UNIV. BEOGRAD. PUBL. ELEKTROTEHN. FAK.
Ser. Mat. 13 (2002), 30–35.

THE VARIANCE OF THE VERTEX DEGREES OF RANDOMLY GENERATED GRAPHS

Ivan Gutman, Peter Paule

We consider graphs with n vertices and m edges constructed from n isolated vertices by selecting among them, uniformly at random, m pairs and connecting them by m edges. The variance of the vertex degrees of such graphs is shown to be equal to $2m\,(n^2 - n - 2m)/(n^3 + n^2)$. In order to arrive at this formula some combinatorial identities are verified.

1. INTRODUCTION

In connection with our recent studies of some spectral properties of randomly generated graphs [**3**, **4**] we encountered the problem of finding the expected value and the variance of the vertex degrees of such graphs.

A graph possessing n vertices and m edges will be referred to as an (n, m)-graph. The random (n, m)-graphs examined by us were constructed by starting with the $(n, 0)$-graph, selecting in it, uniformly at random, m vertex pairs and connecting them by m edges. (This construction produces labelled (n, m)-graphs uniformly at random.)

One should note that under "*random graph*" one usually understands a graph with a fixed number of vertices, in which there is a given probability that an edge exists between any two vertices [**1**]. Hence the number of edges of such a random graph is a random variable [**1**]. On the other hand, the graphs considered in this paper have a fixed number m of edges, and only the placement of these edges between vertices is chosen by random.

Let G be a random (n, m)-graph, constructed as described above. Denote the number of its vertex pairs by N ; evidently $N = n(n-1)/2$.

In view of the fact that m edges in G can be selected in $\binom{N}{m}$ distinct ways, and that the number of choices in which exactly k edges meet at a certain vertex

2000 Mathematics Subject Classification: 05C80, 05A19
Keywords and Phrases: Random graph, vertex degree variance, variance (of graph)

is $\binom{n-1}{k}\binom{N-(n-1)}{m-k}$, the probability that a vertex in G has degree k is

$$P(k|n,m) = \frac{\binom{n-1}{k}\binom{N-(n-1)}{m-k}}{\binom{N}{m}} \quad . \tag{1}$$

This, of course, applies to any vertex of G .

The expected value $E(d)$ of the degree of a vertex (any vertex) of G is thus

$$E(d) = \sum_{k=0}^{n-1} k\, P(k|n,m) \tag{2}$$

the expected value of the square of the respective degree is

$$E(d^2) = \sum_{k=0}^{n-1} k^2\, P(k|n,m) \tag{3}$$

and the variance

$$\mathrm{Var}\,(d) = E(d^2) - E(d)^2 \ . \tag{4}$$

2. THE MAIN RESULTS

The average vertex degree in any (n,m)-graph is $2m/n$. Consequently $E(d) = 2m/n$, which, in view of Eqs. (1) and (2), is tantamount to

$$\sum_{k=1}^{n-1} k \binom{n-1}{k} \binom{(n-1)(n-2)/2}{m-k} = \frac{2m}{n} \binom{n(n-1)/2}{m}$$

i. e.,

$$\sum_{k=1}^{n} k \binom{n}{k} \binom{n(n-1)/2}{m-k} = \frac{2\,m}{n+1} \binom{n(n+1)/2}{m} \ . \tag{5}$$

The above reasoning provides thus an elementary and utmost simple graph–theory–based proof of the combinatorial identity (5).

The next problem is to find a similar expression for $E(d^2)$ in terms of the parameters n and m . This could be achieved by calculating the sum S_2

$$S_2 = \sum_{k=1}^{n-1} k^2 \binom{n-1}{k} \binom{(n-1)(n-2)/2}{m-k} \ .$$

In what follows we show that

$$S_2 = \frac{2m}{n} \cdot \frac{2m+n-1}{n+1} \binom{n(n-1)/2}{m} \ . \tag{6}$$

From this formula and Eqs. (3) and (4) it is immediate that

$$E(d^2) = \frac{2m(2m+n-1)}{n(n+1)}$$

and

$$\text{Var}\,(d) = \frac{2m\big(n(n-1)-2m\big)}{n^2\,(n+1)} \ . \tag{7}$$

3. PROOF OF FORMULA (6)

The expression (6) for S_2 is deduced by using two elementary identities that are, for instance, found in RIORDAN's seminal book [**6**], namely

$$\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1} \qquad (k \neq 0) \tag{8}$$

and VANDERMONDE's formula

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n} \qquad (n \geq 0) \ . \tag{9}$$

By repeated application of (8) and (9) we obtain

$$\begin{aligned}
S_2 &= (n-1)\sum_{k=0}^{n-2} (k+1)\binom{n-2}{k}\binom{(n-1)(n-2)/2}{m-1-k} \\
&= (n-1)\sum_{k=0}^{n-2} k\binom{n-2}{k}\binom{(n-1)(n-2)/2}{m-1-k} + (n-1)\binom{(n-2)(n+1)/2}{m-1} \\
&= (n-1)(n-2)\sum_{k=1}^{n-2} \binom{n-3}{k-1}\binom{(n-1)(n-2)/2}{m-1-k} + (n-1)\binom{(n-2)(n+1)/2}{m-1} \\
&= (n-1)(n-2)\sum_{k=0}^{n-3} \binom{n-3}{k}\binom{(n-1)(n-2)/2}{m-2-k} + (n-1)\binom{(n-2)(n+1)/2}{m-1} \\
&= (n-1)(n-2)\binom{n-3+(n-1)(n-2)/2}{m-2} + (n-1)\binom{(n-2)(n+1)/2}{m-1} \\
&= \frac{(n-1)(n-2)(m-1)}{n-2+(n-1)(n-2)/2}\binom{n-2+(n-1)(n-2)/2}{m-1} \\
&\qquad\qquad + (n-1)\binom{(n-2)(n+1)/2}{m-1} \\
&= (n-1)\binom{n-2+(n-1)(n-2)/2}{m-1}\cdot\left(\frac{(n-2)(m-1)}{(n-2)(1+(n-1)/2)} + 1\right)
\end{aligned}$$

