PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 2, Fasc. 2 (1982), p. 197-205

ON THE NUMBER OF k-TREES IN A RANDOM GRAPH

BY

MICHAŁ KAROŃSKI (POZNAŃ)

Abstract. Let $K_{n,p}$ denote a random graph obtained from a complete labelled graph K_n on n vertices by independent deletion of its edges with the prescribed probability $q = 1-p$, $0 < p < 1$. Moreover, let $p = p(n)$ and let $X_{n,r}^{(k)}$ denote the number of r-vertex subgraphs ($r \geqslant k+1$) of a random graph $K_{n,p}$ being k-trees. In this paper we prove that, under some conditions imposed on probability $p(n)$ as $n \to \infty$, the random variable $X_{n,r}^{(k)}$ has asymptotically the Poisson or normal distribution. We generalize earlier results of Erdös and Rényi [2] dealing with the distribution of the number of trees (i.e. random variable $X_{n,r}^{(1)}$) as well as the results of Schürger [7] on the number of cliques in $K_{n,p}$ (i.e. random variable $X_{n,k+1}^{(k)}$).

1. Introduction. Let us consider a random graph (r.g.) $K_{n,p}$ obtained from a labelled complete graph K_n by means of the following procedure:

Each of $\binom{n}{2}$ edges of K_n is independently deleted with the prescribed probability q, $0 < q < 1$, i.e. each edge remains in K_n with probability $p = 1-q$ and the expected number of edges of an r.g. $K_{n,p}$ equals $\binom{n}{2}p$.

Here we shall consider the asymptotic distribution of the number of subgraphs being k-trees of size r in an r.g. $K_{n,p(n)}$, where the edge probability p depends on the size n of an r.g. and $n \to \infty$.

The notion of *k-tree* ($k = 1, 2, \ldots$) was introduced in [3] and can be defined either as a k-dimensional simplicial complex with certain properties or as a graph. We shall use here an inductive definition of k-tree (see [5] or [6]). A k-tree of size $k+1$ is a complete graph on $k+1$ vertices. A k-tree $T_{r+1}^{(k)}$ of size $r+1$ ($r \geqslant k+1$) is obtained from an arbitrary k-tree $T_r^{(k)}$ of size r by adding a new vertex and joining it to those k points of $T_r^{(k)}$

which form a complete graph. A k-tree of size r consists of $r-k$ complete subgraphs of size $k+1$, $k(r-k)+1$ complete subgraphs of size k and has $kr-\binom{k+1}{2}$ edges. It has been shown (see [1] and [4]) that the total number of k-trees which can be formed on r labelled vertices equals

$$\binom{r}{k}[k(r-k)+1]^{r-k-2}.$$

In this paper we generalize the results of Erdös and Rényi [2] on the number of trees and the results of Schürger [7] on cliques. The methods of proofs are mainly those of [7].

2. The Lemma. Suppose that $G_i = (V(G_i), E(G_i))$, $i = 1, 2, \ldots, t$, are graphs whose vertex sets $V(G_i)$ as well as edge families $E(G_i) \subset V(G_i) \times V(G_i)$ are not necessarily disjoint. Then by their union $\bigcup_{i=1}^{t} G_i$ we mean a graph $G = (V(G), E(G))$, where

$$V(G) = \bigcup_{i=1}^{t} V(G_i) \quad \text{and} \quad E(G) = \bigcup_{i=1}^{t} E(G_i).$$

Denote by $|X|$ the cardinality of a set X and by a,

$$a = a(r, k) = kr-\binom{k+1}{2},$$

the number of edges of a k-tree of size r. In this section we state the following purely graph-theoretic result:

LEMMA. *Suppose that $T_{r,1}^{(k)}, T_{r,2}^{(k)}, \ldots, T_{r,t}^{(k)}$ are pairwise different k-trees of size r, $r \geqslant k+1$, not all of which are pairwise vertex disjoint, $G_t = \bigcup_{i=1}^{t} T_{r,i}^{(k)}$ is their union, $|V(G_t)| = v_t$, and $|E(G_t)| = e_t$. Then $v_t/e_t \leqslant \max\{b_0, b_1, b_2\}$, where*

$$b_0 = b_0(t) = \frac{(t-1)r}{(t-1)a+1}, \quad b_1 = b_1(t) = \frac{tr-1}{ta}, \tag{1}$$
$$b_2 = b_2(t) = \frac{(t-1)r+1}{(t-1)a+k}.$$

