Shahid Bahonar
University of Kerman

Journal of Mahani Mathematical Research

Print ISSN: 2251-7952 Online ISSN: 2645-4505

JMMR
2012

THE DEGREE VARIANCE AND GEOMETRIC DEGREE VARIANCE OF RANDOM TREES

R. Kazemi

Article type: Research Article

(Received: 12 October 2025, Received in revised form 17 December 2025)

(Accepted: 02 February 2026, Published Online: 02 February 2026)

Abstract. The degree variance and the geometric mean of degrees of the vertices for a graph G are defined as $\mathrm{DVar}(G) = \frac{1}{n}\sum_{i=1}^{n}\left(d(v_i) - \frac{2m}{n}\right)^2$ and $\mathrm{GM}(G) = \sqrt[n]{\prod_{i=1}^{n} d(v_i)}$, respectively, where n, m and $d(v)$ represent the number of vertices, edges and degree of vertex v. Also, the geometric degree variance of graph G is defined as $\mathrm{GDVar}(G) = \frac{1}{n}\sum_{i=1}^{n}(d(v_i) - \mathrm{GM}(G))^2$. By *direct* calculations, we determine the first two moments of degree variance in (uniform) random trees. We also show a convergence in probability associated with this quantity. Finally, we present bounds for the expected value of the geometric mean of the degrees and geometric degree variance.

Keywords: Degree variance, Geometric degree variance, Random tree
2020 MSC: Primary 05C05, 60F05

1. Introduction

Nowadays, the study of structures (or systems) with random behavior is an undeniable issue. In these studies, we often encounter variables that are not independent of each other. The dependence of these variables (quantities or topological indices) is due to their dependence on the size of the structure. Many of these structures are tree-like. Therefore, random tree structures are considered a special type. Among the applications of these structures, we can mention providing a model for the generation of systems, the spread of infectious diseases, the study of family trees preserved from ancient or medieval texts, pyramid schemes, algorithms used in the generation of convex hulls in higher dimensions, for maps of Internet connectivity, some evolving physical systems, the stochastic growth of networks, and the law of the spread of some infectious diseases with an unknown agent such as SARS and covid-19.

Every order-n tree (tree with n vertices) can be obtained uniquely by attaching nth vertex to one of the specific $n-1$ vertices in a tree of order $n-1$. As mentioned earlier, it is of particular interest in applications to assume the random tree model. In this article, a random tree is a tree that is constructed as follows: Initially the tree is empty. When the first vertex arrives, a tree of

✉ r.kazemi@sci.ikiu.ac.ir, ORCID: 0000-0001-5496-7565

https://doi.org/10.22103/jmmr.2026.26096.1885

Publisher: Shahid Bahonar University of Kerman

How to cite: R. Kazemi, *The degree variance and geometric degree variance of random trees*, J. Mahani Math. Res. 2026; 15(2): 219-227.

one vertex is created. The second vertex arrives and a child vertex is adjoined to it. Upon arrival of the third vertex, it joins the tree by selecting as its parent one of the two vertices in it with equal probability $\frac{1}{2}$. After $i-1$ steps, there is a random tree of order $i-1$. When the ith vertex appears, it joins the tree by randomly selecting as its parent any of the existing vertices with equal probability $\frac{1}{i-1}$ [5,6]. Let $d(v)$ denote the degree of vertex v. Bell [2] defined the degree variance of graph G of order n and size m as

$$\mathrm{DVar}(G) = \frac{1}{n}\sum_{i=1}^{n}\left(d(v_i) - \frac{2m}{n}\right)^2,$$

where $\frac{2m}{n}$ is the average degree of the vertices in graph G [3]. It is well known that for two extremal trees, i.e., path P_n and star S_n,

$$\mathrm{DVar}(P_n) = \frac{2(n-2)}{n^2}$$

and

$$\mathrm{DVar}(S_n) = \frac{(n-1)(n-2)^2}{n^2}.$$

Smith and Escudero [8] defined the normalised degree variance as

$$\overline{\mathrm{DVar}}(G) = \frac{n-1}{mn(1-d)}\mathrm{DVar}(G),$$

where

$$d = \frac{2m}{n(n-1)} \in (0,1)$$

is the density of graph G. This valid analytical normalization of the degree variance replaces previous normalizations that were either invalid or not applicable to all networks. They showed that this normalization yields equality values for graphs and their complements; that this value is maximal in star graphs (and their complements); and that its expected value is constant with respect to density for Erdos-Renyi (ER) random graphs of equal size. We have

