, automorphisms, and the diameter. Moreover, we restrict ourselves to the main core of the theory and focus on the basic models of random graphs, making no attempt to present such rapidly developing areas as random walks on graphs, randomized algorithms or complexity of Boolean functions. Likewise, we exclude random cubes, directed graphs and percolation.

It has been our goal to make the book accessible to graduate students in mathematics and computer science. This has led to simplifications of some statements and proofs, which, we hope, result in better clarity of exposition. The book may be used as a textbook for a graduate course or an honors course for undergraduate senior mathematics and computer science majors. Although we do not provide problems and exercises separately, we often leave to the reader to complete parts of proofs or to provide proofs of results analogous to those proven. These instances, marked by the parenthetic phrase "(Exercise!)", can easily be picked up by the instructor and turned into homework assignments. The prerequisites are limited to basic courses in graph theory or combinatorics, elementary probability and calculus. We believe that the book will also be used by scientists working in the broad area of discrete mathematics and theoretical computer science. It is both an introduction for newcomers and a source of the most recent developments for those working in the field for many years.

We would like to thank several friends and colleagues, without whom this book would be a.a.s. worse than it is. Among those whose insightful remarks and suggestions led to improvements of earlier drafts are: Andrzej Czygrinow, Dwight Duffus, Ehud Friedgut, Johan Jonasson, Michał Karoński, Yoshiharu Kohayakawa, Michael Krivelevich, Justyna Kurkowiac, Jiří Matoušek, Brendan Nagle, Yuejian Peng, Joanna Polcyn, Vojtěch Rödl, Jozef Skokan, Joel Spencer, Edyta Szymańska, Michelle Wagner, and Julie White.
Special thanks are due to Penny Haxell and Izolda Gorgol. Penny spent several days correcting our English. Without her tedious work the text would probably need subtitles to be understood by an American reader. Izolda generously exercised her editing skills providing us with electronic files of all figures.

Jessica Downey, the Wiley editor, earned our deep appreciation for her continuous enthusiasm and support of the project.

Finally, the three authors would like to thank each other for patience, mutual encouragement and persistence in negotiations, the compromising effect of which is now in your hands.

Svante Janson
Tomasz Łuczak
Andrzej Ruciński

Uppsala, Poznań, and Atlanta
# Contents

 Preface \hspace{12cm} v  

 1 Preliminaries \hspace{1cm} 1  
  1.1 Models of random graphs \hspace{1cm} 1  
  1.2 Notes on notation and more \hspace{1cm} 6  
  1.3 Monotonicity \hspace{1cm} 12  
  1.4 Asymptotic equivalence \hspace{1cm} 14  
  1.5 Thresholds \hspace{1cm} 18  
  1.6 Sharp thresholds \hspace{1cm} 20  

 2 Exponentially Small Probabilities \hspace{1cm} 25  
  2.1 Independent summands \hspace{1cm} 26  
  2.2 Binomial random subsets \hspace{1cm} 30  
  2.3 Suen's inequality \hspace{1cm} 34  
  2.4 Martingales \hspace{1cm} 37  
  2.5 Talagrand's inequality \hspace{1cm} 39  
  2.6 The upper tail \hspace{1cm} 48  

 3 Small Subgraphs \hspace{1cm} 53  
  3.1 The containment problem \hspace{1cm} 53  

  ix
## CONTENTS

3.2 Leading overlaps and the subgraph plot 62  
3.3 Subgraph count at the threshold 66  
3.4 The covering problem 68  
3.5 Disjoint copies 75  
3.6 Variations on the theme 77  

4 Matchings 81  
4.1 Perfect matchings 82  
4.2 G-factors 89  
4.3 Two open problems 96  

5 The Phase Transition 103  
5.1 The evolution of the random graph 103  
5.2 The emergence of the giant component 107  
5.3 The emergence of the giant: A closer look 112  
5.4 The structure of the giant component 121  
5.5 Near the critical period 126  
5.6 Global properties and the symmetry rule 128  
5.7 Dynamic properties 134  

6 Asymptotic Distributions 139  
6.1 The method of moments 140  
6.2 Stein's method: The Poisson case 152  
6.3 Stein's method: The normal case 157  
6.4 Projections and decompositions 162  
6.5 Further methods 176  

7 The Chromatic Number 179  
7.1 The stability number 179  
7.2 The chromatic number: A greedy approach 184  
7.3 The concentration of the chromatic number 187  
7.4 The chromatic number of dense random graphs 190  
7.5 The chromatic number of sparse random graphs 192  
7.6 Vertex partition properties 196  

8 Extremal and Ramsey Properties 201  
8.1 Heuristics and results 202  
8.2 Triangles: The first approach 209
8.3 The Szemerédi Regularity Lemma 212
8.4 A partition theorem for random graphs 216
8.5 Triangles: An approach with perspective 222

9 Random Regular Graphs 233
  9.1 The configuration model 235
  9.2 Small cycles 236
  9.3 Hamilton cycles 239
  9.4 Proofs 247
  9.5 Contiguity of random regular graphs 256
  9.6 A brief course in contiguity 264

10 Zero-One Laws 271
  10.1 Preliminaries 271
  10.2 Ehrenfeucht games and zero-one laws 273
  10.3 Filling gaps 285
  10.4 Sums of models 292
  10.5 Separability and the speed of convergence 301

References 307

Index of Notation 327

Index 331
1

Preliminaries

1.1 MODELS OF RANDOM GRAPHS

The notion of a random graph originated in a paper of Erdős (1947), which is considered by some as the first conscious application of the probabilistic method. It was used there to prove the existence of a graph with a specific Ramsey property.

