Mathematics Interdisciplinary Research **4** (2019) 253 – 261

Probabilistic Properties of F-Indices of Trees

Hadis Morovati, Ramin Kazemi * *and Akram Kohansal*

Abstract

The aim of this paper is to introduce some results for the F-index of the tree structures without any information on the exact values of vertex degrees. Three martingales related to the first Zagreb index and F-index are given.

Keywords: Tree structures, F-indices, martingale.

2010 Mathematics Subject Classification: 05C05, 60F05.

How to cite this article
H. Morovati, R. Kazemi and A. Kohansal, Probabilistic properties of F-indices of trees, *Math. Interdisc. Res.* **4** (2019) 253–261.

1. Introduction

A graph G is a collection of points and lines connecting some pairs of them. The points and lines of a graph are called vertices and edges of that graph, respectively. The vertex set and the edge set of G are denoted by $V(G)$ and $E(G)$, respectively. Let G be a simple connected graph. Two vertices in G which are connected by an edge are called adjacent vertices. The number of vertices adjacent to a given vertex v is the degree of v and is denoted by $d(v)$ (or d_v).

A topological index for a (chemical) graph G is a numerical quantity invariant under automorphisms of G. Topological indices and graph invariants based on the vertex degrees are widely used for characterizing molecular graphs, establishing relationships between structure and properties of molecules, predicting biological activity of chemical compounds, and making their chemical applications. As one of the well-known topological indices, the Zagreb index was introduced by the chemists Gutman and Trinajstić [6]. This index is an important molecular descriptor and has been closely correlated with many chemical properties. In chemistry,

*Corresponding author (E-mail: r.kazemi@sci.ikiu.ac.ir)
Academic Editor: Janez Žerovnik
Received 27 April 2019, Accepted 03 Auguest 2019
DOI: 10.22052/mir.2019.183327.1130

chemical graphs are generated from molecules by replacing atoms with vertices and bonds with edges, or represent only bare molecular skeletons, that is, molecular skeletons without hydrogen atoms. The Zagreb index of a graph G is defined as the sum of the squares of the degrees of all vertices in G:

$$Z_2(G) = \sum_{v \in V(G)} d_v^2.$$

It was immediately recognized that this term increase with the increasing extent of branching of the carbon-atom skeleton and this provide quantitative measure of molecular branching. Ten years later, Z_2 was included among topological indices and was named as *Zagreb group index*. The name Zagreb group index was soon abbreviated to *Zagreb index*, and nowadays Z_2 indicates to the first Zagreb index.

Furtula and Gutman [4] introduced F-index (also called forgotten topological index) which was defined as

$$Z_3(G) = \sum_{v \in V(G)} d_v^3.$$

This topological index has not been further studied till now. Furtula and Gutman [4] raised that the predictive ability of F-index is almost similar to that of first Zagreb index and for the acentric factor and entropy, and both of them obtain correlation coefficients larger than 0.95. This facts show some reasons that why F-index is useful for testing the chemical and pharmacological properties of drug molecular structures. Sun *et al.* [8] deduced some basic nature of F-index and reported that this index can reinforce the physico-chemical flexibility of Zagreb indices. Gao *et al.* [5] computed the F-index of some significant drug molecular structures. Che and Chen [2] provided new lower and upper bounds of the F-index in terms of graph irregularity, Zagreb indices, graph size, and maximum/minimum vertex degrees. They characterized all graphs that attain these new bounds of F-index and they showed that their bounds are better than the bounds given in [4] for all benzenoid systems with more than one hexagon. As corollaries, various upper bounds of F-index easily follow. Moreover, upper bounds for connected K_{r+1}-free graphs are also obtained.

Let F be a real valued function defined on non negative integers. Došlić *et al.* [3] studied a general relation for topological indices of the form $T(G) = \sum_{v \in V(G)} F(d_v)$ as follows. Let G be a connected graph, F be a real valued function defined on non negative integers and $T(G) = \sum_{v \in V(G)} F(d_v)$ be a graph invariant. Then

$$T(G) = \sum_{uv \in E(G)} \Big(\frac{F(d_u)}{d_u} + \frac{F(d_v)}{d_v} \Big),$$

where $E(G)$ is the edge set of G. Thus

$$Z_3(G) = \sum_{uv\in E(G)} (d_u^2 + d_v^2).$$

Trees are defined as connected graphs without cycles, and their properties are basics of graph theory. The structures of many molecules such as dendrimers, alkanes and acyclic molecules are tree like. Trees have wide applications in chemical graph theory such as enumeration and coding problems of chemical structures. Structures of chemical compounds can be synthesized and categorized through mathematical means. Chemists have a long tradition of using atomic valences (vertex degrees) to find molecular structures graphically. However, if two tree structures (molecular graphs) have different values for F-indiex, then their structures are different and hence, they have some different physico-chemical properties.