$$\overline{\mathrm{DVar}}(P_n) = \frac{2}{n^2}$$

and

$$\overline{\mathrm{DVar}}(S_n) = \frac{(n-1)(n-2)}{n^2}.$$

Sayadi et al. [7] define the geometric mean of degrees of the vertices in the graph G as:

$$\mathrm{GM}(G) = \sqrt[n]{\prod_{i=1}^{n} d(v_i)}.$$

They also defined the geometric degree variance of graph G as follows:

$$\mathrm{GDVar}(G) = \frac{1}{n}\sum_{i=1}^{n}(d(v_i) - \mathrm{GM}(G))^2.$$

Table 1. Notations

notation	definition
T_n	random tree of order n
P_n	path of order n
S_n	star of order n
Z_n	the first Zagreb index
D_n	degree variance of T_n
$\overline{D}_n$	normalised degree variance of T_n
G_n	geometric mean of the degrees of T_n
W_n	geometric degree variance of T_n
$M(t)$	the moment generating function

These metrics are fundamental for characterizing graph irregularity and have significant applications in network analysis and the social sciences. Their findings seek to advance the current understanding of graph irregularity and provide a solid foundation for future research in this area [7]. We refer again to the results for extremal trees. For a path P_n,

$$\mathrm{GM}(P_n) = 2^{\frac{n-2}{n}}$$

and for a star S_n,

$$\mathrm{GM}(S_n) = (n-1)^{\frac{1}{n}}.$$

Also for a path P_n [7],

$$\mathrm{GDVar}(P_n) = \frac{1}{n}\left(n(\sqrt[n]{2^{n-2}}-2)^2 + 4\sqrt[n]{2^{n-2}} - 6\right)$$

and for a star S_n,

$$\mathrm{GDVar}(S_n) = (\sqrt[n]{n-1})^2 - 4\frac{n-1}{n}\sqrt[n]{n-1} + n - 1.$$

The lack of independence leads to computational complexity because classical statistical methods based on the independence condition are useless. For example, in most cases we are unable to determine the probability function of the random variable in question, and therefore it seems difficult to determine other statistical properties. One of the goals of this article is to determine the expected value, variance, and in some cases, convergence of the variables under study without using the probability function. Therefore, it is desirable to comment on the above metrics without knowing the degrees of the vertices. Throughout the article, the following variables, which have not been previously studied from a probabilistic perspective, are examined. For better readability, all notations in the article are listed in Table 1.

2. Degree variance

In this section, we will use the martingale approach to determine the expected value and variance of variable of inetrest without having its probability function. This requires the use of a suitable martingale sequence. It is also impossible to use the direct definition to determine the variance because we cannot determine the expected value of square of the random variable. Let D_n be the degree variance of a random tree T_n.

Theorem 2.1. *The expected value of D_n is given by*

$$\mathbb{E}(D_n) = \frac{2(n^2+n-2)}{n^2} - \frac{4}{n}H_{n-1},\ n \geq 3$$

where H_n is the n-th harmonic number.

Proof. Let $Z_n = \sum_{i=1}^{n} d^2(v_i)$ be the first Zagreb index of T_n [3]. By definition,

$$\begin{aligned} n\mathrm{D}_n &= \sum_{i=1}^{n}\left(d(v_i) - \frac{2(n-1)}{n}\right)^2 \\ &= Z_n - \frac{4}{n}(n-1)^2. \qquad (1) \end{aligned}$$

We define $\mathcal{F}_n$ to be the sigma-field generated by the first n stages of the random tree T_n [1,4]. If V_n is a randomly chosen vertex belonging to the n-order tree, then by definition,

$$Z_n = Z_{n-1} + 2d(V_{n-1}) + 2. \qquad (2)$$

Hence,

$$\begin{aligned} \mathbb{E}(Z_n|\mathcal{F}_{n-1}) &= Z_{n-1} + 2\mathbb{E}(d(V_{n-1})|\mathcal{F}_{n-1}) + 2 \\ &= Z_{n-1} + \frac{2}{n-1}\sum_{i=1}^{n-1} d(v_i) + 2 \\ &= Z_{n-1} + 6 - \frac{4}{n-1},\ Z_1 = 0. \qquad (3) \end{aligned}$$

By iteration,

$$\mathbb{E}(Z_n) = (n-1)6 - 4H_{n-1}. \qquad (4)$$

Now, from (1),

$$\mathbb{E}(\mathrm{D}_n) = \frac{2(n^2+n-2)}{n^2} - \frac{4}{n}H_{n-1}.$$

□

For some values of n, $\mathbb{E}(\mathrm{D}_n)$ is presented in the table below.

n	5	10	15	25	30
$\mathbb{E}(\mathrm{D}_n)$	$\frac{43}{75}$	$\frac{6479}{6300}$	$\frac{1687123}{1351350}$	$\frac{81955769933}{55773217500}$	$\frac{26796205393253}{17468171721000}$

In the following theorem, the variance of the variable under investigation is presented without having a probability function and the variance formula.

