MATEMATIČKI VESNIK
МАТЕМАТИЧКИ ВЕСНИК
73, 1 (2021), 55–62
March 2021

research paper
оригинални научни рад

THE MOMENTS OF THE SACKIN INDEX OF RANDOM d-ARY INCREASING TREES

Ramin Kazemi and Ali Behtoei

Abstract. For any fixed integer $d \geq 2$, the d-ary increasing tree is a rooted, ordered, labeled tree where the out-degree is bounded by d, and the labels along each path beginning at the root increase. Total path length, or search cost, for a rooted tree is defined as the sum of all root-to-node distances and the Sackin index is defined as the sum of the depths of its leaves. We study these quantities in random d-ary increasing trees.

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 for convenience). A path in a graph is a sequence of adjacent edges, which do not pass through the same vertex more than once, and the length of the path is the number of edges in it. Trees are defined as connected graphs without cycles, and their properties are basics of graph theory. The analysis of the length of paths in trees has received a lot of attention mostly because of their importance in the analysis of algorithms. A rooted tree is a tree with a countable number of nodes, in which a particular node is distinguished from the others and called the root node. In a rooted tree, the number of immediate descendants of a vertex v is called its out-degree and is denoted by $d^+(v)$ (or d_v^+).

Increasing trees are labelled trees where the nodes of a tree of order n are labelled by distinct integers of the set $\{1, \ldots, n\}$ in such a way that each sequence of labels along any branch starting at the root is increasing. Several important tree models, e.g., recursive trees, plane-oriented recursive trees (also known as non-uniform recursive trees or heap ordered trees) and binary increasing trees (and more generally d-ary

2020 Mathematics Subject Classification: 05C05, 60F05
Keywords and phrases: d-ary increasing tree; total path length; Sackin index; covariance.

increasing trees) are members of the family of increasing trees. In fact, for any fixed integer $d \geq 2$, the d-ary increasing tree is a rooted, ordered, labeled tree where the out-degree is bounded by d, and the labels along each path beginning at the root increase [2].

There is a simple growth rule for the class of d-ary increasing trees. In this class, a random tree T_n, of order n, is obtained from T_{n-1}, a random tree of order $n-1$, by choosing a parent in T_{n-1} and adjoining a node labeled n to it.

We explain the following evolution processes for random d-ary increasing trees of order n, which turns out to be appropriate when studying the Sackin index and total path length of these trees. The possible insertion positions to join a new node to a d-ary increasing tree, are called external nodes. In a d-ary increasing tree, the number of nodes can be attached to node v of out-degree $d^+(v)$ is $d - d^+(v)$. Therefore the number of all external nodes in a d-ary increasing tree T_n of order n is $\sum_{v \in V(T_n)}(d - d^+(v)) = (d-1)n + 1$.

At the step 1 the process starts with the root. At the step i the i-th node is attached to a previous node v of the already grown d-ary increasing tree T_{i-1} of order $i-1$ with probability $p_i(v) = \frac{d-d^+(v)}{(d-1)(i-1)+1}$. It is obvious that $d^+_{root} = d_{root}$ and for other vertices $d^+_v = d_v - 1$. Thus we have $p_i(root) = \frac{d-d_{root}}{(d-1)(i-1)+1}$ and for other vertices $p_i(v) = \frac{d-d_v+1}{(d-1)(i-1)+1}$. This fact specially implies that the higher out-degree vertices possess a lower attraction for new neighbors and there exists no vertex of out-degree greater than d. The distance $D_{n,j}$ (or D_j for convenience) between the root and node j (the depth of j-th node) in a random d-ary increasing tree of order n has been studied by Panholzer and Prodinger [9]. They proved that ($d = 2$):

$$\mathbb{E}(D_{n,j}) = 2H_j - 2, \quad \mathbb{V}ar(D_{n,j}) = 2H_j + 2 - 4H_j^{(2)},$$

where H_n is the n-th harmonic number and $H_n^{(2)}$ is the n-th harmonic number of order 2.

The Sackin index S_n of a tree of order n is defined as the sum of the depths of its leaves. Clearly, if two tree structures (molecular graphs) have different values for Sackin index, then their structures are different and hence, they have some different properties. Let S_n be the Sackin index of a random d-ary increasing tree of order n and $\mathcal{F}_n$ be the sigma-field generated by the first n stages of these trees [3,5]. Moradian et al. [7] have studied the Sackin index in random recursive trees. They showed that $\mathbb{E}(S_n) = \frac{n}{2}\left(H_n - \frac{1}{2}\right)$. Let U_n be a randomly chosen node belonging to a d-ary tree of order n.

