Finite Markov Information-Exchange processes

David Aldous

March 17, 2011

What we have called the *simple epidemic* process could also be called the *SI process*, writing the evolution rules as

- Initially one agent is **I**nfected, others are **S**usceptible.
- When an Infected agent meets a Susceptible agent, the latter becomes infected.

We will now study two related models, which maintain the rules above but add extra rules. An extra rule that both models share is

- An Infected agent becomes un-infected at rate λ, that is after an Exponential(λ) time.

They differ in the consequences of being un-infected. In the *SIS model* or equivalently the *contact process* the un-infected agent becomes susceptible again. In the *SIR model* there is a new state "**R**ecovered" that the un-infected agent enters; subsequently that agent stays in Recovered state, and cannot infect others.

Terminology. The names SIS, SIR come from classical epidemic modelling, with somewhat different background assumptions, e.g. non-Exponential infective times. But let us use the same names in our FMIE setting.

The name *contact process* was introduced in mathematical probability and statistical physics, classically in the infinite lattice setting. Recall our default assumption is a finite number of agents.

To appreciate the distinction between the three processes, consider what the final configurations must be.

- SI: all agents are Infected.
- SIR: some agents are Susceptible (i.e. were never infected), and all the other are Recovered.
- SIS: all agents are Susceptible, because this is the only absorbing configuration.

Where our finite meeting model can be "matched" with an infinite model (the torus $\mathbb{Z}_m^d$ to the lattice $\mathbb{Z}^d$, or the r-regular random graph to the r-regular infinite tree), there are intuitively clear relations between the qualitative behaviors of the two models. In the infinite model, the SIR process may be supercritical (positive chance of an infinite number infected) or subcritical (zero chance). One expects these alternatives to correspond, in the finite setting, to the number of infected (and then recovered) agents being $\Omega(n)$ or $O(1)$.

For an infinite SIS process, typical alternatives are, starting from a configuration with an arbitrary non-zero density ρ_0 of infectives,

(subcritical) the density of infectives $\to 0$
(supercritical) the density of infectives $\to \rho_\infty > 0$.
One expects these alternatives to correspond, in the finite setting, to the alternatives

(subcritical) the infection dies out in $O(1)$ time with $O(1)$ agents ever infected.
(supercritical) the process reaches a "quasistationary" distribution and keeps this distribution for a very long time before (relatively suddenly) dying out.

Digress to some math theory:

General 1-type branching process.

- Parent has lifetime L, and a random number $M \geq 0$ of offspring at random ages $0 < \xi_1 \leq \xi_2 \leq \ldots \leq \xi_M \leq L$.
- The above RVs have arbitrary distribution, except for moment conditions.
- Childen behave as IID copies.

If we ignore clock time and just consider generation structure, this is just a Galton-Watson BP. But, in the supercritical case, what is the behavior of

$$M(t) = \text{Number of births before time } t?$$

$$N(t) = \text{population at time } t?$$

Answer. Under minor assumptions, and on the "non-extinction:" event,

$$M(t) \sim Z \exp(\theta t) \text{ as } t \to \infty$$

where Z is random and $\theta > 0$ is the *Malthusian parameter*.

Proof is technical, but the bottom line formula for θ is easy to derive, assuming the result. Recall births are at ages

$$0 < \xi_1 \leq \xi_2 \leq \ldots \leq \xi_M \leq L.$$

Write

$$\rho(t)dt = \mathbb{E}(\text{number births at age } [t, t + dt])$$

Consider the population growth rate at a large time t. This is

$$Z\theta \exp(\theta t).$$

But population change is caused by a birth to a parent of some age s. The growth rate at time $t - s$ was $Z\theta \exp(\theta(t - s))$, and so the growth rate at time t is

$$\int_0^\infty Z\theta \exp(\theta(t - s)) \times \rho(s) \, ds.$$

Equating these two expressions gives the Malthusian equation

$$\int_0^\infty \rho(s)\, e^{-\theta s}\, ds = 1.$$

As a quick check, the Yule process has $\rho(s) \equiv 1$ and $\theta = 1$.

