Stochastic Processes and their Applications 26 (1987) 1-30
North-Holland

POISSON CONVERGENCE AND POISSON PROCESSES WITH APPLICATIONS TO RANDOM GRAPHS

Svante JANSON
Department of Mathematics, Uppsala University, Uppsala, Sweden

Received 18 March 1987

We give a new sufficient condition for convergence to a Poisson distribution of a sequence of sums of dependent variables. The condition allows each summand to depend strongly on a few of the other variables and to depend weakly on the remaining ones.

As a consequence we obtain sufficient conditions for the convergence of point processes, constructed as sets of (weakly) dependent random points in some space S, to a Poisson process.

The main applications are to random graph theory. In particular, we solve the problem (proposed by Erdös) of finding the size of the first cycle in a random graph.

Poisson limits * Poisson processes * point processes * random graphs * cycles * subgraph statistics

Introduction

It is well known that random variables that can be written as sums of a large number of indicator (zero-one) variables, each having a small probability of being non-zero, generally tend to be approximately Poisson distributed [20]. Results of this type, either limit theorems or quantitative estimates of the distance to a Poisson distribution, have been proved under various conditions by many authors. In the present paper two limit theorems of this type are given. In one of them (Theorem 1.2) it is assumed that each of the indicator variables is independent of all but a relatively small number of the others. Typical examples are U-statistics and some generalizations thereof. Earlier results of this type, wholly or partly contained in Theorem 1.2, have been given by Silverman and Brown [21], Eagleson [9], Berman and Eagleson [4], Barbour and Eagleson [2, 3], Jammalamadaka and Janson [12].

We also give a more general result (Theorem 1.1), partly inspired by the methods of Barbour [1], where the independence assumption is replaced by a certain condition saying that the variables are close to being independent. In typical applications (as in Sections 9 and 10), each variable depends rather strongly on a small number of the others, and weakly on the remaining ones.

The indicator variables in these theorems are not assumed to be identically distributed. The dependence assumptions are not stated in terms of "time" or any other ordering; i.e. they are not of Markov or mixing type.

These theorems have powerful generalizations to point processes. Consider a large number of weakly dependent random points in some set. We then expect their distribution to be approximated by a Poisson process. Limit theorems of that type are proved in Section 3, using the basic theorems in Section 1 and background material on point processes collected in Section 2.

Section 4 contains two generalizations of the Poisson convergence theorems in Section 1, both proved as applications of the theorems in Section 3 on convergence of point processes. Theorem 4.1 is a vector-valued version of Theorem 1.1, while Theorem 4.2 gives a limit theorem for sums of more general real-valued (i.e. not necessarily 0–1) random variables.

Section 5 discusses the special case of dissociated variables, which are important for applications. This also illuminates the relation to the earlier results referred to above.

The second part of the paper gives applications to random graphs. Section 6 contains some pertinent definitions.

Section 7 studies the number and lengths of the cycles in a random graph. One result (Corollary 7.4) is an exact formula for the asymptotic probability that the chromatic number of a random graph equals 2.

Section 8 gives a solution to the following problem by Erdös (communicated by Edgar Palmer to the Second Seminar on Random Graphs in Poznán, August 1985):

What is the size of the first cycle in a random graph?

(Edges are added one by one at random.)

It is shown (Theorem 8.1) that the distribution of the size of the first cycle converges (as the number of vertices tends to ∞) to an explicitly given limit distribution. The average size, however, tend tends to ∞.

Isolated cycles are studied in Section 9. It is shown (Corollary 9.4) that usually they do not appear at all.

We also consider the number of the first cycle (if any) that is isolated. We find (Theorem 9.2) another unusual asymptotic behaviour: about 42% of the mass vanishes off to infinity.

Section 10 gives an example of a subgraph such that the number of copies of it in a random graph is not asymptotically Poisson distributed, although a limit distribution exists. In this example, the limit distribution can be represented as a sum of a random number of Poisson variables.

It is also shown that, in contrast, the number of isolated copies of the graph is asymptotically Poisson distributed.

1. The fundamental Poisson convergence theorem

We consider a triangular array $\{X_{nj}\}_{j\in\mathcal{J}_n}$ of indicator variables. For convenience, we will usually omit n from the notations in the sequel, although (usually) the variables X_j and the index set $\mathcal{J}$, as well as other objects such as $\tilde{X}_{jk}$ in Theorem 1.1 and D_j in Theorem 1.2, depend on n.

We will prove convergence to a Poisson distribution of the sums $\sum X_j$ under conditions involving modifications $\tilde{X}_{kj}$ of X_k, chosen such that $\tilde{X}_{kj}$ is independent of X_j. While this makes the conditions somewhat complicated, we will show below that a simple choice of $\tilde{X}_{kj}$ gives a corollary (Theorem 1.2) which is easily applied in many situations, including several of those studied in the references cited earlier. We will also (in Sections 9 and 10) give applications where the full strength of Theorem 1.1 is used, and show how $\tilde{X}_{kj}$ may be chosen in such cases.

Theorem 1.1. *Let, for each n, $\{X_j\}_{j\in\mathcal{J}}$ be a (finite or infinite) family of indicator variables and assume that there exist random variables $\tilde{X}_{kj}$, j, $k\in\mathcal{J}$, such that X_j is independent of $\{\tilde{X}_{kj}\}_{k\neq j}$ for every $j\in\mathcal{J}$ and, as $n\to\infty$,*

$$\sum_{j\in\mathcal{J}} P(X_j=1)\to\lambda, \quad \textit{where } 0\leq\lambda<\infty, \tag{1.1}$$

$$\sup_{j\in\mathcal{J}} P(X_j=1)\to 0, \tag{1.2}$$

$$\sum_{j\in\mathcal{J}} P(X_j=1 \textit{ and there exists } k\neq j \textit{ with } \tilde{X}_{kj}\neq X_k)\to 0, \tag{1.3}$$

$$\sum_{j\in\mathcal{J}} P(X_j=1)\cdot P\ (\textit{there exists } k\neq j \textit{ with } \tilde{X}_{kj}\neq X_k)\to 0. \tag{1.4}$$

Then

$$\sum_{j\in\mathcal{J}} X_j \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda) \quad \textit{as } n\to\infty. \tag{1.5}$$

Remark 1.1. The conditions may be formulated in various ways. Since $EX_j=P(X_j=1)$, (1.1) may be written

$$E\sum_{j\in\mathcal{J}} X_j=\sum_{j\in\mathcal{J}} EX_j\to\lambda, \quad \text{where } 0\leq\lambda<\infty. \tag{1.1$'$}$$

Since $\sum_j P(X_j=1)^2\leq \sup_j P(X_j=1)\sum_j P(X_j=1)$, (1.2) is equivalent to

$$\sum_{j\in\mathcal{J}} P(X_j=1)^2\to 0, \tag{1.2$'$}$$

provided (1.1) holds. It is sometimes convenient to write (1.3) as

$$\sum_{j\in\mathcal{J}} P(X_j=1)P(\tilde{X}_{kj}\neq X_k \text{ for some } k\neq j\mid X_j=1)\to 0. \tag{1.3$'$}$$

Furthermore, (1.3) may be replaced by the stronger assumption

$$\sum_{j\neq k}\sum P(X_j=1 \text{ and } \tilde{X}_{kj}\neq X_k)\to 0 \tag{1.6}$$

or by, cf. (1.3′) and (1.1),

$$\sup_{j} P(\tilde{X}_{kj}\neq X_k \text{ for some } k\neq j\mid X_j=1)\to 0, \tag{1.7}$$

and (1.4) may be similarly modified.

Proof. We use Stein's method, cf. Chen [7]. Let $S=\sum_j X_j$, $S_j=\sum_{k\neq j}X_k$, $\tilde{S}_j=\sum_{k\neq j}\tilde{X}_{kj}$, $\varphi(t)=E\exp(\mathrm{i}tS)$ and $\tilde{\varphi}_j(t)=E\exp(\mathrm{i}t\tilde{S}_j)$. Since $E|S|=ES<\infty$, $\varphi(t)$ is differentiable and

$$\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}=E(\mathrm{i}S\,\mathrm{e}^{\mathrm{i}tS})=\mathrm{i}\sum_j E(X_j\,\mathrm{e}^{\mathrm{i}tS}). \tag{1.8}$$

Furthermore, since X_j is an indicator variable and X_j is independent of $\{\tilde{X}_{kj}\}_{k\neq j}$,

$$\begin{aligned}|E(X_j\,\mathrm{e}^{\mathrm{i}tS})-P(X_j=1)\,\mathrm{e}^{\mathrm{i}t}\tilde{\varphi}_j(t)|&=|EX_j\,\mathrm{e}^{\mathrm{i}t(X_j+S_j)}-E(X_j\,\mathrm{e}^{\mathrm{i}tX_j})\tilde{\varphi}_j(t)|\\&=|EX_j\,\mathrm{e}^{\mathrm{i}tX_j}(\mathrm{e}^{\mathrm{i}tS_j}-\mathrm{e}^{\mathrm{i}t\tilde{S}_j})|\\&\leqslant E|X_j(\mathrm{e}^{\mathrm{i}tS_j}-\mathrm{e}^{\mathrm{i}t\tilde{S}_j})|\\&\leqslant 2P(X_j=1\text{ and }\tilde{S}_j\neq S_j).\end{aligned} \tag{1.9}$$

Similarly,

$$\begin{aligned}|\tilde{\varphi}_j(t)-\varphi(t)|&=|E(\mathrm{e}^{\mathrm{i}t\tilde{S}_j}-\mathrm{e}^{\mathrm{i}t(S_j+X_j)})|\\&\leqslant 2P(\tilde{S}_j\neq S_j+X_j)\leqslant 2P(\tilde{S}_j\neq S_j)+2P(X_j\neq 0).\end{aligned} \tag{1.10}$$

Consequently,

$$\begin{aligned}&\left|\frac{1}{\mathrm{i}}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}-\sum_j P(X_j=1)\,\mathrm{e}^{\mathrm{i}t}\varphi(t)\right|\\&\quad\leqslant\sum_j|E(X_j\mathrm{e}^{\mathrm{i}tS})-P(X_j=1)\,\mathrm{e}^{\mathrm{i}t}\tilde{\varphi}_j(t)|+\sum_j P(X_j=1)|\tilde{\varphi}_j(t)-\varphi(t)|\\&\quad\leqslant 2\sum_j P(X_j=1\text{ and }\tilde{X}_{kj}\neq X_k\text{ for some }k\neq j)\\&\quad+2\sum_j P(X_j=1)P(\tilde{X}_{kj}\neq X_k\text{ for some }k\neq j)+2\sum_j P(X_j=1)^2.\end{aligned} \tag{1.11}$$

The sums on the right hand side tend to zero as $n\to\infty$ by (1.3), (1.4) and (1.2′). Hence, (1.11) together with (1.1) show that

$$\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}-\mathrm{i}\lambda\,\mathrm{e}^{\mathrm{i}t}\varphi(t)\to 0\quad\text{as } n\to\infty,\quad\text{uniformly in } t. \tag{1.12}$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\varphi(t)\exp(\lambda(1-\mathrm{e}^{\mathrm{i}t})))=\left(\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}-\mathrm{i}\lambda\,\mathrm{e}^{\mathrm{i}t}\varphi(t)\right)\exp(\lambda(1-\mathrm{e}^{\mathrm{i}t}))\to 0$$

uniformly in t, whence by integration

$$\varphi(t)\exp(\lambda(1-\mathrm{e}^{\mathrm{i}t}))-1\to 0,\quad\text{for every } t,\ \text{as } n\to\infty.$$

Consequently, $\varphi(t)\to\exp(\lambda(\mathrm{e}^{\mathrm{i}t}-1))$ and $S\xrightarrow{\mathrm{d}}Po(\lambda)$ as $n\to\infty$. □

In the simplest case, $\{X_j\}$ are independent and we may choose $\tilde{X}_{kj}=X_k$ for all k and j. Then (1.3) and (1.4) are trivially satisfied, and we recover the standard result that sums of independent indicator variables converge to a Poisson distribution provided (1.1) and (1.2) holds.