Proof (by induction on t). Put

$$h_t = \left|V\Big(T_{r,t}^{(k)} - \bigcup_{i=1}^{t-1} T_{r,i}^{(k)}\Big)\right|, \quad t \geqslant 2.$$

It is easy to check that if $1 \leqslant h_t \leqslant r-k$, then $e_t \geqslant e_{t-1}+kh_t$, whereas

$$e_t \geqslant e_{t-1}+a-\binom{r-h_t}{2} \quad \text{if } r-k+1 \leqslant h_t \leqslant r-1.$$

One can also check that the inequality

$$\text{(2)} \qquad \frac{v_{t-1}+r-i}{e_{t-1}+a-\binom{i}{2}} \leqslant \frac{v_{t-1}+r-1}{e_{t-1}+a}$$

holds for $i = 1, 2, \ldots, k-1$. Moreover, it should be noticed that $b_0(t)$, $b_1(t)$, and $b_2(t)$ are increasing functions of t. Having these facts in mind we begin our proof by considering $t = 2$.

Case 1. Let $h_2 = 0$ $(r \geqslant k+2)$. Then $v_2 = v_1 = r$, $e_2 \geqslant e_1+1$ and, consequently,

$$\frac{v_2}{e_2} \leqslant \frac{r}{a+1} = b_0(2).$$

Case 2. If $1 \leqslant h_2 \leqslant r-k$, then

$$\frac{v_2}{e_2} = \frac{r+h_2}{e_2} \leqslant \frac{r+h_2}{a+kh_2} \leqslant \frac{r+1}{a+k} = b_2(2).$$

Case 3. Let $r-k+1 \leqslant h_2 \leqslant r-1$. Now, by (2) we have

$$\frac{v_2}{e_2} \leqslant \frac{r+h_2}{e_1+a-\binom{r-h_2}{2}} \leqslant \frac{2r-1}{2a} = b_1(2)$$

and we arrive at the thesis for $t = 2$.

Let us assume that our thesis is true for some $t \geqslant 3$. Suppose that G_{t-1} is the union of $t-1$ k-trees of size r not all of which are pairwise vertex disjoint and

$$\frac{v_{t-1}}{e_{t-1}} \leqslant \max\{b_0(t-1), b_1(t-1), b_2(t-1)\}$$

holds. Let us take now the union of G_{t-1} with a k-tree also of size r.

Case 1. If $h_t = 0$, then $v_t/e_t \leqslant v_{t-1}/e_{t-1}$ and the thesis follows immediately.

Case 2. Let $1 \leqslant h_t \leqslant r-k$. Assume first that, for G_{t-1}, $b_0(t-1)$ exceeds $b_1(t-1)$ and $b_2(t-1)$. Then

$$\frac{v_t}{e_t} \leqslant \frac{v_{t-1}+h_t}{e_{t-1}+kh_t} \leqslant \frac{(t-2)r\{v_{t-1}+h_t\}}{v_{t-1}\{(t-2)a+1\}+kh_2(t-2)r}$$

$$= \frac{(t-2)r\{v_{t-1}+h_t\}}{\{(t-2)a+1\}\left\{v_{t-1}+h_t\dfrac{kr(t-2)}{a(t-2)+1}\right\}} \leqslant \frac{(t-2)r}{(t-2)a+1} = b_0(t-1) \leqslant b_0(t).$$

Similarly, if $b_1(t-1)$ is maximal, then

$$\frac{v_t}{e_t} \leqslant \frac{\{(t-1)r-1\}\{v_{t-1}+h_t\}}{a(t-1)\left\{v_{t-1}+h_t\frac{k[(t-1)r-1]}{a(t-1)}\right\}} \leqslant b_1(t-1) \leqslant b_1(t),$$

and if, for G_{t-1}, $b_2(t-1)$ is greater than $b_0(t-1)$ and $b_1(t-1)$, then

$$\frac{v_t}{e_t} \leqslant b_2(t).$$

Case 3. Let $r-k+1 \leqslant h_t \leqslant r-1$. As before we assume first that, for G_{t-1}, $b_0(t-1)$ exceeds $b_1(t-1)$ and $b_2(t-1)$. Then by formula (2) and the induction assumption we obtain

$$\frac{v_t}{e_t} \leqslant \frac{v_{t-1}+h_t}{e_{t-1}+a-\binom{r-h_t}{2}} \leqslant \frac{v_{t-1}+r-1}{e_{t-1}+a}$$

$$\leqslant \frac{r(t-2)\{v_{t-1}+r-1\}}{\{a(t-2)+1\}\left\{v_{t-1}+\frac{ar(t-2)}{a(t-2)+1}\right\}} \leqslant b_0(t-1)$$

if $r \leqslant a(t-2)+1$, which is true for all $t \geqslant 3$ and $k \geqslant 1$. Therefore $v_t/e_t \leqslant b_0(t)$.