Note. The behavior of

$$N(t) = \text{ population living at time } t$$

can be deduced [board] as

$$N(t)/M(t) \to \mathbb{E}\exp(-\theta L).$$

Random graphs with prescribed degree distributions

Specify (d_i). Can define models $\mathcal{G}_n$ of n-vertex graph, interpretable as being "random" subject to the following constraint. Write D_n for degree of a uniform random vertex of $\mathcal{G}_n$, then

$$D_n \xrightarrow{d} D \text{ where } \mathbb{P}(D = i) = d_i.$$

Such models have the following "local GWBP approximation". The structure of $\mathcal{G}_n$ within some fixed graph-distance r from a uniform random vertex U_n converges in distribution, as $n \to \infty$, to the random tree comprising generations 0 to r of the following modified Galton-Watson BP. The root has offspring distribution D; in subsequent generation the offspring distribution is the size-biased distribution D^* where $\mathbb{P}(D^* = i) = (i+1)d_{i+1}/\mathbb{E}D$.

Assuming $\mathbb{E}D^{2+\varepsilon} < \infty$ then $\mathbb{E}(D^*)^{1+\varepsilon} < \infty$ and the Kesten-Stigum theorem says that the size Y_r of generation r grows at a particular rate: $Y_r/(\mathbb{E}D^*)^r \to W$ a.s. and L^1.

SIR epidemic on a random graph

Take a typical realization of such a graph $\mathcal{G}_n$. Define a meeting rate matrix $\mathcal{N}$ as the adjacency matrix of $\mathcal{G}_n$:

$$\nu_{ij} = 1 \text{ if } (i, j) \text{ is an edge of } \mathcal{G}_n.$$

We shall analyze the initial phase of a supercritical SIR epidemic. Note that the limiting case where the recovery rate $\lambda = 0$ is just the simple epidemic.

Key point [board]: combining the local GWBP approximation for the geometry and the evolution rule for SIR, we see that (in the initial phase, after the first generation) the process of infectives evolves as the 1-type branching process with

$$0 < \xi_1 \leq \xi_2 \leq \ldots \leq \xi_M \leq L$$

as follows. $L \overset{d}{=}$ Exponential(λ) is infective period; D^* is number of offspring in the GWBP approximation, $(\zeta_j, 1 \leq j \leq D^*)$ are the first meeting times with these neighbors, and the (ξ_i) are those ζ_j for which $\zeta_j < L$.

Conditional on $D^* = d^*$ we have

$$\rho(s) = d^* e^{-(1+\lambda)s}$$

and so unconditionally

$$\rho(s) = m^* e^{-(1+\lambda)s}$$

for $m^* = \mathbb{E}D^* = \mathbb{E}[D(D-1)]/\mathbb{E}[D]$. Note that the mean number of infections passed on by a typical infective is $m^*/(1+\lambda)$ and so the condition for supercriticality of the SIR epidemic is $\lambda < m^* - 1$. The Malthusian equation becomes

$$\frac{m^*}{1 + \lambda + \theta} = 1$$

and so

$$\theta = \frac{m^*}{1 + \lambda} - 1.$$

A full study of "proportion of population infected before time t" is rather complicated, so let us just study the final proportion (π, say) of population never infected. For this, the key point is to consider the probability (x, say) that, for a typical directed edge (v, w), the event

> *v and w meet while v is infected, the infection having reached v from some neighbor other than w*

does **not** happen. We argue informally that

$$\pi = \mathbb{E}x^D \qquad (1)$$

where x is the solution of

$$(1-x)(1+\lambda) = 1 - \phi(x); \quad \phi \text{ the p.g.f. of } D^*. \qquad (2)$$

Recall configuration model. Recall in GWBP approx. a typical vertex w has D branches. Heuristic idea is that, for an epidemic started far away, there is chance x of non-infection from a neighbor v, independently for different v. (This implicitly assumes the proportion ever infected is first-order constant). The independence now gives (1).

To derive the formula for x, consider a neighbor w of v, who has D^* other neighbors y. Now [board] the chance that w is never infected from some other neighbor is (conditionally) x^{D^*} and so is (unconditionally) $\phi(x)$. So w will be sometime infected, with chance $1 - \phi(x)$, and if so then there is chance $1/(1+\lambda)$ to infect v before recovery. In this story (w, v) is a typical edge, so we have shown

$$1 - x = (1 - \phi(x))/(1 + \lambda)$$

which is (2).

