POISSON LIMIT THEOREM FOR A NUMBER OF GIVEN VALUE TREES IN GALTON-WATSON FOREST WITH BOUNDED NUMBER OF VERTEXES

Chuprunov A.N.

Chuvash State University, email: achuprunov@mail.ru

Abstract

We consider a random forest with N root vertexes and not more then n non-root vertexes defined by trajectories Galton-Watson process with Poisson distribution of number of direct descendants which has N particles in beginner. That is a subset of trajectories for which the number of non-root vertexes is not more n. We prove Poisson limit theorem for the number of trees from the first K trees which contains r non-root vertexes. The limit Poisson random variable is described.

Ключевые слова: Galton-Vatson forest, Poisson random variable, gaussian randon variable, limit theorem.

Mathematical Subject Classification: 60C05, 60F05

1 INTRODUCTION AND MAIN RESULT

Let $\xi_1, \xi_2, \dots$ be independent identically distributed non-negative integer valued random variables. We say that the random variables $\eta_1, \dots, \eta_N$ satisfy the generalized allocation scheme of not more n particles by N cells, if there joint distribution is of the form

$$\mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \mathbf{P}\left\{\xi_1 = k_1, \dots, \xi_N = k_N \;\middle|\; \sum_{i=1}^{N} \xi_i \le n\right\},$$

for all non-negative integer numbers $k_1, \dots, k_N$ such that $k_1 + k_2, \dots + k_N \le n$.

The generalized allocation scheme of not more n particles by N cells was introduced in [1]. In [1, 2] it obtained limit theorems which connected with the generalized allocation scheme of not more n particles by N cells.

In [3, 4] it considered the set $F_{N,n}$ of forests with the pointed vertexes which contain N root vertexes and n non-root vertexes. On this set it considered uniform distribution of probabilities. For such random forests it obtained various limit theorems which connected with values of trees and proved for different method of $N, n \to \infty$. In [3] the same problems was solved for Galton-Watson forests. That is random forests which generated by subcretecal or cretical Galton-Watson process which has N particles in the beginner. This random process desintegrates on N independent random processes which begin with one particle. The set of all trajectories of such process is infinity. In [5] it considered a subset of this set in which particles exist during the time of the evolution $N + n$. In such subset the number of realizations of the process is finite and the distribution of probabilities defined by the natural method.

Consider Galton-Watson process which begin from N particles and a number of right descendants each particle has Poisson distribution with the parameter λ, $0 < \lambda \le 1$.

Observe that the subset of trajectories of such process which contains $n+N$ edges coincieds with $F_{N,n}$. Denote by $\xi_1, \xi_2, \dots \xi_N$ the numbers of particles which exist during the time of the evolution in subprocesses which begin from the particle $1, \dots N$, correspondingly. Then (see [6]) $\xi_1, \xi_2, \dots \xi_N$ are independent random variables with the distribution

$$p_k(\lambda) = \mathbf{P}\{\xi_i = k\} = \frac{(\lambda k)^{k-1}}{k!} e^{-\lambda k}, \quad k = 1, 2 \dots, \quad 0 < \lambda \le 1,$$

We will consider a subset of trajectories of the process such that $\xi_1 + \xi_2 + \dots + \xi_N \le n$. Denote by $\eta_1, \dots, \eta_N$ the random variables which are values of trees in the forest from this subset. Then $(\eta_1, \dots, \eta_N)$ has the distribution

$$\mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \mathbf{P}\left\{\xi_1 = k_1, \dots, \xi_N = k_N \;\middle|\; \sum_{i=1}^{N} \xi_i \le n\right\}.$$

So $\eta_1, \dots, \eta_N$ is the generalizeg allocation scheme of allocation of not more n particles by N cells.

We will study the convergence in distribution of the random variables

$$\mu_r(n, K, N) = \sum_{i=1}^{K} I_{\{\eta_i = r\}}, \quad \text{where} \quad 0 < K \le N, \quad r = 1, 2 \dots .$$

to Poisson random variable. Observe that $\mu_r(n, K, N)$ is a number of trees from the first K trees which contain r non-root vertexes.

