PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

ASYMPTOTIC ESTIMATES FOR THE INVERSE EPIDEMIC PROCESS ON A RANDOM GRAPH

LJUBEN R. MUTAFČIEV

An epidemic model on a random graph with n vertices, suggested by Gertsbakh, is considered. The infection is delivered from the initially infected m elements inversely to the directions of the arcs of the random graph. An asymptotic formula when $n\to\infty$ and $m\to\infty$ in terms of regular threshold function for the probability that the infected area arising from these m vertices exceeds a fixed α-ratio $(1/2<\alpha\leqq 1)$ of all n elements conditioned upon the event $A_n=\{$the greatest component of the random graph exceeds $\alpha n\}$ is derived. The formula shows the relation between the epidemic model and the number of cyclic vertices contained in the great components of the random graph.

Consider the set M_n of all mappings of the finite set $X=\{1, 2, \dots, n\}$ into itself, which satisfy tne condition: $Tx\neq x$, for each $T\in M_n$ and $x\in X$. There are $(n-1)^n$ different mappings in M_n. Each mapping $T\in M_n$ is a digraph G_T, whose points belong to the set X; the points x and y are joined by an arrow iff $y=Tx$. G_T may consist of disjoined components and each component includes only one cycle. We classify the components of G_T corresponding to their size, i. e. to the number of points they consist of. Let the uniform probability distribution on M_n be given (each mapping $T\in M_n$ has probability $(n-1)^{-n}$). The random mappings just described are the second type mappings studied by Harris [1].

We shall consider a scheme of an epidemic process on the random graphs G_T, which is introduced in the paper of Gertsbakh [2].

Define $T^k x$ to be the k-th iteration of $T\in M_n$ on $x\in X$, where k is integer, i.e. $T^k x=(T^{k-1}x)$ and $T^0x=x$. If for some $k\leq 0$, $T^kx=y$, y is said to be a k-th inverse of x in T. The set of all k-th inverses of x in T is denoted by $T^k(x)$ and $P_T(x)=\bigcup_{k=-n}^{0} T^{(k)}(x)$ is the set of all inverses (or predecessors) of x.

Let m bacteria be placed at the elements $x_1, \dots, x_m$, where $x_i\in X$, $i=1,\dots,m$. All $\binom{n}{m}$ different occupations are equally probable. An inverse epidemic process (IEP) [2] is defined by the infection being delivered from the infected points to all their predecessors. The area which will be infected is the set of all inverses $P_T(m)=\bigcup_{i=1}^{m} P_T(x_i)$ of $x_1, \dots, x_m$.

Denote the number of distinct elements in the set $P_T(m)$ by $|P_T(m)|$. Consider the function $v_m: M_n\to R^1$ which maps each $T\in M_n$ into the integer $|P_T(m)|$.

Many asymptotic properties of IEP are described in [3]. The aim of this note is to fill out some conclusions of the papers [4] and [5] in which the relation between the IEP and the number of the cyclic vertices contained in the great components of G_T is shown. We need the notion of the "regular threshold function" which comes from Erdös and Renyi [6]. Let us remark that the threshold function for the IEP (the definition see in [2]) is obtained by Burtin in [3] and by the author in [5].

Let $A_n \subset M_n$ be arbitrary event and $0<\alpha<1$.

Definition. *The function* $\varphi(n)$ *is called regular threshold for the IEP conditioned upon* A_n *if* (i) $\varphi(n)\to\infty (n\to\infty)$; (ii) *there exists a probability distribution function* $F_A(t)$ *so that if* $0<t<\infty$ *is a point of continuity of* $F_A(t)$ *then* $\mathsf{P}\{v_m\geq \alpha_n \mid A_n\}\to F_A(t)$ *for* $m\sim tn$ $(m, n\to\infty)$. $F_A(t)$ *is called threshold conditional probability distribution function.*

Let $\nu_n=\nu_n(T)$, $T\in M_n$, be the size of the greatest component of G_T. The following theorem gives the asymptotic of the probability that the infected area arising from m bacteria exceeds a fixed α-ratio $(1/2<\alpha\leq 1)$ of all elements in the population X conditioned upon the event that the size of the greatest component exceeds also αn.