If $b_1(t-1)$ is maximal, then

$$\frac{v_t}{e_t} \leqslant \frac{\{r(t-1)+1\}\{v_{t-1}+r-1\}}{a(t-1)\{v_{t-1}+r-1/(t-1)\}} \leqslant b_1(t-1) \leqslant b_1(t).$$

Finally, if, for G_{t-1}, $b_2(t-1)$ is greater than $b_0(t-1)$ and $b_1(t-1)$, then

$$\frac{v_t}{e_t} \leqslant \frac{\{r(t-2)+1\}\{v_{t-1}+r-1\}}{\{a(t-2)+k\}\left\{v_{t-1}+\frac{a[r(t-2)+1]}{a(t-2)+k}\right\}} \leqslant b_2(t-1)$$

whenever $a(t-1) \geqslant k(r-1)$, which holds for all $t \geqslant 3$. Consequently, $v_t/e_t \leqslant b_2(t)$.

To complete the proof of our lemma we consider also the situation when G_{t-1} consists of $t-1$ pairwise vertex disjoint k-trees of size r and next we form the union of G_{t-1} with some k-tree of size r.

Case 1. If $h_t = 0$ $(r \geqslant k+2)$, then

$$\frac{v_t}{e_t} \leqslant \frac{v_{t-1}}{e_{t-1}+1} = \frac{r(t-1)}{a(t-1)+1} = b_0(t).$$

Case 2. If $1 \leqslant h_t \leqslant r-k$, then

$$\frac{v_t}{e_t} \leqslant \frac{r(t-1)+h_t}{a(t-1)+kh_t} \leqslant b_2(t).$$

Case 3. If $r-k+1 \leqslant h_t \leqslant r-1$, then

$$\frac{v_t}{e_t} \leqslant \frac{r(t-1)+h_t}{a(t-1)+a-\binom{r-h_t}{2}} \leqslant b_1(t).$$

Thus the proof is complete.

3. Asymptotic distribution of the number of k-trees. Let us denote by $X_{n,r}^{(k)}$ the number of k-trees of size r in an r.g. $K_{n,p(n)}$. We shall prove the following theorem:

THEOREM 1. *Suppose that $r \geqslant k+1$, $k = 1, 2, \ldots$, and that*

(3) $$\lim_{n\to\infty} p(n)n^{r/a} = \varrho \in (0, \infty) \text{ exists.}$$

Then

$$\lim_{n\to\infty} P(X_{n,r}^{(k)} = i) = \frac{\lambda^i}{i!}e^{-\lambda}, \quad i = 0, 1, 2, \ldots,$$

where

(4) $$\lambda = \frac{1}{r!}\varrho^a\binom{r}{k}[k(r-k)+1]^{r-k-2},$$

and

$$a = a(r, k) = kr-\binom{k+1}{2}.$$

Proof. Let $\mathscr{T}_r^{(k)}$ denote the family of all k-trees of size r which can be formed on the set of n labelled vertices $\{1, 2, \ldots, n\}$. Suppose that $T_{r,0}^{(k)} \in \mathscr{T}_r^{(k)}$ and

$$I_1(T_{r,0}^{(k)}) = \begin{cases} 1 & \text{if } T_{r,0}^{(k)} \subset K_{n,p(n)}, \\ 0 & \text{otherwise,} \end{cases}$$

whereas

$$I_2(T_{r,0}^{(k)}) = \begin{cases} 1 & \text{if } T_{r,0}^{(k)} \subset K_{n,p(n)} \text{ and } T_{r,0}^{(k)} \text{ is vertex disjoint with} \\ & \text{all other } k\text{-trees of size } r \text{ contained in } K_{n,p(n)}, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, let

$$S_t^{(i,j)} = \sum\nolimits^{(i)} \mathrm{E}\big(I_j(T_{r,1}^{(k)})I_j(T_{r,2}^{(k)}) \ldots I_j(T_{r,t}^{(k)})\big), \quad i,j = 1, 2,$$

where $\mathrm{E}(\cdot)$ denotes the expected value, and the summation is over all combinations of t $(t \geqslant 1)$ different k-trees from the family $\mathcal{T}_r^{(k)}$ which for $i = 1$ are pairwise vertex disjoint and for $i = 2$ are not pairwise vertex disjoint. First, we show that if (3) holds, then $S_t^{(2,1)} \to 0$ as $n \to \infty$.

