BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 79, Number 4, July 1973

THE PROBABILITY OF CONNECTEDNESS OF A LARGE UNLABELLED GRAPH[1]

BY E. M. WRIGHT

Communicated by Gian-Carlo Rota, December 18, 1972

An (n, q) graph is one with n nodes and q edges, in which any two different nodes are or are not joined by a single edge. We write $T = T(n, q)$ for the number of different (n, q) graphs with unlabelled nodes and t for the number of these graphs which are connected, so that $\beta = t/T$ is the probability that an unlabelled (n, q) graph is connected. We write F, f and α for the corresponding numbers for (n, q) graphs whose nodes are labelled. We write also $N = n(n - 1)/2$, $B(h, k) = h!/\{k!(h - k)!\}$ and $\gamma = (2q - n \log n)/n$. Clearly $q \leqq N$. In what follows, A (not always the same at each occurrence) is a fixed positive number at our choice and all statements are true only for $n > n_0, q > q_0$, where n_0 and q_0 depend on the A.

Erdös and Renyi [1] put $q = [n(\log n + a)/2]$, where a is independent of n and q, and showed that, for these q, we have

$$\alpha \to \exp(e^{-a}) \tag{1}$$

as $n \to \infty$. For given n, it can be shown trivially that α increases steadily (in the nonstrict sense) as q increases. Hence, from (1), it can be at once deduced that, as $n \to \infty$, we have $\alpha \sim \exp(e^{-\gamma})$ and, in particular, that

$$\alpha \to 1 \quad (\gamma \to +\infty), \qquad \alpha \to 0 \quad (\gamma \to -\infty).$$

Elsewhere [4] I have shown that, if $\gamma \to +\infty$, then f has an asymptotic expansion of which the first two terms are

$$f = B(N, q) - nB(N - n + 1, q) - \cdots.$$

Now $F = B(N, q)$ and

$$\frac{nB(N - n + 1, q)}{B(N, q)} = n \prod_{s=0}^{q-1} \frac{N - n + 1 - s}{N - s} \leqq n(N - n + 1)^q N^{-q}$$

and the logarithm of this is less than $\log n - \{q(n - 1)/N\} = -\gamma$. Hence my result leads to $\alpha = 1 - O(e^{-\gamma})$, a statement which is only nontrivial

AMS (MOS) subject classifications (1970). Primary 05C30.

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

[1] The research reported herein was sponsored in part by the United States Government.

when $\gamma \to +\infty$. Thus, for this range of q, I obtain a much more detailed result than Erdös and Renyi. On the other hand, my method (depending on Gilbert's [**2**] generating functions identity) appears incapable of extension to obtain (1), as indeed Erdös and Renyi remark.

My first theorem here gives a result for β corresponding to, but differing from, Erdös and Renyi's result for α. The proof depends heavily on the results of [**5**] and [**7**].

THEOREM 1. *As $n \to \infty$, we have*

$$\beta \sim 1 - e^{-\gamma} \qquad (A < \gamma < A),$$

$$\beta \to 0 \qquad (\overline{\lim}\, \gamma \leqq 0),$$

$$\beta \to 1 \qquad (\gamma \to +\infty).$$

These results are in striking contrast to Erdös and Renyi's. They imply that, when $-A < \gamma < A$, a substantially higher proportion of the labelled graphs are connected than of the unlabelled, at least in the limit as $n \to \infty$.

But there is another, and much more interesting difference in the *proof* required when $\beta \to 0$ or $\beta \to 1$. Erdös and Renyi [**1**] did not need to consider the corresponding cases for α since, for fixed n, the number α increases (nonstrictly) with q. No such result is known for β and indeed, as I showed in [**6**], no such result is true.

The behavior of β for fixed n as q increases presents an interesting problem. Obviously $\beta = 0$ for $q \leqq n - 2$ and $\beta = 1$ for $N - n + 2 \leqq q \leqq N$. What appears to be true otherwise (by calculations based on the table [**3**]) is that, for fixed $n \geqq 6$ and some $q_1 = q_1(n)$, we have

$$\beta(n, q) < \beta(n, q + 1) \qquad (n - 2 \leqq q < q_1),$$

$$\beta(n, q) > \beta(n, q + 1) \qquad (q_1 \leqq q \leqq N - n).$$

All that I can prove, however, is the following theorem.

THEOREM 2. *For $n > n_0$ and some $q_1 = q_1(n)$, we have*

$$\beta(n, q) < \beta(n, q + 1) \qquad (n(A + \log n)/2 < q < q_1), \tag{2}$$

$$\beta(n, q) > \beta(n, q + 1) \qquad (q_1 \leqq q \leqq N - n). \tag{3}$$

We can calculate the integer q_1 with a possible error of 1.

It is surprising that we can define so precisely the range of validity of the unexpected result (3). On the other hand, I cannot prove (2) for $\gamma \leqq 0$, i.e. for $2q \leqq n \log n$, although the tables [**3**] and common sense (that dubious guide) combine to indicate that it must be true. In fact, the proof of (2) for $N/2 \leqq q < q_1$ is easier than that for $q \leqq N/2$ and, in particular, my present proof of (2) for $A < \gamma < A$ is not at all simple.

REFERENCES

1. P. Erdös and A. Renyi, *On random graphs.* I, Publ. Math. **6** (1959), 290–297. MR **22** #10924.
2. E. N. Gilbert, *Enumeration of labelled graphs*, Canad. J. Math **8** (1956), 405–411.
3. M. L. and P. R. Stein, *Enumeration of linear graphs and connected linear graphs up to* $P = 18$ *points*, Los Alamos Scientific Laboratory, 1963.
4. E. M. Wright, *Asymptotic enumeration of connected graphs*, Proc. Roy. Soc. Edinburgh A68 (1970), 298–308.
5. ———, *Graphs on unlabelled nodes with a given number of edges*, Acta Math. **126** (1970), 1–9. MR **42** #2975.
6. ———, *The probability of connectedness of an unlabelled graph can be less for more edges*, Proc. Amer. Math. Soc. **35** (1972), 21–25.
7. ———, *The number of unlabelled graphs with many nodes and edges*, Bull. Amer. Math. Soc. **78** (1972), 1032–1034.

PRINCIPAL'S OFFICE, UNIVERSITY OF ABERDEEN, ABERDEEN, UNITED KINGDOM