The model introduced by Erdős is very natural and can be described as choosing a graph at random, with equal probabilities, from the set of all \(2^{\binom{n}{2}}\) graphs whose vertex set is \([n] = \{1, 2, \ldots, n\}\). In other words, it can be described as the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\Omega\) is the set of all graphs with vertex set \([n]\), \(\mathcal{F}\) is the family of all subsets of \(\Omega\), and for every \(\omega \in \Omega\)

\[\mathbb{P}(\omega) = 2^{-\binom{n}{2}}.\]

This probability space can also be viewed as the product of \(\binom{n}{2}\) binary spaces. In simple words, it is a result of \(\binom{n}{2}\) independent tosses of a fair coin, where "turning up heads" means "drawing an edge".

Generally speaking, a random graph is a graph constructed by a random procedure. In accordance with standard definitions in probability theory, this is formalized by representing the "random procedure" by a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and the "construction" by a function from the probability space into a suitable family of graphs. The distribution of a random graph is the induced probability distribution on the family of graphs; for many purposes this is the only relevant feature of the construction and we usually do not distinguish between different random graphs with the same distribution. Indeed, it is
often convenient to define a random graph by specifying its distribution; that is, we specify a family of graphs and a probability distribution on it. Note, however, that it is not sufficient to formally define a random graph as a probability distribution only, as is sometimes done in the literature; an important case in which this would not do is when several random graphs are considered at once, for example, in the two-round exposure described at the end of this section.

The word "model" is used rather loosely in the theory of random graphs. It may refer to a specific class of random graphs, defined as above, or perhaps to a specific distribution. Usually, however, there is also a parameter involved which measures the size of the graphs and typically it tends to infinity; there may also be other parameters. Needless to say, the whole theory of random graphs is thus asymptotic in its nature.

Two basic models

Nowadays, among several models of random graphs, there are two basic ones, the binomial model and the uniform model, both originating in the simple model introduced by Erdős (1947). In this book we will mainly restrict ourselves to studying these two models.

Given a real number \( p, 0 \leq p \leq 1 \), the binomial random graph, denoted by \( \mathbb{G}(n, p) \), is defined by taking as \( \Omega \) the set of all graphs on vertex set \( [n] \) and setting

\[
\mathbb{P}(G) = p^{e_G} (1 - p)^{\binom{n}{2} - e_G},
\]

where \( e_G = |E(G)| \) stands for the number of edges of \( G \). It can be viewed as a result of \( \binom{n}{2} \) independent coin flippings, one for each pair of vertices, with the probability of success (i.e., drawing an edge) equal to \( p \). For \( p = 1/2 \) this is the model of 1947. However, most of the random graph literature is devoted to cases in which \( p = p(n) \to 0 \) as \( n \to \infty \).

The binomial model is a special case of a reliability network. In this more general model, \( \Omega \) is the family of all spanning subgraphs of a given graph \( F \) and \( \mathbb{P}(G) = p^{e_G} (1 - p)^{e_F - e_G} \). By a spanning subgraph we mean a graph \( G \) such that \( V(G) = V(F) \) and \( E(G) \subseteq E(F) \). Thus, in a reliability network, the edges of a given graph (network) are independently destroyed, each with failure probability \( 1 - p \). One can generalize this model even further, by allowing different probabilities of failure at different edges. (Binomial models are sometimes called Bernoulli.)

Taking \( F = K_n \), the complete graph on \( n \) vertices, we obtain the model \( \mathbb{G}(n, p) \). Taking \( F = K_{m,n} \), the complete bipartite graph (here either \( m \) is a function of \( n \), or they are two independent parameters, typically both tending to infinity), we obtain the bipartite random graph \( \mathbb{G}(m, n, p) \). Other popular models, not discussed here, are those in which the initial graph \( F \) is the hypercube or the \( n \times n \) square lattice. The reliability network based
on the infinite square lattice belongs to percolation theory (Grimm
nett 1992a) which too, as all infinite models, is beyond the scope of
this book.

The main advantage of the binomial model $\mathbb{G}(n, p)$ is the indepen-
dence of presence of edges, but the drawback is that the number of edges is not
fixed; it varies according to a binomial distribution with expectation $\binom{n}{2}p$. If
one conditions on the event that $|E(\mathbb{G}(n, p))| = M$, then a uniform space is
obtained. This space can be defined directly.

Given an integer $M, 0 \leq M \leq \binom{n}{2}$, the uniform random graph, denoted
by $\mathbb{G}(n, M)$, is defined by taking as $\Omega$ the family of all graphs on the vertex set
$[n]$ with exactly $M$ edges, and as $\mathbb{P}$ the uniform probability on $\Omega$,

$$
\mathbb{P}(G) = \left( \frac{\binom{n}{2}}{M} \right)^{-1}, \quad G \in \Omega.
$$

This model, closely related to enumerative combinatorics, was apparently
considered already in 1939 in an unpublished work of Erdős and Whitney on
the connectedness of almost all graphs with $n$ vertices and about $M = \frac{1}{2}n \log n$
edges. This was the model used throughout by Erdős and Rényi in their series of
papers between 1959 and 1968, which gave rise to the theory of random
graphs. (For an account of the contents of these eight fundamental papers,
see Karoński and Ruciński (1997).)

