A gentle introduction to combinatorial stochastic processes IV

Enrico Scalas

Department of Mathematics, University of Sussex, UK

Stochastic Models for Complex Systems
10 June - 7 July, 2021

Outline

1. Motivation
2. Previous models
3. A recent stylized model

Modelling complex systems

I shall now apply the ideas discussed in the previous lectures to simple models for the distribution of wealth. While modelling a complex system, I suggest the following approach:

- Identify objects and categories.
- Write a (stochastic) model for time evolution (it will often turn out to be a Markov chain): How do objects move within categories?
- Characterize the Markov chain: Does it lead to statistical equilibrium? Does it show metastability? What about mixing times? What are the Kolmogorov equations? ....
- Consider scaling limits of the Markov chain: Is there convergence? To what limiting process? Can you derive integro-differential equations related to the problem (e.g. for the probability density function of the limiting process)? ....

I shall illustrate an example of this research program.

Wealth and income

Wealth and income are often confused by non-economists. **Wealth** is a *stock* and it is given by *total assets* in a balance sheet (liabilities + equity). It can be negative if there is indebtedness and liabilities exceed equity. *Wealth is determined by taking the total market value of all physical and intangible assets owned, then subtracting all debts* according to Investopedia. **Income** is a flow and it is given by *money or the equivalent value that an individual or business receives, usually in exchange for providing a good or service or through investing capital* again according to Investopedia.

The models we discuss below do work for wealth, but not for income.

Wealth inequality

- Why is there wealth inequality?
- Pareto's opinion: La répartition de la richesse peut dépendre de la nature des hommes dont se compose la societé, de l'organisation de celle-ci, et aussi, en partie, du *hasard* (les *conjonctures* de Lassalle), [...]. (V. Pareto, Cours d'économie politique, Tome II Livre III. F. Pichon, Imprimeur-Éditeur, Paris, France, 1897).
- As mathematicians and physicists, we may be able to answer this question.
- My provisional answer is: Chance is a major determinant of inequality.

Literature review I

Champernowne 1952, 1953, Simon 1955, Wold and Whittle 1957 as well as Mandelbrot 1961 used random processes to derive distributions for income and wealth. Starting from the late 1980s and publishing in the sociological literature, Angle introduced the so-called inequality process, a continuous-space discrete time Markov chain for the distribution of wealth based on the surplus theory of social stratification (Angle 1986).

Literature review II

However, the interest of physicists and mathematicians was triggered by a paper written by Drăgulescu and Yakovenko in 2000 and explicitly relating random exchange models with statistical physics. Among other things, they discussed a simple random exchange model already published in Italian by Bennati 1988. An exact solution of that model was published in Scalas 2006 and is outlined below. Lux wrote an early review of the statistical physics literature up to 2005. An extensive review was written by Chakrabarti and Chakrabarti in 2010. Boltzmann-like kinetic equations for the marginal distribution of wealth were studied by Cordier et al. in 2005 and several other works, we refer to the review article by Düring et al. 2009 and the book by Pareschi and Toscani 2014, and the references therein.

Hard spheres: A prototypical model

In a *microcanonical* fluid of hard spheres, the total number of particles N is conserved and the total energy E is conserved. One finds that the normalised particle energies $\varepsilon_i = E_i/E$ follow a Dirichlet distribution with density:

$$f_{\varepsilon}(\mathbf{u}) = \frac{\Gamma(dN/2)}{[\Gamma(d/2)]^N} \prod_{i=1}^{N} x_i^{d/2-1} \mathbb{I}_S(\mathbf{u}),$$

where $\mathbb{I}_S(\cdot)$ is the indicator function of the simplex defined by $\sum_{i=1}^{N} \varepsilon_i = 1$.

Hard spheres: Marginal distribution I

Particles are exchangeable. After marginalising the Dirichlet, one finds that the normalised energy of a single particle follows a Beta distribution with density:

$$f_\varepsilon(u) = \frac{\Gamma(dN/2)}{\Gamma(d/2)\Gamma(d(N-1)/2)} u^{d/2-1}(1-u)^{d(N-1)/2-1}\,\mathbb{I}_{[0,1]}(u).$$

Energy can be seen as wealth. For large N we have a skewed distribution of energy.

