Study of a Community Evolution Model

Anusua Paul

Roll No : **241080056**

Project Report under the Guidance of

Dr. Soumyarup Sadhukhan

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Certificate

This is to certify that the work presented in the thesis entitled , authored by **Anusua Paul**, has been carried out under my supervision. To the best of my knowledge, this work has not been submitted elsewhere, either in part or in full, for the award of any degree.

Soumyarup Sadhukhan

Soumyarup Sadhukhan
Assistant Professor
Department of Mathematics and Statistics
IIT Kanpur

Acknowledgements

I would like to sincerely thank Dr. Soumyarup Sadhukhan for his guidance, support, and insightful discussions throughout the course of this project. His mentorship was crucial in shaping my understanding of community evolution model and related works. I also thank IIT Kanpur for providing this valuable research opportunity. Lastly, I am grateful to my family and peers for their constant encouragement during this project.

Abstract

We introduce a specific community model where we have finite number of communities and in each of them we have finite population. At each discrete time point a person joins any of the communities based on his or her type and a particular adaption strategy. Our goal is to study how the proportion of the population within this community model evolves over time and to analyze its limiting behavior.

Keywords :

Urn model, Graph theory, Random Walk, Probability.

Introduction

Community theory has wide-ranging applications in environmental science, conservation biology, social evolution theory, behavioral science, and related fields. It can be studied using different approach, such as using urn model theory, graph and network based models, game theory etc. One such implication is done in [5], where game theoritical approach is used to study a disease propagation model.

In our case consider a situation where we have finite number of communities. At each discrete time point one person comes and joins any one of the communities based on his or her type and some adaption rule. For example the communities may be political parties and the type may be the person's political belief. Assume that each community initially has a finite, non-zero population. That means the population proportion in each community is strictly positive. As $t \to \infty$ we want to study how this proportion varies. For this study we will use aspects like urn model, graph and network based modelling and a very few simulation.

Model Description

Initially let us define some notations. Let $m \in \mathbb{N}$ be the initial population over all the communities, $l \in \mathbb{N}$ denote the number of communities, $k \in \mathbb{N}$ denote the number of types, both finite. The set of communities is given by $\mathscr{C} = \{c_1, c_2, \cdots, c_l\}$ and the set of types is given by $\{\theta_1, \theta_2, \cdots, \theta_k\}$. At each discrete time point $t \in \mathbb{N}$ one person enters to any one of these communities based on his or her type and adaption strategy. Suppose the type of the person coming at time t is denoted by a random variable $\theta^{(t)} \sim \lambda_t$ where λ_t is a discrete distribution with support $\{\theta_1, \theta_2, \cdots, \theta_k\}$. Note that λ_t may or may not be identical from time to time.

Define a random vector $X_t(\omega) : (\Omega, \mathscr{F}, \mathbb{P}) \to ((0,1)^l, \mathbb{B}((0,1)^l), \mathbb{P} \circ X_t^{-1})$ where $t \in \mathbb{N}_0$. Here $X_t(\omega) = (X_t^1(\omega), X_t^2(\omega), \cdots, X_t^k(\omega))$ is such that $X_t^j(\omega)$ denotes the population proportion in the community c_j. Define $\mathscr{P}_{m+t}(\mathscr{C})$=$\{x = (x_1, x_2, \cdots, x_l) \in \{0, \frac{1}{m+t}, \cdots, 1\}$ such that $\Sigma x_s = 1\}$. Note that $X_t(\omega) \in \mathscr{P}_{m+t}(\mathscr{C})$. Define $Y_t(\omega) = (m+t) \times X_t(\omega)$. Let $\rho_{ij}(X_t(\omega))$ denote the probability that a person of type i will join community j at time point $t+1$ based on the current state of $X_t(\omega)$.

Different Forms of Adaptation Strategies

We may have various types of adaption rule. It may be based completely on probability that is a person at time point $(t+1)$ with type θ_i will join community c_j with probability $\rho_{ij}(X_t(\omega))$. So after selection of type there will be one more randomization stage where the person will choose the community to join.