The total path length of a d-ary increasing tree, namely, $I_n = \sum_{j=1}^{n} D_{n,j}$, is defined as the sum of all root-to-node distances (in which nodes consists of leaf and non-leaf nodes). This random variable can be served as a global measure of the cost of constructing the tree. The expectation and variance of the external path length in a random d-ary increasing tree of order n is investigated in [4]. Linearity of expectation gives $\mathbb{E}(I_n) = \sum_{j=1}^{n} \mathbb{E}(D_{n,j})$.

For more information about these subjects or related subjects see [1,4,6].

The aim of this paper is to determine some distributional propeties of Sackin index and total path length for each randomly chosen d-ary increasing tree. Also, we provide some useful relations related to these indices or other parameters.

2. The main results

THEOREM 2.1. *Let I_n be the total path length of the random d-ary increasing tree of order n. Then*

$$\mathbb{E}(I_n)=\sum_{i=1}^{n-1}\Big(\mu_i\prod_{j=i+1}^{n-1}\lambda_j\Big),$$

where
$$\lambda_i=\frac{(i+1)(d-1)+1}{i(d-1)+1},\quad \mu_i=\frac{id}{i(d-1)+1},\qquad i=1,2,3,\dots.$$

Proof. Let $D(v)$ (or D_v for convenience) be the depth of node v. It is not difficult to show that

$$\sum_{v\in V(T_n),d^+(v)\geq 1} d^+(v)\,D(v)=I_n-(n-1).$$

By stochastic growth role of the tree, $I_n=I_{n-1}+D_{U_{n-1}}+1$. Thus

$$\begin{aligned}
\mathbb{E}(I_n|\mathcal{F}_{n-1}) &= I_{n-1}+\sum_{v\in V(T_{n-1})} p_n(v)D(v)+1\\
&= I_{n-1}+\frac{1}{(d-1)(n-1)+1}\Big(\sum_{v\in V(T_{n-1})}(d-d_v^+)D(v)\Big)+1\\
&= I_{n-1}+\frac{1}{(d-1)(n-1)+1}\Bigg[d\sum_{v\in V(T_{n-1})}D(v)-\sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} d_v^+D(v)\Bigg]+1\\
&= I_{n-1}+\frac{1}{(d-1)(n-1)+1}\big[dI_{n-1}-(I_{n-1}-(n-2))\big]+1\\
&= \frac{n(d-1)+1}{(n-1)(d-1)+1}I_{n-1}+\frac{d(n-1)}{(n-1)(d-1)+1}.
\end{aligned}$$

Since $\mathbb{E}(I_n)=\mathbb{E}(\mathbb{E}(I_n|\mathcal{F}_{n-1}))$, for each $n\geq 1$ we can obtain $\mathbb{E}(I_n)$ by iteration as follows:

$$\mathbb{E}(I_n)=\sum_{i=1}^{n-1}\Big(\mu_i\prod_{j=i+1}^{n-1}\lambda_j\Big).\qquad\square$$

For each integer $n\geq 1$, we define

$$c[n]:=\begin{cases}\frac{\Gamma(n-1)}{\Gamma\left(n+\frac{1}{d-1}\right)}, & n\geq 2\\ 0, & n=1,\end{cases}$$

where $\Gamma(\cdot)$ is the gamma function. This is well defined because $n+\frac{1}{d-1}>0$ and hence

$\Gamma(n+\frac{1}{d-1})>0$. Also, for each $j \geq 1$ let $\alpha(j)=\frac{(d-1)\mathbb{E}(I_j)+d\,j}{(d-1)j+1}, \quad j\geq 1$.