The main result of the paper is the following theorem.

Theorem 1. *Let r be a fixed number. Suppose $K, n, N \to \infty$ such that*

$$K p_r(\lambda) \to \alpha,$$

where $0 \le \alpha < \infty$ and one of the following conditions is valid:

$$(A) \qquad \lambda^3(1-\lambda)N \to \infty \quad \text{and} \quad n(1-\lambda)^{3/2} - N\sqrt{1-\lambda} \ge C\sqrt{N\lambda} \quad \text{for} \quad \text{some} \quad C < 0;$$

$$(B) \qquad (1-\lambda)N \to \nu \quad \text{and} \quad n \ge CN^2 \quad \text{for} \quad \text{some} \quad C > 0, \;\; 0 < \nu < \infty.$$

Then we have

$$\mu_r(n, K, N) \xrightarrow{d} \pi(\alpha).$$

The proof of Theorem 1 founded on Poisson limit theorem for exchangeable random variables. Recall that the random variables $\eta'_1, \eta'_2, \dots \eta'_K$ are called exchangeable if the distribution of $(\eta'_1, \eta'_2, \dots \eta'_N)$ coincides with the distribution of $(\eta'_{i_1}, \eta'_{i_2}, \dots \eta'_{i_K})$ fo any permutation $(i_1, i_2, \dots i_N)$ of $(1, 2, \dots K)$.

The following known elementary limit theorem will play fundamental role in our paper (see Theorem II in [7]; we mention the Benczúr presented as lightly more general result without proof, see Theorem 1 in [8]).

Theorem A. *Let the array of random variables η'_{Ki}, $1 \le i \le K$, $K = 1, 2, \dots$, be row-wise exchangeable. Let $A_i = A_{Kri} = \{\omega \in \Omega : \eta'_{Ki}(\omega) = r\}$, where r is a fixed and let $S_K = \sum_{i=1}^{K} I_{A_i}$. Suppose that the following condition is valid.*

There exists β ($\le \beta < \infty$) such that for any $k = 1, 2, \dots$

$$K^k \mathbf{P}(A_{K1} \cap A_{K2} \cap \cdots \cap A_{Kk}) \to \beta^k, \quad \text{as} \quad K \to \infty. \tag{1.1}$$

Then

$$S_K \xrightarrow{d} \pi(\beta), \quad \text{as} \quad K \to \infty.$$

Distribution of $\mu_r(n, K, N)$ coincieds with the distribution of the random variable $\mu_r(n, A, N) = \sum_{i \in A} I_{\{\eta_i = r\}}$, where A is a pointed subset of the set $\{1, \dots N\}$ such that $|A| = K$. So Theorem 1 we can consider as a theorem for number of trees from a pointed set.

Limit theorems for a number of empty cells in a pointed set of cells in the schem of allocation of distingushing particals by different cells obtained in [9]. In [10] it obtained limit theorems for a maximal number of a tree in Galton-Watson forest with bounded number of vertexes.

We will denote: γ is a gaussian random variable with the expectation 0 and the variance 1, Φ is a distribution function of γ, $\pi(\alpha)$ is a poissonian random variable with the parameter α, $\overset{d}{=}$ is the equality by dstribution, $\xrightarrow{d}$ is the convergence by distribution.

2 Proof of Theorem 1

In order to check (1.1) we will use the following lemma.

Lemma 1. *Let $\eta_1, \dots, \eta_N$ be a generalized allocation scheme of not more n particles by N cells. Then Тогда $\eta_1, \dots, \eta_N$ are row-wise exchangeable random variables and we have*

$$\mathbf{P}(A_1 \cap A_2 \cap \cdots \cap A_k) = (p_r(\lambda))^k \frac{\mathbf{P}\{\zeta_{N-k} \le n - kr\}}{\mathbf{P}\{\zeta_N \le n\}}, \tag{2.1}$$

where $A_i = A_{ri} = \{\omega \in \Omega : \eta_i(\omega) = r\}$, $\zeta_l = \xi_1 + \xi_2 + \cdots + \xi_l$, $l \in \{N, N-k\}$.