$$= (n-1)\binom{(n-2)(n+1)/2}{m-1}\left(\frac{2(m-1)}{n+1}+1\right)$$
$$= (n-1)\frac{2m}{n(n-1)}\binom{n(n-1)/2}{m}\frac{2m-2+n+1}{n+1}$$

and Eq. (6) follows.

Thus we proved a combinatorial identity

$$\sum_{k=1}^{n} k^2 \binom{n}{k}\binom{n(n-1)/2}{m-k} = \frac{2m(2m+n)}{(n+1)(n+2)}\binom{n(n+1)/2}{m}. \tag{10}$$

Remark. One can prove this also automatically by using Zeilberger's algorithm. We demonstrate this by using the Paule/Schorn package [**5**]:

```
In[1]:= <<zb.m
Fast Zeilberger by Peter Paule and Markus Schorn. (V 2.4test)
Systembreaker = ENullspace

In[4]:= g = k^2 Binomial[n-1,k] Binomial[(n-1)(n-2)/2,m-k]/
                        Binomial[n(n-1)/2,m]

          2                                (-2 + n) (-1 + n)
         k  Binomial[-1 + n, k] Binomial[-----------------, m - k]
                                                   2
Out[4]= ----------------------------------------------------------
                                     (-1 + n) n
                            Binomial[----------, m]
                                         2

In[5]:= Zb[g,{k,0,m},m,1]
If 'm' is a natural number, then:

Out[5]= {-((1 + m) (-2 + n) (1 + 2 m + n) SUM[m]) +

      m (-2 + n) (-1 + 2 m + n) SUM[1 + m] == 0}
```

It is easy to verify that $\mathrm{SUM}\,(m) = 2m(2m+n-1)/\big(n(n+1)\big)$ satisfies the recurrence. By checking that also the initial value of both sequences in m is the same, we complete the proof of the evaluation.

More generally, the identities (5) and (10) are special instances of

$$\sum_{k=1}^{n} k\binom{r}{k}\binom{s}{n-k} = \frac{n\,r}{r+s}\binom{r+s}{n}$$

and

$$\sum_{k=1}^{n} k^2 \binom{r}{k}\binom{s}{n-k} = \frac{n\,r}{r+s} \cdot \frac{n(r-1)+s}{r-1+s}\binom{r+s}{n}$$

respectively. Also these variants of VANDERMONDE's formula can be easily obtained in the same automatic fashion.

4. DISCUSSION

Observing that the complement $\overline{G}$ of the graph G possesses $\overline{m} = n(n-1)/2 - m$ edges, we can rewrite Eq. (7) as

$$\text{(11)} \qquad \operatorname{Var}(d) = \frac{1}{n+1}\left(\frac{2m}{n}\right)\left(\frac{2\overline{m}}{n}\right) .$$

Clearly, $2m/n$ and $2\overline{m}/n$ are the average vertex degrees in G and $\overline{G}$, respectively.

The maximum value of $\operatorname{Var}(d)$ is achieved when $m = \overline{m}$ and is equal to $(n-1)^2/\big(4(n+1)\big) \sim n/4$.

If m is much smaller than $\overline{m}$ (namely, when the graph G is sparse), then $\operatorname{Var}(d) \approx 2m/n$, i.e.,

$$\operatorname{Var}(d) \approx E(d) .$$

In various graph–theoretical considerations we often encounter graphs that are said to be "almost regular". Formula (11) provides now a measure of this "almost–regularity": an (n,m)-graph may be viewed as "almost–regular" if the variance of its vertex degrees is sufficiently smaller than the right–hand side of (11).

Finally, we wish to point out that in contemporary graph theory another problem related to the variance of the vertex degrees is much examined, namely the characterization of graphs in which this variance is large or the largest possible, see [**2**] and the references cited therein. This direction of research has, however, little in common with the present work.

REFERENCES

1. B. BOLLOBÁS: *Random Graphs.* Academic Press, New York, 1984.
2. Y. CARO, R. YUSTER: *Graphs with large variance.* Ars Combinatoria, **57** (2000), 151–162.
3. I. GUTMAN, T. SOLDATOVIĆ, D. VIDOVIĆ: *The energy of a graph and its size dependence. A Monte Carlo approach.* Chem. Phys. Lett., **297** (1998), 428–432.
4. I. GUTMAN, T. SOLDATOVIĆ, D. VIDOVIĆ: *Modeling the dependence of the π-electron energy on the size of conjugated molecules. A Monte Carlo approach.* ACH Models in Chem., **136** (1999), 599–608.

5. P. Paule, M. Schorn: *A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities.* J. Symbolic Computation, **20** (1995), 673–698.

6. J. Riordan: *Combinatorial Identities.* Wiley, New York, 1968.

Faculty of Science,
University of Kragujevac,
P. O. Box 60,
YU–34000 Kragujevac,
Yugoslavia

(Received March 1, 2000)

Research Institute for Symbolic Computation,
Johannes Kepler University Linz,
A–4040 Linz,
Austria