Every tree structure of order n (or with n vertices) can be obtained uniquely by attaching n-th vertex to one of the $n-1$ vertices in a tree of order $n-1$. It is one of particular interests in applications to assume the random tree model and to speak about a random tree with n vertices, which means that all trees of order n are considered to appear equally likely. Equivalently, one may describe random trees via the following tree evolution process, which generates random trees of arbitrary order n. At step 1 the process starts with a vertex. At step i the i-th vertex is attached to a previous vertex v of the already grown tree T of order $i-1$ with probability $p_i(v) = \frac{1}{i-1}$. For applicability of our own results and specially for some connections with the chemical relevance, see [7].

2. Results

Let $Z_{2,n} = \sum_{i=1}^{n} d_{v_i}^2$ be the first Zagreb index of a tree structure of order n. Also, let $Z_{3,n} = \sum_{i=1}^{n} d_{v_i}^3$ be the its F-index.

Theorem 2.1. For each random tree structure of order $n \geq 3$,

$$\mathbb{E}(Z_{2,n}) = 6n + \mathcal{O}(\log n).$$

Proof. Let U_n be a randomly chosen vertex belonging to a tree of order n. Let $\mathcal{F}_n$ be the sigma-field generated by the first n stages of these trees [1]. By definition,

$$Z_{2,n} = Z_{2,n-1} + (d_{U_{n-1}} + 1)^2 - d_{U_{n-1}}^2 + 1 = Z_{2,n-1} + 2d_{U_{n-1}} + 2.$$

Hence,

$$\begin{aligned}
\mathbb{E}(Z_{2,n}|\mathcal{F}_{n-1}) &= Z_{2,n-1} + 2\mathbb{E}(d_{U_{n-1}}|\mathcal{F}_{n-1}) + 2 \\
&= Z_{2,n-1} + 2\,\frac{1}{n-1}\sum_{i=1}^{n-1} d_{v_i} + 2 \\
&= Z_{2,n-1} + 6 - \frac{4}{n-1},
\end{aligned}$$

because $Z_{2,n-1}$ is $\mathcal{F}_{n-1}$-measurable [1]. We have

$$\begin{aligned}\mathbb{E}(Z_{2,n}) &= \mathbb{E}(Z_{2,n-1}) + 6 - \frac{4}{n-1}\\ &= \left(\mathbb{E}(Z_{2,n-2}) + 6 - \frac{4}{n-2}\right) + 6 - \frac{4}{n-1}\\ &\vdots\\ &= (n-1)6 - 4H_{n-1},\end{aligned}$$

note that $Z_{2,1} = 0$ and H_n is the n-th harmonic number. □

Theorem 2.2. For each random tree structure of order $n \geq 3$,

$$\mathbb{E}(Z_{3,n}) = 26n + \mathcal{O}(\log n),$$
$$\mathbb{E}(Z_{4,n}) = 150n + \mathcal{O}(\log n),$$

where $Z_{4,n} = \sum_{i=1}^{n} d_{v_i}^4$.

Proof. By stochastic growth role of the random tree structure,

$$Z_{3,n} = Z_{3,n-1} + (d_{U_{n-1}} + 1)^3 - d_{U_{n-1}}^3 + 1 = Z_{3,n-1} + 3d_{U_{n-1}}^2 + 3d_{U_{n-1}} + 2,$$

and

$$\begin{aligned}Z_{4,n} &= Z_{4,n-1} + (d_{U_{n-1}} + 1)^4 - d_{U_{n-1}}^4 + 1\\ &= Z_{4,n-1} + 4d_{U_{n-1}}^3 + 6d_{U_{n-1}}^2 + 4d_{U_{n-1}} + 2.\end{aligned}$$

This implies that

$$\begin{aligned}\mathbb{E}(Z_{3,n}|\mathcal{F}_{n-1}) &= Z_{3,n-1} + 3\mathbb{E}(d_{U_{n-1}}^2|\mathcal{F}_{n-1}) + 3\mathbb{E}(d_{U_{n-1}}|\mathcal{F}_{n-1}) + 2\\ &= Z_{3,n-1} + \frac{3}{n-1}\sum_{k=1}^{n-1} d_{v_k}^2 + \frac{3}{n-1}\sum_{k=1}^{n-1} d_{v_k} + 2\\ &= Z_{3,n-1} + \frac{3}{n-1}\ Z_{2,n-1} + 8 - \frac{6}{n-1}.\end{aligned}$$

