J. Math. Anal. Appl. 434 (2016) 1–11

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

The asymptotic number of non-isomorphic rooted trees obtained by rooting a tree ☆

Xueliang Li *, Yiyang Li, Yongtang Shi

Center for Combinatorics and LPMC-TJKLC, Nankai University, Tianjin 300071, China

ARTICLE INFO

Article history:
Received 23 August 2013
Available online 5 September 2015
Submitted by V. Pozdnyakov

Keywords:
Tree
Rooted tree
Pattern
Counting, generating function
Limiting distribution
Automorphism

ABSTRACT

Let $\mathcal{T}_n$ denote the set of trees with n vertices. Suppose that each tree in $\mathcal{T}_n$ is equally likely. We show that the number of non-isomorphic rooted trees obtained by rooting a tree equals $(\mu_r + o(1))n$ for almost every tree of $\mathcal{T}_n$, where μ_r is a constant. As an application, we show that in $\mathcal{T}_n$ the number of any given pattern, which is a fixed small tree with internal vertices specified, is asymptotically normally distributed with mean $\sim \mu_M n$ and variance $\sim \sigma_M n$, where μ_M and σ_M are some constants related to the given pattern. This solves an open question claimed in Kok's thesis.

1. Introduction

A *pattern* M is a prescribed tree. We say that M *occurs* in a tree T if M is a subtree of T in the sense that the degree of each internal vertex (of degree more than one) of M matches the degree of the corresponding vertex in T, while each external vertex (of degree one) of M matches a vertex of T with an arbitrary degree. Let $\mathcal{T}_n$ denote the set of trees with n vertices. If we use $X_{n,M}(T)$ to denote the number of occurrences of a given pattern M in $T \in \mathcal{T}_n$, then $X_{n,M}(T)$ is a random variable with probability

$$P(X_{n,M} = k) = \frac{t_{n,k}}{t_n},$$

where $t_{n,k}$ denotes the number of those trees in $\mathcal{T}_n$ that the number of occurrences of the pattern M in each of the trees is k, and $t_n = |\mathcal{T}_n|$.

Moreover, let $\mathcal{R}_n$ denote the set of rooted trees. We can also consider the number of occurrences of a given pattern in $\mathcal{R}_n$. Denote the corresponding random variable by $X_{n,M}(R)$.

☆ Supported by NSFC No. 11371205 and 11531011, the "973" program No. 2013CB834204, and PCSIRT.

* Corresponding author.
E-mail addresses: lxl@nankai.edu.cn (X. Li), liycldk@mail.nankai.edu.cn (Y. Li), shi@nankai.edu.cn (Y. Shi).

http://dx.doi.org/10.1016/j.jmaa.2015.09.007
0022-247X/

The main work of this paper is to show that some random variable Y_n in $\mathcal{T}_n$ (or $\mathcal{R}_n$) satisfies

$$\frac{Y_n - \mathbf{E}(Y_n)}{\sqrt{\mathbf{Var}(Y_n)}} \to_w \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ is the random variable with standard normal distribution and $\to_w$ means weak convergence. We then call this random variable Y_n *asymptotically normal*. Moreover, if

$$\frac{Y_n - \mu n}{\sqrt{\sigma n}} \to_w \mathcal{N}(0,1),$$

then Y_n is asymptotically normal with mean $\sim \mu n$ and variance $\sim \sigma n$. We refer to [10] for details.

In fact, it was shown in [3] that in $\mathcal{R}_n$ the number $X_{n,M}(R)$ of occurrences of any given pattern is asymptotically normal with mean $\sim \mu_M n$ and variance $\sim \sigma_M n$, where μ_M and σ_M are some constants corresponding to the given pattern. But, for the set $\mathcal{T}_n$ there is no such a result on normal distribution. In [9], the authors proved that for any given pattern in $\mathcal{T}_n$ the limiting distribution has a density $(a + bt^2)e^{ct^2}$, where a, b, c are some constants. The mean and variance of the number of occurrences of any given pattern are still asymptotically $\mu_M n$ and $\sigma_M n$ where the constants are the same as in $\mathcal{R}_n$. Clearly, if one shows that $b = 0$, then the distribution is normal. For some special patterns, such as a star pattern (or a node with a given degree) [5], a double-star pattern [11], and a path pattern [10], the corresponding limiting distributions were proved to be normal. For some previous work we refer to Robinson and Schwenk [16]. For more details, we refer to [3,9,10,16]. Moreover, Gittenberger [7], Panagiotou and Sinha [14] considered the case for the growing star pattern (the number of vertices of the star tending to infinity with n), which yields a non-normal limiting distribution. However, Kok claimed in his thesis [10] that for any given pattern it seems much more difficult to demonstrate the normality. In this paper, we will solve this problem from a new point of view which is different from the existing ones. We study the number of non-isomorphic rooted trees obtained by rooting a tree and get that for almost every tree of order n the number of corresponding non-isomorphic rooted trees is $(\mu_r + o(1))n$ (the authors of [2] and [12] computed this number for other sorts of trees). Based on this result and the normal limiting property on the number of occurrences of any given pattern in $\mathcal{R}_n$, we will show that in $\mathcal{T}_n$ the limiting distribution is also normal.

We organize this paper as follows. In Section 2, we will introduce some basic knowledge that will be used in our proofs. In Section 3, we will present the detailed proofs. Our focus is on studying the number of non-isomorphic rooted trees obtained by rooting a tree. Section 4 is devoted to study the limiting distribution for any given pattern.

2. Preliminaries

Let T_n be a tree in $\mathcal{T}_n$. We say that two vertices u and v of T_n are in the same vertex class if u can be mapped to v by an automorphism of T_n. Clearly, this establishes an equivalent relation on the vertex set of T_n, and hence the vertices in T_n are partitioned into some classes. If u and v are in the same vertex class of T_n, then rooting T_n at u and v, respectively, yields the same rooted tree.

Hence, the number of non-isomorphic rooted trees obtained by rooting a tree is exactly the number of vertex classes of the tree. Let $X_n(T)$ represent the number of vertex classes of the tree T. Clearly, $X_n(T)$ is also a random variable on $\mathcal{T}_n$. Therefore, we can similarly introduce the random variable $X_n(R)$ of vertex classes in the space of rooted trees $\mathcal{R}_n$.

Notice that $X_n(T) \geq 1$. In analogy to the enumeration of patterns in [9], we introduce the following two functions:

$$t(x) = \sum_{n \geq 1} t_n x^n,$$