Hard spheres: Marginal distribution II

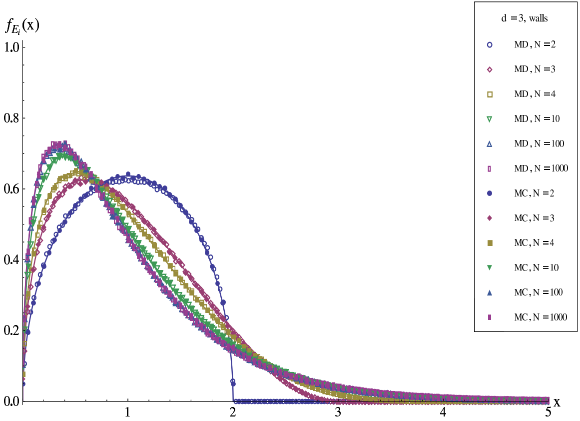

Figure: This is the distribution of non-normalised energy per particle for $d = 3$ and $E = N\bar{\varepsilon}$ when $\bar{\varepsilon} = 1$.

Finitary models: The Ehrenfest-Brillouin model

The Ehrenfest-Brillouin model is a Markov chain in which n objects can move into g categories or boxes according to the following transition probability

$$\mathbb{P}(\mathbf{n}_i^k|\mathbf{n}) = \frac{n_i}{n}\frac{\alpha_k + n_k - \delta_{k,i}}{\alpha + n - 1}$$

where the α_i's are category weights such that $\sum_{i=1}^{g} \alpha_i = \alpha$. The invariant distribution which is also an equilibrium distribution is a generalised g-dimensional Pólya distribution

$$\pi(\mathbf{n}) = \frac{n!}{\alpha^{[n]}} \prod_{i=1}^{g} \frac{\alpha_i^{[n_i]}}{n_i!},$$

where $\alpha^{[n]} = \alpha(\alpha+1)\cdots(\alpha+n-1)$. This was used as a toy model for taxation and redistribution.

Finitary models: The continuum limit

Assume that all the $\alpha_i = \theta$ for every i. The marginal distribution on a category is

$$\pi(k) = \frac{n!}{k!(n-k)!} \frac{\theta^{[k]}((n-1)\,\theta)^{[n-k]}}{(n\theta)^{[n]}}$$

whose continuum limit is (for $u = k/n$ with $k < n$ and k and n large) the density

$$\pi(u) = \frac{\Gamma(n\,\theta)}{\Gamma(\theta)\Gamma((n-1)\theta)} u^{\theta-1}(1-u)^{(n-1)\theta-1}\, \mathbb{I}_{[0,1]}(u).$$

The identification $\theta = d/2$ and $n = N$ gives the same distribution as for normalised energies in the hard-sphere fluid.

Finitary models: The BDY model I

We have g agents with n/g coins, each. We play the following game:

1. At each step a *loser* is selected by chance from all the agents with at least one coin;
2. the loser gives one of his/her coins to a *winner* randomly selected among all the agents.

This can be represented by the following transition probability

$$\mathbb{P}(\mathbf{n}'|\mathbf{n}) = \frac{1-\delta_{n_i,0}}{g-z_0(\mathbf{n})}\frac{1}{g},$$

where $z_0(\mathbf{n})$ represents the number of agents without coins.

Finitary models: The BDY model II

The invariant and equilibrium distribution is

$$\pi(\mathbf{n}) = C(g - z_0(\mathbf{n})).$$

The marginalisation is not trivial even if there is agent exchangeability. Consider the partition vector $\mathbf{Z} = (Z_0, \ldots, Z_n)$ where Z_0 represents the number of agents with zero coins, Z_1 the number of agents with one coin, and so on, with $\sum_{i=1}^{n} Z_i = g$ and $\sum_{i=1}^{n} iZ_i = n$. We cannot use naive maximum entropy to find the most probable value of $\mathbf{Z}$ ($\pi(\mathbf{n})$ is not uniform), but we have the multivariate distribution of $\mathbf{Z}$:

$$\mathbb{P}(\mathbf{Z} = \mathbf{z}) = \frac{g!}{z_0! z_1! \cdots z_n!} \pi(\mathbf{n}) = \frac{g!}{z_0! z_1! \cdots z_n!} C(g - z_0(\mathbf{n})).$$