Of more interest is the following situation. Suppose that X_j is independent of all X_k except some (preferably relatively small) set $\{X_k\}_{k\in D_j}$. We then define

$$\tilde{X}_{kj} = \begin{cases} X_k, & k \notin D_j, \\ 0 & k \in D_j, \end{cases} \tag{1.13}$$

and obtain the following result.

Theorem 1.2. *Let, for each n, $\{X_j\}_{j\in\mathscr{J}}$ be a family of indicator variables and assume that for every $j\in\mathscr{J}$ there exists a subset D_j of $\mathscr{J}$ (with $j\in D_j$) such that X_j is independent of $\{X_k: k\notin D_j\}$ and, as $n\to\infty$,*

$$\sum_{j\in\mathscr{J}} P(X_j=1)\to\lambda, \quad \textit{where } 0\leqslant\lambda<\infty, \tag{1.14}$$

$$\sum_{j\in\mathscr{J}}\sum_{k\in D_j} P(X_j=1)P(X_k=1)\to 0, \tag{1.15}$$

$$\sum_{j\in\mathscr{J}}\sum_{k\in D_j\setminus\{j\}} P(X_j=1 \textit{ and } X_k=1)\to 0. \tag{1.16}$$

Then

$$\sum_{j\in\mathscr{J}} X_j \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda) \textit{ as } n\to\infty.$$

Proof. (1.14) is the same as (1.1). (1.16) is (assuming (1.13)) the same as (1.6), which implies (1.3). (1.15) is a combination of (1.2′) and

$$\sum_{j\neq k}\sum P(X_j=1)P(\tilde{X}_{kj}\neq X_k)\to 0.$$

Thus (1.1)–(1.4) hold and the conclusion follows by Theorem 1.1. □

Remark 1.2. (1.14) can be written $E\sum X_j\to\lambda$ while (1.16) (which often is the condition that is most difficult to verify) is equivalent to $\mathrm{Var}(\sum X_j)\to\lambda$ (assuming (1.14) and (1.15)). Note also that if the variables X_j are equi-distributed (and (1.14) holds with $\lambda>0$), (1.15) says that the (average) relative size of D_j tends to zero.

Various applications of Theorems 1.1 and 1.2 will be given in later sections. Other applications include e.g. sums of m-dependent indicator variables (also with multi-dimensional index sets).

Remark 1.3. A related result, with normal convergence, holds when $\lambda=\infty$. Viz., if (1.1) is replaced by

$$\mu = ES = \sum P(X_j=1)<\infty \quad\text{and}\quad \mu\to\infty \text{ as } n\to\infty, \tag{1.17}$$

and (1.2) is replaced by (1.2′) (which no longer is equivalent), then we obtain as in the proof above

$$\mathrm{d}\varphi(t)/\mathrm{d}t - \mathrm{i}\mu\, \mathrm{e}^{\mathrm{i}t}\varphi(t)\to 0 \text{ as } n\to\infty, \quad \text{uniformly in } t.$$

An argument similar to the one above then shows that $\varphi(t/\sqrt{\mu})\,\mathrm{e}^{-it\sqrt{\mu}} \to \mathrm{e}^{-t^2/2}$, i.e.

$$(S-\mu)/\sqrt{\mu} \xrightarrow{\mathrm{d}} N(0,1) \quad \text{as } n\to\infty.$$

(In fact, it suffices that the left hand sides of (1.2′), (1.3) and (1.4) are $o(\mu^{1/2})$.)

Remark 1.4. A vector-valued version of Theorem 1.1 may be proved by the same method. We will give another proof, based on Theorem 1.1 and the theory of point processes, in Section 4.

2. Point processes

We review some basic facts on point processes and refer to Kallenberg [14] for further details. The processes take place in some set $\mathscr{S}$ which we assume to be a locally compact second countable Hausdorff topological space. (For example $\mathscr{S}$ may be a closed or open subset of $\mathbb{R}^d$.) A Radon measure on $\mathscr{S}$ is a Borel measure μ such that $\mu(K)<\infty$ for every compact $K\subset\mathscr{S}$; we are only interested in such measures. Point processes are defined as random integer valued Radon measures, i.e. Radon measures that can be written as

$$\xi=\sum_1^N \delta_{X_j}, \tag{2.1}$$

where X_j are random variables with values in $\mathscr{S}$, N is a finite or infinite random variable, and δ_x is the Dirac measure

$$\delta_x(A)=I(x\in A), \quad A\subset\mathscr{S}. \tag{2.2}$$

Informally, we may regard ξ as the random (unordered) set $\{X_j\}$ (but note that multiple points may occur); $\xi(A)=\sum_1^N I(X_j\in A)$ is the number of points of this set that fall in A.

Let λ be a Radon measure on $\mathscr{S}$. The Poisson process with intensity λ then is the (unique) point process ξ such that the random variable $\xi(A)$ is Poisson distributed with parameter $\lambda(A)$ for every Borel set $A\subset\mathscr{S}$, and $\xi(A_1),\dots,\xi(A_k)$ are independent for any disjoint Borel sets $A_1,\dots,A_k$. A simple, but useful, example is when $\mathscr{S}$ is a finite or infinite discrete set; then a Poisson process on $\mathscr{S}$ is a collection of independent Poisson variables.

A λ-continuity set is a Borel set A such that $\lambda(\partial A)=0$. Similarly, if ξ is a point process, A is a ξ-continuity set if $\xi(\partial A)=0$ a.s. If ξ is a Poisson process with intensity λ, the ξ-continuity sets are exactly the λ-continuity sets. Note that the ξ-continuity sets form a ring.

We will discuss convergence of point processes in two topologies, viz. the vague topology defined on the set of all Radon measures and the weak topology defined on the subset of finite measures. We use $\xrightarrow{\mathrm{vd}}$ and $\xrightarrow{\mathrm{wd}}$ to denote convergence in distribution in these topologies, respectively, as $n\to\infty$. (The phrase "as $n\to\infty$" will usually be omitted from the formulae). The following lemma gives useful characterizations. (Parts (a) and (c) are contained in Kallenberg [14, Theorems 4.2 and 4.9]; part (b) follows similarly.)

Lemma 2.1. *Let ξ, $\xi_1, \ldots$ be point processes on $\mathscr{S}$.*

(a) $\xi_n \xrightarrow{\text{vd}} \xi$ *iff* $(\xi_n(A_1), \ldots, \xi_n(A_k)) \xrightarrow{\text{d}} (\xi(A_1), \ldots, \xi(A_k))$ *for all integers k and all relatively compact ξ-continuity sets $A_1, \ldots, A_k$.*

(b) $\xi_n \xrightarrow{\text{wd}} \xi$ *iff* $(\xi_n(A_1), \ldots, \xi_n(A_k)) \xrightarrow{\text{d}} (\xi(A_1), \ldots, \xi(A_k))$ *for all integers k and all ξ-continuity sets $A_1, \ldots, A_k$.*

(c) $\xi_n \xrightarrow{\text{wd}} \xi$ *iff* $\xi_n \xrightarrow{\text{vd}} \xi$ *and* $\xi_n(\mathscr{S}) \xrightarrow{\text{d}} \xi(\mathscr{S})$. □

Warning 2.1. The notion of vague convergence depends in an essential manner on the space $\mathscr{S}$. For example, vague convergence in $(0, \infty)$ does not imply vague convergence in $[0, \infty)$ (because mass may disappear at 0). Hence some care is required in specifying $\mathscr{S}$ in applications. However, $\mathscr{S}$ may always be replaced by a larger space of which it is a *closed* subspace, without affecting vague convergence.

If ξ is a point process and A a Borel set in $\mathscr{S}$, let $A\xi$ denote the restriction of ξ to A defined by $A\xi(B) = \xi(A \cap B)$, B a Borel set in $\mathscr{S}$. (We thus keep all points in A and eliminate the others.) Lemma 2.1 easily yields the following.

Lemma 2.2. *Let ξ, $\xi_1, \ldots$ be point processes on $\mathscr{S}$ and A a ξ-continuity set.*

(a) *If* $\xi_n \xrightarrow{\text{vd}} \xi$, *then* $A\xi_n \xrightarrow{\text{vd}} A\xi$.
(b) *If* $\xi_n \xrightarrow{\text{wd}} \xi$, *then* $A\xi_n \xrightarrow{\text{wd}} A\xi$.
(c) *If* $\xi_n \xrightarrow{\text{vd}} \xi$, *then* $A\xi_n \xrightarrow{\text{wd}} A\xi$ *iff* $\xi_n(A) \xrightarrow{\text{d}} \xi(A)$. □

We see from Lemma 2.1 that the vague topology is weaker than the weak topology since we only get conclusions on the distribution of $\xi_n(A)$ for relatively compact sets A. It is important for applications to have criteria that enable us to conclude that $\xi_n(A) \xrightarrow{\text{d}} \xi(A)$ (and thus, by Lemma 2.2(c), $A\xi_n \xrightarrow{\text{wd}} A\xi$) also for sets that are not relatively compact. The following is a useful criterion. Define the measure $E\xi_n$ by $(E\xi_n)(B) = E(\xi_n(B))$; $E\xi_n \leq \mu$ means $E\xi_n(B) \leq \mu(B)$ for every B.

Lemma 2.3. *Let ξ, $\xi_1, \ldots$ be point processes on $\mathscr{S}$ and let μ be a Borel measure such that $E\xi_n \leq \mu$ for every n. If $\xi_n \xrightarrow{\text{vd}} \xi$ and A is a ξ-continuity set with $\mu(A) < \infty$, then $\xi_n(A) \xrightarrow{\text{d}} \xi(A)$.*

Proof. Let K be a compact set in $\mathscr{S}$. Then

$$\limsup_{n\to\infty} P(\xi_n(A\backslash K) > 0) = \limsup_{n\to\infty} P(\xi_n(A\backslash K) \geq 1)$$
$$\leq \limsup_{n\to\infty} E\xi_n(A\backslash K) \leq \mu(A\backslash K). \tag{2.2}$$

Since $\mu(A)<\infty$, and $\mathcal{S}$ is σ-compact, we can choose K such that the right-hand side is arbitrarily small. The result follows by Kallenberg [14, Theorem 4.9] applied to $A\xi_n$ and $A\xi$. (Alternatively, one may use Billingsley [5, Theorem 4.2].) $\square$

We will also study the behaviour under continuous mappings. Let ψ be a continuous mapping of $\mathcal{S}$ into $\mathcal{S}'$ (another space of the same type). A Borel measure μ on $\mathcal{S}$ induces the measure $\mu\psi^{-1}$ on $\mathcal{S}'$. In particular, a point process $\xi=\sum \delta_{X_i}$ induces $\xi\psi^{-1}=\sum \delta_{\psi(X_i)}$, which is a point process provided it is Radon (i.e. only a finite number of points lie in each compact set). The following result is evident.

$$\text{If } \xi_n \xrightarrow{\text{wd}} \xi \text{ then } \xi_n\psi^{-1} \xrightarrow{\text{wd}} \xi\psi^{-1}. \tag{2.3}$$

The corresponding result for vague convergence, however, holds if ψ is a proper map (i.e. the inverse image of a compact set is compact), but in general not otherwise. Lemma 2.3 yields the following complement.