The two basic models are in many cases asymptotically equivalent, provided
$\binom{n}{2}p$ is close to $M$ (see Section 1.4).

The uniform random graph $\mathbb{G}(n, M)$ belongs to a broad family of uniform
random graphs defined by taking the uniform distribution over a family of
graphs $\mathcal{F}$. The pioneering model from Erdős (1947) belongs here too, with
$\mathcal{F}$ being the family of all graphs on a given set of vertices. Other popular
models of this type are random trees (not studied in this book), where $\mathcal{F}$
is the family of all $n^{n-2}$ trees on $n$ labeled vertices, and random $r$-regular
graphs (see Chapter 9), where $\mathcal{F}$ is the family of all graphs on $n$ vertices of
equal degree $r$, provided $nr$ is even. We will use $\mathbb{G}(n, r)$ to denote a uniform
random $r$-regular graph. It may look dangerous to use the notation $\mathbb{G}(n, p)$,
$\mathbb{G}(n, M)$ and $\mathbb{G}(n, r)$ for three different things: What is $\mathbb{G}(n, 1)$? In practice,
however, the correct meaning is always clear from the context. (As for the
three models: $\mathbb{G}(n, p)$ with $p = 1$, $\mathbb{G}(n, M)$ with $M = 1$, and $\mathbb{G}(n, r)$ with
$r = 1$, each one is rather dull.)

Both the binomial and the uniform model have their counterparts for di-
rected graphs. Besides these, there are interesting, natural random directed
graphs which do not have analogues in the undirected case. Let us mention
the $k$-out model, in which every vertex independently chooses $k$ out-neighbors
(including or excluding itself); the case of random mappings (i.e., $k = 1$) is
well studied (Kolchin 1986, Aldous and Pitman 1994). Random tournaments,
in which every edge of a complete graph assumes randomly one of the two pos-
sible orientations, have a broad literature too (Moon 1968, Gruszka, Łuczak
There are still other random graphs which do not fall into either category (binomial or uniform). For instance, in some reliability networks the vertices but not the edges are destroyed. Furthermore, some random graphs result from more complex probabilistic experiments, and here the sky is the limit. Restricted random graph processes constitute an interesting class of such experiments, but we should better define the unrestricted case first.

Random graph processes

In general, a random graph process is a stochastic process that describes a random graph evolving in time. In other words, it is a family \( \{G(t)\}_t \) of random graphs (defined on a common probability space) where the parameter \( t \) is interpreted as time; the time can be either discrete or continuous. The processes studied here will have a fixed vertex set (typically \([n]\)), and they will start without any edges and grow monotonically by adding edges according to some rule but never deleting any.

A simple and important random graph process \( \{G(n,M)\}_M \) (sometimes called the random graph process) was introduced by Erdős and Rényi (1959) and has been well studied since then. It begins with no edges at time 0 and adds new edges, one at a time; each new edge is selected at random, uniformly among all edges not already present. Hence this random graph process is a Markov process, with time running through the set \( \{0, 1, \ldots, \binom{n}{2}\} \). The \( M \)-th stage of this process can be identified with the uniform random graph \( G(n,M) \). The process, however, allows one to study the random graph \( G(n,M) \) as it evolves with \( M \) growing from 0 to \( \binom{n}{2} \). For example, a typical result, meaningful only for random graph processes, says that, with probability approaching 1 as \( n \to \infty \), the very edge which links the last isolated vertex with another vertex makes the graph connected (Bollobás and Thomason (1985); see also Bollobás (1985)).

A related continuous time random graph process can be defined by assigning a random variable \( T_e \) to each edge \( e \) of the complete graph \( K_n \), such that the \( \binom{n}{2} \) variables \( T_e \) are independent with a common continuous distribution, and then defining the edge set of \( \{G(t)\}_t \) to consist of all \( e \) with \( T_e \leq t \). Clearly, the resulting random graph \( \{G(t)\}_{t_0} \) at a fixed time \( t_0 \) can be identified with the binomial random graph \( G(n,p) \), where \( p = \mathbb{P}(T_e \leq t_0) \). Furthermore, since almost surely no two values of the random variables \( T_e \) coincide, we may define \( T_{(i)} \) as the random time at which the \( i \)-th edge is added. Then, by symmetry, \( G(T_{(i)}) \) is the uniform random graph \( G(n,i) \), and the sequence \( \{G(T_{(i)})\} \) for \( i = 1, \ldots, \binom{n}{2} \), equals the ordinary random graph process \( \{G(n,M)\}_M \) defined above. Hence, this continuous time random graph process is a joint generalization of the binomial random graph, the uniform random graph and the standard discrete time random graph process.

Clearly, different choices of the distribution of \( T_e \) affect the model only trivially, by a change in the time variable. The continuous time evolving
model was introduced by Stepanov (1970) with $T_e$ exponentially distributed; we prefer the uniform distribution over the interval $[0, 1]$, in which case $p = \mathbb{P}(T_e \leq t) = t$, $0 \leq t \leq 1$. Thus, we may unambiguously use the notation $\{G(n, t)\}_t$.