The normalization constant C is given by

$$C = \left[\sum_{k=1}^{g} k \binom{n}{k} \binom{n-1}{k-1} \right]^{-1}$$

We can find $\mathbb{E}(Z_i)$ as

$$\mathbb{E}(Z_i) = \sum_{k=0}^{g} \mathbb{E}(Z_i|k)\mathbb{P}(k) = \sum_{k=0}^{g} g\,\mathbb{P}(n_1 = i|k)\mathbb{P}(k),\; k = g - z_0.$$

$$\mathbb{P}(k) = Ck\binom{n}{k}\binom{n-1}{k-1}$$

and

$$\mathbb{E}(Z_0|k) = g - k$$

$$\mathbb{E}(Z_i|k > 1) = k\frac{\binom{n-i-1}{k-2}}{\binom{n-1}{k-1}},\; i = 1, \ldots, n-1$$

$$\mathbb{E}(Z_i|k = 1) = \delta_{i,n},\; i = 1, \ldots, n$$

$$\mathbb{E}(Z_i|k) = 0,\; \text{for}\; n - i - 1 < k - 2 \;\text{and}\; i = n.$$

Only for $n \gg g \gg 1$, we get

$$\frac{\mathbb{E}(Z_i)}{g} \approx \frac{g}{n}\left(1 - \frac{g}{n}\right)^i,$$

a geometric distribution coinciding with the naive maximum entropy solution. In fact, in this limit, the probability of finding agents without coins is negligible.

General framework

Distributional problems in Economics can be presented in a rather general form. Assume one has N economic agents, each one endowed with his/her stock (for instance wealth) $w_i \geq 0$. Let $W = \sum_{i=1}^{N} w_i$ be the total wealth of the set of agents. Consider the random variable W_i, i.e. the stock of agent i. One is interested in the distribution of the vector $(W_1, \ldots, W_N)$ as well as in the marginal distribution W_1 if all agents are on a par (exchangeable). The transformation $X_i = W_i/W$, normalises the total wealth of the system to be equal to one since $\sum_{i=1}^{N} X_i = 1$ and the vector $(X_1, \ldots, X_N)$ is a finite random partition of the interval $(0, 1)$. The X_is are called *spacings* of the partition.

Remarks

The following remarks are useful and justify simplified modelling of wealth distribution.

1. If the stock w_i represents wealth, it can be negative due to indebtedness. In this case, one can always shift the wealth to non-negative values by subtracting the negative wealth with largest absolute value.
2. A mass partition is an infinite sequence $\mathbf{s} = (s_1, s_2, \ldots)$ such that $s_1 \geq s_2 \geq \ldots \geq 0$ and $\sum_{i=1}^{\infty} s_i \leq 1$.
3. Finite random interval partitions can be mapped into mass partitions, just by ranking the spacings and adding an infinite sequence of 0s.
4. In principle, the total wealth W can change in time. Here we assume it is constant.

Research questions

The vector $\mathbf{X} = (X_1, \ldots, X_N)$ lives on the $N-1$ dimensional simplex Δ_{N-1}, defined by

Definition (The simplex Δ_{N-1})

$$\Delta_{N-1} = \left\{ \mathbf{x} = (x_1, \ldots, x_N) : \, x_i \geq 0 \; \forall i = 1, \ldots, N, \, \sum_{i=1}^{N} x_i = 1 \right\}.$$

There are two natural questions that immediately arise from defining such a model.

1. Which is the distribution of the vector $(X_1, \ldots, X_N)$ at a given time?
2. Which is the distribution of the random variable X_1, the proportion of the wealth of a single individual?

Random dynamics on the simplex I

To define our simple model, we first introduce two types of moves on the simplex.

Definition (Coagulation)

By coagulation, we denote the aggregation of the stocks of two or more agents into a single stock. This can happen in mergers, acquisitions and so on.

Definition (Fragmentation)

By fragmentation, we denote the division of the stock of one agent into two or more stocks. This can happen in inheritance, failure and so on.