Thus,

$$\mathbb{E}(Z_{3,n}) = \mathbb{E}(Z_{3,n-1}) + \frac{3}{n-1}\ \mathbb{E}(Z_{2,n-1}) + 8 - \frac{6}{n-1}.$$

From Theorem 2.1, $\mathbb{E}(Z_{3,n}) = 26n + +\mathcal{O}(\log n)$ since $Z_{3,1} = 0$. With the same approach we can obtain $\mathbb{E}(Z_{4,n}) = 150n + \mathcal{O}(\log n)$. □

Lemma 2.3. Let $\mathbb{C}ov(Z_{3,n}, Z_{3,n-1}) = \mathbb{E}((Z_{3,n} - \mathbb{E}(Z_{3,n}))(Z_{3,n-1} - \mathbb{E}(Z_{3,n-1})))$ be the covariance between $Z_{3,n}$ and $Z_{3,n-1}$. Then

$$\mathbb{C}ov(Z_{3,n-1}, Z_{3,n}) = \mathbb{V}(Z_{3,n-1}) + r_{n-1},$$

where

$$r_i = \frac{3}{i}\mathbb{C}ov(Z_{2,i}, Z_{3,i}), \quad i = 1, 2, 3, \ldots.$$

Proof. We have

$$\mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n})|\mathcal{F}_{n-1}) = Z_{3,n-1} - \mathbb{E}(Z_{3,n-1}) + \frac{3}{n-1}(Z_{2,n-1} - \mathbb{E}(Z_{2,n-1})).$$

Then

$$\begin{aligned}\mathbb{C}ov(Z_{3,n}, Z_{3,n-1}) &= \mathbb{E}(\mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}))(Z_{3,n-1} - \mathbb{E}(Z_{3,n-1})))|\mathcal{F}_{n-1})) \\ &= \mathbb{E}((Z_{3,n-1} - \mathbb{E}(Z_{3,n-1}))\mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n})|\mathcal{F}_{n-1})) \\ &= \mathbb{V}(Z_{3,n-1}) + r_{n-1}.\end{aligned}$$

□

Set

$$b_i := \frac{1}{i}(9\mathbb{E}(Z_{4,i}) + 18\mathbb{E}(Z_{3,i}) + 9\mathbb{E}(Z_{2,i})), \quad i = 1, 2, \ldots.$$

Theorem 2.4. For each random tree structure of order $n \geq 3$,

$$\mathbb{V}(Z_{3,n}) = \sum_{i=1}^{n-1}\left(b_i + \frac{3}{i}\mathbb{C}ov(Z_{2,i}, Z_{3,i}) - \left(26 - \frac{26}{i} + 4H_{i-1}\right)^2\right).$$

Proof. We have

$$\begin{aligned}& \mathbb{E}(Z_{3,n} - Z_{3,n-1} - 2)^2 \\ = \; & \mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}) - Z_{3,n-1} + \mathbb{E}(Z_{3,n-1}) + \mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)^2 \\ = \; & \mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}) - Z_{3,n-1} + \mathbb{E}(Z_{3,n-1}))^2 \\ + \; & \mathbb{E}(\mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)^2 \\ + \; & 2\mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}) - Z_{3,n-1} + \mathbb{E}((Z_{3,n-1}))(\mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)) \\ = \; & \mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}))^2 + \mathbb{E}(Z_{3,n-1} - \mathbb{E}(Z_{3,n-1}))^2 \\ - \; & 2\mathbb{E}((Z_{3,n} - \mathbb{E}(Z_{3,n}))(Z_{3,n-1} - \mathbb{E}(Z_{3,n-1}))) \\ + \; & \mathbb{E}(\mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)^2,\end{aligned}$$

since

$$\mathbb{E}(Z_{3,n} - \mathbb{E}(Z_{3,n}) - Z_{3,n-1} + \mathbb{E}(Z_{3,n-1}))(\mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)) = 0.$$

From Theorem 2.2,

$$\begin{aligned}\mathbb{E}(\mathbb{E}(Z_{3,n}) - \mathbb{E}(Z_{3,n-1}) - 2)^2 &= \left(\frac{3}{n-1}\ \mathbb{E}(Z_{2,n-1}) + 8 - \frac{6}{n-1}\right)^2 \\ &= \left(\frac{3}{n-1}\ (6(n-2) + 4H_{n-2}) + 8 - \frac{6}{n-1}\right)^2 \\ &= \left(26 - \frac{26}{n-1} + 4H_{n-2}\right)^2.\end{aligned}$$