Theorem. *The regular threshold function for the IEP conditioned upon the event* $A_n=\{\nu_n\geq \alpha n\}$, $1/2<\alpha\leq 1$, *is* $\varphi(n)=\sqrt{n}$. *The threshold conditional probability distribution function is*

$$F_A(t)=[\ln\frac{1+\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}-I_\alpha(t)]/\ln\frac{1+\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}},$$

where

$$I_\alpha(t)=\sqrt{\frac{2}{\pi}}\int_\alpha^1\int_0^\infty\frac{e^{-tu\sqrt{x}-u^2/2}}{x\sqrt{1-x}}\,du\,dx. \tag{1}$$

Proof. Let $\mu_{s,n}=\mu_{s,n}(T)$ be the number of the cyclic points of G_T, contained in components of sizes exceeding $s(\geq 2)$, and $\eta_{m,s}$ be the number of bacteria placed in cyclic elements satisfying the above condition. According to the formula of the total probability, for the distribution of $\eta_{m,s}$, we obtain

$$\mathsf{P}\{\eta_{m,s}=k\}=\sum_{l=2}^{n}\frac{\binom{l}{k}\binom{n-l}{m-k}}{\binom{n}{m}}\mathsf{P}\{\mu_{s,n}=l\},\quad 0\leq k\leq m. \tag{2}$$

(By the definition of the mappings $T\in M_n$ it follows that $\mathsf{P}\{\mu_{s,n}=1\}=0$).

In [4] was found the limit distribution of $\mu_{s,n}/\sqrt{n}$ for $s\sim \alpha n$, $0<\alpha<1$, in terms of Laplace-Stieltjes transforms. For a given $\alpha\in(1/2, 1]$ the result which we shall use is

$$\lim_{n\to\infty, s\sim\alpha n}\mathsf{E}\{\exp(-t\mu_{s,n}/\sqrt{n})\}=1-\frac{1}{2}F_A(t)\ln\frac{1+\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}. \tag{3}$$

In order to be able to apply the distribution of $\eta_{m,s}$ from (2) *for* $s=\alpha n$ and the limit relation (3) we shall note that

$$\mathsf{P}\{v_m\geq\alpha n,\ \nu_n\geq\alpha n\}=\mathsf{P}\{\eta_{m,\alpha n}\geq 1,\ \nu_n\geq \alpha n\}=[1-\mathsf{P}\{\eta_{m,\alpha n}=0\mid \nu_n\geq\alpha n]\mathsf{P}(\nu_n\geq\alpha n\}$$
$$=[1-\mathsf{P}\{\eta_{m,\alpha n}=0\mid \tilde\mu_{\alpha,n}\geq 2\}]\mathsf{P}\{\tilde\mu_{\alpha,n}\geq 2\}=\mathsf{P}\{\tilde\mu_{\alpha,n}\geq 2\}-\mathsf{P}\{\eta_{m,\alpha n}=0,\ \tilde\mu_{\alpha,n}\geq 2\}$$
$$=\mathsf{P}\{\nu_n\geq\alpha n\}-[\mathsf{P}\{\eta_{m,\alpha n}=0\}-\mathsf{P}\{\eta_{m,\alpha n}=0,\ \tilde\mu_{\alpha,n}=0\}].$$

Here $\tilde\mu_{\alpha,n}=\mu_{\alpha n,n}$. Since $\{\tilde\mu_{\alpha,n}=0\}\subset\{\eta_{m,\alpha n}=0\}$ and $\{\tilde\mu_{\alpha,n}=0\}=\bar A_n$ then

$$\mathsf{P}\{v_m\geq\alpha n,\ \nu_n\geq\alpha n\}=\mathsf{P}(A_n)-\mathsf{P}\{\eta_{m,\alpha n}=0\}+\mathsf{P}(\bar A_n)=1-\mathsf{P}\{\eta_{m,\alpha n}=0\} \tag{4}$$

So it suffies to find the limit of the $\mathsf{P}\{\eta_{m,\alpha n}=0\}$ when $m\sim t\sqrt{n}$, $n\to\infty$. By (2) it follows that

$$\mathsf{P}\{\eta_{m,\alpha n}=0\}=\sum_{l=2}^{n}\binom{n-l}{m}\binom{n}{m}^{-1}\mathsf{P}\{\tilde{\mu}_{\alpha,n}=l\}. \tag{5}$$

Since $\binom{n-l}{m}\binom{n}{m}^{-1}=(1-\frac{m}{n})(1-\frac{m}{n-1})\dots(1-\frac{1}{n-l+1})$

then

$$(1-\frac{m}{n-l+1})^{l}<\binom{n-l}{m}\binom{n}{m}^{-1}<(1-\frac{m}{n})^{l}. \tag{6}$$

Now, combining (5) and (6), we obtain

$$\mathsf{E}\{(1-\frac{m}{n-\tilde{\mu}_{\alpha,n}+1})^{\tilde{\mu}_{\alpha,n}}\}\leqq\mathsf{P}\{\eta_{m,\alpha n}=0\}\leqq\mathsf{E}\{(1-\frac{m}{n})^{\tilde{\mu}_{\alpha,n}}\}. \tag{7}$$

We shall use also the asymptotic of $\mathsf{E}\tilde{\mu}_{\alpha,n}$ [5], which has the form

$$\mathsf{E}\tilde{\mu}_{\alpha,n}=\sqrt{\frac{2n}{\pi}}(\frac{\pi}{2}-\arcsin\sqrt{\alpha})+O(1). \tag{8}$$