Lemma 2.4. *Let $\xi, \xi_1, \ldots$ be point processes on $\mathcal{S}$ and let μ be a Borel measure on $\mathcal{S}$ such that $E\xi_n \leqslant \mu$ for every n. If $\xi_n \xrightarrow{\text{vd}} \xi$ and ψ is a continuous mapping of $\mathcal{S}$ into $\mathcal{S}'$ such that $\mu\psi^{-1}$ is a Radon measure, then $\xi_n\psi^{-1} \xrightarrow{\text{vd}} \xi\psi^{-1}$.* $\square$

($\xi_n\psi^{-1}$ and $\xi\psi^{-1}$ are point processes because, if $K\subset\mathcal{S}'$ is compact, then e.g. $E\xi_n\psi^{-1}(K)\leqslant \mu\psi^{-1}(K)<\infty$ whence $\xi_n\psi^{-1}(K)<\infty$ a.s.).

We also consider integration. Note that, if $\xi=\sum\delta_{X_j}$, then $\int f\,\mathrm{d}\xi=\sum f(X_j)$ for any function f.

Lemma 2.5. *Let $\xi, \xi_1, \ldots$ be point processes on $\mathcal{S}$.*

(a) *If $\xi_n \xrightarrow{\text{vd}} \xi$ and μ is a Borel measure on $\mathcal{S}$ such that $E\xi_n\leqslant\mu$ for every n, then $\int f\,\mathrm{d}\xi_n \xrightarrow{\text{d}} \int f\,\mathrm{d}\xi$ for every real-valued continuous function f on $\mathcal{S}$ with $\int \min(|f|,1)\,\mathrm{d}\mu<\infty$.*

(b) *If $\xi_n \xrightarrow{\text{wd}} \xi$, then $\int f\,\mathrm{d}\xi_n \xrightarrow{\text{d}} \int \mathrm{d}\xi$ for every real-valued continuous function f.*

Proof. (a) If f has compact support, this holds without any μ, see e.g. [14, Theorem 4.2]. The general case follows by approximating f by such functions; we omit the details.

(b) It suffices to prove that the mapping $\xi\to\int f\,\mathrm{d}\xi$ is continuous in the weak topology, which is easily verified. $\square$

In applications, it is inconvenient to check a condition for all Borel sets; it is preferable if it suffices to check the condition for some smaller family of sets. We introduce two types of families that often will do. Choose some separable and complete metric on $\mathcal{S}$. (This is always possible, and the definitions below do not depend on the chosen metric.)

A DC-semiring $\mathcal{I}$ is a semiring of Borel sets such that for any $\varepsilon>0$, any compact subset of $\mathcal{S}$ may be covered by a finite number of elements of $\mathcal{I}$ having diameter less than ε. A DC-ring is a DC-semiring that is a ring.

The family of finite disjoint unions of sets in a given DC-semiring is a DC-ring.

Examples 2.1. Convenient DC-semirings on $\mathbb{R}$ are the families of half-open intervals $\{[a, b)\}$ and $\{(a, b]\}$. On $[0, \infty)$ we may e.g. take $\{[a, b): a \geq 0\}$ and on $(0, \infty)$ $\{[a, b): a > 0\}$. On $\mathbb{R}^d$ we can take the half-open rectangles $\{[a, b)\}$.

The following result is useful in applications of Lemmas 2.3–2.5.

Lemma 2.6. *Let $\mathcal{I}$ be a* DC*-semiring on $\mathcal{S}$. Let ξ be a point process on $\mathcal{S}$ and μ a Borel measure such that $E\xi(B) \leq \mu(B)$ for every $B \in \mathcal{I}$. Then $E\xi \leq \mu$.*

Proof. The class of sets B such that $E\xi(B) \leq \mu(B)$ is closed under disjoint unions. Hence we may assume that $\mathcal{I}$ is a DC-ring. The result now follows by a monotone class argument (cf. [14, 1.2 and 15.2.2]). □

Finally, we give a criterion for vague convergence contained in Kallenberg [14, Theorem 4.7].

Lemma 2.7. *Let $\xi, \xi_1, \ldots$ be point processes on $\mathcal{S}$ and assume that ξ is a.s. simple. Further, suppose that $\mathcal{U}$ is a* DC*-ring and $\mathcal{I}$ a* DC*-semiring, both consisting of ξ-continuity sets. If*

$$P(\xi_n(U) = 0) \to P(\xi(U) = 0) \quad \textit{for every } U \in \mathcal{U}$$

and

$$\limsup_{n \to \infty} E\xi_n(I) \leq E\xi(I) < \infty \quad \textit{for every } I \in \mathcal{I},$$

then $\xi_n \xrightarrow{\text{vd}} \xi$. □

3. Convergence to a Poisson process

Let $\mathcal{S}$ be as above and consider a sequence ξ_n of point processes on $\mathcal{S}$. Each ξ_n has a representation (2.1); however, for technical reasons, we prefer a representation with a non-random (finite or infinite, and possibly depending on n) number of terms. We achieve that by the following device. Let $\mathcal{S}^*$ be a space that contains $\mathcal{S}$ as a subspace and consider representations (2.1) where X_j are random variables with values in $\mathcal{S}^*$, but δ_{X_j} are regarded as measures on $\mathcal{S}$. Thus $\delta_{X_j} = 0$ if $X_j \in \mathcal{S}^* \setminus \mathcal{S}$, which means that we may add any number of "ghosts" X_j with values in $\mathcal{S}^* \setminus \mathcal{S}$. Evidently we then may fix the total number of terms (e.g. to be infinite).

Note that the actual values taken by X_j outside $\mathcal{S}$ are irrelevant, because all points in $\mathcal{S}^* \setminus \mathcal{S}$ are treated as non-existent. Hence we could fix $\mathcal{S}^*$ to be an extension of $\mathcal{S}$ by a single ("infinite") point, but we will keep the more general version (which also allows $\mathcal{S}^* = \mathcal{S}$ when the total mass of the point process is non-random).

As in Section 1, we present two versions of the limit theorem, one more general and one simpler that is more convenient for applications.

Theorem 3.1. *Let $\mathcal{S}$ be as in Section 2 and let λ be a Radon measure on $\mathcal{S}$. Let, for each n, ξ_n be a point process $\sum_{j\in\mathcal{J}}\delta_{X_j}$ on $\mathcal{S}$, where $\{X_j\}_{j\in\mathcal{J}}$ is a family of random variables with values in $\mathcal{S}^*\supset\mathcal{S}$. ($\mathcal{J}$ and X_j depend on n.) Assume that, for each n, there exist random variables $\tilde{X}_{kj}$, $j, k\in\mathcal{J}$, with values in $\mathcal{S}^*$ such that X_j is independent of $\{\tilde{X}_{kj}\}_{k\neq j}$ for every $j\in\mathcal{J}$ and, for every U and U' in a fixed DC-semiring $\mathcal{I}$ (on $\mathcal{S}$) of λ-continuity sets, as $n\to\infty$,*

$$\sum_{j\in\mathcal{J}} P(X_j\in U)\to\lambda(U), \tag{3.1}$$

$$\sup_{j\in\mathcal{J}} P(X_j\in U)\to 0, \tag{3.2}$$

$$\sum_{j\in\mathcal{J}} P(X_j\in U \text{ and for some } k\neq j, \text{ either } X_k\in U' \text{ and } \tilde{X}_{kj}\notin U' \text{ or } X_k\notin U' \text{ and } \tilde{X}_{kj}\in U')\to 0, \tag{3.3}$$

$$\sum_{j\in\mathcal{J}} P(X_j\in U)P\,(\text{for some } k\neq j, \text{ either } X_k\in U' \text{ and } \tilde{X}_{kj}\notin U' \text{ or } X_k\notin U' \text{ and } \tilde{X}_{kj}\in U')\to 0. \tag{3.4}$$

Then $\xi_n\xrightarrow{\text{vd}}\xi$ as $n\to\infty$, where ξ is a Poisson process on $\mathcal{S}$ with intensity λ.

Proof. Let $\mathcal{U}$ be the set of finite disjoint unions of relatively compact sets belonging to $\mathcal{I}$. $\mathcal{U}$ is a DC-ring of λ-continuity sets, and (3.1)-(3.4) hold for all $U, U'\in\mathcal{U}$. Hence, if $U\in\mathcal{U}$ is fixed we may apply Theorem 1.1. to the variables $I(X_j\in U)$ (and $I(\tilde{X}_{kj}\in U)$) and obtain

$$\xi_n(U)=\sum_j I(X_j\in U)\xrightarrow{\text{d}}\text{Po}(\lambda(U)). \tag{3.5}$$

In particular, since $\xi(U)$ has the distribution $Po(\lambda(U))$,

$$P(\xi_n(U)=0)\to P(\xi(U)=0),\quad U\in\mathcal{U}. \tag{3.6}$$

Furthermore,

$$E\xi_n(U)=\sum_j P(X_j\in U)\to\lambda(U)=E\xi(U)<\infty,\quad U\in\mathcal{U}. \tag{3.7}$$

If λ is non-atomic, ξ is a.s. simple and the conclusion follows by Lemma 2.7.

If λ has atoms, let $\{Y_j\}_{j\in\mathcal{J}}$ be a family of random variables, uniformly distributed on $[0, 1]$, that are independent of each other and of $\{X_j\}$. Let ξ'_n be the point process $\sum_j\delta_{(X_j, Y_j)}$ on $\mathcal{S}\times[0, 1]$. Since $\lambda\times\mathrm{d}x$ is a non-atomic measure on $\mathcal{S}\times[0, 1]$, the argument just given (with $\tilde{X}_{kj}$ replaced by $(\tilde{X}_{kj}, Y_k)$ and $\mathcal{I}$ replaced by $\{U\times([a, b)\cap[0, 1])\colon U\in\mathcal{I}, a<b\}$) shows that $\xi'_n\xrightarrow{\text{vd}}\xi'$ as $n\to\infty$, where ξ' is a Poisson process on $\mathcal{S}\times[0, 1]$ with intensity $\lambda\times\mathrm{d}x$. Since the projection $\mathcal{S}\times[0, 1]\to\mathcal{S}$ is proper, this implies that $\xi_n\xrightarrow{\text{vd}}\xi$. □

Theorem 3.2. *Let $\mathcal{S}$ be as above and let λ be a Radon measure on $\mathcal{S}$. Let, for each n, ξ_n be a point process $\sum_{j\in\mathcal{J}}\delta_{X_j}$ on $\mathcal{S}$, where $\{X_j\}_{j\in\mathcal{J}}$ is a family of random variables with values in $\mathcal{S}^*\supset\mathcal{S}$. ($\mathcal{J}$ and X_j depend on n.) Assume that, for each n, for every*

$j \in \mathcal{J}$ there exists a subset D_j of $\mathcal{J}$ (with $j \in D_j$) such that X_j is independent of $\{X_k\colon k \notin D_j\}$. Assume further that, for every U and U' in a fixed DC*-semiring $\mathcal{I}$ of λ-continuity sets, as $n \to \infty$,*

$$\sum_{j \in \mathcal{J}} P(X_j \in U) \to \lambda(U), \tag{3.8}$$

$$\sum_{j \in \mathcal{J}} \sum_{k \in D_j} P(X_j \in U) P(X_k \in U') \to 0, \tag{3.9}$$

$$\sum_{j \in \mathcal{J}} \sum_{k \in D_j \setminus \{j\}} P(X_j \in U \text{ and } X_k \in U') \to 0. \tag{3.10}$$

Then $\xi_n \xrightarrow{\text{vd}} \xi$ as $n \to \infty$, where ξ is a Poisson process on $\mathcal{S}$ with intensity λ.

