

## 5.11 Toy models for social networks

There has been a lot of theoretical work using the basic “random graph with prescribed degree distribution  $\mathbf{p}^*$ ” mathematical model (and many variant models) as a background “toy model” for social networks, and then adding another level of “toy modeling” to study one of the questions outlined in section 5.10.2. I will describe three simple instances of the second level but little imagination is required to see innumerable variations. This background model is hardly realistic (instead of forming relationships at random there is invariably some degree of clustering, via geography or occupation or activities etc) but is used for mathematical tractability. Also, if the population under consideration is an entire country, then the only data easy to obtain may be an estimate of  $\mathbf{p}^*$  by sampling.

Bear in mind that our theoretical analysis leans upon the local GWBP approximation, and so assumes the population size  $n$  is large.

### 5.11.1 An epidemic model

The mathematical setup in the previous section can be re-interpreted as an epidemic model. In jargon, a discrete-time S-I-R model, standing for *susceptible - infective - recovered*. Measure time in discrete steps  $t = 0, 1, 2, \dots$  and make the rule:

If an individual  $v$  is infective at time  $t$ , then each susceptible neighbor becomes infective at time  $t + 1$  with probability  $\theta$ , independently for different neighbors; individual  $v$  is recovered at time  $t + 1$  and all following times.

So an individual, initially susceptible, might never get infective, or else gets infected from a neighbor at some random time  $T$  but then recovers at time  $T + 1$  and stays recovered. So the individual progresses through states  $S \rightarrow I \rightarrow R$ . Suppose initially ( $t = 0$ ) one random individual  $J$  is infective. Write  $\mathcal{I}_t$  for the set of infective individuals at time  $t$ .

Now reconsider the random subgraph studied in the previous section, and write  $\mathcal{J}_t$  for the set of vertices at distance (= number edges on shortest path)  $t$  from  $J$ . A little thought reveals

the processes  $(\mathcal{I}_t, t = 0, 1, 2, \dots)$  and  $(\mathcal{J}_t, t = 0, 1, 2, \dots)$  are mathematically the same (xxx same distribution).

In particular, the total number of individuals infected (“epidemic size”) over time equals (has the same distribution as) the size of the component of the subgraph containing  $J$ . Appealing to the results of the previous section:

- if  $\theta\mu < 1$  then the epidemic size is small ( $O(1)$ )
- whereas if  $\theta\mu > 1$  then there is some chance of an epidemic whose size is order  $n$

the latter corresponding to the usual English language meaning of “epidemic”.

xxx if  $1 \ll m \ll n$  infected initially, or succession of arriving infectives, then such an epidemic will occur.

### 5.11.2 Adopting a new product

A second type of process (S-I, in the jargon above) is best viewed as a model for adopting a new idea or product; once adopted, the individual does not change back to unadopted. In the previous kind of model (at each time, there is some chance each friend leads you to adopt), everyone will eventually adopt (assuming the background social network is connected). Instead, let us model adoption as depending on an intrinsic preference for the product and the number of friends who have adopted it already. The model is:

(i) each person  $v$  has a “threshold”  $\xi_v$  (possible values  $0, 1, 2, \dots$ ), modeled as IIDRVs.

(ii) For each individual  $v$ :

if  $\xi_v = 0$  then adopt at time  $t = 0$ ;

otherwise wait until the first time  $t$  (if any) that at least  $\xi_v$  friends have adopted, and then adopt at time  $t + 1$ .

Assuming  $P(\xi = 0) > 0$ , this proportion of the population adopt at time 0, so some non-zero proportion ultimately adopt – what is that proportion?

To analyze this problem, recall that we write  $D^*$  and  $D$  for random variables with distributions  $\mathbf{p}^*$  and  $\mathbf{p}$ . Write

$$q_t(d) = \text{chance an individual with } d \text{ friends adopts by time } t.$$

The key formula, derived below, is

$$q_{t+1}(d) = P(\text{Bin}(d, E q_t(D)) \geq \xi). \quad (5.9)$$

This could be solved numerically, inductively in  $t$ . Analytically it is most convenient to study the *size-biased* mean adoption proportion

$$x_t = E q_t(D)$$

because (5.9) gives

$$x_{t+1} = H(x_t) \text{ where } H(x) = P(\text{Bin}(D, x) \geq \xi). \quad (5.10)$$

Note that what we know about the function  $H : [0, 1] \rightarrow [0, 1]$  is

(i)  $H$  is increasing;

(ii)  $H(0) = P(\xi = 0) > 0$  by assumption;

(iii)  $H(1) = P(D \geq \xi)$ ; let’s assume this is  $< 1$ .

We are free to choose the distributions of  $D$  and  $\xi$  rather arbitrarily, so the function  $H$  is otherwise rather arbitrary.

xxx simplest kind of dynamical system; there are (generically) an odd number of solutions of  $x = H(x)$ , and  $x_t \rightarrow x_\infty =$  the smallest solution.

**Derivation of formula (5.9).** Suppose “you” are a typical individual with  $d$  friends, and you have not adopted by time  $t$ . Then you cannot have affected your friends’ adoption or non-adoption; each friend  $w$  with  $d(w)$  other friends has chance  $q_t(d(w))$  to have adopted. Since the distribution of “other friends” is  $D$ ,

$$P(\text{ friend } w \text{ has adopted } ) = E q_t(D) = p, \text{ say.}$$

So the number of your friends who have adopted has  $\text{Bin}(d, p)$  distribution, and you will adopt at time  $t+1$  (if not before) with chance  $P(\text{Bin}(d, p) \geq \xi_{you})$ . But you are a typical individual with  $d$  friends, so this chance must equal  $q_{t+1}(d)$ .