Theorem 2.2. *The variance of* D_n *is given by*

$$\mathbb{V}(D_n) = \frac{1}{n^2}\sum_{i=2}^{n-1}\xi(i),\ n \geq 3$$

where

$$\xi(i) = \frac{4\left(5(i+1)^2 - 12(i+1) + 2\right)}{i(i+1)} - \frac{16}{i}H_{i-1},\ i \geq 1.$$

Proof. Define

$$\mathrm{D}_{n,n-1} := n\mathrm{D}_n - (n-1)\mathrm{D}_{n-1} + \frac{4}{n(n-1)} - 6.$$

From (2),

$$\begin{aligned} \mathbb{E}(\mathrm{D}^2_{n,n-1}) &= \mathbb{E}\Big(\mathbb{E}(\mathrm{D}^2_{n,n-1})|\mathcal{F}_{n-1}\Big) \\ &= 4\mathbb{E}(\mathbb{E}(d(V_{n-1}))^2|\mathcal{F}_{n-1}) \\ &= \frac{4\mathbb{E}(Z_{n-1})}{n-1} \\ &= \frac{24(n-2)}{n-1} - \frac{16}{n-1}H_{n-2}. \qquad (5) \end{aligned}$$

We have

$$\begin{aligned} &\mathbb{E}\Big(\Big(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n) - (n-1)\mathrm{D}_{n-1} + \mathbb{E}((n-1)\mathrm{D}_{n-1})\Big)(\mathbb{E}(\mathrm{D}_{n,n-1}))\Big) \\ =\ & \mathbb{E}(\mathrm{D}_{n,n-1}) \times \mathbb{E}\Big(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n) - (n-1)\mathrm{D}_{n-1} + \mathbb{E}((n-1)\mathrm{D}_{n-1})\Big) \\ =\ & \mathbb{E}(\mathrm{D}_{n,n-1}) \times 0 \\ (6) =\ & 0. \end{aligned}$$

From relations (3) and (4), the sequence of $(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n))_{n\geq 1}$ is a martingale [1]. Then

$$\begin{aligned} &\mathbb{E}\Big((n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n))((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))\Big) \\ =\ & \mathbb{E}(\mathbb{E}((n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n))((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1})))|\mathcal{F}_{n-1})) \\ =\ & \mathbb{E}(((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))\mathbb{E}(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n)|\mathcal{F}_{n-1})) \\ =\ & \mathbb{E}\Big(((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))\Big) \\ (7) =\ & \mathbb{V}((n-1)\mathrm{D}_{n-1}). \end{aligned}$$

Now, from (6) and (7),

$$
\begin{aligned}
\mathbb{E}(\mathrm{D}^2_{n,n-1}) &= \mathbb{E}\Big(\mathrm{D}^2_{n,n-1} \pm \mathbb{E}(n\mathrm{D}_n) \pm \mathbb{E}((n-1)\mathrm{D}_{n-1})\Big)^2 \\
&= \mathbb{E}(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n) - (n-1)\mathrm{D}_{n-1} + \mathbb{E}((n-1)\mathrm{D}_{n-1}))^2 \\
&+ \mathbb{E}\Big(\mathbb{E}(n\mathrm{D}_n) - \mathbb{E}((n-1)\mathrm{D}_{n-1}) + \frac{4}{n(n-1)} - 6\Big)^2 \\
&+ 2\mathbb{E}\Big((n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n) - (n-1)\mathrm{D}_{n-1} + \mathbb{E}((n-1)\mathrm{D}_{n-1})) \\
&\times (\mathbb{E}(n\mathrm{D}_n) - \mathbb{E}((n-1)\mathrm{D}_{n-1}) + \frac{4}{n(n-1)} - 6)\Big) \\
&= \mathbb{E}(n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n))^2 + \mathbb{E}((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))^2 \\
&- 2\mathbb{E}((n\mathrm{D}_n - \mathbb{E}(n\mathrm{D}_n))((n-1)\mathrm{D}_{n-1} - \mathbb{E}((n-1)\mathrm{D}_{n-1}))) \\
&+ \mathbb{E}\Big(\mathbb{E}(n\mathrm{D}_n) - \mathbb{E}((n-1)\mathrm{D}_{n-1}) + \frac{4}{n(n-1)} - 6\Big)^2 \\
&= \mathbb{V}(n\mathrm{D}_n) - \mathbb{V}((n-1)\mathrm{D}_{n-1}) + \frac{16(n^2-2)^2}{n^2(n-1)^2}. \qquad (8)
\end{aligned}
$$