Proof. We may assume that $\mathcal{S}^* \neq \mathcal{S}$. Let $*$ be a point in $\mathcal{S}^* \setminus \mathcal{S}$ and apply Theorem 3.1 with

$$\tilde{X}_{kj} = \begin{cases} X_k, & k \notin D_j \\ *, & k \in D_j. \quad \square \end{cases} \tag{3.11}$$

Convergence in the weak topology (on $\mathcal{S}$ or on a subset) can be obtained under suitable conditions by combining these theorems with Lemmas 2.1–2.3. For example:

Corollary 3.1. *Assume that the conditions of Theorem* 3.1 (*or Theorem* 3.2) *are satisfied and, furthermore, that $\lambda(S) < \infty$ and that* (3.1)–(3.4) ((3.8)–(3.10)) *hold also for $U = U' = \mathcal{S}$. Then $\xi_n \xrightarrow{\text{wd}} \xi$ as $n \to \infty$, where ξ is a Poisson process with intensity λ.*

Proof. $\xi_n \xrightarrow{\text{vd}} \xi$ by Theorem 3.1 (3.2), and $\xi_n(\mathcal{S}) \to \xi(\mathcal{S})$ by Theorem 1.1 (1.2). The result follows by Lemma 2.1(c). $\square$

Corollary 3.2. *Assume that the conditions of Theorem* 3.1 (*or Theorem* 3.2) *are satisfied and, furthermore, that μ is a Borel measure such that, for every $U \in \mathcal{I}$ and every n,*

$$\sum_{j \in \mathcal{J}} P(X_j \in U) \leqslant \mu(U). \tag{3.12}$$

Then, for every ξ-continuity set $A \subset \mathcal{S}$ with $\mu(A) < \infty$, $A\xi_n \xrightarrow{\text{wd}} A\xi$ (with ξ as above), in particular

$$\xi_n(A) \xrightarrow{\text{d}} \text{Po}(\lambda(A)). \tag{3.13}$$

Proof. $\xi_n \xrightarrow{\text{vd}} \xi$ by Theorem 3.1 (3.2). Lemma 2.6 and (3.12) imply that $E\xi_n \leqslant \mu$ for every n. Hence the result follows by Lemmas 2.3 and 2.2. $\square$

4. Two corollaries

A theorem yielding convergence of some random variables to a *normal* distribution can usually be immediately extended to vector-valued cases by the Cramér–Wold device. The Cramér–Wold device fails, however, for *Poisson* convergence results like our Theorem 1.1. One way to obtain vector-valued results (i.e. convergence of joint distributions) is to prove a basic result on convergence to more general infinitely divisible distributions, such that the Cramér–Wold device applies (e.g. Theorem 4.2 below). Here we will instead use Lemma 2.7 as a substitute for the Cramér–Wold device by deriving a vector-valued generalization of Theorem 1.1 as a corollary to Theorem 3.1; in fact, it is essentially the case of a finite set $\mathcal{S}$.

Theorem 4.1. *Let, for each n, $\{X_j\}_{j\in\mathcal{J}}$ be a family of d-dimensional random vectors $(X_j^{(1)},\dots,X_j^{(d)})$ (d is independent of n) and assume that there exist random vectors $\tilde{X}_{kj}$, $j, k\in\mathcal{J}$, such that X_j is independent of $\{\tilde{X}_{kj}\}_{k\neq j}$ for every $j\in\mathcal{J}$ and, as $n\to\infty$,*

$$\sum_j P(X_j^{(i)}=1)\to\lambda_i,\quad i=1,\dots,d,\quad \text{where } \lambda_i<\infty, \tag{4.1}$$

$$\sup_j P(X_j\neq 0)\to 0, \tag{4.2}$$

$$\sum_j P(X_j\neq 0 \text{ and there exists } k\neq j \text{ with } \tilde{X}_{kj}\neq X_k)\to 0, \tag{4.3}$$

$$\sum_j P(X_j\neq 0)\cdot P\ (\text{there exists } k\neq j \text{ with } \tilde{X}_{kj}\neq X_k)\to 0, \tag{4.4}$$

$$\sum_j P(X_j\notin\{0, e_1, e_2,\dots,e_d\})\to 0,\quad \text{where } e_i \text{ is the ith unit vector.} \tag{4.5}$$

Then

$$\sum_j X_j=\Big(\sum_j X_j^{(1)},\dots,\sum_j X_j^{(d)}\Big)\xrightarrow{\text{d}}(Y^{(1)},\dots,Y^{(d)})\quad \text{as } n\to\infty, \tag{4.6}$$

where $\{Y^{(i)}\}$ are independent Poisson distributed random variables with expectations $\lambda_1,\dots,\lambda_d$.

Proof. Let $\mathcal{S}$ be the finite set $\{e_1,\dots,e_d\}$ of unit vectors in $\mathbb{R}^d$, let $\mathcal{J}$ be the DC-semiring of all one-point subsets of $\mathcal{S}$, and let λ be the measure $\sum_1^d \lambda_i\delta_{e_i}$ on $\mathcal{S}$. Apply Theorem 3.1 (with $\mathcal{S}^*=\mathbb{R}^d$) and interpret the result (using Cramér's theorem). □

The reader can easily write down the special case that corresponds to Theorems 1.2 and 3.2.

We can also obtain generalizations of Theorems 1.1 and 1.2 where X_j no longer are assumed to be indicator variables. This time we give, for a change, only the version corresponding to Theorem 1.2.

Theorem 4.2. *Let λ and μ be two Borel measures on $\mathbb{R}\backslash\{0\}$ such that $\int_{\mathbb{R}}\min(1,|x|)\,\mathrm{d}\mu(x)<\infty$. Let, for each n, $\{X_j\}_{j\in\mathcal{J}}$ be a family of real-valued random variables and assume that for every $j\in\mathcal{J}$ there exists a subset D_j of $\mathcal{J}$ such that X_j is independent of $\{X_k\colon k\notin D_j\}$. ($\mathcal{J}$, X_j and D_j may depend on n.) Assume further that*

(3.8)–(3.10) *and* (3.12) *hold for all sets* U, U' *of the type* $(a, b]$, *with* $0<a<b$ *or* $a<b<0$. *Then*

$$\sum_{\mathcal{J}} X_j \xrightarrow{\mathrm{d}} Y \text{ as } n\to\infty, \tag{4.7}$$

where Y *has an infinitely divisible distribution with characteristic function*

$$E(\mathrm{e}^{\mathrm{i}tY}) = \exp\left(\int (\mathrm{e}^{\mathrm{i}tx}-1)\,\mathrm{d}\lambda(x)\right). \tag{4.8}$$

Proof. It follows by Theorem 3.2, with $\mathcal{S}^*=\mathbb{R}$ and $\mathcal{S}=\mathbb{R}\setminus\{0\}$, that $\xi_n=\sum \delta_{X_j} \xrightarrow{\mathrm{vd}} \xi$, where ξ is a Poisson process on $\mathbb{R}\setminus\{0\}$ with intensity λ. By Lemma 2.6, $E\xi_n \leqslant \mu$ for every n, whence Lemma 2.5(a), with $f(x)=x$, yields

$$\sum X_j = \int x\,\mathrm{d}\xi_n \xrightarrow{\mathrm{d}} \int x\,\mathrm{d}\xi. \tag{4.9}$$

It is easy to show that the characteristic function of $\int x\,\mathrm{d}\xi$ is given by (4.8). □

Theorem 4.2 does not cover all cases of convergence to infinitely divisible distributions. In particular, note that Y has no normal component. Cf. the related theorems by Chen [7, Theorem 4.1] and Jammalamadaka and Janson [12, Theorem 3.1]. (In e.g. the Poisson case, those theorems, however, use stronger conditions than the ones here.)

5. Dissociated variables

Let $m\geqslant 1$ be a fixed integer. Let (for a given n) $\mathcal{J}$ be the set of the $\binom{n}{m}$ unordered m-tuples $J=\{j_1,\dots,j_m\}$ of distinct positive integers $j_i\leqslant n$. We say that $\{X_J\}_{J\in\mathcal{J}}$ is a family of dissociated random variables if each X_J is independent of the family $\{X_K\colon K\cap J=\varnothing\}$ of all variables that are indexed by m-tuples not having any element in common with J (McGinley and Sibson [17]). (We may similarly consider some related situations, e.g. when $\mathcal{J}$ is the set of ordered m-tuples, or when repetitions are allowed among $j_1,\dots,j_m$.) We are in particular interested in the sum

$$S_n=\sum_{\mathcal{J}} X_J. \tag{5.1}$$

We give two prominent examples.

Example 5.1. *U-statistics.* Let $Y_1, Y_2, \dots$ be independent, identically distributed random variables and let

$$X_J = g(Y_{j_1},\dots,Y_{j_m}),\quad J=(j_1,\dots,j_m), \tag{5.2}$$

where g is a fixed symmetric function of m arguments. Then S_n is known as a U-statistic. We may here let g depend on n, and obtain a "triangular array" of U-statistics.

Example 5.2. *Incomplete U-statistics.* An incomplete U-statistic is obtained by taking X_J as in (5.2), but summing only over a subset (fixed or random) of $\mathcal{J}$. This is put in the form (5.1) by redefining X_J: Let X_J be as in (5.2) when J belongs to the select subset of $\mathcal{J}$, and $X_J = 0$ otherwise.

The theory of random graphs furnishes other examples of dissociated variables, see e.g. Sections 6 and 7, and (for a different type of example) Janson [13, the proofs of Theorems 1 and 3].

Theorem 1.2 (with $D_J = \{K\colon K \cap J \neq \emptyset\}$) yields the following result by Barbour and Eagleson [3]. (They used a different method that also gives estimates on the rate of convergence.)

Corollary 5.1. *Let, for each n, $\{X_J\}$ be a family of dissociated indicator variables and put $p_J = EX_J$. Suppose that, as $n \to \infty$,*

$$\sum_{\mathcal{J}} p_J \to \lambda, \quad \text{with } 0 \leq \lambda < \infty, \tag{5.3}$$

$$\sum_{0<|J \cap K|} p_J p_K \to 0, \tag{5.4}$$

$$\sum_{0<|J \cap K|<m} EX_J X_K \to 0. \tag{5.5}$$

Then $S_n \xrightarrow{\text{d}} \text{Po}(\lambda)$ as $n \to \infty$. □

Some special cases had been proved earlier, e.g. in [21] (U-statistics) and [4, 2] (incomplete U-statistics); [21] and [4] also contain results on convergence to Poisson processes (special cases of our Theorem 3.2 with $\mathcal{S} = [0, \infty)$).

It should be clear that Theorem 1.2 is a natural generalization of Corollary 5.1. The difference is that we do not assume any special structure on the index set in Theorem 1.2. This added flexibility is convenient in some applications. For example, we may study families of dissociated variables for several values of m simultaneously, or with $m = m(n) \to \infty$.

6. Random graphs

In the remaining sections, the general results above will be applied to random graphs. We give here some pertinent definitions, see e.g. Erdös and Renyi [10], Bollobás [6] or Palmer [19] for further information.

The random graph $G_{n,p}$ $(0 \leq p \leq 1)$ has n vertices and the $\binom{n}{2}$ possible edges occur independently of each other, each with probability p.

We will consider the evolution of random graphs when the edges are added (at random) sequentially. This is best done as follows. Let $\{T_e\}$, where e ranges over the set of edges in the complete graph K_n, be $\binom{n}{2}$ independent random variables with a common continuous distribution on $[0, \infty)$ (or, for some problems, on

$(-\infty, \infty)$). Let $G_n(t)$ be the random graph with n vertices and all edges e for which $T_e \leq t$. Hence T_e is interpreted as the time the edge e appears.

$G_n(t)$ is a random graph $G_{n,p}$ with $p = P(T_e \leq t)$; the construction above nests $G_{n,p}$ for different values of p. Furthermore, as t increases, new edges are added (at random) at the random times $\{T_{(i)}\}_{i=1}^{\binom{n}{2}}$ (the order statistics of $\{T_e\}$), which are a.s. distinct. Hence the random graph $G_{n,N}$ $(0 < N \leq \binom{n}{2})$, which has n vertices and N edges (all possible sets of N edges having the same probability), can be constructed as $G_n(T_{(N)})$. Hence results for both $G_{n,p}$ and $G_{n,N}$ can be obtained from results for the process $G_n(t)$.