Another can be utility based. Suppose we have defined the utility of person with θ_i joining c_j at time $t+1$ by $u_{ij}(X_t(\omega)) = u_{ij} \times f(X_t^j(\omega))$ where u_{ij} is fixed utility and $f(.)$ is the function which captures the information regarding the current status of the communities. Then a person with a particular type chooses among only those communities for which $u_{ij}(X_t(\omega))$ is maximized. If we take $l = k$ and $U = \mathbb{I}_k$ then for θ_i the utility is maximized uniquely at c_i. This scenario will lead to the case when selecting type i is equivalent with selection of community i. Beyond this we may have another adaption rule. We will elaborate on some of them in this paper.

Community Evolution Theory via Urn-Model

The building block of reinforced processes is urn model [7]. This theory was primarily introduced in [2]. One further advancement of this framework were done in [3]. We now attempt to align our community evolution model with this theoretical formulation.

In general framework we have a concept of “Rich gets richer” that is the community which have highest population proportion, people are likely to be inclined towards that community. In other words they will choose their types so that it allows them to join the most populated community. This is the general trend as per studies. For example in today’s job structure data science is playing an important role. So most of the students are trying to build their career based on it so that they can easily join the industry based on data science.

Assume that $X_0^j(\omega) > 0 \; \forall j = 1(1)l$ and $l = k$. Define λ_{t+1}, which will vary over time, as follows:

$$P\big[\theta^{(t+1)} = \theta_i\big] = X_t^i(\omega), \quad i = 1, \dots, l, \quad t \in \mathbb{N}_0.$$

Now, we define $\rho_{ij}\big(X_t(\omega)\big)$ as the probability that a person of type θ_i will join community c_j, as follows:

$$\rho_{ij}\big(X_t(\omega)\big) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \quad \text{for all } i, j = 1, \dots, l.$$

This set up implies that a person with type θ_i always joins community c_i. This model of ours is quite similar to urn model of [2]. This is our primary set up here. Now we are interested to study the behaviour of $X_t(\omega)$ as $t \to \infty$.

We start with the simplest case when $l = k = 2$. Here $Y_t(\omega)$ has two components $(Y_t^1(\omega), Y_t^2(\omega))$. Let $Y_0^1(\omega) = a$ and $Y_t^2(\omega) = b$ where $a, b > 0$ and $a + b = m$. Then, we can express the reinforcement in the communities according to the following structure :

$$Y_{t+1}^1(\omega) = Y_t^1(\omega) + I_{\{U_{t+1} \leq X_t^1(\omega)\}}$$

$$Y_{t+1}^2(\omega) = Y_t^2(\omega) + I_{\{U_{t+1} > X_t^1(\omega)\}}$$

Here, $Y_t^1(\omega)$ represents the number of individuals present in community 1 at time t and U_{t+1} is a random number from $U(0,1)$ independent of t, using which we decide which community to join. This structure ensures the previous model set up. The main result concerning to this model is given by the theorem below.

Theorem 1. *[4] The random variable $X_t^1(\omega)$ converges almost surely to a random variable (limit) $X^1(\omega)$ where $X^1(\omega) \sim Beta_I(a, b)$ with support in $(0, 1)$. In particular when $a = b = 1$ then $X^1(\omega)$ has $U(0, 1)$ distribution.*

Proof. Define $Z_t = Y_t^1(\omega) - Y_0^1(\omega)$. Then Z_t denotes the number of times among t in which the person joins community 1. The probability that at 1st k consecutive times persons will join community 1 is

$$= \frac{a(a+1)\cdots(a+k-1)\, b(b+1)\cdots(b+t-k-1)}{m(m+1)\cdots(m+k-1)\,(m+k)(m+k+1)\cdots(m+t-1)}$$

$$= \frac{\mathrm{Beta}(a+k,\, b+t-k)}{\mathrm{Beta}(a,\, b)} \tag{1}$$

Z_t can take the value k at any of the k time points among the total of t steps. So the probability Z_t taking value $0 \leq k \leq t$ is given by,

$$P[Z_t = k] = \binom{t}{k} \frac{Beta(a+k, b+t-k)}{Beta(a, b)}$$

This is the pmf of a Beta Binomial distribution. For any $r \in \mathbb{R}$ we can show the following. (Change of integration and summation is needed to prove this).