The proof of Lemma 1 is the same as the proof of Lemma 1.2.1 from [11]. In order to estimate the numerator and the denominator in (2.1) we will use the following lemmas which obtained in [10] (see Lemma 6 and Lemma 9).

Lemma 2. *Let (A) be valid. Then we have*

$$\frac{\zeta_N - \frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \xrightarrow{d} \gamma, \quad \text{as} \quad N \to \infty.$$

Lemma 3. *Let (B) be valid. Then we have*

$$\frac{\zeta_N}{N^2} \xrightarrow{d} \delta, \quad \text{as} \quad N \to \infty,$$

where δ is a random variable with a distribution function difined by the density

$$g(x) = \frac{1}{\sqrt{2\pi x^3}} \exp\left\{\nu - \frac{\nu^2 x}{2} - \frac{1}{2x}\right\}, \quad x > 0, \ \ g(x) = 0, \ \ x \le 0.$$

Proof of Theorem 1. Let (A) be valid. Using (2.1) we have

$$K^k \mathbf{P}(A_1 \cap A_2 \cap \dots \cap A_k) = (Kp_r(\lambda))^k \frac{\mathbf{P}\left\{\frac{\zeta_{N-k} - \frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}} \le \frac{n-kr-\frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}}\right\}}{\mathbf{P}\left\{\frac{\zeta_N - \frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\}} =$$

$$= (Kp_r(\lambda))^k \frac{\mathbf{P}\left\{\sqrt{\frac{N-k}{N}}\frac{\zeta_{N-k} - \frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} - \frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} + \frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}}\right\}}{\mathbf{P}\left\{\frac{\zeta_N - \frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\}}. \tag{2.2}$$

Let $0 < \varepsilon < 1/2$. Choose $C_1 > 0$ such that $\Phi(C_1) > 1 - \varepsilon$. Since

$$\frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} - \frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}} \to 0,$$

as $C \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \le C_1$, by Lemma 1 we have

$$\frac{\mathbf{P}\left\{\sqrt{\frac{N-k}{N}}\frac{\zeta_{N-k} - \frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} - \frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} + \frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}}\right\}}{\mathbf{P}\left\{\frac{\zeta_N - \frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\}} =$$

$$= \frac{\Phi\left\{\frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} - \frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} + \frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}}\right\} + o(1)}{\Phi\left\{\frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\} + o(1)} = 1 + o(1).$$

Let $C_1 < \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}$. By Lemma 1 we have

$$1 - \varepsilon + o(1) < \frac{\mathbf{P}\left\{\sqrt{\frac{N-k}{N}}\frac{\zeta_{N-k} - \frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} - \frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} + \frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}}\right\}}{\mathbf{P}\left\{\frac{\zeta_N - \frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}} \le \frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\}} < \frac{1}{1 - \varepsilon + o(1)}.$$

Therefore we obtain

$$\frac{\mathbf{P}\left\{\sqrt{\frac{N-k}{N}}\frac{\zeta_{N-k}-\frac{N-k}{1-\lambda}}{\sqrt{\frac{(N-k)\lambda}{(1-\lambda)^3}}}\le\frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}-\frac{kr}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}+\frac{k}{\sqrt{\frac{N\lambda}{1-\lambda}}}\right\}}{\mathbf{P}\left\{\frac{\zeta_N-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\le\frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}\right\}}=1+o(1), \tag{2.3}$$

as $C<\frac{n-\frac{N}{1-\lambda}}{\sqrt{\frac{N\lambda}{(1-\lambda)^3}}}$. Using (2.3) in (2.2) we obtain (1.1). So we can applicate Theorem A.