For the applications in the following sections, we let T_e be uniformly distributed on $[0, n]$. (We would obtain the same results with the exponential distribution $\mathrm{Exp}(n)$, a choice that has some advantages but gives slightly more complicated formulae.) We also define, for any subgraph $H \subset G_n$, $T_H = \max\{T_e\colon e \in H\}$, i.e. the time at which the subgraph H appears. Thus, if H has $\|H\|$ edges,

$$P(T_H \leq t) = (t/n)^{\|H\|}, \quad 0 \leq t \leq n. \tag{6.1}$$

The statistic that counts the number of subgraphs of the random graph $G_{n,p}$ that are isomorphic to a given graph can be written as a sum of dissociated variables as in Section 5 (with m equal to the order of the given graph), whence we may apply our theorems. (Nowicki [18] makes a different approach that has the same effect; he writes the subgraph count statistic as an incomplete U-statistic based on the $\binom{n}{2}$ indicator variables I (edge e exists).) Theorem 1.2 (or Corollary 5.1) yields Poisson convergence results, under appropriate conditions on p, for the number of copies of any strictly balanced graph in $G_{n,p}$. (We may also consider a family of balanced graphs, or let the graph grow slowly with n.) See Bollobás [6] and Karoński [15] for definitions and earlier proofs of such results. (The simplest cases are already in Erdös and Renyi [10].) We will in the remaining sections give some slightly more complicated examples, where the theorems in Section 3 are useful.

7. Cycles

Let, for each n, $\mathscr{J} = \bigcup_3^\infty \mathscr{J}_l$, where $\mathscr{J}_l$ is the set of cycles of length l in the complete graph K_n. A cycle in $\mathscr{J}_l$ can be represented as a sequence of l distinct vertices in K_n by choosing a starting point and a direction. Since that choice may be made in $l \cdot 2$ ways,

$$\# \mathscr{J}_l = \frac{1}{2l} (n)_l. \tag{7.1}$$

Let $\mathscr{S} = [0, \infty) \times \{3, 4, 5, \ldots\}$; thus $\mathscr{S}$ is the disjoint union of infinitely many half-lines $\mathscr{S}_l$, $l \geq 3$. Define, for any cycle $J \in \mathscr{J}$,

$$X_J = (T_J, l) \quad \text{when } J \in \mathscr{J}_l, \tag{7.2}$$

where T_J is as in Section 6. Let $\xi_n = \sum_{\mathscr{J}} \delta_{X_J}$. Thus, $\xi_n([0, t]\times\{3\})$, $\xi_n([0, t]\times\{4\}), \dots$ are the numbers of cycles of length $3, 4, \dots,$ respectively in the random graph $G_n(t) = G_{n,p}$, with $p = t/n$ $(t \leq n)$. The idea of introducing the space $\mathscr{S}$ is that it allows us to consider cycles of all lengths simultaneously.

It is evident from the definitions that, if $J \in \mathscr{J}_l$,

$$P(X_J \in [0, t) \times \{l\}) = P(T_J < t) = (t/n)^l, \quad t \leq n, \tag{7.3}$$

and thus, as $n \to \infty$,

$$\sum_{\mathscr{J}} P(X_J \in [0, t) \times \{l\}) = \sum_{\mathscr{J}_l} \cdots = \# \mathscr{J}_l \cdot (t/n)^l \to \frac{1}{2l} t^l. \tag{7.4}$$

Consequently, we define a Radon measure λ on $\mathscr{S}$ by $\lambda([0, t)\times\{l\}) = F_l(t)$, $t \geq 0$, $l \geq 3$, with

$$F_l(t) = \frac{1}{2l} t^l, \quad t \geq 0; \tag{7.5}$$

i.e. λ equals $f_l(t)\,\mathrm{d}t$ on $\mathscr{S}_l$, where

$$f_l(t) = \frac{\mathrm{d}}{\mathrm{d}t} F_l(t) = \tfrac{1}{2} t^{l-1}, \quad t \geq 0. \tag{7.6}$$

We will show that the conditions of Theorem 3.2 are satisfied.

Let, for each cycle J, D_J be the set of all cycles with at least one edge in common with J. Then X_J and $\{X_K : K \notin D_J\}$ are independent.

Let $\mathscr{I} = \{[a, b)\times\{l\},\ 0 \leq a < b < \infty,\ l \geq 3\}$. (A set in $\mathscr{I}$ is thus a half-open interval on one of the half-lines in $\mathscr{S}$.) It is easily seen that $\mathscr{I}$ is a DC-semiring on $\mathscr{S}$. Clearly, $\mathscr{I}$ consists of λ-continuity sets. (3.8) holds by (7.4) and additivity.

It remains to verify (3.9) and (3.10). Since their left-hand sides are monotone in U and U', it suffices to consider the case $U = [0, t)\times\{l\}$, $U' = [0, t)\times\{l'\}$ for $t > 0$ and $l, l' \geq 3$ (possibly equal). It then follows by (7.3), since any $K \in D_J \cap \mathscr{J}_{l'}$ has at least two vertices in common with J, whence there are at most $l^2 n^{l'-2}$ such K, that the left-hand side of (3.9) is $O(n^{-2})$. This proves (3.9); for (3.10) we have to be a tiny bit more careful.

Let, for $1 \leq m \leq l-1$, $D_{J,m}$ be the set of all cycles $K \in \mathscr{J}_{l'}$ that have exactly m edges in common with J. Since each such K has at least $m+1$ vertices in common with J,

$$\begin{aligned}
&\sum_{K \in D_J \setminus \{J\}} P(X_J \in U \text{ and } X_K \in U') \\
&= \sum_{m=1}^{l-1} \sum_{K \in D_{J,m}} P(T_J \leq t \text{ and } T_K \leq t) \\
&= \sum_{m=1}^{l-1} \sum_{D_{J,m}} P(T_{J \cup K} \leq t) = \sum_{m=1}^{l-1} \# D_{J,m} \cdot (t/n)^{l+l'-m} \\
&\leq \sum_{1}^{l-1} \binom{l}{m+1} \binom{n}{l'-m-1} \frac{l'!}{2l'} \left(\frac{t}{n}\right)^{l+l'-m} = O(n^{-l-1}) \quad \text{as } n \to \infty,
\end{aligned} \tag{7.7}$$

for every $J \in \mathcal{J}_l$. (3.10) follows by (7.1). We have proved

Theorem 7.1. *$\xi_n \xrightarrow{\text{vd}} \xi$, where ξ is a Poisson process on $\mathcal{S}$ with intensity λ.* $\square$

Note that ξ can be regarded as a collection of independent Poisson processes on $[0, \infty)$ with the intensities $f_3(t), f_4(t), \ldots$ given by (7.6).

Furthermore, by (7.1) (cf. (7.4)),

$$\sum_{\mathcal{J}} P(X_J \in [a, b) \times \{l\}) = \#\mathcal{J}_l \cdot ((b/n)^l - (a/n)^l) \leq \lambda([a, b) \times \{l\}). \tag{7.8}$$

By Lemma 2.6, with $\mu = \lambda$,

$$E\xi_n \leq \lambda \quad \text{for every } n, \tag{7.9}$$

and Corollary 3.2 yields the following extension of Theorem 7.1.

Theorem 7.2. *If A is a λ-continuity set in $\mathcal{S}$ with $\lambda(A) < \infty$, then $A\xi_n \xrightarrow{\text{wd}} A\xi$, in particular*

$$\xi_n(A) \xrightarrow{\text{d}} \text{Po}(\lambda(A)). \quad \square \tag{7.10}$$

Remark 7.1. It follows easily that (7.9) holds also when $\lambda(A) = \infty$, in the sense that $\xi_n(A) \xrightarrow{\text{p}} \infty$.

Let $C_l(G)$ denote the number of cycles of length l in the graph G. Theorem 7.1 (or (7.2)) immediately yields

$$C_l(G_{n,c/n}) = \xi_n([0, c] \times \{l\}) \xrightarrow{\text{d}} \text{Po}\left(\frac{1}{2l} c^l\right) \quad \text{as } n \to \infty, \ 0 \leq c < \infty. \tag{7.11}$$

More generally, we obtain the following result by Erdös and Rényi [10].

Corollary 7.1. *Let $0 \leq c < \infty$. If $n \to \infty$ and $np \to c$, then*

$$C_l(G_{n,p}) \xrightarrow{\text{d}} \text{Po}\left(\frac{1}{2l} c^l\right). \tag{7.12}$$

If $n \to \infty$ and $2N/n \to c$, then

$$C_l(G_{n,N}) \xrightarrow{\text{d}} \text{Po}\left(\frac{1}{2l} c^l\right). \tag{7.13}$$

Proof. Observe that $np \to c$ and $\xi_n \xrightarrow{\text{vd}} \xi$ implies

$$\xi_n([0, np] \times \{l\}) \xrightarrow{\text{d}} \xi([0, c] \times \{l\}),$$

which yields (7.12). (7.13) follows similarly because $C_l(G_{n,N}) = \xi_n([0, T_{(N)}] \times \{l\})$ (see Section 6), and $T_{(N)} \xrightarrow{\text{p}} c$. $\square$

Note that (7.12) can be proved directly using Theorem 1.2; (7.13) can be proved using Theorem 1.1 (see Barbour [1] for a similar argument in a related situation), but the method above seems simpler. Theorem 7.1 adds the information that $C_l(G_{n,p})$ (or $C_l(G_{n,N})$) for different values of l are asymptotically independent [6]. Lemma 2.5 (or Theorem 4.2) gives the asymptotic distribution of the number of vertices in cycles when $c<1$.

Since

$$\sum_{3}^{\infty} F_l(t) = -\tfrac{1}{2}\log(1-t) - \tfrac{1}{2}t - \tfrac{1}{4}t^2 < \infty \tag{7.14}$$

for $0 \leqslant t < 1$, Theorem 7.2 yields (cf. [10])

Corollary 7.2. *If $n \to \infty$ and $np \to c$, $0 \leqslant c < 1$, then*

$$\textit{The number of cycles in } G_{n,p} \xrightarrow{\mathrm{d}} \mathrm{Po}(-\tfrac{1}{2}\log(1-c) - \tfrac{1}{2}c - \tfrac{1}{4}c^2). \tag{7.15}$$

In particular, for the probability that $G_{n,p}$ is a forest,

$$P(\textit{there are no cycles in } G_{n,p}) \to (1-c)^{1/2}\,\mathrm{e}^{c/2+c^2/4}. \quad \square \tag{7.16}$$

A corresponding result holds for $G_{n,N}$. It follows also, by Theorem 7.2 or Lemma 2.4, that the times the successive cycles (of any length) appear, asymptotically form a Poisson process on $[0, 1)$ with intensity $\sum_3^\infty \lambda_k(t) = \frac{1}{2}t^2/(1-t)$. The limit process has a.s. infinite mass on $[0, 1)$, in particular we see that if $np \to c \geqslant 1$, then $P(G_{n,p}$ has cycles$) \to 1$ [10]. We also obtain

Corollary 7.3. *If N_{cycl} is the number of edges when the first cycle appears, then $2N_{\mathrm{cycl}}/n \xrightarrow{\mathrm{d}} Z$ as $n \to \infty$, where Z has the distribution*

$$P(Z \leqslant t) = 1 - (1-t)^{1/2}\,\mathrm{e}^{t/2+t^2/4}, \quad 0 \leqslant t \leqslant 1. \quad \square \tag{7.17}$$

Of course, corresponding results hold for cycles with lengths in a given subset of $\{3, 4, \ldots\}$. In particular, this gives a result on the chromatic number $\chi(G_{n,N})$ of the random graphs.