THEOREM 2.2. *Let $n \geq 1$. The mean of the Sackin index S_n of a random d-ary increasing tree T_n of order n is given by*

$$\mathbb{E}(S_n)=c[n]\sum_{j=1}^{n-1}\frac{\alpha(j)}{c[j+1]}.$$

Proof. Since the result holds for the case $n\in\{1,2\}$, we assume that $n \geq 3$. For each $v \in V(T_n)$ define the indicator $\tilde{I}(v)$ as below:

$$\tilde{I}(v)=\begin{cases}0, & d^+(v)=0\\ D(v), & d^+(v)\geq 1.\end{cases}$$

Note that $\tilde{I}(root)=D(root)=0$. Assume that v_n is attached to a randomly chosen vertex U_{n-1} in T_{n-1}. Hence $S_n=S_{n-1}+\tilde{I}(U_{n-1})+1$. Thus

$$\begin{aligned}\mathbb{E}(S_n|\mathcal{F}_{n-1})&=S_{n-1}+\mathbb{E}(\tilde{I}(U_{n-1})|\mathcal{F}_{n-1})+1\\
&=S_{n-1}+\sum_{v\in V(T_{n-1})}\frac{d-d_v^+}{(d-1)(n-1)+1}\tilde{I}(v)+1\\
&=S_{n-1}+\frac{1}{(d-1)(n-1)+1}\sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}}(d-d_v^+)D(v)+1\\
&=S_{n-1}+\frac{1}{(d-1)(n-1)+1}\left[d\sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}}D(v)-\sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}}d_v^+\,D(v)\right]+1\\
&=S_{n-1}+\frac{1}{(d-1)(n-1)+1}\big[d(I_{n-1}-S_{n-1})-(I_{n-1}-(n-2))\big]+1\\
&=\frac{(d-1)(n-2)}{(d-1)(n-1)+1}\;S_{n-1}+\frac{d-1}{(d-1)(n-1)+1}\;I_{n-1}+\frac{d(n-1)}{(d-1)(n-1)+1}\;.\end{aligned}$$

Since $\frac{(d-1)(n-2)}{(d-1)(n-1)+1}=\frac{n-2}{(n-1)+\frac{1}{d-1}}=\frac{c[n]}{c[n-1]}$, we have

$$\mathbb{E}(S_n)=\frac{c[n]}{c[n-1]}\mathbb{E}(S_{n-1})+\alpha(n-1), \quad n\geq 3. \tag{1}$$

Since $\mathbb{E}(S_2)=1$, the recurrence (1) implies that $\mathbb{E}(S_n)=c[n]\sum_{j=2}^{n}\frac{\alpha(j-1)}{c[j]}$, and this completes the proof. □

LEMMA 2.3. *Let $n\geq 2$ and let $\mathbb{C}ov(S_n,S_{n-1})$ be the covariance between two random variables S_n and S_{n-1}. Then*

$$\mathbb{C}ov(S_n,S_{n-1})=\frac{(d-1)(n-2)}{(d-1)(n-1)+1}\mathbb{V}ar(S_{n-1})+\frac{d-1}{(d-1)(n-1)+1}\mathbb{C}ov(S_{n-1},I_{n-1}).$$

Proof. For $n=2$ the result follows directly because $\mathbb{C}ov(S_1,I_1)=0$. For $n\geq 3$ from equation (1) we see that

$$\begin{aligned}\mathbb{E}(S_n - \mathbb{E}(S_n)|\mathcal{F}_{n-1}) &= \mathbb{E}(S_n|\mathcal{F}_{n-1}) - \mathbb{E}(S_n)\\ &= \frac{c[n]}{c[n-1]}(S_{n-1} - \mathbb{E}(S_{n-1})) + \frac{(d-1)(I_{n-1} - \mathbb{E}(I_{n-1}))}{(d-1)(n-1)+1}.\end{aligned}$$

Then

$$\begin{aligned}\mathbb{C}ov(S_n, S_{n-1}) &= \mathbb{E}(\mathbb{E}((S_n - \mathbb{E}(S_n))(S_{n-1} - \mathbb{E}(S_{n-1}))|\mathcal{F}_{n-1}))\\ &= \mathbb{E}((S_{n-1} - \mathbb{E}(S_{n-1}))\mathbb{E}(S_n - \mathbb{E}(S_n)|\mathcal{F}_{n-1}))\\ &= \frac{c[n]}{c[n-1]}\mathbb{V}ar(S_{n-1}) + \frac{d-1}{(d-1)(n-1)+1}\mathbb{C}ov(S_{n-1}, I_{n-1}).\end{aligned}$$

□

For each $i \geq 1$, set $a_i = \mathbb{E}(S_i)$, $b_i = \mathbb{E}(I_i)$, $c_i = \frac{\mu_i - \lambda_i}{1-\lambda_i}\mathbb{E}(\tilde{I}(U_i)) - \frac{\lambda_i}{1-\lambda_i}\mathbb{E}(\tilde{I}(U_i)D(U_i))$ and $e_i = \mu_i a_i + b_{i+1} + c_i$.