Random dynamics on the simplex II

Let $\mathbf{X} = \mathbf{x}$ be the current value of the random variable $\mathbf{X}$. For any ordered pair of indices i, j, $1 \leq i, j \leq N$, chosen uniformly at random, define the coagulation application

$$\mathrm{coag}_{ij}(\mathbf{x}) : \Delta_{N-1} \to \Delta_{N-2}$$

by creating a new agent with stock $x = x_i + x_j$ while the proportion of wealth for all others remain unchanged. Next enforce a random fragmentation application

$$\mathrm{frag}(\mathbf{x}) : \Delta_{N-2} \to \Delta_{N-1}$$

that takes x defined above and splits it into two parts as follows. Given $u \in (0, 1)$ drawn from the uniform distribution $U[0, 1]$, set $x_i = ux$ and $x_j = (1 - u)x$.

Random dynamics on the simplex III

The sequence of coagulation and fragmentation operators defines a time-homogeneous Markov chain on the simplex Δ_{N-1}. Let $\mathbf{x}(t) = (x_1(t), \ldots, x_i(t), \ldots, x_j(t), \ldots, x_N(t))$ be the state of the chain at time t with i and j denoting the selected indices. Then the state at time $t+1$ is

$$\mathbf{x}(t+1) = (x_1(t+1) = x_1(t), \ldots, x_i(t+1) = u(x_i(t)+x_j(t)), \ldots,$$
$$x_j(t+1) = (1-u)(x_i(t)+x_j(t)), \ldots, x_N(t+1) = x_N(t)).$$

The Markov transition kernel for this process is degenerate because each step only affects a zero-measure Lebesgue set of the simplex.

Invariant distribution

General state space discrete time Markov chains are difficult to study as time is changing in discrete steps and the chain cannot explore the whole state space given that real numbers cannot be put in 1-to-1 correspondence with integers.

Proposition (Duality of coagulation and fragmentation)

Let $\mathbf{X}(t)$ denote the coagulation-fragmentation Markov chain defined above. If $\mathbf{X}(t) \sim U[\Delta_{N-1}]$ then $\mathbf{X}(t+1) \sim U[\Delta_{N-1}]$, as well.

This proposition means that the uniform distribution on the simplex Δ_{N-1} is an invariant distribution for the coagulation-fragmentation chain. Is it unique? Is it the equilibrium distribution?

φ-irreducibility

Definition (φ-irreducibility)

Let $(S, \mathcal{B}(S), \varphi)$ be a measured Polish space. A discrete time Markov chain $\mathbf{X}$ on S is φ-irreducible if and only if for any Borel set A the following implication holds:

$$\varphi(A) > 0 \Longrightarrow L(u, A) > 0, \qquad \text{for all } u \in S,$$

where

$$L(u, A) = \mathbb{P}_u\{\mathbf{X}(n) \in A \text{ for some } n\} = \mathbb{P}\{\mathbf{X}(n) \in A \text{ for some } n | \, \mathbf{X}(0) = u\}.$$

This replaces the notion of irreducibility for discrete Markov chains and means that the chain is visiting any set of positive measure with positive probability.

Conditions for equilibrium I

Definition (Foster-Lyapunov function)

For a *petite* set C, a function $V \geq 0$ with $\rho > 0$ such that for all $x \in S$

$$\int P(x, dy) V(y) \leq V(x) - 1 + \rho \mathbb{I}_C(x),$$

where P is the transition kernel of a Markov chain is called Foster-Lyapunov function (if it exists).

The existence of a Foster-Lyapunov function implies convergence of the kernel P of a φ-irreducible, aperiodic chain to a unique equilibrium measure π coinciding with the invariant measure

$$\sup_{A \in \mathcal{B}(S)} |P^n(x, A) - \pi(A)| \to 0, \text{ as } n \to \infty.$$

(see Meyn&Tweedie 1993) for all x for which $V(x) < \infty$.

Conditions for equilibrium II

If we define τ_C to be the number of steps it takes the chain to return to the set C, the existence of a Foster-Lyapunov function (and therefore convergence to a unique equilibrium) is equivalent to τ_C having finite expectation, i.e.

$$\sup_{x \in C} \mathbb{E}_x(\tau_C) < M_C$$

which in turn is implied when τ_C has geometric tails.
In our case, φ is the Lebesgue measure and the role of the petite set C is played by any set with positive Lebesgue measure. This useful simplification of the mathematical technicalities is due to the compact state space (Δ_{N-1}) and the fact that the uniform distribution on the simplex is invariant for the chain.

