1 Introduction

Why are people still outside playing sports? Why are people still having parties?

Core idea: different kinds of people have different structure to their social connections. The nature of this structure may be important for thinking about whether outbreaks of a contagion break out into epidemics or dissapear. May also be important for reasoning about how people respond to risk induced by social connections.

2 Basic assumptions of how the contagion spreads

Let's start by supposing that you are host to some contagion – a virus, an idea, etc. And suppose that you have a fixed number n of neighbors, and for each neighbor, you have an independent chance r of transmitting some contagion to that neighbor. (For simplicity, none of your neighbors are immune to the contagion.)

What is the expected number of neighbors whom you transmit the contagion to? The answer is of course $n \cdot r$

And what if there are multiple people like you who can transmit the contagion, such that each other such person also has n neighbors with a chance r of transmitting the contagion? Then in expectation, if there are I such people, the group of you will transmit Inr new cases.

And if $nr > 1$, then on average, each of you will transmit to more than one person, and $Inr > I$ so the number of new infectious people has grown.

So far, this resembles the early stages of a standard epidemic model, where the disease is growing exponentially.

2.1 Adding some structure

Next, let's impose some additional structure on the network of conections between people. In principle, we could have contagion spread randomly through any arbitray graph, with individuals as nodes and the edges as social connections along which the disease can spread.

To keep things simple, I'll make the following assumptions about the structure of the social graph: There are some number of distinct 'types' of people, indexed by i. Each type is defined by the number of connections is has to each other type. Let n_{ij} represent the number of connections a type i person has to type j people. Suppose that each of these groups is very large and that the connections are chosen at random, save for the fact about which group they connect.

Let I_i be the number of infectious people of type i. And let Φ_i be the portion of infectious people who are of type i. So $\Phi_i \equiv \frac{I_i}{\sum_k I_k}$.

What is the total number of expected new tranmission events?

$$\sum_k \left[I_k \sum_j r n_{kj} \right]$$

What is the number of new transmission events divided by the number of currently infectious?

$$\frac{\sum_k \left[I_k \sum_j r n_{kj} \right]}{\sum_h I_h} = \sum_k \left[\frac{I_k}{\sum_h I_h} \sum_j r n_{kj} \right] = \sum_k \left[\Phi_k \sum_j r n_{kj} \right]$$

And what is the expected number of new transmission events where the contagion is specifically transmitted to a person of type i?

$$\sum_k \left[I_k r n_{ki} \right]$$

Therefore the portion of the newly infected who are of type i will be

$$\begin{aligned}
\Phi_i' &= \frac{\sum_k \left[I_k r n_{ki}\right]}{\sum_k \left[I_k \sum_j r n_{kj}\right]} \\
&= \frac{\sum_k \left[I_k r n_{ki}\right]}{\sum_k \left[I_k \sum_j r n_{kj}\right]} \cdot \frac{\left(\frac{1}{r\sum_h I_h}\right)}{\left(\frac{1}{r\sum_h I_h}\right)} \\
&= \frac{\sum_k \left[\left(\frac{I_k}{r\sum_h I_h}\right) n_{ki}\right]}{\sum_k \left[\left(\frac{I_k}{r\sum_h I_h}\right) \sum_j r n\right]} \\
&= \frac{\sum_k \left[\Phi_k n_{ki}\right]}{\sum_k \left[\Phi_k \sum_j r n\right]}
\end{aligned}$$

In the later stages of a contagion, some of the neighbors of an infestious individual may be immune to the contagion. But in the early stages, the dynamics are described above.

3 Fixed Point results and other dynamics

Note that, holding the parameters $\eta_{..}$ constant, the system of equations describing each Φ_i' is a mapping from a probability distribution to itself, which doesn't depend on the transmission rate r.

In other words, I'm splitting up two questions about how the contagion spreads: how does the distribution of types amongst the infected evolve?, and does the incidence of the infection grow or shrink?

First let's note a few things about this mapping that describes how the probability distribution evolves. Because it's a continous mapping from a simplex to the same simplex, there must be at least one fixed point. This fixed point is not necessarily unique. If the connecton parameters $\{\eta\}$ are chosen such that $\eta_{ij} \neq 0$ for any $i \neq j$, then the fixed point must be in the interior of the probability. simplex.

In the latter case, subsequent generations of infectious will evolve to have a distribution of types which converges to the fixed distribution.

Then the question of whether the contagion can spread to a large portion of the network depends on the average number of new tranmissions per person among this distribution of infectious people. So if the distribution evolves in such a way as to concentrate the contagion among the highly connected, then an outbreak is more likely to occur than would be suggested by the transmissions caused by a randomly selected person.