By definition of the variance and Lemma 2.3,

$$\begin{aligned}
\mathbb{E}(Z_{3,n}-Z_{3,n-1}-2)^2 &= \mathbb{V}(Z_{3,n})+\mathbb{V}(Z_{3,n-1}) \\
&- 2\mathbb{E}((Z_{3,n}-\mathbb{E}(Z_{3,n}))(Z_{3,n-1}-\mathbb{E}(Z_{3,n-1}))) \\
&+ \Big(26-\frac{26}{n-1}+4H_{n-2}\Big)^2 \\
&= \mathbb{V}(Z_{3,n})-\mathbb{V}(Z_{3,n-1})-r_{n-1} \\
&+ \Big(26-\frac{26}{n-1}+4H_{n-2}\Big)^2. \qquad (1)
\end{aligned}$$

Also,

$$\begin{aligned}
\mathbb{E}(Z_{3,n}-Z_{3,n-1}-2)^2 &= 9\mathbb{E}(d^2_{U_{n-1}}+d_{U_{n-1}})^2 \\
&= 9\mathbb{E}(d^4_{U_{n-1}}+2d^3_{U_{n-1}}+d^2_{U_{n-1}}) \\
&= 9\sum_{i=1}^{n-1}\frac{\mathbb{E}(d_i^4)}{n-1}+18\sum_{i=1}^{n-1}\frac{\mathbb{E}(d_i^3)}{n-1}+9\sum_{i=1}^{n-1}\frac{\mathbb{E}(d_i^2)}{n-1} \qquad (2) \\
&= \frac{9}{n-1}\mathbb{E}(Z_{4,n-1})+\frac{18}{n-1}\mathbb{E}(Z_{3,n-1})+\frac{9}{n-1}\mathbb{E}(Z_{2,n-1}).
\end{aligned}$$

Now, from (1) and (2),

$$\mathbb{V}(Z_{3,n})=c_{n-1}+\mathbb{V}(Z_{3,n-1}),$$

where

$$c_i=b_i+\frac{3}{i}\mathbb{C}ov(Z_{2,i},Z_{3,i})-\Big(26-\frac{26}{i}+4H_{i-1}\Big)^2, \qquad i=1,2,\ldots.$$

By iteration, $\mathbb{V}(Z_{3,n})=\sum_{i=1}^{n-1}c_i$ and proof is completed. □

The sequence $(X_n)_{n\geq 1}$ of random variables is said to be a *martingale* relative to the sigma-field $\mathcal{F}_n$ if and only if for all $n=1,2,\ldots$, $\mathbb{E}(X_{n+1}|\mathcal{F}_n)=X_n$ (*a.e.*) [1]. From Theorem 2.1, the sequence $(Z_{2,n}-\mathbb{E}(Z_{2,n}))_{n\geq 1}$ is a martingale. Also, From Theorem 2.2, the sequence $(Z_{3,n}-\mathbb{E}(Z_{3,n}))_{n\geq 1}$ is not a martingale. In passing, we introduce three martingale structures related to $Z_{2,n}$ and $Z_{3,n}$. These martingales can be important to study the asymptotic normality of these indices.

Theorem 2.5. Assume

$$Z_n^*=Z_{3,n}-3H_{n-1}Z_{2,n}+18nH_{n-1}+6H_{n-1}-6(H^2_{n-1}+3H^{(2)}_{n-1})-26(n-1),$$

where $H_n^{(2)}$ is the n-th harmonic number of order 2. Then, the process $\{Z_n^*,\mathcal{F}_n\}_{n\geq 1}$ is a martingale.

Proof. Set

$$M_n = Z_{3,n} - \alpha_n Z_{2,n} + \beta_n$$

for yet-to-be computed deterministic functions that will render M_n a martingale. Recall that we also denote the sigma-field generated by the first n evolutionary steps by $\mathcal{F}_n$. According to the fundamental property of martingales, we want to have

$$\begin{aligned}\mathbb{E}(M_n|\mathcal{F}_{n-1}) &= \mathbb{E}(Z_{3,n}|\mathcal{F}_{n-1}) - \alpha_n \mathbb{E}(Z_{2,n}|\mathcal{F}_{n-1}) + \beta_n \\ &= M_{n-1} \\ &= Z_{3,n-1} - \alpha_{n-1} Z_{2,n-1} + \beta_{n-1}.\end{aligned}$$

