PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 35, No. 1, September 1972

THE PROBABILITY OF CONNECTEDNESS OF AN UNLABELLED GRAPH CAN BE LESS FOR MORE EDGES

E. M. WRIGHT[1]

ABSTRACT. We write $\beta=\beta(n, q)$ for the probability that a graph on n unlabelled nodes with q edges is connected; that is β is the ratio of the number of connected graphs to the total number of graphs. We write $N=n(n-1)/2$. For fixed n we might expect that β would increase with q, at least nonstrictly. On the contrary, we show that, for any given integer s, we have $\beta(n, q+1)<\beta(n, q)$ for $N-n-s\leqq q\leqq N-n$ and $n>n_0(s)$. We can show that $\beta(n, q+1)<\beta(n, q)$ for a much longer range, but this requires much more elaborate arguments.

An (n, q) graph has n nodes and q edges, where each edge is a different unordered pair of nodes. We write $N=n(n-1)/2$, so that $0\leqq q\leqq N$. We write F_{nq} (resp. T_{nq}) for the number of (n, q) graphs in which the nodes are labelled (resp. unlabelled) and f_{nq} (resp. t_{nq}) for the number of these graphs which are connected. Then $\alpha_{nq}=f_{nq}/F_{nq}$ (resp. $\beta_{nq}=t_{nq}/T_{nq}$) is the probability that a labelled (resp. unlabelled) (n, q) graph is connected. It is natural to expect that, for fixed n, these probabilities will increase (in the nonstrict sense) as q increases. For α_{nq}, this is true and can be proved trivially. For β_{nq} it is false; the simplest counterexample is that $\beta_{69}=20/21$ and $\beta_{6,10}=14/15$ (see, for example, the diagrams of all (6, 9) and (6, 10) graphs in [**6**]).

The slightly surprising phenomenon that β_{nq} can decrease as q increases seems worth further study. Here we prove the following theorem.

THEOREM. *For any positive integer v, there is an $n_0=n_0(v)$, such that*

$$\beta_{n,N-n-s} > \beta_{n,N-n-s+1} \qquad (0 \leqq s \leqq v) \tag{1}$$

when $n>n_0$ and, indeed, such that

$$1 - \beta_{n,N-n-s+1} > Cn^{1/2}(1 - \beta_{n,N-n-s}). \tag{2}$$

Presented to the Society, September 15, 1971 under the title *Decreasing probability of connectedness of a graph as number of edges increases*; received by the editors October 9, 1971.

AMS 1969 *subject classifications*. Primary 0565.

Key words and phrases. Unlabelled graphs, probability of connectedness.

[1] The research reported herein has been sponsored by the United States Government.

Thus, for these values of q, β_{nq} retreats rapidly from 1 as q increases. Every statement in what follows is subject to the condition "for large enough n." The positive number C, not always the same at each occurrence, is independent of n and q.

For the moment we consider only unlabelled graphs. Any disconnected (n, q) graph consists of a (k, q_1) graph together with a $(n-k, q-q_1)$ graph for some q_1 and for some k (not necessarily unique) such that $1 \leqq k \leqq n/2$. Compared with the complete (n, N) graph, the disconnected (n, q) graph lacks at least $k(n-k)$ edges, so that $q \leqq N-k(n-k)$. If $q=N-n-s$, where $-1 \leqq s \leqq n-5$, we have $k=1$. Hence a disconnected $(n, N-n-s)$ graph has just 2 components, one an isolated node and the other an $(n-1, N-n-s)$ graph. The latter graph, compared with the complete $(n-1, (n-1)(n-2)/2)$ graph, lacks $\{(n-1)(n-2)/2\}-(N-n-s)=s+1$ edges. Hence it is the complement of an $(n-1, s+1)$ graph. The relation is $(1, 1)$ and so

$$T_{n,N-n-s} - t_{n,N-n-s} = T_{n-1,s+1} \qquad (-1 \leqq s \leqq n-5). \tag{3}$$

Clearly $T_{n-1,0}=T_{n-1,1}=1$ and so $\beta_{n,N-n+1}<\beta_{n,N-n}$ provided $T_{n,N-n+1}<T_{n,N-n}$, i.e. provided $T_{n,n-1}<T_{n,n}$. The smallest n for which this is true is 6, where $T_{65}=15$, $T_{66}=21$ (see [**6**]). Hence $\beta_{6,10}<\beta_{69}$, the counterexample given above.