[Insert image of hyperplane and tranmission threshold]

Then after the distribution evolves in a way determined by the structure of connections between types, the number of new transmissions per infectious person will be a weighted sum of the average number of transmissions per type of person

$$\sum_k \left[\Phi_k \sum_j r n_{kj}\right]$$

And the contagion will increase in prevalence iff this quantity is > 1.

4 Continuous example with two distinct types.

your percentage Φ_i is equal to $\frac{I_i}{I}$. Thus the rate of change of your percentage $\frac{\partial \Phi_i}{\partial t}$ is

$$\frac{\partial}{\partial t}\frac{I_i}{I} = \frac{\left(\frac{\partial}{\partial t} I_i\right)(I) - (I_i)\left(\frac{\partial}{\partial t} I\right)}{(I^2)}$$

If we assume that the rate of change of a population is proportional to the size of the group it is spreading from, then when $S_i \approx 1$, and that recovery happens randomly at rate γ

$$\frac{\partial}{\partial t} I_i = \sum_j \left[r \eta_{ji} I_j \right] - \gamma I_i$$

Because $I = \sum_k I_k$, we get that

$$\frac{\partial}{\partial t} I = \sum_k \left(\sum_j \left[r \eta_{jk} I_j \right] - \gamma I_k \right) = \sum_k \left(\sum_j \left[r \eta_{jk} I_j \right] \right) - \gamma I$$

And so

$$\frac{\partial}{\partial t} \frac{I_i}{I} = \frac{\left(\sum_j \left[r \eta_{ji} I_j \right] - \gamma I_i \right) (I) - (I_i) \left(\sum_k \left(\sum_j \left[r \eta_{jk} I_j \right] \right) - \gamma I \right)}{(I^2)}$$

$$\begin{aligned}
\frac{\partial}{\partial t} \frac{I_i}{I} &= r \frac{\left(\sum_j \left[\eta_{ji} I_j \right] - \gamma I_i \right) (I)}{(I^2)} - r \frac{I_i \left(\sum_k \left(\sum_j \left[\eta_{jk} I_j \right] \right) - \gamma I \right)}{(I^2)} \\
&= r \frac{\sum_j \left[\eta_{ji} I_j \right] - \gamma I_i}{I} - r \frac{I_i}{I} \frac{\sum_k \left(\sum_j \left[\eta_{jk} I_j \right] \right) - \gamma I}{I} \\
&= r \sum_j \left[\eta_{ji} \frac{I_j}{I} \right] - r \gamma \frac{I_i}{I} - r \frac{I_i}{I} \left[\sum_k \left(\sum_j \left[\eta_{jk} \frac{I_j}{I} \right] \right) - \gamma \right] \\
&= r \sum_j \left[\eta_{ji} \frac{I_j}{I} \right] - \cancel{r \gamma \frac{I_i}{I}} - r \frac{I_i}{I} \sum_k \left(\sum_j \left[\eta_{jk} \frac{I_j}{I} \right] \right) + \cancel{r \frac{I_i}{I} \gamma}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial t} \Phi_i &= r \sum_j \left[\eta_{ji} \Phi_j \right] - r \Phi_i \sum_k \sum_j \left[\eta_{jk} \Phi_j \right] \\
&= r \sum_j \left[\eta_{ji} \Phi_j - \Phi_i \sum_k \left[\eta_{jk} \Phi_j \right] \right] \\
&= r \sum_j \left[\Phi_j \left[\eta_{ji} - \Phi_i \sum_k \eta_{jk} \right] \right] \\
&= r \left[\sum_j \eta_{ji} \left[\Phi_j - \Phi_j \Phi_i \right] - \Phi_i \sum_j \sum_{k \neq i} \Phi_j \eta_{jk} \right] \\
&= r \sum_j \left[\Phi_j \left[(1 - \Phi_i) \eta_{ji} - \Phi_i \sum_{k \neq i} \eta_{jk} \right] \right]
\end{aligned}$$

(Weird how the recovery rate cancels out. It's still implicitly there in the form of the rate of transmission r. If β chance of transmission each period and $\frac{1}{\gamma}$ periods of infectiveness on average then overall chance of transmission is only $\frac{\beta}{\gamma}$. So recovery rate only matters for the evolution of distribution insomuch as it affects transmission rate, and tranmission rate only effects speed of adjustment.)