$$E[(Z_t)_r] = (t)_r \frac{Beta(a+r,b)}{Beta(a,b)}$$

Now we have,

$$\frac{(Z_t)_r}{(t)_r} = \frac{Z_t(Z_t-1)\cdots(Z_t-r+1)}{t(t-1)\cdots(t-r+1)} \quad = \frac{Z_t^r}{t^r+O(t^{r-1})} + \frac{O(Z_t^{r-1})}{t^r+O(t^{r-1})} \tag{2}$$

Now as $t \to \infty$ then $O(Z_t^{r-1})$ is dominated by t^r in the denominator, as a result the second quantity term to 0 as $t \to \infty$. Using this fact we can write in limiting sense $\frac{(Z_t)_r}{(t)_r} \approx \frac{Z_t^r}{t^r}$ As $\frac{Z_t}{t}$ is bounded between $(0,1)$ and it converges in r th moment to a $Beta(a,b)$ distribution from there it follows that it is also convergent in distribution to that beta distribution [6]. This proves the convergence in distribution. To prove almost sure convergence a random walk approach is made in [1].

□

Note that it converges to a random limit and not a particular value. Now we generalize this for $l = k \in \mathbb{N}$

Theorem 2. *[4] Suppose we have l communities and k types where $l = k$. In $X_0(\omega)$ we have each and every component strictly positive. We have $P[\theta^{(t+1)} = \theta_i] = X_t^i(\omega)$ and $\rho_{ii}(X_t(\omega)) = 1$ for all $i = 1(l)l$ and 0 otherwise. Under this community evolution scheme the following almost sure convergence holds.*

$$X_t(\omega) \to (S_1, S_2, \cdots, S_{l-1}, S_l)$$

where, $(S_1, S_2, \cdots, S_l)$ is a random vector with dirichlet density

$$\Gamma\left(\sum_{j=1}^{l} Y_0^j(\omega)\right) \prod_{j=1}^{l} \frac{s_j^{Y_0^j(\omega)-1}}{\Gamma(Y_0^j(\omega))},$$

This model can indeed be connected to real-life situations, as it is reasonable to assume that with an increasing population size, more individuals become attracted to a particular community and adapt their preferences accordingly. Moreover, the convergence to a random limit reflects the dynamic nature of such processes, highlighting their relevance to real-world phenomena.

Fixed-Probability Community Evolution

Now we intend to study what happens in long run if $P[\theta^{(t+1)} = \theta_i] = \frac{1}{l}$ for all $i = 1(l)l$ and $t \in \mathbb{N}$. Further assume we have $l = k$. In $X_0(\omega)$ we have each and every component strictly positive and $\rho_{ii}(X_t(\omega)) = 1$ for all $i = 1(l)l$ and 0 otherwise. Under this background we have the process X_t is given by

$$X_{t+1}^s(\omega) = \begin{cases} \dfrac{X_t^s(\omega)(m+t)+1}{m+t+1}, & \text{with probability } p_s\big(X_t(\omega)\big), \\ \dfrac{X_t^s(\omega)(m+t)}{m+t+1}, & \text{with probability } 1 - p_s\big(X_t(\omega)\big). \end{cases}$$

where

$$\begin{aligned} p_s\big(X_t(\omega)\big) &= P\Big[\text{person at time } (t+1) \text{ joins community } s \\ &\qquad \text{depending on the state of } X_t(\omega)\Big] \\ &= \sum_{i=1}^{k} P\Big[\text{person at time } (t+1) \text{ joins community } s \text{ based on } X_t(\omega) \\ &\qquad |\theta^{(t+1)} = \theta_i]\, P\big[\theta^{(t+1)} = \theta_i\big] \\ &= \sum_{i=1}^{k} \rho_{is}\big(X_t(\omega)\big)\, P\big[\theta^{(t+1)} = \theta_i\big] \\ &= \frac{1}{l} \end{aligned}$$

That is the probability remains fixed from time to time. Also it does not depend on either on the current state of community or the choice of type and selection of which community to join. From here it is clear that the process $Y_t(\omega)$ is a markov process as it only depends upon $Y_{t-1}(\omega)$ but the question remains is whether $X_t(\omega)$ is also a Markov Process? However as for now we are interested in the limiting state of $X_t(\omega)$. The first question what arises is whether it exist or not. Before going theoretically let us try to justify it with simulation.