Recently, a number of restricted random graph processes have been studied. In general, such a process can be defined as a random graph process in which edges are chosen one by one uniformly from a dynamically modified set of available pairs of vertices until this set becomes empty. More formally, consider a Markov chain of random edge sets $E_0 = \emptyset, E_1, \ldots, E_s$, where $E_i = \{e_1, \ldots, e_i\}$ and $e_i$ is chosen uniformly from a set $A_i$ which depends only on the set $E_{i-1}$.

In one of these restricted models, studied by Ruciński and Wormald (1992), the maximum degree is bounded from above by a given integer $d$. Thus, the set $A_i$ contains only those pairs whose addition to the set $E_{i-1}$ does not create a vertex of degree $d + 1$. The graph at the end of the process may not be $d$-regular, though it is shown to be so with probability approaching 1. See also Wormald (1999a), where, moreover, further related processes are defined and studied.

Another restricted process is studied by Erdős, Suen and Winkler (1995), in which it is not allowed to create a triangle. In this model it is even an open problem to determine the length of a typical process, measured by the number of edges in the final graph. It is only known that with high probability the process takes more than $c_1 n^{3/2}$ but fewer than $c_2 n^{3/2} \log n$ steps, where $c_1$ and $c_2$ are positive constants. Recently, this result was generalized to a wide class of forbidden subgraphs by Osthus and Taraz (2000+).

By forbidding cycles, one obtains a process which creates a non-uniform random tree (Aldous 1990), while forbidding components with more than one cycle leads to a random graph which still is to be studied.

**Random subsets**

The two basic models of random graphs fall into the framework of random subsets of a set. Monotonicity, equivalence and threshold behavior of the probabilities of properties of random graphs can often been proved at no extra cost in this general setting. Other principal examples of random subsets of a set include random sets of integers and random hypergraphs. In the remaining sections of this chapter (as well as in parts of Chapter 2) we will mainly study this more general random set framework. For an arbitrary set $X$ and an integer $k$, let $[X]^k$ stand for the family of all $k$-element subsets of $X$. If $X = [n]$, we will simplify this notation to $[n]^k$.

Let $\Gamma$ be a finite set, $|\Gamma| = N$, let $0 \leq p \leq 1$ and $0 \leq M \leq N$. Then the random subset $\Gamma_p$ of $\Gamma$ is obtained by flipping a coin, with probability $p$ of success, for each element of $\Gamma$ to determine whether the element is to be included in $\Gamma_p$; the distribution of $\Gamma_p$ is the probability distribution on $\Omega = 2^\Gamma$ given by $\mathbb{P}(F) = p^{|F|}(1 - p)^{|\Gamma| - |F|}$ for $F \subseteq \Gamma$. Similarly, let $\Gamma_M$ be
a randomly chosen element of \([\Gamma]^M\); that is, \(\Gamma_M\) has the uniform distribution \(\mathbb{P}(F) = \binom{M}{F}^{-1}\) for \(F \in [\Gamma]^M\).

Taking \(\Gamma = [n]^2\) we obtain the two basic models of random graphs defined above, \(\mathbb{G}(n,p)\) and \(\mathbb{G}(n,M)\).

The binomial model \(\Gamma_p\) can be generalized to \(\Gamma_{p_1 \ldots p_N}\), where the element \(i\) is included with probability \(p_i\), independently for all \(i = 1, \ldots, N\).

Two-round exposure

The two-round exposure is a successful proof technique applicable to the binomial model. It relies on viewing \(\Gamma_p\) as a union of two independent random subsets \(\Gamma_{p_1}\) and \(\Gamma_{p_2}\), where \(p_1\) and \(p_2\) are such that \(p = p_1 + p_2 - p_1 p_2\). (It is easy to see that this union indeed is distributed as \(\Gamma_p\) – Exercise!) In the special case of random graphs we first generate a random graph \(\mathbb{G}(n,p_1)\) and then, independently, another random graph \(\mathbb{G}(n,p_2)\) on the same vertex set. By replacing double edges by single ones, we obtain \(\mathbb{G}(n,p)\).

An argument typically used in applications of the two-round exposure can be expressed in the following general form. Let \(\mathbb{P}_1\) be the probability distribution associated with \(\Gamma_{p_1}\), and let \(\mathbb{P}_F\) be the conditional probability in \(\Gamma_p\) under the condition \(\Gamma_{p_1} = F\). Then for any two families \(A\) and \(B\) of subsets of \(\Gamma\)

\[
\mathbb{P}(A) \geq \sum_{F \in B} \mathbb{P}_F(A) \mathbb{P}_1(F) \geq \mathbb{P}_{F_0}(A) \mathbb{P}_1(B),
\]

where \(F_0\) minimizes the probability \(\mathbb{P}_F(A)\) over all \(F \in B\). Thus, knowing that \(\mathbb{P}_1(B) \to 1\), in order to prove that also \(\mathbb{P}(A) \to 1\), it is enough to show that \(\mathbb{P}_F(A) \to 1\), for every \(F \in B\). In practice, computing the last probability means fixing an instance of \(\Gamma_{p_1} \in B\) and throwing in new elements independently with probability \(p_2\) (the second round of exposure).