Let (B) be valid. Using (2.1) we have

$$K^k\mathbf{P}(A_1\cap A_2\cap\cdots\cap A_k)=(Kp_r(\lambda))^k\frac{\mathbf{P}\left\{\frac{\zeta_{N-k}}{(N-k)^2}\le\frac{n-kr}{(N-k)^2}\right\}}{\mathbf{P}\left\{\frac{\zeta_N}{N^2}\le\frac{n}{N^2}\right\}}. \tag{2.4}$$

Let $0<\varepsilon<1/2$. Choose $C_2>0$ such that $\mathbf{P}\{\delta>C_1\}<1-\varepsilon$. Since

$$\frac{kr}{N^2}\to 0,\quad \frac{k}{N}\to 0,$$

as $C\le\frac{n}{N^2}\le C_2$, by Lemma 3 we have

$$\frac{\mathbf{P}\left\{\frac{\zeta_{N-k}}{(N-k)^2}\le\frac{n-kr}{(N-k)^2}\right\}}{\mathbf{P}\left\{\frac{\zeta_N}{N^2}\le\frac{n}{N^2}\right\}}=\frac{\mathbf{P}\left\{\delta\le\frac{n-kr}{(N-k)^2}\right\}+o(1)}{\mathbf{P}\left\{\delta\le\frac{n}{N^2}\right\}+o(1)}=1+o(1).$$

For $C_2<\frac{n}{N^2}$ by Lemma 3 we have

$$1-\varepsilon+o(1)<\frac{\mathbf{P}\left\{\frac{\zeta_{N-k}}{(N-k)^2}\le\frac{n-kr}{(N-k)^2}\right\}}{\mathbf{P}\left\{\frac{\zeta_N}{N^2}\le\frac{n}{N^2}\right\}}<\frac{1}{1-\varepsilon+o(1)}.$$

Therefore we obtain

$$\frac{\mathbf{P}\left\{\frac{\zeta_{N-k}}{(N-k)^2}\le\frac{n-kr}{(N-k)^2}\right\}}{\mathbf{P}\left\{\frac{\zeta_N}{N^2}\le\frac{n}{N^2}\right\}}=1+o(1), \tag{2.5}$$

as $C\le\frac{n}{N^2}$. Using (2.5) in (2.4), we obtain (1.1). So we can applicate Theorem A. □

References

[1] Chuprunov A.N., Fazekas I. *An analogue of the generalized allocation scheme: limit theorems for the number of cells containing a given number of particles* // Discrete Math. Apple. – 2012. – V. 22. – No. 1. – P. 101-122.

[2] Chuprunov A.N., Fazekas I. *An analogue of the generalized allocation scheme: limit theorems for the maximal cell load.*// Discrete Math. Apple. – 2012. – V. 22. – No. 3. – P. 307-314.

[3] Pavlov Yu. L. *The assimptotic distribution of a maximal tree in a random forect.* // Theory Probab. Apple. – 1977. – V. 22. – No. 3. – P. 509-524.

[4] Pavlov Yu. L. *Limit theorems for the number of trees of a given size in a random forest.* // Math. USSR-Sb. – 1977. – V. 32. – No. 3. – P. 335-345.

[5] Pavlov Yu. L. *Random forests.* // VSP – 2000.

[6] Sevast'yanov B. A. *Branching processes.* // Nauka. Moskov. –1971.(In Russian.)

[7] Kendall D. G. *On finite and infinite seqquences of exchangeable events.* // Studia Sci. Math. Hanggar. – 1967 – V. 2.– P. 319-327.

[8] Benczúr A. *On sequences of equivalent events and the comcompound Poisson process.* // Studia Sci. Math. Hanggar. – 1968 – V. 3.– P. 451-458.

[9] Khakimullin E.R., Enatskaya N.Yu. *Limit theorems for the number of empty cells.* // Discrete Math. Apple. – 1997. – V. 7. – No. 2.– P. 209-219.

[10] Pavlov Yu. L., Hvorostianskaya E. V. *On the maximal size of a tree in the Galton-Watson forest with a bounded number of vertices.* // Discrete Math. Apple. – 2014. – V. 24. – No. 6. – P. 101-122.

[11] Kolchin, V. F. *Random Graphs.* Cambridge University Press. Cambridge. 1999.