This is possible, if we equate the coefficients of $Z_{2,n-1}$ and also the constant terms. Equating the coefficients of $Z_{2,n-1}$ gives the recurrence for α_n. For $n \geq 2$, we obtain

$$\alpha_n = \frac{3}{n-1} + \alpha_{n-1},$$

which has the solution $\alpha_n = 3H_{n-1}$ if $\alpha_1 = 0$. Equating the constant terms gives

$$\beta_n = \beta_{n-1} - 8 + 18H_{n-1} + \frac{6}{n-1} - \frac{12H_{n-1}}{n-1}.$$

Setting $\beta_1 = 0$,

$$\beta_n = -8(n-1) + 18\sum_{i=1}^{n-1} H_j + 6H_{n-1} - 12\sum_{i=1}^{n-1} \frac{H_j}{j}.$$

Since

$$\sum_{i=1}^{n-1} H_j = n(H_n - n), \qquad \sum_{i=1}^{n-1} \frac{H_j}{j} = \frac{1}{2}(H_{n-1}^2 + H_{n-1}^{(2)}),$$

proof is completed. □

Corollary 2.6. If

$$W_n^* = Z_{2,n} - 3H_{n-1}Z_{3,n} - 18nH_{n-1} - 6H_{n-1} + 6(H_{n-1}^2 + 3H_{n-1}^{(2)}) + 26(n-1),$$

then, the process $\{W_n^*, \mathcal{F}_n\}_{n\geq 1}$ is a martingale.

Theorem 2.7. Assume

$$S_n^* = Z_{3,n} - \mathbb{E}(Z_{3,n}) - (a + 3H_{n-1})(Z_{2,n} - \mathbb{E}(Z_{2,n})), \qquad a \in \mathbb{R}.$$

Then, the process $\{S_n^*, \mathcal{F}_n\}_{n\geq 1}$ is a zero-mean martingale.

Proof. Set

$$L_n = \frac{\alpha_n}{n}(Z_{3,n} - \mathbb{E}(Z_{3,n})) + \frac{\beta_n}{n}(Z_{2,n} - \mathbb{E}(Z_{2,n})).$$

According to the fundamental property of martingales, we want to have

$$\begin{aligned}\alpha_n &= \frac{n}{n-1}\alpha_{n-1},\\ \beta_n &= n\left(\frac{\beta_{n-1}}{n-1} - \frac{3\alpha_n}{n(n-1)}\right).\end{aligned}$$

Setting $\alpha_1 = 1$ and $\beta_1 = a$, proof is completed. □

As $n \to \infty$, the coefficients in the martingale structure have the following asymptotic values:

$$\begin{aligned}\alpha_n &\sim n,\\ \beta_n &\sim -a - 3\log n.\end{aligned}$$

3. Conclusion

In this paper, we studied the F-index of random tree structures. We obtained the exact and asymptotic values of the mean of this index. Also, we introduced a relation for the variance of the F-index in terms of the covariance between the two indices. With the approach presented here, the study of another topological indices is possible. Three martingales introduced that can be important to study the asymptotic normality of this index.

Conflicts of Interest. The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, 1885.

[2] Z. Che and Z. Chen, Lower and upper bounds of the forgotten topological index, *MATCH Commun. Math. Comput. Chem.* **76** (2016) 635–648.

[3] T. Došlić, T. Réti and D. Vukičević, On the vertex degree indices of connected graphs, *Chem. Phys. Lett.* **512** (2011) 283–286.

[4] B. Furtula and I. Gutman, A forgotten topological index, *J. Math. Chem.* **53** (2015) 1184–1190.

[5] W. Gao, M. R. Farahani and L. Shi, Forgotten topological index of some drug structures, *Acta Med. Medit.* **32** (2016) 579–585.

[6] I. Gutman and N. Trinajstić, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, *Chem. Phys. Lett.* **17** (1972) 535–538.

[7] R. Kazemi, The eccentric connectivity index of bucket recursive trees, *Iranian J. Math. Chem.* **5** (2014) 77–83.

[8] Y. C. Sun, Z. Lin, W. X. Peng, T. Q. Yuan, F. Xu, Y. Q. Wu, J. Yang, Y. S. Wang and R. C. Sun, Chemical changes of raw materials and manufactured binderless boards during hot pressing: lignin isolation and characterization, *BioResources* **9** (2014) 1055–1071.

Hadis Morovati
Department of Statistics,
Imam Khomeini International University,
Qazvin, I. R. Iran
E-mail: hadis72mo@gmail.com

Ramin Kazemi
Department of Statistics,
Imam Khomeini International University,
Qazvin, I. R. Iran
E-mail: r.kazemi@sci.ikiu.ac.ir

Akram Kohansal
Department of Statistics,
Imam Khomeini International University,
Qazvin, I. R. Iran
E-mail: kohansal@sci.ikiu.ac.ir