If $n \geqq 2q$, the (n, q) graphs can be put into $(1, 1)$ correspondence with the $(2q, q)$ graphs by removing $n-2q$ of the isolated nodes from each of the former and conversely. Hence $T_{nq}=T_{2q,q}$ and, from (3), we have

$$T_{n,N-n-s} - t_{n,N-n-s} = T_{2(s+1),s+1} \qquad (-1 \leqq s \leqq v).$$

To prove (2), we have then to show that

$$T_{2s,s}/T_{n,N-n-s+1} > Cn^{1/2}T_{2s+2,s+1}/T_{n,N-n-s}.$$

Since $T_{2s+2,s+1}/T_{2s,s}=C$, it is enough to show that $T_{n,N-n-s}>Cn^{1/2}T_{n,N-n-s+1}$, that is $T_{n,n+s}>Cn^{1/2}T_{n,n+s-1}$.

We shall prove a little more, viz. the following lemma.

LEMMA. *If* $Cn<q<Cn$, *then* $T_{n,q+1}>Cn^{1/2}T_{nq}$.

Henceforth we suppose the hypothesis of the lemma to be satisfied. We now use Polya's famous counting theorem [**1**], [**5**], [**8**]. S_n is the symmetric group (of degree n and order $n!$) of all the permutations π of the n labelled nodes in a labelled (n, q) graph, and $F_\pi=F_\pi(n, q)$ is the number of labelled (n, q) graphs invariant under π. Then Polya's theorem tells us that

$$n!\, T_{nq} = \sum_{\pi \in S_n} F_\pi. \tag{4}$$

We are henceforth concerned with the enumeration of labelled graphs. The permutation π has a unique expression as a product of disjoint cycles, which contains p_j node-cycles of length j, where $1 \leqq j \leqq n$. The node permutation π induces a permutation on the N possible edges in the graph which contains P_j edge-cycles of length j. We have

$$\sum_{j=1}^{n} jp_j = n, \qquad \sum_{j=1}^{n} jP_j = N. \tag{5}$$

Oberschelp [7] gives formulae for P_j in terms of the p_j. All we require here is to note that

$$P_1 = \tfrac{1}{2}p_1(p_1 - 1) + p_2. \tag{6}$$

Any graph invariant under π must contain all the edges of a particular edge-cycle or none. Hence

$$F_\pi(n, q) = \sum{}' \prod_{j=1}^{n} B(P_j, s_j), \tag{7}$$

where $B(h, k) = h!/\{k!(h-k)!\}$, $B(h, 0) = 1$ and $\sum'$ denotes summation over all $s_1, \cdots, s_n$ such that $\sum js_j = q$. We write $M = n^{3/4}$ and separate the sum on the right-hand side of (4) into two parts, so that

$$n!\, T_{nq} = T_1(q) + T_2(q), \tag{8}$$

where $T_1(q)$ contains all the F_π for which $p_1 > M$ and $T_2(q)$ those for which $p_1 \leqq M$.

First we take $p_1 > M$, so that, $P_1 > Cn^{3/2}$. In (7), we replace q by $q+1$, discard those terms on the right in which $s_1 = 0$ and replace s_1 by $s_1 + 1$ in the others. We have then

$$F_\pi(n, q + 1) \geqq \sum{}' \lambda(P_1, s_1) \prod B(P_j, s_j) \geqq Cn^{1/2} F_\pi(n, q),$$

where

$$\lambda(P_1, s_1) = (P_1 - s_1)/(s_1 + 1) \geqq (Cn^{3/2} - q)/(q + 1) \geqq Cn^{1/2}$$

under the hypothesis of the lemma. Hence

$$T_1(q + 1) > Cn^{1/2} T_1(q). \tag{9}$$

We write $[R(X)]_q$ for the coefficient of X^q in the polynomial $R(X)$. From (7),

$$F_\pi(n, q) = \left[\prod (1 + X^j)^{P_j}\right]_q \leqq Y^{-q} \prod (1 + Y^j)^{P_j},$$

where Y is any positive number. Again, if $j > 2$, we have $(1 + Y^j)^2 \leqq (1 + Y^2)^j$. Hence

$$\prod (1 + Y^j)^{P_j} \leqq (1 + Y)^{P_1}(1 + Y^2)^{(N-P_1)/2}$$

by (5). We now put $Y=(\{q/(N-q)\})^{1/2}$ and have $F_\pi(n, q) \leqq V^{P_1} W$, where

$$V^2 = (1 + Y)^2/(1 + Y^2) \leqq 1 + Cn^{-1/2},$$

and

$$W = Y^{-q}(1 + Y^2)^{N/2} = (N/q)^{q/2}\{N/(N - q)\}^{(N-q)/2}.$$

If $p_1 \leqq M$ we have $P_1 \leqq Cn^{3/2}$ and $P_1 \log V < Cn$. Again

$$\begin{aligned} \log W &\leqq (q/2)\log(N/q) + \{(N - q)/2\}\log(1 + \{q/(N - q)\}) \\ &\leqq (q/2)\log q + Cn. \end{aligned}$$