Conditions for equilibrium III

Proposition

The discrete chain $\mathbf{X}$ *as defined above is* φ*-irreducible, where* $\varphi \equiv \lambda_{N-1}$ *is the Lebesgue measure on the simplex.*

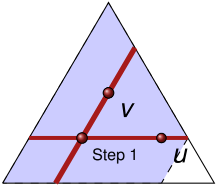

Figure: Schematic of a possible coagulation-fragmentation route from u to v in two steps. Starting from point $u \in \Delta_2$, fix z_u. Then on the line $x + y = 1 - z_u$, pick the point $(x_v, 1 - z_u - x_v, z_u)$. From there, fix x_v and choose (y_v, z_v) on the line $1 - x_v = y_v + z_v$. The shaded region are all points v that can be reached with this procedure from u, first by fixing z_u and then by fixing x_v. Points in the white region can be reached from u first by fixing z_u and then y_v.

Conditions for equilibrium IV

Proposition (Existence of a Foster-Lyapunov function)

The return times τ_A to any set $A \in \mathcal{B}(\Delta_{N-1})$ of positive measure, have at most geometric tails. As a consequence, the Foster-Lyapunov function exists.

Theorem

Let $\mathbf{X}(t)$ denote the coagulation-fragmentation Markov chain defined above with initial distribution π_0 on Δ_{N-1}. Let π_t denote the distribution of $\mathbf{X}(t)$ at time $t \in \mathbb{N}_0$. Then the uniform distribution on Δ_{N-1} is the unique invariant distribution that can be found as the weak limit of the sequence π_t.

Proof.

We have that $U[\Delta_{N-1}]$ is an invariant distribution for the process. Since the chain is φ-irreducible as shown before, uniqueness of the equilibrium follows from the existence of a Foster-Lyapunov function, as a consequence of the previous proposition. $\square$

Wealth distribution

Corollary

The distribution of X_1 is given by the following Beta density

$$\pi_{X_1}(u) = \frac{\Gamma(N)}{\Gamma(N-1)}(1-u)^{(N-2)}\,\mathbb{I}_{[0,1]}(u) = (N-1)u^{(N-2)}\,\mathbb{I}_{[0,1]}(u).$$

Proof.

This can be obtained via direct marginalization of the symmetric Dirichlet distribution with all the $\alpha_i = \theta = 1$ (it is the uniform distribution on Δ_{N-1}). □

Further steps

- Not discussed here: Fulfilling Boltzmann's program: From discrete state space models to continuous state space models.
- Not discussed here: Relation with kinetic equations of Boltzmann type.
- Not discussed here: Preliminary empirical work (more to come).
- Not discussed here: How to get fatter tails.
- Not discussed here: Rate of convergence to equlibrium (mixing times, spectral gaps, etc.).

Acknowledgments

This is joint work with Bertram Düring and Nicos Georgiou.

Thank you for your kind attention!

This work was partially funded by the University of Sussex SDF.

Exercise I

Consider a homogeneous Markov chain whose space state is given by occupation vectors of n objects in g categories with the following transition matrix

$$\mathbb{P}(\mathbf{n}_i^j|\mathbf{n}) = \frac{1-\delta_{n_i,0}}{g-z_0(\mathbf{n})}\frac{1}{g}; \quad (1)$$

this means that an object is selected at random from any category containing at least one object and randomly moves to any one of the other categories. Explicitly consider the case $n = g = 3$; write and diagonalize the transition matrix in order to find the stationary (or invariant) distribution.

Exercise II

Write a Monte Carlo simulation implementing the chain described by the transition probability (1) in the case $n = g = 3$.

For Further Reading

[1] U. Garibaldi and E. Scalas
Finitary Probabilistic Methods in Econophysics
Cambridge University Press, 2010.

[2] B. Düring, N. Georgiou and E. Scalas
A stylized model for wealth distribution
arXiv:1609.08978 [math.PR]