Corollary 7.4. *If $n \to \infty$, $N \geqslant 1$ and $2N/n \to c$ with $0 \leqslant c \leqslant 1$, then*

$$P(\chi(G_{n,N}) = 2) \to \left(\frac{1-c}{1+c}\right)^{1/4} \mathrm{e}^{c/2}. \tag{7.18}$$

Proof. Since $N \geqslant 1$, $\chi(G_{n,N}) \geqslant 2$. On the other hand, $\chi(G_{n,N}) \geqslant 3$ iff $G_{n,N}$ has a cycle of odd length, and the number of such cycles converge in distribution to a Poisson distribution with expectation

$$\sum_{l \text{ odd}} F_l(c) = -c/2 - \tfrac{1}{4}\log(1-c) + \tfrac{1}{4}\log(1+c). \quad \square$$

The same result holds for $G_{n,p}$ provided $np \to c \leqslant 1$ and $n^2 p \to \infty$ so that P(no edge)$\to 0$. The assumptions in Corollary 7.4 imply that $P(\chi(G_{n,N}) \leqslant 3) \to 1$ as $n \to \infty$ (at least when $c<1$), see [10], whence the asymptotic distribution of $\chi(G_{n,N})$ is completely known for these cases. No corresponding exact result is known for larger values of c, but a lot of information exists, see McDiarmid [16].

8. The first cycle

Andiam, chè la via lunga ne sospigne.
Così si mise a così mi fe' entrare.
Nel primo cerchio che l'abisso cigne.
[8, IV. 22–24]

We are now prepared to solve the problem by Erdös stated in the introduction.

Theorem 8.1. *Let L_n denote the length of the first cycle that appears when edges are randomly added between n vertices. Then*

$$L_n \xrightarrow{\mathrm{d}} L \quad \text{as } n \to \infty, \tag{8.1}$$

where L is a random variable with the distribution

$$P(L=l) = \tfrac{1}{2} \int_0^1 t^{l-1}(1-t)^{1/2} \, \mathrm{e}^{t/2+t^2/4} \, \mathrm{d}t, \quad l = 3, 4, \ldots. \tag{8.2}$$

I.e. $P(L_n = l) \to p_l$ as $n \to \infty$, with $p_l = P(L=l)$ given by (8.2).

Proof. We may assume that the edges are added at random times as in Section 6, because these times do not enter the statement of the theorem. We continue to use the notation of Section 7. If η is any integer-valued Radon mesure on $\mathcal{S}$, let $\varphi(\eta)$ be the second coordinate of the first point in η, if such a point exists. More formally, $\varphi(\eta) = l$ if there exists $t \geqslant 0$ such that $\eta([0, t] \times \{l\}) = 1$ and $\eta([0, t] \times \{k\}) = 0$ for every $k \neq l$; if no such t exists we put $\varphi(\eta) = 0$. Hence $L_n = \varphi(\xi_n)$, and we put $L = \varphi(\xi)$. It is easily seen that $\varphi(\xi) > 0$ a.s. and, using simple properties of the Poisson process (and (7.14)),

$$\begin{aligned} P(\varphi(\xi) = l) &= \int_0^1 f_l(t) \, \mathrm{e}^{-\sum_3^\infty F_k(t)} \, \mathrm{d}t \\ &= \int_0^1 \tfrac{1}{2} t^{l-1} (1-t)^{1/2} \, \mathrm{e}^{t/2+t^2/4} \, \mathrm{d}t, \quad l = 3, 4, \ldots. \end{aligned}$$

Consequently, it only remains to verify that $\varphi(\xi_n) \xrightarrow{\mathrm{d}} \varphi(\xi)$. Unfortunately, this does not follow directly from Theorem 7.1, because the functional φ is not vaguely continuous. One way around this obstacle is as follows.

Let $\bar{\mathcal{N}}=\{3,4,\ldots,\infty\}$ be the one-point compactification of $\mathcal{N}=\{3,4,\ldots\}$, and let $\bar{\mathcal{S}}=[0,1)\times\bar{\mathcal{N}}$. Let $\bar{\xi}$ and $\bar{\xi}_n$ be the restrictions of ξ and ξ_n to $[0,1)\times\mathcal{N}$, regarded as point processes on $\bar{\mathcal{S}}$. By Lemmas 2.2(a) and 2.4 (with $\mu=\lambda$, cf. (7.9), and ψ the identity mapping), $\bar{\xi}_n \xrightarrow{\text{vd}} \bar{\xi}$. Defining φ as above also for integer valued Radon measures on $\bar{\mathcal{S}}$, it is easily seen that φ is vaguely continuous at any η such that $\varphi(\eta)\neq 0$. (If l and t are as in the definition of φ, then for some $t'>t$, $\{\eta'$: $\eta'([0,t')\times\{l\})=\eta'([0,t']\times\bar{\mathcal{N}}\}=1\}$ is a neighbourhood of η where φ is constant.) Since $\varphi(\bar{\xi})\neq 0$ a.s., $\varphi(\bar{\xi}_n)\xrightarrow{\text{d}}\varphi(\bar{\xi})$ by Billingsley [5, Theorem 5.1]. Finally, $\varphi(\bar{\xi})=\varphi(\xi)$ a.s. and

$$P(\varphi(\bar{\xi}_n)\neq\varphi(\xi_n))=P(\varphi(\bar{\xi}_n)=0)=P(\xi_n([0,1)\times\mathcal{N})=0)\to 0$$

(by Remark 7.1). Hence $\varphi(\xi_n)\xrightarrow{\text{d}}\varphi(\xi)$ by Cramér's theorem. □

Remark 8.1. The reason for introducing $\bar{\mathcal{S}}$ in this proof is that it contains more compact sets than $\mathcal{S}$ or $[0,1)\times\mathcal{N}$, whence its vague topology is stronger. Cf. Warning 2.1. In fact, the proof can be reformulated to involve point processes on $\mathcal{S}$ only, and the topology "weak convergence on $[0,t_0]\times\mathcal{N}$ for every $t_0<1$" (cf. Theorem 7.2), but we prefer to remain within the framework of Section 2. (An alternative proof, which fixes l and then maps $[0,1)\times\mathcal{N}$ to $[0,1)\times\{0,1\}$ by $\psi(t,k)=(t,\delta_{kl})$, using Lemma 2.4, is also possible.)

Some numerical values are given in Table 8.1.

Comparing (8.2) to the beta-integral, we see that

$$P(L=l)\sim cl^{-3/2}\quad\text{as } l\to\infty \tag{8.3}$$

(with $c=\frac{1}{4}\sqrt{\pi}\,\mathrm{e}^{3/4}\approx 0.94$), whence

$$P(L>l)\sim 2cl^{-1/2}\quad\text{as } l\to\infty. \tag{8.4}$$

In particular, $EL=\infty$.

Table 8.1
The asymptotic distribution of the length of the first cycle

l	$P(L=l)$	$P(L\leqslant l)$
3	0.1216	0.1216
4	0.0849	0.2065
5	0.0638	0.2704
6	0.0503	0.3207
7	0.0410	0.3617
8	0.0343	0.3961
9	0.0293	0.4253
10	0.0254	0.4507
20	0.0096	0.5973
50	0.0026	0.7391
100	0.0009	0.8140

Fatou's lemma shows that $EL_n \to \infty$ as $n \to \infty$, but our methods give no information on the rate of growth of EL_n. (EL_n is obviously finite, because $L_n \leq n$.) We thus have the somewhat unexpected situation that the average length of the first cycle tends to infinity with n, while the distribution of the length converges (without any normalization).

Philippe Flajolet [11] has found the asymptotic value of EL_n, which turns out to be $O(n^{1/6})$, by a combinatorial method.

The asymptotic distribution of the number of edges required to complete the first cycle was given in Corollary 7.3. It should be obvious how to find the asymptotic joint distribution of the required number of edges and the length of the first cycle. (They are not asymptotically independent. Cycles that come early tend to be smaller than cycles that are late.)

Furthermore, we may study the second cycle, etc. In fact, if the lengths of the consecutive cycles are denoted $L_n^{(1)} = L_n$, $L_n^{(2)}$, $L_n^{(3)}, \ldots$, the following result comes forth.

Theorem 8.2. *If* $m \geq 1$ *and* $l_1, \ldots, l_m \geq 3$, *then*

$$P(L_n^{(1)} = l_1, \ldots, L_n^{(m)} = l_m) \to \prod_{j=1}^{m-1} (2(l_1 + \ldots + l_j))^{-1} \cdot p_{l_1 + \ldots + l_m}, \tag{8.5}$$

with p_l *given by* (8.2).

Proof. For simplicity we take $m = 2$. The same method as above yields by (7.5) and (7.6), with $F(t) = \sum_3^\infty F_i(t)$,

$$\begin{aligned} P(L_n^{(1)} = l_1, L_n^{(2)} = l_2) &\to \iint_{0<s<t<1} f_{l_1}(s) f_{l_2}(t)\, e^{-F(t)}\, dt \\ &= \int_0^1 F_{l_1}(t) f_{l_2}(t)\, e^{-F(t)}\, dt \\ &= \frac{1}{2l_1} \int_0^1 f_{l_1+l_2}(t)\, e^{-F(t)}\, dt = \frac{1}{2l_1} p_{l_1+l_2}. \qquad \square \end{aligned}$$

Several curious consequences follow. The cycles tend to increase in size, e.g. if $l_1 < l_2$,

$$P(L_n^{(1)} = l_1, L_n^{(2)} = l_2) / P(L_n^{(1)} = l_2, L_n^{(2)} = l_1) \to l_2 / l_1 > 1,$$

whence

$$\lim P(L_n^{(1)} < L_n^{(2)}) > \lim P(L_n^{(1)} > L_n^{(2)}).$$

Also,

$$\lim_{n\to\infty} P(L_n^{(m)} = l) \sim c_m (\log l)^{m-1} / l^{3/2} \quad \text{as } l \to \infty.$$

On the other hand, the asymptotic probabilities decrease with l, so the most probable length of $L_n^{(m)}$ (at least for large n) is 3 for every m. Note also that $P(L_n^{(1)}=l \mid L_n^{(1)}+L_n^{(2)}=k) \to c_k'/l$ and that (by summing in (8.5) with $m=2$)

$$p_l = (2l)^{-1} \sum_{l+3}^{\infty} p_k.$$

Finally, we remark that the method above applies to some other types of random graphs as well. For example, for random directed graphs without loops (the $n(n-1)$ edges being added in random order), the analogue of Theorem 7.1 holds with $F_l(t)=t^l/l$ and $f_l(t)=t^{l-1}$, where now $l \geq 2$. Hence Theorem 8.1 holds with (8.2) replaced by

$$p_l = \int_0^1 f_l(t) \exp\left(-\sum_2^{\infty} F_k(t)\right) \mathrm{d}t = \int_0^1 t^{l-1}(1-t)\, \mathrm{e}^t \,\mathrm{d}t, \quad l=2,3,\ldots.$$

($p_l \sim e/l^2$ as $l \to \infty$.) For random directed graphs with loops,

$$p_l = \int_0^1 t^{l-1}(1-t)\,\mathrm{d}t = 1/l(l+1), \quad l=1,2,\ldots.$$

In both cases, Theorem 8.2 holds (with these p_l) if the factor 2 is deleted from (8.5).