Assume that $t \geqslant 2$, $n \geqslant tr$, $T_{r,1}^{(k)}, T_{r,2}^{(k)}, \ldots, T_{r,t}^{(k)}$ are different k-trees from $\mathcal{T}_r^{(k)}$ which are not pairwise vertex disjoint, and put

$$w_j = \Big|\bigcup_{i=1}^{j-1} \big(V(T_{r,i}^{(k)}) \cap V(T_{r,j}^{(k)})\big)\Big|.$$

Therefore

$$v_j = \Big|V\Big(\bigcup_{i=1}^{j} T_{r,i}^{(k)}\Big)\Big| = jr - w_2 - \ldots - w_j, \quad j = 2, 3, \ldots, t.$$

Thus, for every $r \geqslant k+1$ and given t we get

$$S_t^{(2,1)} \leqslant \frac{1}{t!} \sum_{\substack{0 \leqslant w_2, \ldots, w_t \leqslant r \\ 1 \leqslant w_2 + \ldots + w_t \leqslant (t-1)r}}' \binom{n}{r}\binom{r}{w_2}\binom{n-r}{r-w_2} \times \ldots \times$$

$$\times \binom{(t-1)r - w_2 - \ldots - w_{t-1}}{w_t}\binom{n-(t-1)r+w_2+\ldots+w_{t-1}}{r-w_t} \times$$

$$\times \left\{\binom{r}{k}[k(r-k)+1]^{r-k-2}\right\}^t p^{e_t},$$

where e_t is the number of edges of the graph $\bigcup_{i=1}^{t} T_{r,i}^{(k)}$. Consequently, we obtain

$$S_t^{(2,1)} \leqslant \frac{1}{t!\,r!}\left\{\binom{r}{k}[k(r-k)+1]^{r-k-2}\right\}^t \times$$

$$\times \sum{}' n^{tr-w_2-\ldots-w_t} p^{e_t} \left\{\prod_{j=1}^{t-1} \binom{jr - w_2 - \ldots - w_j}{w_{j+1}} [(r-w_{j+1})!]^{-1}\right\}.$$

By the Lemma we have

$$e_t \geqslant \frac{tr - w_2 - \ldots - w_t}{\max\{b_0, b_1, b_2\}}, \tag{5}$$

where b_0, b_1, and b_2 are determined by formula (1). It follows from (5) and (3) that for every $t \geqslant 2$

$$S_t^{(2,1)} = o(1). \tag{6}$$

On the other hand, for $t \geqslant 1$ we get

$$0 \leqslant S_t^{(1,1)} - S_t^{(1,2)} \leqslant (t+1) S_{t+1}^{(2,1)}.$$

Consequently, for $t \geqslant 2$ we have

$$S_t^{(1,2)} = S_t^{(1,1)} + o(1).$$

But

$$S_t^{(1,1)} = \frac{1}{t!}\binom{n}{r}\binom{n-r}{r}\cdots\binom{n-(t-1)r}{r}\left\{\binom{r}{k}[k(r-k)+1]^{r-k-2}p^a\right\}^t$$

$$\sim \frac{1}{t!}\left\{\frac{1}{r!}(pn^{r/a})^a\binom{r}{k}[k(r-k)+1]^{r-k-2}\right\}^t.$$

Therefore

(7) $$\lim_{n\to\infty} S_t^{(1,2)} = \frac{\lambda^t}{t!}, \quad t \geqslant 1,$$

where λ is given by (4).

If we denote by $Y_{n,r}^{(k)}$ the number of all k-trees of size r being vertex disjoint with all other k-trees of size r in $K_{n,p(n)}$, then using (7) and Bonferroni's inequalities

$$\sum_{j=i}^{i+2s-1} (-1)^{j-i}\binom{j}{i} S_j^{(1,2)} \leqslant P(Y_{n,r}^{(k)} = i) \leqslant \sum_{j=i}^{i+2s} (-1)^{j-i}\binom{j}{i} S_j^{(1,2)}, \quad s \geqslant 1,$$

we obtain

$$\lim_{n\to\infty} P(Y_{n,r}^{(k)} = i) = \frac{\lambda^i}{i!} e^{-\lambda},$$

i.e. $Y_{n,r}^{(k)}$ has the Poisson distribution with the expectation λ. The thesis of our theorem follows immediately from the fact that, by (6),

$$P(X_{n,r}^{(k)} \neq Y_{n,r}^{(k)}) \to 0 \quad \text{as } n \to \infty.$$

Consider now two specific k-trees, namely a 1-tree and a k-tree of size $k+1$.