1.2 NOTES ON NOTATION AND MORE

Graph theory

All graphs are simple and undirected, unless otherwise stated. We use standard notation for graphs. For example, \(V(G)\) is the vertex set of a graph \(G\), \(E(G)\) is the edge set, \(v_G = |V(G)|\) is the number of vertices and \(e_G = |E(G)|\) is the number of edges; for typographical reasons we sometimes write the latter two as \(v(G)\) and \(e(G)\). In this book the size of \(G\) always means \(v(G)\) (and not \(e(G)\) as sometimes used by other authors). However, we also will call \(v(G)\) the order of \(G\).

Moreover, let \(d(G) = e_G/v_G\) be the density and \(m(G) = \max_{H \subseteq G} d(H)\) the maximum density of \(G\). (Note that \(d(G)\) equals half the average degree of \(G\), and that some authors define \(d(G)\) as the average degree, which is twice
Another measure of the density of a graph $G$, ranging between 0 and 1, is defined as $\rho(G) = e(G)/\binom{|V(G)|}{2}$. (It is sometimes called the relative density of $G$.)

Furthermore, $\delta(G)$ is the minimum degree, $\Delta(G)$ is the maximum degree, $\chi(G)$ is the chromatic number, $D(G) = \max_{H \subseteq G} \delta(H)$ is the degeneracy number, $\alpha(G)$ is the stability number (the size of the largest stable, or independent, set of vertices), and $\text{aut}(G)$ is the number of automorphisms of $G$.

We let $N(v) = N_G(v)$ denote the neighborhood of a vertex $v$ in $G$, that is, the set $\{w \in V(G) : vw \in E(G)\}$. Its size is called the degree of $v$ and is denoted by $\deg(v) = \deg_G(v)$. Similarly, if $S \subseteq V(G)$, its neighborhood $N_G(S) = \bigcup_{v \in S} N_G(v) \setminus S$ is the set of all vertices outside $S$ adjacent to at least one vertex in $S$. Moreover, we let $\overline{N}_G(v) = N_G(v) \cup \{v\}$ and $\overline{N}_G(S) = N_G(S) \cup S$ denote the corresponding closed neighborhoods, which include $v$ and $S$, respectively.

Any graph without edges will be called empty, while the graph with no vertices (and thus no edges) will be called the null graph and denoted by $\emptyset$.

Some special graphs are: the complete graph $K_n$ on $n$ vertices, the complete bipartite graph $K_{m,n}$ on $m + n$ vertices, the cycle $C_k$ with $k$ vertices, and the path $P_k$ with $k$ edges and thus $k + 1$ vertices. A star is any graph $K_{1,n}$, $n \geq 0$. We let $jG$ denote the union of $j$ vertex-disjoint copies of $G$. A matching is a forest consisting of isolated edges only (i.e., a graph of the form $jK_2$, $j \geq 0$).

If $G$ is a graph and $V \subseteq V(G)$, then $G[V]$ denotes the restriction of $G$ to $V$, defined as the graph with vertex set $V$ and edge set $E(G) \cap [V]^2$; similarly, if $E \subseteq [V(G)]^2$, $G[E]$ denotes the graph with vertex set $V(G)$ and edge set $E(G) \cap E$. A subgraph of $G$ of the type $G[V]$ is called induced or spanned by $V$, while a subgraph of the type $G[E]$ is called spanning. The number of edges in the subgraph $G[V]$ is sometimes denoted by $e_G(V) = e(V)$, while for two disjoint subsets $A, B \subseteq V(G)$, the quantity $e_G(A, B)$ counts the number of edges of $G$ with one endpoint in $A$ and the other in $B$.

By a copy of a given graph $G$ inside another graph $F$ we mean any, not necessarily induced, subgraph of $F$ which is isomorphic to $G$. If the subgraph happens to be induced, we call it an induced copy of $G$.

Although we define our random graphs as labelled, we are mainly interested in properties that are independent of the labelling, that is, properties that depend on the isomorphism type only. Such properties are called graph properties. (In contrast, "vertex 1 is isolated" is not a graph property; such properties will occasionally be studied too.)

**Probability**

We use $\text{Bi}(n, p)$, $\text{Be}(p) = \text{Bi}(1, p)$, $\text{Po}(\lambda)$ and $\text{N}(\mu, \sigma^2)$ to denote the binomial, Bernoulli, Poisson and normal distributions, respectively. We further write $X \in \mathcal{L}$, meaning that $X$ is a random variable with distribution $\mathcal{L}$ (e.g., $X \in$...
The distribution of a random variable $X$ is occasionally denoted by $\mathcal{L}(X)$.

We denote by $\mathbf{1}[\mathcal{E}]$ the indicator function of the event $\mathcal{E}$, which equals 1 if $\mathcal{E}$ occurs and 0 otherwise. We will often consider random variables that are the indicator functions of some events; such random variables will be called indicator or zero-one random variables. They clearly have Bernoulli distributions with $p = \mathbb{P}(\mathcal{E})$, where $\mathcal{E}$ is the corresponding event.