9. Isolated cycles

We will now study isolated cycles in the random graphs. For simplicity, we fix the length l (≥ 3), although we might as well have treated all lengths simultaneously as in Section 7. Hence we let $\mathscr{J}$ be the set of cycles of length l in the complete graph K_n (this set was denoted $\mathscr{J}_l$ in Section 7). Let $\mathscr{S}=[0,\infty)\times[0,\infty)$ and define for any $J \in \mathscr{J}$,

$$U_J = \min\{T_e\colon e \text{ is an edge not in } J, \text{ but with at least one endpoint in } J\}, \tag{9.1}$$

$$X_J = (T_J, U_J), \tag{9.2}$$

and write, as usually, $\xi_n = \sum_{\mathscr{J}} \delta_{X_J}$. Thus, if J is an isolated cycle at some stage of the development, U_J is the time it stops being one. The number of isolated cycles in $G_{n,p}$ equals $\xi_n([0,np]\times(np,\infty))$. We will use Theorem 3.1 with the DC-semiring

$$\mathscr{I} = \{[a,b)\times[c,d)\colon 0 \leq a < b < \infty,\ 0 \leq c < d < \infty\}.$$

Since the variables T_J and U_J are independent for every J,

$$\begin{aligned}\sum_{\mathscr{J}} P(X_J \in [0,t)\times[0,u)) &= \#\mathscr{J}\cdot P(T_J<t)\cdot P(U_J<u)\\ &= \frac{1}{2l}(n)_l\left(\frac{t}{n}\right)^l\left(1-\left(1-\frac{u}{n}\right)^{l(n-l)+l(l-3)/2}\right)\\ &\to \frac{1}{2l}t^l(1-\mathrm{e}^{-lu}) \quad \text{as } n\to\infty.\end{aligned} \tag{9.3}$$

By additivity, (3.1) holds for $U \in \mathscr{I}$ and the measure λ given by

$$\lambda([0, t) \times [0, u)) = \frac{1}{2l} t^l (1 - e^{-lu}), \quad t, u \geq 0, \tag{9.4}$$

i.e.

$$d\lambda = \frac{l}{2} t^{l-1} e^{-lu} \, dt \, du, \quad t, u \geq 0. \tag{9.5}$$

(3.2) is obvious. We define $\tilde{X}_{KJ}$ as follows. Let

$$\tilde{U}_{KJ} = \min\{T_e: e \text{ is an edge, not in } K, \text{ with at least one endpoint in } K \text{ and no endpoint in } J\}, \tag{9.6}$$

$$\tilde{X}_{KJ} = \begin{cases} (T_K, \tilde{U}_{KJ}) & \text{when } J \text{ and } K \text{ are disjoint,} \\ * & \text{when } J \text{ and } K \text{ have a common vertex,} \end{cases} \tag{9.7}$$

where $*$ is any point in $\mathscr{S}^* \setminus \mathscr{S}$ ($\mathscr{S}^*$ can be any set which strictly contains $\mathscr{S}$). Then X_J is independent of $\{\tilde{X}_{KJ}\}$ for every J. In order to verify (3.3), it suffices to show that

$$\sum_{K \neq J}\sum P(X_J \in U \text{ and either } X_k \in U' \text{ and } \tilde{X}_{KJ} \notin U' \text{ or } X_k \in U' \text{ and } \tilde{X}_{KJ} \in U') \to 0, \tag{9.8}$$

when $U = [a, b) \times [c, d)$, $U' = [a', b') \times [c', d')$. We divide this sum into two parts. The sum, $\sum\sum'$ say, over all J and K that have a common vertex is (because then $\tilde{X}_{KJ} = * \notin U'$)

$$\begin{aligned} \sum\sum{}' P(X_J \in U \text{ and } X_K \in U') &\leq \sum\sum{}' P(T_J < b \text{ and } T_K < b') \\ &\leq \sum\sum{}' P(T_{J \cup K} < \max(b, b')), \end{aligned} \tag{9.9}$$

which is $O(1/n)$ by an argument similar to (7.7). The sum, $\sum\sum''$ say, over all disjoint pairs J, K in (9.8) is at most

$$\begin{aligned} &\sum\sum{}'' P(X_J \in U \text{ and } T_K \in [a', b') \text{ and } U_K \neq \tilde{U}_{KJ}) \\ &\quad \leq \sum\sum{}'' P(T_J < b \text{ and } T_K < b' \text{ and } U_K \neq \tilde{U}_{KJ}). \end{aligned} \tag{9.10}$$

For any disjoint J and K, the three events $\{T_J < b\}$, $\{T_K < b'\}$ and $\{U_K \neq \tilde{U}_{KJ}\}$ are independent. Furthermore (a.s.), $U_K \neq \tilde{U}_{KJ}$ iff $U_K = T_e$ for some edge e that connects K to J. By symmetry,

$$P(U_K \neq \tilde{U}_{KJ}) = l^2/(l(n-l) + l(l-3)/2) \leq l/(n-l).$$

Thus the right-hand sum in (9.10) equals

$$\begin{aligned} &\sum\sum{}'' P(T_J < b) P(T_K < b') P(U_K \neq \tilde{U}_{KJ}) \leq \left(\frac{1}{2l} n^l\right)^2 \left(\frac{b}{n}\right)^l \left(\frac{b'}{n}\right)^l \frac{l}{n-l} \\ &\quad \leq \frac{(bb')^l}{n-l} \to 0 \end{aligned}$$

as $n\to\infty$. This proves (3.3), and (3.4) is proved similarly. Thus Theorem 3.1 applies. Furthermore, if $n>3$, $E\xi_n$ has a density which is concentrated on $\{(t,u): u\leq n\}$ and there equal to (cf. (9.3))

$$\frac{1}{2l}(n)_l\frac{l}{n}\left(\frac{t}{n}\right)^{l-1}\frac{l(n-l+(l-3)/2)}{n}\left(1-\frac{u}{n}\right)^{l(n-l+(l-3)/2)-1}$$

$$\leq\frac{l}{2}t^{l-1}\,\mathrm{e}^{-(nl-l^2-1)u/n}$$

$$\leq\frac{l}{2}\mathrm{e}^{l^2+1}t^{l-1}\,\mathrm{e}^{-lu}.$$

Thus Corollary 3.2 also applies, with μ a constant (depending on l) times λ, and we obtain the following result.

Theorem 9.1. *$\xi_n\xrightarrow{\text{vd}}\xi$ as $n\to\infty$, where ξ is a Poisson process on $\mathcal{S}=[0,\infty)^2$ with intensity $\mathrm{d}\lambda=\frac{1}{2}lt^{l-1}\,\mathrm{e}^{-lu}\,\mathrm{d}t\,\mathrm{d}u$. If $A\subset\mathcal{S}$ and $\lambda(A)<\infty$, then $\xi_n(A)\xrightarrow{\text{d}}\mathrm{Po}(\lambda(A))$.* □

Taking $A=[0,c]\times(c,\infty]$, we obtain the following result by Erdös and Rényi [10].

Corollary 9.1. *If $np\to c$, $0<c<\infty$, then the number of isolated cycles of length l in $G_{n,p}$ converges in distribution to $\mathrm{Po}(c^l\,\mathrm{e}^{-cl}/2l)$.* □

Taking $A=\{(t,u): 0\leq t<u<\infty\}$, we count the isolated cycles that appear at some time when the edges are added one by one.

Corollary 9.2. *The number of isolated cycles of length l that appear during the evolution converges in distribution to $\mathrm{Po}((l-1)!/2l^l)$. In particular, the probability that no isolated cycles of length l appear converges to $\exp(-(l-1)!/2l^l)$.*

Proof. $\displaystyle\iint_{t<u}\frac{l}{2}t^{l-1}\,\mathrm{e}^{-lu}\,\mathrm{d}u=\frac{1}{2}\int_0^\infty t^{l-1}\,\mathrm{e}^{-lt}\,\mathrm{d}t=\frac{1}{2}l^{-l}(l-1)!.$ □

As was said earlier, we can also consider isolated cycles of different lengths simultaneously. (Not surprisingly, the joint distributions are asymptotically independent.) Summing over all lengths, we obtain (using Lemma 2.3) asymptotic Poisson distributions for the total number of isolated cycles. In particular, the following results hold; we omit the details.

Corollary 9.3. *If $n\to\infty$ and $np\to c$, $0\leq c<\infty$, then the number of isolated cycles in $G_{n,p}$ converges in distribution to*

$$\mathrm{Po}(-\tfrac{1}{2}\log(1-c\,\mathrm{e}^{-c})-\tfrac{1}{2}c\,\mathrm{e}^{-c}-\tfrac{1}{4}c^2\,\mathrm{e}^{-2c}).\qquad\square$$

Corollary 9.4. *The probability that no isolated cycle ever appears during the evolution converges to*

$$\exp\left(-\sum_{3}^{\infty}\tfrac{1}{2}l^{-l}(l-1)!\right)\approx 0.947. \qquad \square$$

Isolated cycles are not very common.

It is also possible to combine the methods of this section and the preceding one, taking e.g. $\mathcal{S}=[0,1)\times[0,\infty)\times\bar{\mathcal{N}}$ (with $\bar{\mathcal{N}}$ as earlier). We then obtain e.g.

$$P(\text{the first cycle is isolated})\to\sum_{l=3}^{\infty}\int_0^1 \tfrac{1}{2}t^{l-1}\,\mathrm{e}^{-lt}(1-t)^{1/2}\,\mathrm{e}^{t/2+t^2/4}\,\mathrm{d}t$$

$$=\int_0^1 \tfrac{1}{2}t^2(1-t\,\mathrm{e}^{-t})^{-1}(1-t)^{1/2}\,\mathrm{e}^{-5t/2+t^2/4}\,\mathrm{d}t\approx 0.026. \qquad (9.11)$$

Hence, the conditional probability that the first cycle is isolated given that some cycle is, is close to $\frac{1}{2}$.

More generally, we can watch the cycles appear as the graph evolves, number then consecutively (resolving ties e.g. randomly), and see which ones of them (if any) that are isolated when they appear. In particular, let us condition on the event that some isolated cycle appears (which has probability ≈ 0.05 by Corollary 9.4), and define N_n as the number of the first (and usually only) isolated cycle (i.e. the first cycle that is isolated when it appears). The following surprising result then holds.

Theorem 9.2. *There exist positive numbers a_k, $k=1,2,\ldots$ given by* (9.18) *below such that*

$$P\ (\textit{The kth cycle is the first isolated cycle})\to a_k \quad \textit{as } n\to\infty. \qquad (9.12)$$

However, $\sum_1^\infty a_k$ is strictly less than $b=\lim_{n\to\infty}P$ (some isolated cycle appears). Hence

$$P(N_n=k)\to a_k/b>0 \quad \textit{as } n\to\infty \qquad (9.13)$$

for every $k\geqslant 1$, but $\{N_n\}$ does not converge in distribution because part of the mass vanishes off to infinity.

The numbers a_k converge rapidly to zero; in fact $a_k\sim\text{const.}\cdot 3^{-k}$ as may be shown by computing $\sum a_k z^k$ by (9.18) and using residue calculus. The theorem thus implies that a large part of the mass of the distribution of N_n is concentrated on the first few values of k, and that the rest is spread out with a very slowly decreasing tail. In other other words, isolated cycles (if they appear at all) tend to be very early or very late. Numerical values are given in Table 9.1.