For example, with only two groups, we have that

$$\begin{aligned}\frac{\partial}{\partial t}\Phi_1 &= r\sum_{j=1,2}\left[\Phi_j\left[\eta_{j1}-\Phi_1\sum_{k=1,2}\eta_{jk}\right]\right]\\ &= r\left[\Phi_1\left[\eta_{11}-\Phi_1\left[\eta_{11}+\eta_{12}\right]\right]+\Phi_2\left[\eta_{21}-\Phi_1\left[\eta_{21}+\eta_{22}\right]\right]\right]\\ &= r\left[\Phi_1\eta_{11}-\Phi_1\Phi_1\eta_{11}-\Phi_1\Phi_1\eta_{12}+\Phi_2\eta_{21}-\Phi_2\Phi_1\eta_{21}-\Phi_2\Phi_1\eta_{22}\right]\\ &= r\left[\eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[\Phi_2-\Phi_2\Phi_1\right]-\eta_{22}\Phi_2\Phi_1\right]\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial t}\Phi_2 &= r\sum_{j=1,2}\left[\Phi_j\left[\eta_{j2}-\Phi_2\sum_{k=1,2}\eta_{jk}\right]\right]\\ &= r\left[-\eta_{11}\Phi_1\Phi_2+\eta_{12}\left[\Phi_1-\Phi_1\Phi_2\right]+\eta_{21}\Phi_2\Phi_2+\eta_{22}\left[\Phi_2-\Phi_2\Phi_2\right]\right]\end{aligned}$$

Note that

$$\frac{\partial}{\partial t}\Phi_1 = r\left[\underbrace{\eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]}_{positive}\underbrace{-\eta_{12}\Phi_1\Phi_1}_{negative}\underbrace{+\eta_{21}\left[\Phi_2-\Phi_2\Phi_1\right]}_{positive}\underbrace{-\eta_{22}\Phi_2\Phi_1}_{negative}\right]$$

Also note that stationary point doesn't depend on r, so r only affects the speed at which convergence happens. Then it also matters for whether the outbreak is above one of course.

Oh wait, forgot the healing factor. People recover. They don't just keep on infecting.

Also, with only two groups, we have that $\Phi_2 = 1-\Phi_1$. Thus

$$\begin{aligned}\frac{\partial}{\partial t}\Phi_1 &= r\left[\eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[(1-\Phi_1)-(1-\Phi_1)\Phi_1\right]-\eta_{22}\left(1-\Phi_1\right)\Phi_1\right]\\ &= r\left[\eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[1-2\Phi_1+\Phi_1\Phi_1\right]-\eta_{22}\left(\Phi_1-\Phi_1\Phi_1\right)\right]\\ &= r\left[\eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[1-\Phi_1\right]^2-\eta_{22}\left(1-\Phi_1\right)\Phi_1\right]\\ &= r\left[\left(\eta_{11}-\eta_{22}\right)\left(1-\Phi_1\right)\Phi_1-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[1-\Phi_1\right]^2\right]\end{aligned}$$

If we are at the stationary point, then $\frac{\partial}{\partial t}\Phi_1 = 0$ so

$$0 = \eta_{11}\left[\Phi_1-\Phi_1\Phi_1\right]-\eta_{12}\Phi_1\Phi_1+\eta_{21}\left[1-\Phi_1\right]^2-\eta_{22}\left(1-\Phi_1\right)\Phi_1$$

$$\eta_{11}\left(1-\Phi_1\right)\Phi_1+\eta_{21}\left[1-\Phi_1\right]^2 = \eta_{12}\Phi_1\Phi_1+\eta_{22}\left(1-\Phi_1\right)\Phi_1$$

$$\eta_{11}+\eta_{21}\frac{1-\Phi_1}{\Phi_1} = \eta_{12}\frac{\Phi_1}{(1-\Phi_1)}+\eta_{22}$$

Analysis of the two-group scenario:

Note how each of the following chunks responds to changes in Φ_1:

- $+\eta_{11}\left[\Phi_1 - \Phi_1\Phi_1\right]$: Increases until Φ_1 reaches $\frac{1}{2}$ then starts decreasing
- $+\eta_{21}\left[1-\Phi_1\right]^2$: decreasing in Φ_1. Higher portion means slower growth from the other portion.
- $-\eta_{12}\Phi_1\Phi_1$: increase in Φ_1 which means the whole term is decreasing of course. Higher portion means more bleeding off.
- $-\eta_{22}\left(1-\Phi_1\right)\Phi_1$: same as first part. Just negative

So two of the pieces decrease when we increase Φ_1 and the other two pieces add up to a parabola with the peak at 0.5.