Hence $\log F_\pi(n, q) \leqq (q/2)\log q + Cn$. There are less than $n!$ terms in $T_2(q)$ and so

$$T_2(q) \leqq n!\, q^{q/2} e^{Cn}. \tag{10}$$

We now write $p = [(2q)/\log(2q)]$ so that $M < p \leqq n$. The number of π for which $p_1 = p$ is $B(n, p)D(n-p)$, where $D(m)$ is the number of permutations of the m numbers $1, 2, \cdots, m$ in which no number remains unmoved, i.e. there are no unit cycles. Then $D(m)$ is Euler's rencontre number and is the nearest integer to $m!/e$ (see, for example, **[3]**, **[4]**, **[9]**, **[12]**). Hence the number of π for which $p_1 = p$ is greater than $Cn!/p!$. Again, by (6), for such a π, we have $P_1 \geqq \frac{1}{2}p(p-1)$ and so

$$F_\pi \geqq B(P_1, q) \geqq B(p(p - 1)/2, q) = p!\, \Lambda_p$$

(say). Now $\log(p!) = O(p \log p) = O(q)$ and so

$$\log \Lambda_p = 2q \log p - q \log q + O(q) = q \log q - 2q \log \log q + O(n).$$

Hence

$$T_1(q) \geqq C(n!)\Lambda_p \geqq C(n!)\{q/(\log q)^2\}^q e^{-Cn}. \tag{11}$$

It follows from (8), (10) and (11) that

$$T_2(q) = o(T_1(q)), \qquad n!\, T_{nq} = T_1(q)\{1 + o(1)\}.$$

This is true when $q+1$ replaces q and so our lemma follows from (9).

We can improve the factor $n^{1/2}$ in (2) to $n/(\log n)^2$. More significantly, we can show that (1) holds over a range of q of length $C_1 n$ for any fixed positive C_1 and large enough n. It is possible that we can replace $C_1 n$ by a constant multiple of $n \log n$. But these results require the development of a much more elaborate theory, and, in particular, a study of the asymptotic behaviour of T_{nq} and of $T_{n,q+1}/T_{nq}$ for large n and large q such that $q < (n \log n)/2$.

References

1. N. G. de Bruijn, *Applied combinatorial mathematics*, (ed. E. F. Beckenback), Wiley, New York, 1964, Chap. 5.

2. P. Erdös and A. Rényi, *On random graphs.* I, Publ. Math. Debrecen **6** (1959), 290–297. MR **22** #10924.

3. L. Euler, *Calcul de la probabilité dans le jeu de rencontre*, Mem. Acad. Sci. Berlin **1753**, 255–270; Opera Omnia (1) **7** (1923), 11–25.

4. ———, *Solutio questionis curiosae ex doctrina combinationum*, Mem. Acad. Sci. St. Petersbourg **3** (1811), 57–64; Opera Omnia (1) **7** (1923), 435–448.

5. F. Harary, *The number of linear, directed, rooted, and connected graphs*, Trans. Amer. Math. Soc. **78** (1955), 445–463. MR **16**, 844.

6. ———, *Graph theory*, Addison-Wesley, Reading, Mass., 1969, pp. 214–223. MR **41** #1566.

7. W. Oberschelp, *Kombinatorische Anzahlbestimmungen in Relationen*, Math. Ann. **174** (1967), 53–78. MR **36** #1342.

8. G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Math. **68** (1937), 147–254.

9. J. Riordan, *An introduction to combinatorial analysis*, Wiley Publ. in Math. Statist., Wiley, New York; Chapman and Hall, London, 1958. MR **20** #3077.

10. E. M. Wright, *Asymptotic enumeration of connected graphs*, Proc. Roy. Soc. Edinburgh Sect. A **68** (1968/69), 298–308. MR **42** #1723.

11. ———, *Graphs on unlabelled nodes with a given number of edges*, Acta Math. **126** (1970), 1–9. MR **42** #2975.

12. ———, *Arithmetical properties of Euler's rencontre number*, J. London Math. Soc. **4** (1972), 437–442.

Principal's Office, University of Aberdeen, Aberdeen, United Kingdom