Proof. We will be somewhat sketchy. Let λ be as in (9.5) (although l now is variable) and let $\xi \sim Po(\lambda)$ and ξ_n have their usual meanings, so that $\xi_n \xrightarrow{\text{vd}} \xi$. Let, for $0<t<1$,

$$A_t = \{(s, u, l) \in \mathcal{S} \colon s<t \text{ and } s<u\}$$

and

$$B_t = \{(s, u, l) \in \mathcal{S} \colon s<t \text{ and } u \leqslant s\}.$$

Then $\xi_n(A_t)$ is the number of isolated cycles (of any lengths) that appear before time t and $\xi_n(B_t)$ is the number of non-isolated cycles that appear before t. Define

$$g(t) = \sum_{l=3}^{\infty} \tfrac{1}{2} t^{l-1} \,\mathrm{e}^{-lt} = \tfrac{1}{2} t^2 \,\mathrm{e}^{-3t} (1 - t\,\mathrm{e}^{-t})^{-1}, \tag{9.14}$$

$$G(t) = \int_0^t g(s)\,\mathrm{d}s, \tag{9.15}$$

$$F(t) = \sum_{l=3}^{\infty} F_l(t) = -\tfrac{1}{2}\log(1-t) - \tfrac{1}{2}t - \tfrac{1}{4}t^2. \tag{9.16}$$

Then $F(t) = \lambda(A_t \cup B_t)$ is the asymptotic expected number of cycles before time t, and $G(t) = \lambda(A_t)$ is the asymptotic expected number of isolated cycles before time t. It also follows that, cf. Corollary 9.4,

$$b = 1 - \exp(-G(\infty)). \tag{9.17}$$

Since $\lambda(A_t) = G(t)$ and $\lambda(B_t) = F(t) - G(t)$, standard arguments yield

$$\begin{aligned} &P(\text{The } k\text{th cycle is the first isolated one and it appears before time } 1) \\ &\quad \to \int_0^1 P(\xi(A_t) = 0 \quad \text{and} \quad \xi(B_t) = (k-1) g(t)\,\mathrm{d}t \\ &\quad = \int_0^1 \mathrm{e}^{-G(t)} \frac{(F(t)-G(t))^{k-1}}{(k-1)!} \,\mathrm{e}^{-(F(t)-G(t))} g(t)\,\mathrm{d}t. \end{aligned}$$

Since $P(\text{The } k\text{th cycle appears after time } 1) \to 0$, see Corollary 7.2, this implies (9.12) with

$$a_k = \int_0^1 \frac{1}{(k-1)!} (F(t)-G(t))^{k-1} \,\mathrm{e}^{-F(t)} g(t)\,\mathrm{d}t. \tag{9.18}$$

The proof is completed by noting that

$$\begin{aligned} \sum_1^{\infty} a_k &= \int_0^1 \mathrm{e}^{F(t)-G(t)} \,\mathrm{e}^{-F(t)} g(t)\,\mathrm{d}t = \int_0^1 \mathrm{e}^{-G(t)} G'(t)\,\mathrm{d}t \\ &= 1 - \mathrm{e}^{-G(1)} < 1 - \mathrm{e}^{-G(\infty)} = b. \qquad \square \end{aligned}$$

Table 9.1
The improper limits of the (improper) unconditional distribution and the (proper) conditional distribution of the number of the first isolated cycle

k	a_k	a_k/b
1	0.0261	0.491
2	0.0035	0.066
3	0.0008	0.016
4	0.0002	0.004
5	0.0001	0.001
6	0.0000	0.000
sum	0.0307	0.579

A numerical integration gives $G(1)\approx 0.0312$, and thus $\sum a_k \approx 0.0307$, while $b\approx 1-\exp(-0.0545)\approx 0.0531$. Hence about 42% of the mass of the distribution of N_n vanishes off to infinity.

10. A non-Poisson limit

In this section we for simplicity study only $G_{n,p}$ with $p=c/n$, where c is fixed, $0<c<\infty$. Fix $l\geqslant 3$, and let H be the comet-like graph with $l+1$ vertices and $l+1$ edges consisting of a cycle of length l and a single edge from the cycle to the last vertex. H is balanced but not strictly balanced. Let Y_n be the number of subgraphs of $G_{n,p}$ that are isomorphic to H. It is easily seen that $EY_n=\frac{1}{2}(n)_{l+1}p^{l+1}\to\frac{1}{2}c^{l+1}$.

Theorem 10.1. *With notations as above,*

$$Y_n \xrightarrow{\mathrm{d}} Y \quad \text{as } n\to\infty, \tag{10.1}$$

where Y has an infinitely divisible distribution with the characteristic function

$$E(\mathrm{e}^{\mathrm{i}tY})=\exp\left(\frac{1}{2l}c^l(\exp(lc(\mathrm{e}^{\mathrm{i}t}-1))-1)\right). \tag{10.2}$$

In other words,

$$Y \stackrel{\mathrm{d}}{=} \sum_1^N Z_i, \tag{10.3}$$

where $N\sim \mathrm{Po}((1/2l)c^l)$ and $Z_1, Z_2, \ldots \sim \mathrm{Po}(lc)$, all being independent. In particular,

$$P(Y_n=0)\to \mathrm{e}^{-(1-\mathrm{e}^{-lc})c^l/2l}. \tag{10.4}$$

Proof. Let again $\mathcal{J}$ be the set of cycles of length l. Let $\mathcal{S}=\{0,1,2,\ldots\}$ and $\mathcal{S}^*=\mathcal{S}\cup\{*\}$. Define

$$X_J=\begin{cases}*, & J\not\subset G_{n,p} \quad \text{(i.e. } T_J>c)\\ m, & \text{if } J\subset G_{n,p} \text{ and exactly } m \text{ edges in } G_{n,p}\\ & \text{have exactly one endpoint in } J,\end{cases} \tag{10.5}$$

and, as always, $\xi_n=\sum \delta_{X_J}$. Then

$$Y_n=\sum_{X_J\in\mathcal{S}} X_J=\int_{\mathcal{S}} x\,\mathrm{d}\xi_n. \tag{10.6}$$

Define $\tilde{X}_{KJ}=*$ if $K\not\subset G_{n,p}$ or $K\cap J\neq\emptyset$, and otherwise $\tilde{X}_{KJ}=$ the number of edges in $G_{n,p}$ with one endpoint in K and the other neither in K nor J. It is easily seen that the assumptions in Theorem 3.1 hold with $\mathcal{I}$ the set of singletons in $\mathcal{S}$ and the measure λ given by

$$\lambda\{m\}=\frac{1}{2l}c^l\frac{(lc)^m}{m!}\mathrm{e}^{-lc}, \quad m\geqslant 0. \tag{10.7}$$

Hence, if ξ is a Poisson process with intensity λ,

$$\xi_n \xrightarrow{\text{vd}} \xi. \tag{10.8}$$

Furthermore, $\xi_n(\mathcal{S})$ is the number of cycles of length l in $G_{n,p}$, whence $\xi_n(\mathcal{S})\to^{\mathrm{d}}\xi(\mathcal{S})$ by Corollary 7.1. Consequently, Lemma 2.1(c) yields

$$\xi_n \xrightarrow{\text{wd}} \xi, \tag{10.9}$$

whence, by Lemma 2.5(b),

$$Y_n=\int x\,\mathrm{d}\xi_n \xrightarrow{\mathrm{d}} \int x\,\mathrm{d}\xi. \tag{10.10}$$

This proves (10.1) with $Y=\int x\,d\xi$. The expressions (10.2), (10.3), (10.4) follow by properties of the Poisson process; e.g.

$$P(Y=0)=P(\xi(\mathcal{S}\backslash\{0\})=0)=\exp(-\lambda(\mathcal{S}\backslash\{0\})). \qquad \square$$

Similar, but more complicated, results may be obtained for some other non-strictly balanced graphs. For example, if we proceed as above but take

$$\mathcal{S}=\{(m_1,\ldots,m_l)\in\mathbb{Z}^l;\ 0\leqslant m_1\leqslant m_2\leqslant\cdots\leqslant m_l\}$$

and let X_J count the number of edges from each vertex in J to the complement of J, we see that the number of subgraphs of $G_{n,p}$ that consist of a cycle of length l with two tails (of length 1) attached to the same vertex converges in distribution to $\sum_1^N\sum_1^l\binom{W_{ij}}{2}$, where N is as before and $W_{ij}\sim\mathrm{Po}(c)$, all being independent. See also Bollobás [6, Chapter IV.2] for related results.

It should be clear why we do not get Poisson convergence in this case (and for other graphs that are not strictly balanced). The reason is that two different copies of H that share a common cycle will appear in $G_{n,p}$ with a rather strong correlation; once the cycle exists it is easy to add several tails to it. In fact, it is easy to see that $\mathrm{Var}(Y_n) \sim EY_n \cdot (1+lc)$.

No such problem exists, however, if we only count isolated copies of H, and we have the following result. A similar result holds for the number of isolated copies of any connected graph, although the only interesting cases are trees and graphs with exactly one cycle.

Theorem 10.2. *The number of isolated subgraphs of $G_{n,p}$ that are isomorphic to H converges in distribution to $\mathrm{Po}(\frac{1}{2}c^{l+1}\,\mathrm{e}^{-(l+1)c})$ as $n \to \infty$.*

Proof. Let $\mathcal{J}$ be the set of all copies of H in the complete graph K_n and let $X_J = I$ (J is an isolated subgraph of $G_{n,p}$). Let $\tilde{X}_{KJ} = I$ (K and J are disjoint and K is an isolated subgraph of $G^J_{n,p}$), where $G^J_{n,p}$ is the subgraph of $G_{n,p}$ obtained by removing the vertices of J and all edges incident upon them. The result follows easily by Theorem 1.1; note that (1.3) now is trivial because $X_J = 1$ and $K \neq J$ imply $\tilde{X}_{KJ} = X_K$. □

Acknowledgements

This study was initiated during the Second Seminar on Random Graphs in Poznán, August 1985. I thank the organizers for their hospitality and several participants, in particular Geoffrey Grimmett, Michał Karoński, Colin McDiarmid, Krzysztof Nowicki and John Wierman, for helpful discussions.

I also thank Michael Thuné for help with numerical integrations and Julie White for general support.

References

[1] A.D. Barbour, Poisson convergence and random graphs, Math. Proc. Camb. Phil. Soc. 92 (1982) 349–359.
[2] A.D. Barbour and G.K. Eagleson, Poisson approximation for some statistics based on exchangeable trials, Adv. Appl. Prob. 15 (1983) 585–600.
[3] A.D. Barbour and G.K. Eagleson, Poisson convergence for dissociated statistics, J. R. Statist. Soc. Ser. B 46 (1984) 397–402.
[4] M. Berman and G.K. Eagleson, A Poisson limit theorem for incomplete symmetric statistics, J. Appl. Probab. 20 (1982) 47–60.
[5] P. Billingsley, Convergence of Probability Measures (John Wiley & Sons, New York, etc., 1968).
[6] B. Bollobás, Random Graphs (Academic Press, London, etc., 1985).
[7] L.H.Y. Chen, Two central limit problems for dependent random variables, Z. Wahrsch. Verw. Geb. 43 (1978) 223–243.
[8] Dante Alighieri, Divina Commedia (Florence, 1321).

[9] G.K. Eagleson, A Poisson limit theorem for weakly exchangeable events, J. Appl. Probab. 16 (1979) 794-802.

[10] P. Erdös and A. Rényi, On the evolution of random graphs, Magyar Tud. Akad. Mat. Kut. Int. Közl. 5 (1960) 17-61.

[11] P. Flajolet, personal communication.

[12] S.R. Jammalamadaka and S. Janson, Limit theorems for a triangular scheme of U-statistics with applications to inter-point distances, Ann. Probab, 14 (1986) 1347-1358.

[13] S. Janson, Random trees in a graph and trees in a random graph, Math. Proc. Camb. Phil. Soc. 100 (1986) 319-330.

[14] O. Kallenberg, Random Measures (Akademie-Verlag, Berlin, 3rd ed., 1983).

[15] M. Karónski, Balanced subgraphs of large random graphs, Adam Mickiewicz University, Poznań, 1984.

[16] C. McDiarmid, Colouring random graphs, Ann. Operations Research 1 (1984) 183-200.

[17] W.G. McGinley and R. Sibson, Dissociated random variables, Math. Proc. Camb. Phil. Soc. 77 (1975) 185-188.

[18] K. Nowicki, Poisson limit distribution of incomplete U-statistics with application to graph theory, Report, University of Lund (1986).

[19] E. Palmer, Graphical Evolution (John Wiley & Sons, New York, etc., 1985).

[20] S.D. Poisson, Recherches sur la probabilité des jugements en matière criminelle et en matière civile, ... (Paris, 1837).

[21] B. Silverman and T. Brown, Short distances, flat triangles and Poisson limits, J. Appl. Probab. 15 (1978) 815-825.