At $\Phi_1 = 0$,

$$\frac{\partial}{\partial t}\Phi_1 = r\eta_{21}$$

At $\Phi_1 = 1$,

$$\frac{\partial}{\partial t}\Phi_1 = -r\eta_{12}$$

Therefore if $\eta_{21} = 0$, we have a steady state at $\Phi_1 = 0$ and if $\eta_{21} = 0$, then we have $\Phi_1 = 1$ as stable state.

No self propagation:

If $\eta_{11} = \eta_{22}$, then

$$\begin{aligned}\frac{\partial}{\partial t}\Phi_1 &= r\left[\eta_{21}\left[1-\Phi_1\right]^2 - \eta_{12}\Phi_1\Phi_1\right] \\ &= r\left[\eta_{21} - 2\eta_{21}\Phi_1 + \left[\eta_{21} - \eta_{12}\right]\Phi_1\Phi_1\right]\end{aligned}$$

Set equal to zero and solve. Then solutions are:

$$\begin{aligned}\Phi_1 &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{2\eta_{21} \pm \sqrt{4\eta_{21}^2 - 4\left[\eta_{21} - \eta_{12}\right]}}{2\left[\eta_{21} - \eta_{12}\right]} \\ &= \frac{2\eta_{21} \pm 2\sqrt{\eta_{21}\eta_{21} - \eta_{12} + \eta_{12}}}{2\left[\eta_{21} - \eta_{12}\right]}\end{aligned}$$

where $a = [\eta_{21} - \eta_{12}]$, $b = -2\eta_{21}$, and $c = 1$.

Note that there will always be two real solutions here. But is one of them inside of the correct range? Well, yes, as long as $\eta_{21} \neq 0$ and $\eta_{12} \neq 0$, then by IVT there is some value of Φ_1 between 0 and 1 such that $\frac{\partial}{\partial t}\Phi_1 = 0$. See above for how $\frac{\partial}{\partial t}\Phi_1 > 0$ at $\Phi_1 = 0$ and $\frac{\partial}{\partial t}\Phi_1 < 0$ at $\Phi_1 = 1$.

Vertex of parabola is at $\frac{\eta_{21}}{\eta_{21}-\eta_{12}}$. If $\eta_{21} > \eta_{12}$, then $a > 0$ parabola is concave up, and fixed point is at $\Phi_1 = \frac{\eta_{21}-\sqrt{\eta_{21}\eta_{21}-\eta_{12}+\eta_{12}}}{[\eta_{21}-\eta_{12}]}$. Likewise if $\eta_{12} > \eta_{21}$, then parabola is concave down, denominator is negative and fixed point is at $\Phi_1 = \frac{\eta_{21}-\sqrt{\eta_{21}\eta_{21}-\eta_{21}+\eta_{12}}}{[\eta_{21}-\eta_{12}]}$.

So either way, fixed point is at $\Phi_1 = \frac{\eta_{21}-\sqrt{\eta_{21}\eta_{21}-\eta_{21}+\eta_{12}}}{[\eta_{21}-\eta_{12}]}$, which is in $(0,1)$, and any other distribution wil evolve towards this distribution.

5 A note about endogenizing the network connections.

The above sections look at the question of how the contagion evolves on a fixed network of social connections.

But another interesting question is how the network of connections will itself evolve in response to the contagion.

Suppose you are an individual with n social contacts. You are susceptible to the contagion, and you know that all else being equal, some portion R of your contacts will eventually get the contagion and become infectious before the end of the outbreak.

The chance that any one of your particular contacts *does not* transmit the disease to you is given by $(1 - rR)$. Thus the probability that none of your contacts transmit to you over the course of the outbreak is $(1 - rR)^n$ where n is the total number of your neighbors. Thus the total exposure risk over the entire outbreak for you personally is

$$1 - (1 - rR)^n$$

Now consider the marginal benefit of reducing your contacts by one. If catching the contagion will impose a cost of C to you, then the benefit of cutting somebody out from your contacts is

$$C\left[1 - (1 - rR)^n - \left[1 - (1 - rR)^{n-1}\right]\right]$$

$$C\left[(1 - rR)^{n-1} - (1 - rR)^n\right]$$

Note that $(1 - rN)$ is a percentage chance, and n is assumed ≥ 1, so $(1 - rR)^{n-1} > (1 - rR)^n$ and the above benefit is positive.

Note also that the above is decreasing in n.

What does this mean? It means that in comparing two people, one with a high number of contacts, and one with a low number of contacts, the highly connected person may be less willing to give up contacts on the margin.