We take $l = k = 5$ and $X_0(\omega) = (0.2, 0.2, 0.2, 0.2, 0.2)$ and $m = 50$ that is we basically start with a symmetric proportion. We take thousand iteration and then check for the limiting proportion. We conduct this simulation six times and the result is listed below.

So even after starting with a symmetric population from the plot and table it is clear that as $t \to \infty$ the limiting proportion does not converge to a particular value. Rather it is very random in nature. Even from the trajectory given below it is clear that all proportions are very close to 0.2 but they exactly do not converge. Rather they oscillate between a particular band. So in this case

Table 1:

SL.NO.	C1	C2	C3	C4	C5
1	0.2105	0.2095	0.2067	0.1914	0.1819
2	0.2029	0.1952	0.2057	0.2076	0.1886
3	0.1943	0.2200	0.1924	0.1838	0.2095
4	0.1886	0.1733	0.2143	0.2067	0.2172
5	0.1952	0.2181	0.1876	0.1895	0.2095
6	0.1933	0.2086	0.2029	0.1933	0.2019

also it converges to a random limit though as before we are not able to capture any structure of the randomness here theoretically.

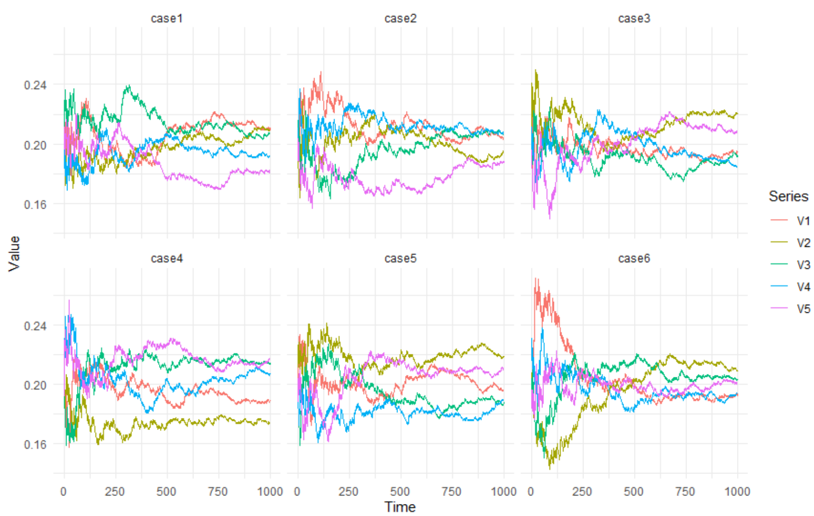

Figure 1: Trajectory of $X_t(\omega)$ wrt t

Graph Based Community Evolution

This is another aspect of studying such community evolution model. Suppose $l = k$, the person with type θ_i always chooses community c_i to join, so selecting the type and community both are equivalent. Now we define Edge Reinforced Random Walk on a graph $G = (V, E)$ where $V = \{1, 2, ..., k\}$. This is a finite graph and any fixed vertex i, it is connected to j for all $j = 1(1)l$, that means self connected edge is also there and note that this is a directed graph. So there is in total k^2 elements in E, $(k^{`}, j) \in E$ for all $j, k^{`} = 1, 2, .., k$. Suppose initially at time point 0 $w(e, 0)$ or $\{W(e, 0)\} = 1$ for all $e \in E$, implying initially all communities are equally likely to join. We start the random walk at $i \in V$, note that at $t = 0$ no person joins any of the communities. Suppose L_t denotes the position of random walk at time t. If at time point t the random walk is on vertex $j \in V$ then we say that at t, one person joins c_j. The probability of transition is given by,

$$P(L_{t+1} = j|\mathscr{F}_t) = \frac{W(\{L_t, j\}, t)}{\sum_z W(\{L_t, z\}, t)}$$

with,

$$W(\{i, j\}, t) = W(\{i, j\}, t-1) + I_{\{(L_t, L_{t-1})=(i,j)\}}$$

and $\mathscr{F}_t = \sigma(L_0, L_1, ..., L_t)$. Note that here at first the joining is done then the updation of $w(e, t)$ is done. Then the following theorem holds.