Let $\delta>0$ be an arbitrary number. Using the representation

$$(1-\frac{m}{n-\tilde{\mu}_{\alpha,n}+1})^{\tilde{\mu}_{\alpha,n}}=(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\sqrt{n}\,\tilde{\mu}_{\alpha,n}/\sqrt{n}} \tag{9}$$

$$=\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}-1}{n})^{-1}\}(1+O(\frac{\tilde{\mu}_{\alpha,n}}{n}))$$

by (8) and the Markov inequality, we have

$$\mathsf{P}\{|(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\sqrt{n}\,\tilde{\mu}_{\alpha,n}/\sqrt{n}}-\exp\{t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}-1}{n})^{-1}\}|>\varepsilon\}$$

$$\leqq\frac{\text{const}}{\varepsilon n}\,\mathsf{E}\{\exp[-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}}{n})^{-1}\,\tilde{\mu}_{\alpha,n}]\}\leqq\frac{\text{const}}{\varepsilon n}\,\mathsf{E}\tilde{\mu}_{\alpha,n}\to 0,\quad n\to\infty.$$

These relations yield the following convergence in probability

$$(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\sqrt{n}\,\tilde{\mu}_{\alpha,n}/\sqrt{n}}-\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}}{n})^{-1}\}\xrightarrow{P}0,\quad n\to\infty.$$

On the other hand from (9) it follows with probability 1, that

$$|(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\sqrt{n}\,\tilde{\mu}_{\alpha,n}/\sqrt{n}}-\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}}{n})^{-1}\}|\leqq\frac{\tilde{\mu}_{\alpha,n}}{n}.$$

Now by the dominated convergence theorem [7, Theorem 5.4] we obtain

$$\lim_{n\to\infty}\mathsf{E}\{(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\tilde{\mu}_{\alpha,n}}\}=\lim_{n\to\infty}\mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}}{n})^{-1}\}\}.$$

Using (8), it is easy to see, that

$$\mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{(1-\delta)\sqrt{n}}\}\}\leqq \mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}(1-\frac{\tilde{\mu}_{\alpha,n}}{n})^{-1}\}\}\leqq \mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}\}\}$$

for an arbitrary small $\delta>0$. Therefore we have

$$\lim_{n\to\infty}\mathsf{E}\{(1-\frac{t+o(1)}{\sqrt{n}-(\tilde{\mu}_{\alpha,n}-1)/\sqrt{n}})^{\tilde{\mu}_{\alpha,n}}\}=\lim_{n\to\infty}\mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}\}\}. \tag{10}$$

Similarly can be proved, that

$$\lim_{\substack{n\to\infty\\ m\sim t\sqrt{n}}}\mathsf{E}\{(1-\frac{m}{n})^{\tilde{\mu}_{\alpha,n}}\}=\lim_{n\to\infty}\mathsf{E}\{\exp\{-t\frac{\tilde{\mu}_{\alpha,n}}{\sqrt{n}}\}\}. \tag{11}$$

Combining (3), (4), (7), (9)—(11), and using the relation [5]

$$\lim_{n\to\infty}\mathsf{P}\{\nu_n\geqq\alpha n\}=\frac{1}{2}\ln\frac{1+\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}=\lim_{n\to\infty}\mathsf{P}\{\tilde{\mu}_{\alpha,n}\geqq 2\}$$

we obtain the assertion of the theorem.

Finally let us remark that the value of $\lim_{\substack{n\to\infty\\ m\sim t\sqrt{n}}}\mathsf{P}\{v_m\geqq\alpha n\}$, $0<t<\infty$, $0<\alpha<1$, can be found by other methods described by Burtin [3].

REFERENCES

1. B. Harris. Probability distributions related to random mappings. *Ann. Math. Statist.*, **31**, 1960, 1045—1062.
2. I. B. Gertsbakh. Epidemic process on a random graph: some preliminary results. *J. Appl. Probab.*, **14**, 1977, 427—438.
3. Y. D. Burtin. On a simple formula for random mappings and its applications. *J. Appl. Probab.*, **17**, 1980, 403—414.
4. L. Mutafchiev. A limit distribution related to random mappings and its application to epidemic process. *Serdica*, **8**, 1982, No 2, 197—203.
5. L. Mutafchiev. Epidemic processes on random graphs and their threshold function. *Serdica*, **7**, 1981, 153—159.
6. P. Erdös, A. Renyi. On the evolution on random graphs. *Maguar. Tud. Acad. Math.* **5**, 1960, 17—61.
7. П. Биллингсли. Сходимость вероятностных мер. Москва, 1977.

Centre for Mathematics and Mechanics
1090 Sofia P. O. Box 373

Received 2. 2. 1982