Theorem 2.4. *Let I_n and S_n be the total path length and Sackin index of a random d-ary increasing tree of order n, respectively. Then*

$$\mathbb{C}ov\,(S_n, I_n) = \begin{cases}\sum\limits_{j=1}^{n-1}\left(e_j \prod\limits_{i=j+1}^{n-1} \lambda_i\right) - \mathbb{E}(S_n)\mathbb{E}(I_n), & n \geq 4\\ 0, & n = 1, 2, 3.\end{cases}$$

Proof. For $n = 1, 2, 3$, $\mathbb{C}ov\,(S_n, I_n) = 0$. Assume $n \geq 4$. By definition,

$$\mathbb{C}ov\,(S_n, I_n) = \mathbb{E}(I_n S_n) - a_n b_n.$$

Since $S_n = S_{n-1} + \tilde{I}(U_{n-1}) + 1$, we have

$$\mathbb{E}\,(S_n I_n) = \mathbb{E}\Big(\big(S_{n-1} + \tilde{I}(U_{n-1}) + 1\big) I_n\Big) = \mathbb{E}\big(S_{n-1} I_n\big) + \mathbb{E}\,\Big(\tilde{I}\big(U_{n-1}\big) I_n\Big) + b_n.$$

We have

$$\begin{aligned}\mathbb{E}\big(\tilde{I}\big(U_{n-1}\big)\big) &= \frac{1}{(n-1)(d-1)+1}\mathbb{E}\Bigg(\sum_{v\in V(T_{n-1})} \tilde{I}(v)(d - d_v^+)\Bigg)\\ &= \frac{1}{(n-1)(d-1)+1}\mathbb{E}\Bigg(d(I_{n-1} - S_{n-1}) - \sum_{v\in V(T_{n-1})} d_v^+ D_v\Bigg)\\ &= \frac{1}{(n-1)(d-1)+1}\big((d-1)\mathbb{E}(I_{n-1}) - \mathbb{E}(S_{n-1}) + n - 2\big).\end{aligned}$$

It is not difficult to show that

$$\sum_{v\in V(T_n)} D_v^2 = 2I_n - (n-1) + \sum_{\substack{v\in V(T_n),\\ d^+(v)\geq 1}} d_v^+ D_v^2.$$

Now, we have

$$\mathbb{E}\,\Big(\tilde{I}\big(U_{n-1}\big) D\big(U_{n-1}\big)\Big) = \frac{1}{(n-1)(d-1)+1}\mathbb{E}\Bigg(\sum_{v\in V(T_{n-1})} (d - d_v^+)\tilde{I}(v) D_v\Bigg)$$

$$= \frac{(d-1) \sum_{v\in V(T_{n-1})} \mathbb{E}(D_v^2) + 2\mathbb{E}(I_{n-1}) - n + 2}{(n-1)(d-1)+1}.$$

But
$$\begin{aligned}\mathbb{E}\left(\tilde{I}(U_{n-1})I_n\right) &= \mathbb{E}\left(\tilde{I}(U_{n-1})\ \mathbb{E}(I_n|\mathcal{F}_{n-1})\right)\\ &= \mathbb{E}\left(\tilde{I}(U_{n-1})\ (\lambda_{n-1}I_{n-1}+\mu_{n-1})\right)\\ &= \lambda_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})(I_n - D(U_{n-1})-1)\right) + \mu_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})\right)\\ &= \lambda_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})I_n\right) - \lambda_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})D(U_{n-1})\right)\\ &\quad -\lambda_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})\right) + \mu_{n-1}\mathbb{E}\left(\tilde{I}(U_{n-1})\right).\end{aligned}$$

Hence,

$$\mathbb{E}\left(\tilde{I}(U_{n-1})I_n\right) = \frac{\mu_{n-1}-\lambda_{n-1}}{1-\lambda_{n-1}}\mathbb{E}\left(\tilde{I}(U_{n-1})\right) - \frac{\lambda_{n-1}}{1-\lambda_{n-1}}\mathbb{E}\left(\tilde{I}(U_{n-1})D(U_{n-1})\right) := c_{n-1},$$

where the two terms on the right of the above equality were previously calculated. Now, from Theorem 2.1,

$$\mathbb{E}\left(S_{n-1}I_n\right) = \mathbb{E}\big(\mathbb{E}\left(S_{n-1}I_n|\mathcal{F}_{n-1}\right)\big) = \mathbb{E}\big(S_{n-1}\ \mathbb{E}\left(I_n|\mathcal{F}_{n-1}\right)\big) = \lambda_{n-1}\mathbb{E}\left(S_{n-1}I_{n-1}\right) + \mu_{n-1}a_{n-1}.$$