Theorem 3. *[7] Suppose L_t is an ERRW with on $G = (V, E)$ then it is a mixture of markov chains means there exist a measure μ such that,*

$$P(L_0 = i_0, .., L_t = i_t) = \int p_{i_0,i_1} p_{i_1,i_2} ....p_{i_{t-1},i_t} d\mu$$

Moreover new weights constructed keeping the constraint that sum of them should be 1, that is $\frac{w(e^*,t)}{\sum_e w(e,t)}$, for fixed $e^* \in E$, it converges to a random limit continuous with respect to the lebesgue measure in distribution.

Now as per our model we can write

$$Y_t^j(\omega) = Y_0^j(\omega) + w((L_{t-1}, j), t-1)$$

this imply,

$$\frac{Y_t^j(\omega)}{\sum_e w(e, t)} = \frac{Y_0^j(\omega)}{\sum_e w(e, t)} + \frac{w((L_{t-1}, j), t-1)}{\sum_e w(e, t)}$$

This imply $\frac{Y_t^j(\omega) - Y_0^j(\omega)}{k^2+t}$ in distribution tends to a random limit. Note that here $\sum_e w(e, 0) = k^2 + t$. So the fraction of joining each community within the time t tends to a random limit as $t \to \infty$.

Now instead of using the weight function $w(e, 0)$ corresponding to edges we define a function which is dependent on vertex, given by, $Z_t(v) = 1 + \sum_0^t I_{\{L_z=v\}}$ and instead of self connected graph we use a graph with no self edge and denote it by $G = (V, E)$. Here the adaption probability is defined by,

$$P[L_{t+1}(\omega) = j|\mathscr{F}_t] = \frac{Z_t(j)}{\sum_1^k Z_t(z)}$$

Then the following theorem holds,

Theorem 4. *[7] Suppose G is a complete graph (no self edges) with $V = \{1, 2, \ldots, l\}$. Then the global centroid of $Y_t'(\omega)$ is given by $(\frac{1}{l}, \frac{1}{l}, \ldots, \frac{1}{l})$, and it is an attractor (a set towards which a system tends to evolve in long run), where the ith component of $Y_t'(\omega)$ is the fraction of time spent by the random walk within t time points. Each permutation of the form $(\frac{1}{c}, \frac{1}{c}, \ldots, \frac{1}{c}, 0, \ldots, 0)$ is linearly unstable.*

Conclusion

In this paper, we attempted to determine the limiting behavior of $X_t(\omega)$. However, in certain cases, our analysis led us only to conclusions regarding the convergence in distribution of the j-th component, expressed as

$$\frac{Y_t^j(\omega) - Y_t^0(\omega)}{t}.$$

We acknowledge that this represents a rather weak form of convergence and does not adequately capture the exact joint structure or the joint limiting behavior of $X_t(\omega)$. Further work can be undertaken to improve and strengthen the results in this direction.

References

[1] Gunnar Blom and Lars Holst. Random walks of ordered elements with applications. *The American Statistician*, 40(4):271–274, 1986.

[2] F. Eggenberger and G. Pólya. Über die statistik verketteter vorgänge. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 3(4):279–289, January 1923.

[3] David A Freedman. Bernard friedman's urn. *The Annals of Mathematical Statistics*, pages 956–970, 1965.

[4] Raúl Gouet. Strong convergence of proportions in a multicolor pólya urn. *Journal of Applied Probability*, 34(2):426–435, 1997.

[5] Sayar Karmakar, Moumanti Podder, Souvik Roy, and Soumyarup Sadhukhan. The spread of an epidemic: a game-theoretic approach, 03 2023.

[6] A. W. Kemp. Urn models and their applications: an approach to modern discrete probability theory, by norman l. johnson and samuel kotz. pp xiii, 402. £16·45. 1977. sbn 0 471 44630 0 (wiley). *The Mathematical Gazette*, 62(420):140–141, 1978.

[7] Robin Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, 4(none), January 2007.