As a quick example, for a 4-regular random graph we get [board]

$$\pi = x^4; \quad x = x(\lambda) = \tfrac{1}{2}(\sqrt{1 + 4\lambda} - 1).$$

The condition for supercriticality is $\lambda < 2$, and we get the correct limits $x(0+) = 0$, $x(2-) = 1$.

See e.g. Karrer - Newman (2010) for more detailed analysis along these lines.

Note: SIR epidemics and percolation. A classical discrete-time SIR model is the Reed-Frost model on a graph G. Here, for an agent v infected at time t, each susceptible neighbor is infected independently with probability p at time $t+1$, and v itself is recovered at time $t+1$. A moment's thought shows this process is closely related to the random subgraph G_p of G in which each edge is retained with probability p; in fact the set of agents who are ever infected in the epidemic started at i is distributed as the component of G_p containing i. In particular, on the complete graph, the sometime-infected set is just the connected component (containing the original infective) of the Erdos-Renyi random graph, and on the lattice $\mathbb{Z}^d$ it is the analogous component in bond percolation.

Note this identification does not hold precisely for our continuous-time model: the events that neighbor w becomes infected from v are no longer independent as w varies, because all are influenced by the random duration that v is infective. However, certain "expectation" results could be rephrased in the percolation setting.

If we seek conditions for subcriticality of the SIR epidemic, it turns out that the simplest arguments give bounds for the (a priori larger) SIS epidemic, so we study that.

Consider the FMIE setting of an SIS epidemic over a meeting model $\mathcal{N}$. In the "normalized" case where $\nu_i := \sum_j \nu_{ij} = 1$ for all i, there is a very simple sufficient condition for subcriticality. In this case, the process

$$N(t) := \text{ number infected at time } t$$

is clearly dominated by the linear birth-and death process $N^*(t)$ with transition rates

$$q_{i,i+1} = i, \ q_{i,i-1} = i\lambda.$$

For the latter process we have

$$\mathbb{E}N^*(t) = \exp((1 - \lambda)t)$$

and, for $\lambda > 1$,
$$\mathbb{E}(\text{total number infected}) = \tfrac{\lambda}{\lambda-1}.$$

Corollary

For the SIS epidemic over a normalized meeting model, and for $\lambda > 1$,

$$\mathbb{E}(\text{total number infected}) \leq \tfrac{\lambda}{\lambda-1}.$$

Corollary

For the SIS epidemic over a normalized meeting model, and for $\lambda > 1$,

$$\mathbb{E}(\text{total number infected}) \leq \tfrac{\lambda}{\lambda-1}.$$

For a sequence of meeting models one can define [board] asymptotically critical values λ_n^{crit}, and the result is saying that

$$\lambda_n^{\text{crit}} \leq 1 \text{ in the normalized case.}$$

This bound is not tight on the lattice. To relate this to our previous analysis, for the SIR model on the r-regular infinite tree with normalized rates we have $\lambda^{\text{crit}} = (r - 2)/r$, agreeing with the specialization of our previous "random graphs with prescribed degree distributions" analysis in the case of random r-regular graphs. Roughly speaking, the bound is good only in the case of high-degree locally tree-like models.

Extending the subcriticality condition to the non-normalized case involves some standard ideas, useful in other contexts too.

Probability and matrix powers

For the transition matrix $\mathbf{P}$ of a discrete-time MC (X_t), we know the interpretation of the entries p_{ij}^t of $\mathbf{P}^t$ as

$$p_{ij}^t = \mathbb{P}_i(X_t = j).$$

For the adjacency matrix $\mathbf{A}$ of a graph, the textbook interpretation of the entries a_{ij}^t of $\mathbf{A}^t$ is

$$a_{ij}^t = \text{ number of length-}t \text{ walks from } i \text{ to } j.$$

To bring these ideas together, consider first a "deterministic branching walk". Initially there is one particle at i. For each particle at state j at time t, it is replaced at time $t+1$ with one particle at each neighbor of j. Writing

