The largest component in a random subgraph of the n-cycle

Gyula O.H. Katona

Mathematical Institute of the Hungarian Academy of Sciences, Budapest, P.O. Box 127, 1364, Hungary

Louis V. Quintas

Mathematics Department, Pace University, New York, NY 10038, USA

Received 31 October 1990
Revised 5 May 1991

Abstract

Katona, G.O.H. and L.V. Quintas, The largest component in a random subgraph of the n-cycle, Discrete Mathematics 121 (1993) 113–116.

Let M denote the order of the largest component in a random subgraph H of the n-cycle C_n, where H has the same vertex set as C_n and its edge set is defined by independently selecting, with the same constant probability, each of the edges of C_n. The probability that M is equal to k is known for $k=1$ and for $n \geqslant k \geqslant \lfloor n/2 \rfloor$. Here we obtain the exact result for $k=2$ and comment on the cases $\lfloor n/2 \rfloor > k > 2$.

1. Introduction

Let M denote the order of the largest component in a random subgraph H of the n-cycle C_n, where H has the same vertex set as C_n and its edge set is defined by independently selecting, with the same constant probability, each of the edges of C_n.

In [2] $P(M=k)$, the probability that $M=k$, was determined for $k=1$ and for $n \geqslant k \geqslant \lfloor n/2 \rfloor$. Here we investigate the cases $1 < k < \lfloor n/2 \rfloor$ which left unresolved in [2, Problem 3.5, p. 249]. Specifically, we obtain the exact solution for $k=2$ and the asymptotic solution for fixed k and large n.

First note that

$$P(M=k)=P(M \leqslant k)-P(M \leqslant k-1) \quad \text{with } P(M \leqslant 1)=(1-p)^n. \tag{1}$$

Thus, it is sufficient to determine $P(M \leqslant k)$.

Now, let $f(s,k,p)$ denote the probability that a random subgraph of a path with s edges has its largest component of order at most k. Here a random subgraph of the path is defined in the same way as was done for the n-cycle and with the same edge probability p.

The event $M \leqslant k$ in the n-cycle will occur if and only if any one of the disjoint events (a) or (b) occurs:

(a) Edge $\{1,2\}$ in the n-cycle is absent and the complement of $\{1,2\}$ in the n-cycle does not contain a component of order greater than k; or

(b) For $i=1$ to $i=k-1$, a path of length i in the n-cycle contains the edge $\{1,2\}$, the two edges contiguous to this path are absent, and the complementary path of $n-2-i$ edges in the n-cycle does not contain a component of order greater than k.

Then, with $q=1-p$,

$$P(M \leqslant k) = qf(n-1,k,p) + \sum_{i=1}^{k-1} ip^i q^2 f(n-2-i,k,p). \tag{2}$$

We next note the recurrence relation for f.

$$f(s,k,p)=1 \quad \text{if } 0 \leqslant s \leqslant k-1$$

and (3)

$$f(s,k,p) = \sum_{i=0}^{k-1} p^i q f(s-1-i,k,p) \quad \text{if } s \geqslant k.$$

This is a recurrence relation for $f(s,k,p)$ in the variable s. Its characteristic equation (see e.g. [3, pp. 210–215]) is

$$x^k - qx^{k-1} - pqx^{k-2} - p^2qx^{k-3} - \cdots - p^{k-1}q = 0. \tag{4}$$

2. The case $k=2$

If $k=2$ then the roots of the characteristic equation (4) can be easily calculated as functions of p. Thus, it is an easy exercise to determine $f(s,2,p)$ as a function of s and p. Finally, (2) and (1) lead to the full solution in the present case.

3. The cases $2 < k < \lfloor n/2 \rfloor$

Lemma 1. *The characteristic equation* (4) *has exactly one positive real solution r_k which is larger in absolute value than each of the other solutions and for*

$$w_k(x) = q(1 + (p/x) + (p/x)^2 + \cdots + (p/x)^{k-1}):$$

(a) *if* $p > k/(k+1)$ *then*

$$\max\{kq, 1-p^k\} < r_k < \min\{w_k(kq), p\},$$

(b) *if* $p < k/(k+1)$ *then*

$$\max\{p, w_k(kq), 1-p^k\} < r_k < \min\{kq, 1\},$$

(c) *if* $p = k/(k+1)$ *then* $r_k = p$.