The expected value and the variance of a random variable $X$ (if they exist) will be denoted by $\mathbb{E}X$ and $\text{Var}X$, respectively. Thus, the well-known Chebyshev’s inequality, which will be frequently used throughout the book, can be stated in the following, standard form. If $\text{Var}X$ exists, then

$$
\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \frac{\text{Var}X}{t^2}, \quad t > 0.
$$

(1.2)

Similarly, Markov’s inequality states that, if $X \geq 0$ a.s., then

$$
\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}, \quad t > 0.
$$

(1.3)

We denote the covariance of two random variables $X$ and $Y$ by $\text{Cov}(X, Y)$. Recall that the variance of a (finite) sum of random variables is given by $\text{Var}(\sum_i X_i) = \sum_i \sum_j \text{Cov}(X_i, X_j)$.

The conditional expectation of $X$ given an event $\mathcal{E}$ is denoted by $\mathbb{E}(X \mid \mathcal{E})$. We similarly write $\mathbb{E}(X \mid Y_1, \ldots, Y_k)$ for the conditional expectation of $X$ given some random variables $Y_1, \ldots, Y_k$; note that this conditional expectation is a function of $(Y_1, \ldots, Y_k)$ and thus itself a random variable. When using martingales (Section 2.4), we will more generally denote by $\mathbb{E}(X \mid \mathcal{G})$ the conditional expectation of $X$ given a sub-$\sigma$-algebra $\mathcal{G}$ of $\mathcal{F}$.

Quite frequently our proofs will rely on the elementary law of total probability which states that for any partition of the probability space $\Omega = \mathcal{E}_1 \cup \mathcal{E}_2 \ldots$ and any random variable $X$ defined on $\Omega$,

$$
\mathbb{E}X = \sum_i \mathbb{E}(X \mid \mathcal{E}_i)\mathbb{P}(\mathcal{E}_i).
$$

In particular, if $X = \mathbf{1}[\mathcal{E}]$, then $\mathbb{P}(\mathcal{E}) = \sum_i \mathbb{P}(\mathcal{E} \mid \mathcal{E}_i)\mathbb{P}(\mathcal{E}_i)$.

If $X_1, X_2, \ldots$ are random variables and $a$ is a constant, we say that $X_n$ converges in probability to $a$ as $n \to \infty$, and write $X_n \xrightarrow{P} a$, if $\mathbb{P}(|X_n - a| > \varepsilon) \to 0$ for every $\varepsilon > 0$; see, for example, Gut (1995, Chapter VI).

One similarly defines $X_n \xrightarrow{P} Y$, where $Y$ is another random variable, but then $Y$ and every $X_n$ have to be defined on the same probability space; this can be reduced to the preceding case, since $X_n \xrightarrow{P} Y$ if and only if $X_n - Y \xrightarrow{P} 0$.

Let $X_1, X_2, \ldots$ and $Z$ be random variables. We say that $X_n$ converges in distribution to $Z$ as $n \to \infty$, and write $X_n \xrightarrow{d} Z$, if $\mathbb{P}(X_n \leq x) \to \mathbb{P}(Z \leq x)$ for every real $x$ that is a continuity point of $\mathbb{P}(Z \leq x)$ (Billingsley 1968, Gut 1995).
If $X_1, X_2, \ldots$ and $Z$ are integer-valued then, equivalently, $X_n \xrightarrow{d} Z$ if and only if $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(Z = k)$ for every integer $k$.

Note that convergence in distribution is really a property of the distributions of the random variables and does not require the variables to be defined on the same probability space. Nevertheless, it is customary (and convenient) to talk about convergence of random variables. We also use hybrid notation such as $X_n \xrightarrow{d} \mathcal{N}(0,1)$, which means $X_n \xrightarrow{d} Z$ for some (and thus every) random variable $Z \in \mathcal{N}(0,1)$.

An important special case is one in which $Z$ is a (non-random) real constant. It is easily shown that convergence in distribution to a constant is the same as convergence in probability, that is, $X_n \xrightarrow{d} a$ if and only if $X_n \xrightarrow{p} a$ for $a \in \mathbb{R}$. A useful fact is that if $X_n \xrightarrow{d} Z$ and $Y_n \xrightarrow{p} a$, where $a$ is a constant, then $X_n + Y_n \xrightarrow{d} Z + a$ and $Y_n X_n \xrightarrow{d} aZ$ (Cramér's theorem), see, for example, Gut (1995, Theorem VI.7.5).

The definition of convergence in distribution extends to random vectors with values in $\mathbb{R}^k$ for every fixed $k$; this is also expressed as joint convergence in distribution of the components of the vectors. A powerful method for extending results on the real random variables to the vector-valued ones is known as the Cramér–Wold device (Billingsley 1968, Theorem 7.7). It states that $(X_{n1}, \ldots, X_{nk}) \xrightarrow{d} (Z_1, \ldots, Z_k)$ if and only if $\sum_i t_i X_{ni} \xrightarrow{d} \sum_i t_i Z_i$ for every sequence of real numbers $t_1, \ldots, t_k$. For more details, as well as for the convergence of random variables with values in even more general spaces, see Billingsley (1968).

**Remark 1.1.** Convergence in distribution does not, in general, imply convergence of the sequence of means or variances. However, in many specific applications we find that these sequences do, in fact, converge to the mean and variance of the limit distribution.