$$Z_{ij}(t) = \text{number of particles at } j \text{ at time } t$$

we have

$$a_{ij}^t = Z_{ij}(t).$$

Changing to a "random branching walk" where instead the replacement is a random (mean 1) number of particles at each neighbor of j, then

$$a_{ij}^t = \mathbb{E}Z_{ij}(t).$$

This line of thought quickly leads to the following interpretation of the powers $\mathbf{W}^t$ of an arbitrary non-negative $n \times n$ matrix. Take a **branching Markov chain**, in which each particle at state j at time t is replaced at time $t+1$ with a random (mean w_{jk}) number of particles at each neighbor k of j. Then, for $Z_{ij}(t)$ as before,

$$w_{ij}^t = \mathbb{E}Z_{ij}(t).$$

Perron-Frobenius theory says that, under the usual "irreducible aperiodic" condition on $\mathbf{W}$ familiar from MC theory,

$$w_{ij}^t \sim \alpha_i \beta_j \theta^t$$

where θ is the eigenvalue of $\mathbf{W}$ of largest modulus (and θ is real) and α and β are the associated eigenvectors, normalized as $\sum_i \alpha_i \beta_i = 1$.

One can apply this fact directly to a discrete-time Reed-Frost type of SIR epidemic over a general geometry, to deduce that the epidemic is subcritical when $\theta < 1$ – see e.g. section 8.2 of Draif - Massoulié. We will instead write out the continuous-time analog, that is our FMIE model of a SIS epidemic.

Recall we are studying the following SIS model over a general geomtery $\mathcal{N} = (\nu_{ij})$.

- Initially one agent is Infected, others are Susceptible.
- When an Infected agent meets a Susceptible agent, the latter becomes infected.
- An Infected agent becomes Susceptible at rate λ.

We couple this to the particle process (branching MC) in which particles have Exponential(λ) lifetimes, and a particle at i has a child at j at rate ν_{ij}. The coupling maintains the relationship

- if i is infective at t in the SIS process then there is at least one particle at i at t in the particle process.

As before, we consider, for the branching MC,

$$Z_{ij}(t) = \text{number of particles at } j \text{ at time } t$$

from one initial particle at i.

Fact: The matrix $\mathbb{E}Z_{ij}(t)$ is the matrix $e^{-\lambda t} \exp(\mathcal{N}t)$.

- Consider a line of descent $i = i_0, i_1, \ldots, i_m = j$ with births at times $0 < t_1 < t_2 < \ldots < t_m < t$.
- The chance of this particular line is $\nu_{i_0,i_1}\nu_{i_1,i_2}\ldots\nu_{i_{m-1},i_m}\, dt_1 \ldots dt_m \times e^{-\lambda t}$.
- Summing over lines gives $(\mathcal{N}^m)_{ij}\, dt_1 \ldots dt_m \times e^{-\lambda t}$
- and integrating over times (t_i) gives $e^{-\lambda t}(\mathcal{N}^m)_{ij}\, t^m/m!$
- so sum over m.

Of course, this is just the continuous-time analog of matrix powers.

We now can almost formalize the following assertion. Write $\theta(\mathcal{N})$ for the Perron-Frobenius eigenvalue of $\mathcal{N}$. [board: P-F extends]. Consider a sequence with $\theta(\mathcal{N}_n) \to \theta \in (0, \infty)$.

- For $\lambda > \theta$ the SIS process is subcritical.

(minor) theory project: cleanest formalization?
[Discuss on board; and see e.g. section 8.3.1 of Draif - Massoulié.]

Diffusion of innovation [coordination game] on random graphs

This lecture follows Lelarge (2011), which formalizes work of Watts (2002). Topic is mathematically a variant of SI epidemics. The setting is discrete time and a graph G of agents (later we'll ponder the FMIE setting). Parameter $0 < q < 1$. Each agent at each time chooses action A ("old") or B ("new"). Each edge (i,j) gives payoffs, to each agent,

$\quad\quad$ 0 if agents choose opposite actions
$\quad\quad$ q if both i and j choose A
$\quad\quad$ $1 - q$ if both i and j choose B.

If i knows that N_i^A neighbors will play A and N_i^B neighbors will play B then

$$\text{(payoff from A) - (payoff from B)} = q(d_i - N_i^B) - (1 - q)N_i^B$$

where $d_i =$ number of neighbors of i.