Proof. We consider the characteristic equation (4) in the form

$$x = w_k(x). \tag{5}$$

The right-hand side of this equation corresponds to a curve $y = w_k(x)$ which is strictly decreasing and concave up for all $x > 0$. Thus, there is only one positive real solution r_k.

Furthermore, if z is any nonpositive root of (5), we have $z = w_k(z)$, and from the triangle inequality applied to (5) that

$$|z| \leqslant q(1 + p/|z| + p^2/|z|^2 + \cdots + p^{k-1}/|z|^{k-1})$$

with equality if and only if z is a positive real. That is, $|z| < w_k(|z|)$. Therefore, if z is not a positive real root of (4), we have $|z| < r_k$.

The conditions separating the cases can be written as $p > kq, p < kq$ and $p = kq$, resp. It is helpful to view the positive real solution as corresponding to the intersection of $y = x$ and $y = w_k(x)$. The value of $w_k(x)$ at $x = p$ is kq. This observation settles case (c).

Suppose that $p > kq$, then $kq < r_k < p$. Furthermore, $w_k(kq) > r_k$. These inequalities prove

$$kq < r_k < \min\{w_k(kq), p\} \tag{6}$$

The value of $w_k(x)$ at 1 is $1 - p^k$. This implies

$$1 - p^k < r_k < 1 \tag{7}$$

(6) and (7) settle case (a). Case (b) can be proved similarly. □

As noted in the proof of the above lemma, for all p and k it follows that $1 - p^k \leqslant r_k \leqslant 1$. However, it is the case that for some k (perhaps for all $k \geqslant 3$) there are values of p such that the bounds for r_k are to be found among p, kq, and $w_k(kq)$. For example, we can make the following observation.

For $k = 3$, so that $k/(k+1) = 0.75$, we have

if $p = 0.5$, then $1 - p^k < r_k < 1$, which yields $0.875 < r_3 < 1$;

if $p = 0.7$, then $w_k(kq) < r_k < kq$, which yields $0.715 < r_3 < 0.9$;

if $p = 0.75$, then $r_3 = p = 0.75$; and

if $p = 0.8$, then $kq < r_k < p$, so that $0.6 < r_3 < 0.8$.

Lemma 2. *If* $k \geqslant 2$, *then* $r_k > r_{k-1}$.

Proof. By definition, r_k is a solution of $x=w_k(x)$ while r_{k-1} is a solution of $x=w_{k-1}(x)$. Since $w_{k-1}(x)<w_k(x)$ for all x, we have $r_{k-1}<r_k$. □

Theorem. *If n is large and k is fixed, then $P(M=k)\sim c(k,p)(r_k)^n$, where r_k is the unique positive real solution of* (4) *and $c(k,p)$ is a constant, independent of n.*

Proof. $f(n,k,p)$ is a linear combination of the nth powers of the roots of the characteristics equation (4) (with the obvious modifications in the case of repeated roots). It is known by Lemma 1 that there is a unique and maximum in absolute value real r_k among these roots. The term containing $(r_k)^n$ in $f(n,k,p)$ will dominate if n tends to infinity, that is, $f(n,k,p)\sim d(r_k)^n$, where d is a constant independent of n and $\sim$ denotes that the ratio of the expressions tends to 1 if n tends to infinity. Using (2) we obtain the analogous statement for $P(M\leqslant k)$ with another constant c in place of d. Thus, in $P(M=k)=P(M\leqslant k)-P(M\leqslant k-1)$, the first expression, $P(M\leqslant k)$ will, by Lemma 2, dominate asymptotically.

Therefore, $P(M=k)\sim c(r_k)^n$ as stated. □

Acknowledgements

Support of this work for G.O.H.K. was provided in part by the Hungarian National Science Foundation grant numbers 1812, 1908 and for L.V.Q. by Dyson College of Arts and Sciences Faculty Research and Pace University Scholarly Research Committee Grants.

We are indebted to the anonymous referee for calling our attention to the existence of the paper of Kordecki [1] which is about to appear. However the contents of this paper is not known by us. *Added in proof* (*June*, 1993): Kordecki's paper appeared in November, 1991.

References

[1] W. Kordecki, Random subgraphs of the n-cycle and the n-wheel, Discrete Math. 93 (1991) 35–53.
[2] Z. Palka and L.V. Quintas, Random subgraphs of the n-cycle, Discrete Math. 52 (1984) 243–251.
[3] F.S. Roberts, Applied Combinatorics (Prentice-Hall, Englewood Cliffs, NJ, 1984).