**Asymptotics**

We will often use the following standard notation for the asymptotic behavior of the relative order of magnitude of two sequences of numbers $a_n$ and $b_n$, depending on a parameter $n \rightarrow \infty$. The same notation is also used in other situations, for example, for functions of a variable $\varepsilon$ that tends to 0. We will often omit the phrase "as $n \rightarrow \infty$" when there is no risk of confusion. For simplicity we assume $b_n > 0$ for all sufficiently large $n$.

- $a_n = O(b_n)$ as $n \rightarrow \infty$ if there exist constants $C$ and $n_0$ such that $|a_n| \leq Cb_n$ for $n \geq n_0$, i.e., if the sequence $a_n/b_n$ is bounded, except possibly for some small values of $n$ for which the ratio may be undefined.

- $a_n = \Omega(b_n)$ as $n \rightarrow \infty$ if there exist constants $c > 0$ and $n_0$ such that $a_n \geq cb_n$ for $n \geq n_0$. If $a_n \geq 0$, this is equivalent to $b_n = O(a_n)$. 
PRELIMINARIES

- \( a_n = \Theta(b_n) \) as \( n \to \infty \) if there exist constants \( C, c > 0 \) and \( n_0 \) such that \( cb_n \leq a_n \leq Cb_n \) for \( n \geq n_0 \), i.e., if \( a_n = O(b_n) \) and \( a_n = \Omega(b_n) \). This is sometimes expressed by saying that \( a_n \) and \( b_n \) are of the same order of magnitude.

- \( a_n \asymp b_n \) if \( a_n = \Theta(b_n) \).

- \( a_n \sim b_n \) if \( a_n / b_n \to 1 \).

- \( a_n = o(b_n) \) as \( n \to \infty \) if \( a_n / b_n \to 0 \), i.e., if for every \( \varepsilon > 0 \) there exists \( n_\varepsilon \) such that \( |a_n| < \varepsilon b_n \) for \( n \geq n_\varepsilon \).

- \( a_n \ll b_n \) or \( b_n \gg a_n \) if \( a_n \geq 0 \) and \( a_n = o(b_n) \).

Since most results in this book are asymptotic, we will be frequently assuming in the proofs that \( n \) is sufficiently large, sometimes without explicitly saying so.

Probability asymptotics

We say that an event \( \mathcal{E}_n \), describing a property of a random structure depending on a parameter \( n \), holds asymptotically almost surely (abbreviated a.a.s.), if \( \mathbb{P}(\mathcal{E}_n) \to 1 \) as \( n \to \infty \).

**Remark 1.2.** In many publications on random structures the phrase “almost surely” or a.s. is used. However, we wish to reserve that phrase for what it normally means in probability theory, i.e. that the probability of an event equals exactly 1. It seems that the first paper where the phrase a.a.s. and not a.s. was used is Shamir and Upfal (1981). (Some authors use the phrase “almost every” or a.e. which we reject for the same reason as “almost surely”. Others write “with high probability”, or whp.)

When discussing asymptotics of random variables, we avoid expressions like “\( X_n = O(1) \) a.a.s.” or “\( X_n = o(1) \) a.a.s.”, which may be ambiguous, since they combine two asymptotic notions. As a substitute we give probabilistic versions of some of the symbols above, denoting them with a subscript \( p \) or \( C \). Let \( X_n \) be random variables and \( a_n \) positive real numbers. We then define:

- \( X_n = O_p(a_n) \) as \( n \to \infty \) if for every \( \delta > 0 \) there exist constants \( C_\delta \) and \( n_0 \) such that \( \mathbb{P}(|X_n| \leq C_\delta a_n) > 1 - \delta \) for every \( n \geq n_0 \).

- \( X_n = O_C(a_n) \) as \( n \to \infty \) if there exists a constant \( C \) such that a.a.s. \( |X_n| \leq Ca_n \).

- \( X_n = \Theta_p(a_n) \) as \( n \to \infty \) if for every \( \delta > 0 \) there exist constants \( \epsilon_\delta > 0 \), \( C_\delta > 0 \) and \( n_0 \) such that \( \mathbb{P}(\epsilon_\delta a_n \leq X_n \leq C_\delta a_n) > 1 - \delta \) for every \( n \geq n_0 \).

- \( X_n = \Theta_C(a_n) \) as \( n \to \infty \) if there exist positive constants \( c \) and \( C \) such that a.a.s. \( ca_n \leq X_n \leq Ca_n \).
• \(X_n = o_p(a_n)\) as \(n \to \infty\) if for every \(\epsilon > 0\), a.a.s. \(|X_n| < \epsilon a_n\).

Note that \(X_n = O_C(a_n)\) implies \(X_n = O_p(a_n)\), but not conversely; indeed, \(X_n = O_C(a_n)\) if and only if the constant \(C_\delta\) in the definition of \(O_p\) can be chosen independently of \(\delta\). For example, any sequence \(X_n\) of identically distributed random variables is \(O_p(1)\), but such a sequence is \(O_C(1)\) only if the common distribution has support in a finite interval.

Similarly, \(X_n = \Theta_C(a_n)\) implies \(X_n = \Theta_p(a_n)\), but not conversely. On the other hand, \(X_n = o_p(a_n)\) implies \(X_n = O_C(a_n)\).