As default we assume agents use "best response dynamics", the myopic strategy of making the choice at time $t + 1$ that would be optimal if neighbors make same choice at $t + 1$ as they did at t. Explicitly,

$$i \text{ chooses B iff } N_i^B > qd_i.$$

So these are deterministic dynamics. Obviously "all A" and "all B" are stable configurations.

We first consider a "single seed model". Start with all agents except v choosing A; agent v chooses B and is forced to choose B forever. Check inductively the process is "monotone" – only changes are $A \to B$ actually happen. So there is a final configuration – write $C(v, q)$ for the set of agents choosing B in the final configuration.

We study this in the random graph model.

Random graphs with prescribed degree distributions

Specify (d_i). Can define models $\mathcal{G}_n$ of n-vertex graph, interpretable as being "random" subject to the following constraint. Write D_n for degree of a uniform random vertex of $\mathcal{G}_n$, then

$$D_n \xrightarrow{d} D \text{ where } \mathbb{P}(D = i) = d_i.$$

Such models have the following "local GWBP approximation". The structure of $\mathcal{G}_n$ within some fixed graph-distance r from a uniform random vertex U_n converges in distribution, as $n \to \infty$, to the random tree comprising generations 0 to r of the following modified Galton-Watson BP. The root has offspring distribution D; in subsequent generation the offspring distribution is the size-biased distribution D^* where $\mathbb{P}(D^* = i) = (i+1)d_{i+1}/\mathbb{E}D$.

Assuming $\mathbb{E}D^{2+\varepsilon} < \infty$ then $\mathbb{E}(D^*)^{1+\varepsilon} < \infty$ and the Kesten-Stigum theorem says that the size Y_r of generation r grows at a particular rate: $Y_r/(\mathbb{E}D^*)^r \to W$ a.s. and L^1.

Define
$$q_{\text{crit}} = \sup\{q : \mathbb{E}[D(D-1)1(D < q^{-1}] > \mathbb{E}[D]\}.$$

Write C^n for the final B-set, from a random initial single seed. Write Π^n for the set of "pivotal" agents, defined as the (largest component of) the subgraph on agents of degree $< q^{-1}$.

> **Proposition**
>
> *(i) For $q > q_{crit}$ we have $n^{-1}|C^n| \to_p 0$.*
> *(ii) For $q < q_{crit}$ we have $n^{-1}|\Pi^n| \to_p \gamma(q) > 0$. And for $v \in \Pi^n$ we have $n^{-1}|C^n(v)| \geq_p (1 - o(1))s(q)$.*
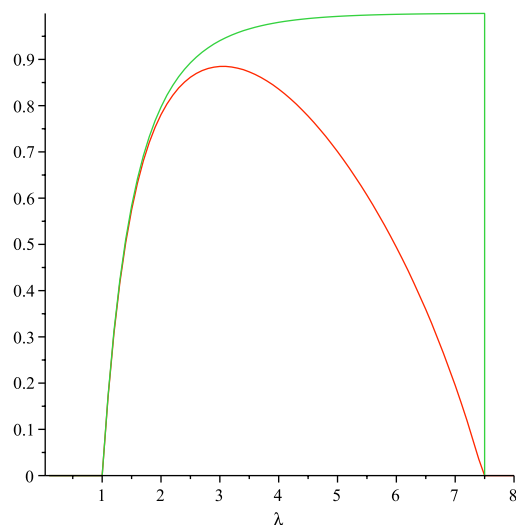
Here $\gamma(q)$ and $s(q)$ depend on q and (p_i).

xxx explain [board]

Now imagine fixing q and varying a parameter in the random graph model. Consider the Erdos-Renyi random graph $\mathcal{G}(n, \lambda/n)$, that is (d_i) is Poisson(λ). Here is a figure of

$s(q, \lambda) = $ proportional size of B-set

$\gamma(q, \lambda) = $ proportional size of pivotal set

for $q = 0.15$.   xxx discuss [board]

This is typical; for small q, as we vary a parameter λ in the random graph model, there is some interval $[\lambda_i(q), \lambda_s(q)]$ in which the **cascade** has size $\Omega(n)$.