Hence

$$\mathbb{E}\left(S_nI_n\right) = \lambda_{n-1}\mathbb{E}\left(S_{n-1}I_{n-1}\right) + \mu_{n-1}a_{n-1} + c_{n-1} + b_n = \lambda_{n-1}\mathbb{E}\left(S_{n-1}I_{n-1}\right) + e_{n-1}.$$

Since $S_1I_1 = 0$, iterating this recurrence completes the proof. □

For each $i \geq 1$, set

$$\begin{aligned}\beta(i) =& 2\frac{d-1}{i(d-1)+1}\mathbb{C}ov(I_i, S_i) - \left(\frac{-dS_i}{(d-1)i+1} + \alpha(i) - 1\right)^2\\ &+ \frac{1}{(d-1)i+1}\left(d\sum_{v\in V(T_i), d^+(v)\geq 1}\mathbb{E}(D_v^2) - \sum_{v\in V(T_i)}\mathbb{E}(D_v^2) + 2\mathbb{E}(I_i) - (i-1)\right),\end{aligned}$$

and for each $n \geq 3$ set $t[n] := \frac{\Gamma\left(n-2-\frac{1}{d-1}\right)}{\Gamma\left(n+\frac{1}{d-1}\right)}$.

Theorem 2.5. *Let $n \geq 1$. The variance of the Sackin index S_n of a random d-ary increasing tree T_n of order n is given by* $\mathbb{V}ar(S_n) = t[n]\sum_{j=3}^{n-1}\frac{\beta(j)}{t[j+1]}$.

Proof. Since the empty sum is zero, the result holds for $n \in \{1, 2, 3\}$ and we can assume that $n \geq 4$. We have

$$\mathbb{E}(S_n - S_{n-1} - 1)^2 = \mathbb{E}\left(\sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} \frac{d-d^+(v)}{(d-1)(n-1)+1}D^2(v)\right)$$

$$= \frac{1}{(d-1)(n-1)+1}\mathbb{E}\Bigg(d \sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} D^2(v) - \sum_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} d^+(v)D^2(v)\Bigg)$$

$$= \frac{\mathbb{E}\Bigg(d \sum\limits_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} D_v^2 - \sum\limits_{v\in V(T_{n-1})} D_v^2 + 2I_{n-1} - (n-2)\Bigg)}{(d-1)(n-1)+1}$$

$$= \frac{d \sum\limits_{\substack{v\in V(T_{n-1}),\\ d^+(v)\geq 1}} \mathbb{E}(D_v^2) - \sum\limits_{v\in V(T_{n-1})} \mathbb{E}(D_v^2) + 2\mathbb{E}(I_{n-1}) - (n-2)}{(d-1)(n-1)+1}. \tag{2}$$

From Lemma 2.3 it follows

$$\begin{aligned}
&\mathbb{E}(S_n - S_{n-1} - 1)^2 = \mathbb{E}(S_n - \mathbb{E}(S_n) - S_{n-1} + \mathbb{E}(S_{n-1}) + \mathbb{E}(S_n) - \mathbb{E}(S_{n-1}) - 1)^2 \\
=&\mathbb{E}(S_n - \mathbb{E}(S_n) - S_{n-1} + \mathbb{E}(S_{n-1}))^2 + \mathbb{E}(\mathbb{E}(S_n) - \mathbb{E}(S_{n-1}) - 1)^2 \\
&+ 2\mathbb{E}\Bigg((S_n - \mathbb{E}(S_n) - S_{n-1} + \mathbb{E}(S_{n-1}))(\mathbb{E}(S_n) - \mathbb{E}(S_{n-1}) - 1)\Bigg) \\
=&\mathbb{V}ar(S_n) + \Big(1 - 2\frac{c[n]}{c[n-1]}\Big)\mathbb{V}ar(S_{n-1}) - 2\mathbb{C}ov(S_{n-1}, I_{n-1})\frac{d-1}{(d-1)(n-1)+1} \\
&+ \Big(\frac{-d\ \mathbb{E}(S_{n-1})}{(d-1)(n-1)+1} + \alpha(n-1) - 1\Big)^2 + 0.
\end{aligned} \tag{3}$$