In the first case ($k=1$) the 1-tree is simply a tree in the usual sense and $X_{n,r}^{(1)}$ denotes the number of trees of size r in $K_{n,p(n)}$. Thus, from Theorem 1 we obtain the following result which was proved earlier by Erdös and Rényi [2]:

COROLLARY 1. *Suppose that* $r \geqslant 2$ *and*

(8) $$\lim_{n\to\infty} p(n)n^{r/(r-1)} = \varrho \in (0, \infty) \text{ exists.}$$

Then the r.v. $X_{n,r}^{(1)}$ *has asymptotically the Poisson distribution with the expectation*

$$\lambda = \frac{1}{r!}\varrho^{r-1} r^{r-2}.$$

In fact, if condition (8) holds, then in an r.g. $K_{n,p(n)}$ all subgraphs of size r being trees are almost surely isolated (see [2], p. 27). Therefore, the r.v. $X_{n,r}^{(1)}$ is the number of isolated trees of size r.

In the second case, where $r = k+1$, the k-tree is the smallest one and is simply a complete graph on $k+1$ vertices. The random variable $X_{n,k+1}^{(k)}$ is now the number of complete subgraphs of size $k+1$ ($k+1$ – cliques) in an r.g. $K_{n,p(n)}$.

Thus, also from Theorem 1, we obtain the following result which was proved earlier by Schürger in [7]:

COROLLARY 2. *Suppose that* $k \geqslant 1$ *and*

$$\lim_{n\to\infty} p(n) n^{2/k} = \varrho \in (0, \infty) \text{ exists.}$$

Then the random variable $X_{n,k+1}^{(k)}$ *has asymptotically the Poisson distribution with the expectation*

$$\lambda = \frac{1}{r!} \varrho^{k(k+1)/2}.$$

Finally, basing on Theorem 1 and using the fact that if a random variable X_λ has the Poisson distribution with the expectation λ, then $(X-\lambda)/\lambda^{1/2}$ has the standardized normal distribution as $\lambda \to \infty$, one can deduce the following

THEOREM 2. *Let* $r \geqslant k+1$ *be fixed and suppose that*

$$\lim_{n\to\infty} p(n) n^{r/a} = \infty,$$

whereas

$$\lim_{n\to\infty} p(n) n^{r/a-\delta} = 0 \quad \text{for all } \delta > 0.$$

Then

$$\lim_{n\to\infty} P\{(X_{n,r}^{(k)} - d)d^{-1/2} < x\} = (2\pi)^{-1/2} \int_{-x}^{x} \exp(-u^2/2) du, \quad -\infty < x < \infty,$$

where

$$d = d(n, r, k, p) = \frac{n^r}{r!} \binom{r}{k} [k(r-k)+1]^{r-k-2} p^a.$$

We omit a proof of this theorem because it follows similar lines as proofs of the respective theorems from [7] and [2].

REFERENCES

[1] L. W. Beineke and R. E. Pippert, *The number of labeled k-dimensional trees*, J. Combinatorial Theory 6 (1969), p. 200-205.

[2] P. Erdös and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960), p. 17-61.

[3] F. Harary and E. M. Palmer, *On acyclic simplicial complexes*, Mathematica 15 (1968), p. 115-122.

[4] J. W. Moon, *The number of labeled k-trees*, J. Combinatorial Theory 4 (1969), p. 206-207.

[5] A. Rényi, *On the number of endpoints of a k-tree*, Studia Sci. Math. Hungar. 5 (1970), p. 5-10.

[6] C. Rényi and A. Rényi, *The Prüfer code for k-trees*, p. 945-971 in: Comb. Theory and its Appl. III, Bolyai János Matem. Tár., Budapest, North Holland Publ. Co., Amsterdam – London 1970.

[7] K. Schürger, *Limit theorems for complete subgraphs of random graphs*, Period. Math. Hungar. 10 (1979), p. 47-53.

Institute of Mathematics
Adam Mickiewicz University
Matejki 48/49, 60-769 Poznań, Poland

Received on 7. 9. 1979