Little difference between starting with 1 seed or $O(1)$ seeds. What about starting with a small proportion of agents as seeds? Fix a measure (α_d) where

$\alpha_d = $ proportion of degree-d agents who are initially B.

Introduce a more general model. Fix a probability π (previous model was the $\pi = 1$ case). For agent i, each time a new neighbor adopts B, agent i increments their "enthusiasm" by 1 with probability π; agent i will adopt B when their enthusiasm reaches qd_i. (Note this model looks more like FMIE.)

Mathematical point: the kind of heuristic GWBP-approximation argument (which we used before for SIR epidemics) continues to work in more elaborate settings like this. One gets an explicit formula for the proportional size of cascade.
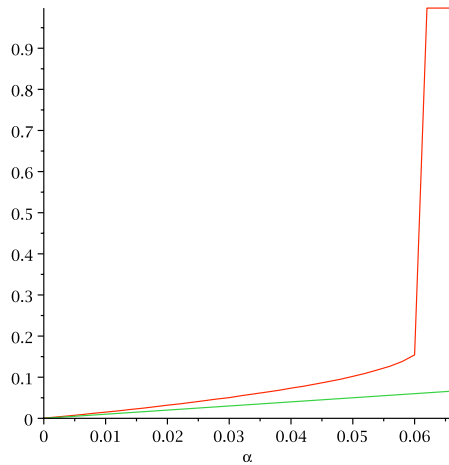
Lower line is size of initial seed.

To quote Lelarge (2011), from which figures and some other text are taken:

If the seed is too small, then each early adopter starts a small cascade which is stopped by high-degree nodes. When the seed reaches the critical mass, then the cascades coalesce and are able to overcome barriers constituted by high-degree nodes so that a large fraction of the nodes in the giant component of the graph adopt. Then increasing the size of the seed has very little effect since most of the time, the new early adopters (chosen at random by the firm) are already reached by the global diffusion.

$\sigma_i = 1(i \text{ initially active})$
$Y_\ell = 1(\ell \text{ active from above})$
$B_{\ell i}$ independent Bernoulli(π).
State of general i:

$$Y_i = 1 - (1 - \sigma_i) \ 1\left(\sum_{\ell \to i} B_{\ell i} Y_\ell \leq q d_i\right), \tag{3}$$

State of the root:

$$X_\emptyset = 1 - (1 - \sigma_\emptyset) \ 1\left(\sum_{i \to \emptyset} B_{i\emptyset} Y_i \leq q d_\emptyset\right). \tag{4}$$

In order to compute the distribution of $X_\emptyset$, we first solve the Recursive Distributional Equation (RDE) associated to the Y_i's: thanks to the tree structure, the random variables Y_ℓ are i.i.d. and have the same distribution as Y_i. Hence their distribution solves the RDE

$$Y \overset{d}{=} 1 - (1 - \sigma(D^* + 1)) \ 1\left(\sum_{i=1}^{D^*} B_i Y_i \leq q(D^* + 1)\right), \tag{5}$$

$$Y \overset{d}{=} 1 - (1 - \sigma(D^* + 1)) \ \mathbf{1}\left(\sum_{i=1}^{D^*} B_i Y_i \le q(D^* + 1)\right), \qquad (6)$$

where for a given d, the random variable $\sigma(d)$ is Bernoulli with parameter α_d, B_i's are independent Bernoulli with parameter π, D^* has distribution p_r^*, Y and the Y_i are i.i.d. copies (with unknown distribution).

To solve the RDE (6), we need to compute only the mean of the Bernoulli random variable Y. Hence taking expectation in (6) directly gives a fixed-point equation for this mean.

Can we devise an interesting FMIE variant?
Recall: envisage unknown rates ν_{ij}, no analog of degree d_i.

Suppose we are told global proportions of B's at each time.
Agent i has record of past meetings with already-met j, and their states.
Agent i makes some (Bayes?) estimate ν_{ij}^* of such meeting rates, and of meeting rate with new agents.
So get estimate of (current) rates $\nu_i^*(A)$ and $\nu_i^*(B)$ of meeting each type.
Switch when estimated payoff rate $(1 - q)\nu_i^*(B) > q\nu_i^*(A)$.

More elaborate: agents have some different opinions about future growth of number of Bs.