From (5) and then (8),

$$\mathbb{V}(n\mathrm{D}_n) = \mathbb{V}((n-1)\mathrm{D}_{n-1}) + \xi(n-1),\ \mathbb{V}(S_1) = \mathbb{V}(S_2) = 0.$$

By iteration, the proof is completed. □

Corollary 2.3. *From* (1),

$$D_n > \frac{1}{n} Z_n - 4.$$

From Theorem 2.2, and by Chebyshev's inequality, it follows that $\frac{D_n}{\mathbb{E}(D_n)} \xrightarrow{P} 1$.

A numerical simulation to illustrate convergence to 1 is presented in Figure 1.

Corollary 2.4. *Let* $\overline{D}_n$ *be the normalised degree variance of random tree* T_n*. We have*

$$\overline{D}_n = \frac{1}{n-2} D_n.$$

Then

$$
\begin{aligned}
\mathbb{E}(\overline{D}_n) &= \frac{2(n^2+n-2)}{n^2(n-2)} - \frac{4}{n(n-2)} H_{n-1}, \\
\mathbb{V}(\overline{D}_n) &= \frac{1}{(n(n-2))^2} \sum_{i=2}^{n-1} \xi(i).
\end{aligned}
$$

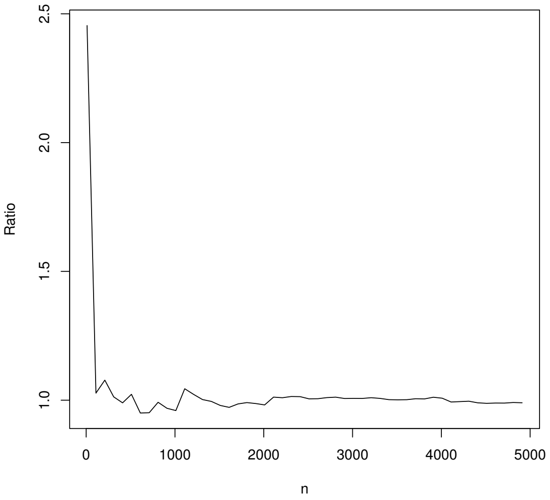

FIGURE 1. A numerical simulation to illustrate convergence to 1

3. Geometric degree variance

In this section, let G_n be the geometric mean of the degrees and $d_n(v)$ denote the degree of vertex v in T_n. By definition of the geometric mean of the degrees and growth rule of the our random tree,

$$G_n = G_{n-1}\left(\frac{d_{n-1}(U)+1}{d_{n-1}(U)}\right)^{\frac{1}{n-1}} \times 1, \tag{9}$$

where U is distributed on the vertex set $\{v_1, v_2, ..., v_{n-1}\}$, uniformly, and independent of $\mathcal{F}_{n-1}$. The recurrence (9) indicates that it is difficult to determine its mathematical expectation. On the other hand, providing lower and upper bounds is sufficient for many studies, including chemical structures [3]. Note that $1 \leq d_n(v) \leq n-1$. Thus, for vertices v_j $(j = 1, ..., n)$,

$$\frac{n}{n-1} \leq \frac{d_n(v_j)+1}{d_n(v_j)} \leq 2, \; j = 1, ..., n. \tag{10}$$

Theorem 3.1. *For $n \geq 3$,*

$$\prod_{i=2}^{n-1}\left(\frac{i}{i-1}\right)^{\frac{1}{i-1}} \leq \mathbb{E}(G_n) \leq 2^{H_{n-1}-1}.$$

Proof. Clearly, from (9),

$$\begin{aligned}\mathbb{E}(G_n|\mathcal{F}_{n-1}) &= \mathbb{E}(G_n|d_{n-1}(v_j),\ j=1,...,n-1)\\ &= \frac{G_{n-1}}{n-1}\sum_{j=1}^{n-1}\left(\frac{d_{n-1}(v_j)+1}{d_{n-1}(v_j)}\right)^{\frac{1}{n-1}}\\ &\leq 2^{\frac{1}{n-1}}G_{n-1},\end{aligned}$$

since G_{n-1} is $\mathcal{F}_{n-1}$-measurable. Then,

$$\mathbb{E}(G_n) \leq 2^{\frac{1}{n-1}}\mathbb{E}(G_{n-1}). \tag{11}$$

It is obvious that $G_2 = 1$. Thus (11) leads to

$$\mathbb{E}(G_n) \leq 2^{H_{n-1}-1}.$$

The lower bound is obtained in a similar way. □

Given the inequality (10) and the unknown degree of the vertices, these bounds are relatively tight.