**Remark 1.3.** It is easy to verify (Exercise!) that \(X_n = O_p(a_n)\) if and only if for every function \(\omega(n) \to \infty\), \(|X_n| \leq \omega(n)a_n\) a.a.s. Similarly, \(X_n = o_p(a_n)\) if and only if for some function \(\omega(n) \to \infty\), \(|X_n| \leq a_n/\omega(n)\) a.a.s.

Such notation with an unspecified sequence \(\omega(n)\) is common in publications on random structures, but we believe that the equivalent notation \(O_p\) and \(o_p\) is clearer.

It is an immediate consequence of the definitions (Exercise!) that \(X_n = o_p(a_n)\) if and only if \(X_n/a_n \xrightarrow{p} 0\). Conversely, \(X_n \xrightarrow{p} a\) if and only if \(X_n = a + o_p(1)\) (and \(X_n \xrightarrow{p} Y\) if and only if \(X_n = Y + o_p(1)\)).

**Remark 1.4.** The symbol \(O_p\) can also be expressed by equivalent standard probabilistic concepts. In fact, a sequence \(X_n\) is *bounded in probability*, or *tight*, if \(X_n = O_p(1)\). Hence, \(X_n = O_p(a_n)\) if and only if the sequence \(X_n/a_n\) is bounded in probability (or tight).

**Dependency graphs**

Let \(\{X_i\}_{i \in \mathcal{I}}\) be a family of random variables (defined on a common probability space). A **dependency graph** for \(\{X_i\}\) is any graph \(L\) with vertex set \(V(L) = \mathcal{I}\) such that if \(A\) and \(B\) are two disjoint subsets of \(\mathcal{I}\) with \(e_{\mathcal{I}}(A, B) = 0\), then the families \(\{X_i\}_{i \in A}\) and \(\{X_i\}_{i \in B}\) are mutually independent.

Dependency graphs will be used several times in this book. They are particularly useful when they are sparse, meaning that there is a lot of independence in the family \(\{X_i\}\).

**Example 1.5.** In a standard situation, there is an underlying family of independent random variables \(\{Y_\alpha\}_{\alpha \in A}\), and each \(X_i\) is a function of the variables \(\{Y_\alpha\}_{\alpha \in A_i}\) for some subset \(A_i \subseteq A\). Let \(S = \{A_i : i \in \mathcal{I}\}\). Then the graph \(L = L(S)\) with vertex set \(\mathcal{I}\) and edge set \(\{ij : A_i \cap A_j \neq \emptyset\}\) is a dependency graph for the family \(\{X_i\}_{i \in \mathcal{I}}\) (Exercise!).

**Example 1.6.** As a special case of the preceding example, let \(\{H_i\}_{i \in \mathcal{I}}\) be given subgraphs of the complete graph \(K_n\) and let \(X_i\) be the indicator that \(H_i\) appears as a subgraph in \(G(n, p)\), that is, \(X_i = 1[H_i \subseteq G(n, p)], i \in \mathcal{I}\). Then \(L(S)\), with \(S = \{E(H_i) : i \in \mathcal{I}\}\), is a natural dependency graph with edge set \(\{ij : E(H_i) \cap E(H_j) \neq \emptyset\}\) (Exercise!).

Remark 1.7. In particular, if $L$ is a dependency graph for $\{X_i\}$, then two variables $X_i$ and $X_j$ are independent unless there is an edge in $L$ between $i$ and $j$. Note, however, that this is only a necessary condition, and does not imply that $L$ is a dependency graph (Exercise!).

Remark 1.8. Another context, outside the scope of this book, in which dependency graphs are used is the Lovász Local Lemma (Erdős and Lovász (1975); see also Alon and Spencer (1992)). There it actually suffices to use a slightly weaker definition, considering only singletons $B$ in the definition above.

Remark 1.9. In our applications, there exists a natural dependency graph, but it should be observed that, in general, there is no canonical choice and the dependency graph is not unique, even if it is required to be minimal (Exercise!).

The subsequence principle

It is often convenient to use the well-known subsequence principle, which states that if for every subsequence of a sequence there is a subsequence converging to a limit $a$, then the entire sequence must converge to the same limit. This holds for sequences of real numbers, vectors, random variables (both for convergence in probability and for convergence in distribution) and, in general, for sequences in any topological space.

For example, this means that if we want to prove a limit theorem for $G(n,p)$, we may without loss of generality assume that an expression such as $n^a p^b$ converges to some $c \leq \infty$ (provided, of course, that the result we want to prove does not depend on the limit $c$).

We will be using this principle throughout the book (see, e.g., the proof of Proposition 1.15), sometimes without explicitly mentioning it.

And finally . . .

The base of all logarithms is $e$, unless specified otherwise.

1.3 MONOTONICITY

A family of subsets $Q \subseteq 2^\Gamma$ is called increasing if $A \subseteq B$ and $A \in Q$ imply that $B \in Q$. A family of subsets is decreasing if its complement in $2^\Gamma$ is increasing, or, equivalently, if the family of the complements in $\Gamma$ is increasing. A family which is either increasing or decreasing is called monotone. A family $Q$ is convex if $A \subseteq B \subseteq C$ and $A, C \in Q$ imply $B \in Q$. We identify properties of subsets of $\Gamma$ with the corresponding families of all subsets having the property; we thus use the same notation and terminology for properties.