Now, from (2) and (3) we see that

$$\mathbb{V}ar(S_n) = \Big(2\frac{c[n]}{c[n-1]} - 1\Big)\mathbb{V}ar(S_{n-1}) + \beta(n-1) = \frac{t[n]}{t[n-1]}\mathbb{V}ar(S_{n-1}) + \beta(n-1).$$

By iteration, proof is completed because $\mathbb{V}ar(S_3) = 0$. □

3. Conclusion

In this paper, we studied the first two moments of the Sackin index in random d-ary increasing trees. Doing longer calculations with the same approach we can obtain the higher moments of this quantity. For example,

$$\mathbb{E}(S_n^3) = c[n]\sum_{j=1}^{n-1}\frac{\xi(j)}{c[j+1]}, \quad \mathbb{E}(S_n^4) = c[n]\sum_{j=1}^{n-1}\frac{\eta(j)}{c[j+1]},$$

where
$$\xi(j) = \frac{(d-1)\mathbb{E}(I_j^3) + 2d\mathbb{E}(I_j^2) + (1+3d)\mathbb{E}(I_j) + dj}{(d-1)j+1}$$

and
$$\eta(j) = \frac{(d-1)\mathbb{E}(I_j^4) + 4d\mathbb{E}(I_j^3) + d\mathbb{E}(I_j^2) + (2+6d)\mathbb{E}(I_j) + dj}{(d-1)j+1}.$$

Using the above results, the measures of skewness and kurtosis of the Sackin index can be obtained as

$$Skewness(S_n) = \frac{\mathbb{E}(S_n^3) - 3\mathbb{E}(S_n^2)\mathbb{E}(S_n) + 2\mathbb{E}^3(S_n)}{\mathbb{V}ar(S_n)^{3/2}},$$

$$Kurtosis(S_n) = \frac{\mathbb{E}(S_n^4) - 4\mathbb{E}(S_n^3)\mathbb{E}(S_n) + 6\mathbb{E}(S_n^3)\mathbb{E}^2(S_n) - 3\mathbb{E}^4(S_n)}{\mathbb{V}ar(S_n)},$$

respectively (see also Naeini et al. [8] for the same calculations for degree-based indices).

Acknowledgement. The authors would like to express their sincere gratitude to the referees for their valuable suggestions and comments.

References

[1] A. R. Ashrafi, T. Dehghan-Zadeh, N. Habibi, *Extremal atom-bond connectivity index of cactus graphs*, Commun. Korean Math. Soc., **30(3)** (2015), 283–295.

[2] F. Bergeron, P. Flajolet, S. Salvy, *Varieties of increasing trees*, Lecture Notes in Computer Science., **581** (1992), 24–48.

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

[4] M. Javanian, *On the external path length of random increasing k-ary trees*, Italian J. Pure. Appl. Math., **31** (2013), 21–26.

[5] R. Kazemi, *The second Zagreb index of molecular graphs with tree structure*, MATCH Commun. Math. Comput. Chem., **72** (2014), 753–760.

[6] D. Knuth, *The Art of Computer Programming, Vol 3: Sorting and Searching*, 2nd ed. Addison-Wesley, Reading, Mass, 1973.

[7] M. Moradian, R. Kazemi, M. H. Behzadi, *The Sackin index of random recursive trees*, U.P.B. Sci. Bull., Series A, **79(2)** (2017), 125–130.

[8] H. Naeini, R. Kazemi, M. H. Behzadi, *The modified forgotten topological index of random b-ary recursive trees*, U.P.B. Sci. Bull., Series A, **80(4)** (2018), 125–132.

[9] A. Panholzer, H. Prodinger, *The level of nodes in increasing trees revisited*, Random Structure Algorithms. **31(2)** (2007), 203–226.

(received 08.06.2019; in revised form 10.09.2019; available online 03.06.2020)

Department of Statistics, Faculty of Science, Imam Khomeini International University, Qazvin, Iran

E-mail: r.kazemi@sci.ikiu.ac.ir

Department of Pure Mathematics, Faculty of Science, Imam Khomeini International University, Qazvin, Iran

E-mail: a.behtoei@sci.ikiu.ac.ir