Corollary 3.2. *From Theorem 3.1 and for large n,*

$$\mathbb{E}(G_n) \leq 2^{\log n - 1}.$$

For an arbitrary $a > 0$ *and by Markov's inequality,*

$$P(G_n > a) \leq \frac{2^{H_{n-1}-1}}{a}.$$

Again, for large n,

$$P(G_n > a) \leq \frac{2^{\log n-1}}{a}.$$

Theorem 3.3. *Let* $M(t) = \mathbb{E}(e^{tG_n^k})$. *Then*

$$\exp\left(t\prod_{i=2}^{n-1}\left(\frac{i}{i-1}\right)^{\frac{1}{i-1}}\right)\right) \leq M(t) \leq \exp(t2^{H_{n-1}-1}),$$

Proof. It is not difficult to show that

$$\prod_{i=2}^{n-1}\left(\frac{i}{i-1}\right)^{\frac{k}{i-1}} \leq \mathbb{E}(G_n^k) \leq 2^{k(H_{n-1}-1)},\ n \geq 3.$$

By definition,

$$M(t) = \sum_{k=0}^{\infty}\frac{\mathbb{E}(G_n^k)}{k!}t^k,$$

and proof is completed. □

Corollary 3.4. *Let W_n be the geometric degree variance of random tree T_n. We have*

$$\mathbb{E}(W_n) \leq \frac{6(n-1)-4H_{n-1}}{n} - \frac{4(n-1)}{n}\prod_{i=2}^{n-1}\left(\frac{i}{i-1}\right)^{\frac{1}{i-1}} + \left(\prod_{i=2}^{n-1}\left(\frac{i}{i-1}\right)^{\frac{1}{i-1}}\right)^2$$

and

$$\mathbb{E}(W_n) \geq \frac{6(n-1)-4H_{n-1}}{n} - \frac{4(n-1)}{n}2^{H_{n-1}-1} + 2^{2H_{n-1}-2}.$$

4. Conclusion

In this paper, by *direct* calculations, two first moments of degree variance in a special random tree are determined. Also, a convergence in probability associated with this quantity and its normalised version are given. Finally, bounds for the expected value of geometric mean of the degrees and geometric degree variance are presented. This approach can be applied to other random trees.

References

[1] Ash, RB. & Doléans-Dade, CA. (2000). Probability and Measure Theory. Second Edition, Academic Press.

[2] Bell, FK. (1992). A note on the irregularity of a graph. Linear Algebra and its Applications, 161, 45–54. https://doi.org/10.1016/0024-3795(92)90004-t

[3] Gutman, I. & Das, KC. (2004). The first Zagreb indices 30 years after. MATCH Communications in Mathematical and in Computer Chemistry, 50, 83–92. https://Match50/83-92.pdf

[4] Kazemi, R. (2014). The eccentric connectivity index of bucket recursive trees. Iranian Journal of Mathematical Chemistry, 5(2), 77-83. https://doi.org/10.22052/ijmc.2014.5671

[5] Kazemi, R. & Behtoei, A. (2017). The first Zagreb and forgotten topological indices of *d*-ary trees. Hacettepe Journal of Mathematics and Statistics, 46(4), 603–611. https://dergipark.org.tr/en/pub/hujms/article/447401?issue_id=38705

[6] Kazemi, R. & Meimondari, L. (2016). Degree distance and Gutman index of increasing trees. Transactions on Combinatorics, 5(2), 23–31. https://doi.org/10.22108/toc.2016.9915

[7] Sayadi, M., Barzegar, H., & Alikhani, S. (2025). More results on degree deviation and degree variance. Journal of Mahani Mathematical Research, 15(1), 111–123. https://doi.org/10.22103/jmmr.2025.24749.1757

[8] Smith, KM., & Escudero, J. (2020). Normalised degree variance. Applied Network Science, 5, 32. https://doi.org/10.1007/s41109-020-00273-3

Ramin Kazemi
Orcid number: 0000-0001-5496-7565
Department of Statistics
Imam Khomeini International University
Qazvin, Iran
Email address: `r.kazemi@sci.ikiu.ac.ir`