A STOCHASTIC MODEL FOR THE SPREAD OF A SEXUALLY TRANSMITTED DISEASE WHICH RESULTS IN A SCALE-FREE NETWORK.

William J. Reed*
Department of Mathematics and Statistics
University of Victoria
PO Box 3045
Victoria B.C.
Canada V8W 3P4.
(**e-mail:** reed@math.uvic.ca)

March, 2004; revised May, 2005.

*Research supported by NSERC grant OGP 7252.

Abstract

A stochastic model for the spread of a sexually transmitted disease (STD) is presented. To reflect varying degrees of promiscuity among individuals it is assumed that the infectivity of any infected individual is proportional to the number of previous contacts the individual has had with other infected individuals. In both the simple single-sex model and in the more complex two-sex model, the tree graphs of the infection exhibit scale-free network behaviour (*i.e.* power-law behaviour in the upper tail of the degree distribution). The distributions of the size of the infection and of the ring number (distance from the original source of the infection) are determined.

Keywords: SFN; STD; preferential attachment; ring number, Yule tree; Reed-Hughes tree

1 Introduction.

It has been widely recognized, *e.g.* [1, 2], that networks of human sexual partners are scale-free *i.e.* the distribution of the number of contacts over all individuals in the network has a long upper tail which exhibits power-law behaviour. The implications of this for the spread of a sexually transmitted disease (STD), and for possible immmunization strategies have been also been noted [3, 2].

In view of the scale-free nature of networks of sexual contacts, one would expect that the evolving tree network of the spread of an STD (two nodes linked if one infected the other or vice versa) would also exhibit scale-free behaviour, although this has not apparently been verified from data. In this article a simple model for the spread of an STD is presented which yields such behaviour. The essential ingredient of the model, which produces the scale-free behaviour, is the phenomenon of preferential attachment [4] in which an individual who has already infected a large number of others is more likely to infect a new individual than one who has previously infected few or no others.

To demonstrate this two simple single-sex stochastic network models (Yule tree and Reed-Hughes tree) whose properties have been discussed elsewhere [5] are compared in Sec. 2. The way the size of the epidemic grows is similar for both models, as is the behaviour of the distance of an infected individual from the primary source of the infection. However the models differ

in the resulting degree distributions, with the Yule tree model producing a geometric degree distribution (which is not scale-free) and the Reed-Hughes tree model producing a scale-free degree distribution, with exponent 3 (independent of model parameters). The Reed-Hughes tree model has preferential attachment, while the Yule tree model does not.

In Sec.3 the Reed-Hughes tree model, with preferential attachment, is extended to include two sexes, leading to a bipartite graph with a tree structure. For this model it is shown that the degree distributions of males and females both are scale-free with the exponents $2 + \rho/\lambda$ and $2 + \lambda/\rho$, where λ and ρ are parameters reflecting the infectivity of males and females, respectively. These results are shown to be consistent with numerical estimates obtained by Liljeros [1] for a network of Swedish sexual contacts.

2 Single-sex models.

Chan *et al.* [5] considered two models of evolving trees which can be considered as models of the spread of a persistent disease in a large population. In both models infected individuals are represented by nodes with two nodes A and B considered connected if A infected B or *vice versa*. It is assumed that one individual introduced the infection into the community at time $t = 0$ and that in the infinitesimal time increment $(t, t + h)$ any infected individual i can infect a new individual with probability $\lambda_i(t)h + o(h)$ and that all new infections are independent. The two models differ in the assumptions concerning the infection rate functions $\lambda_i(t)$.

(i) Yule tree. In this model it is assumed that $\lambda_i(t) \equiv \lambda$ (constant) for all i and t so that all individuals are equally likely to spread the disease at all times. The number of infected individuals $N(t)$ follows a *Yule process* (homogeneous birth process) (see *e.g.* [6]) and so it follows that $N(t)$ is geometrically distributed with

$$\mathrm{E}(N(t)) = e^{\lambda t}, \qquad \mathrm{var}(N(t)) = e^{\lambda t}(e^{\lambda t} - 1)$$

Furthermore Chan *et al.* [5] obtain an expression for the distribution of the degree (connectivity), $K(t)$ of a randomly selected node at time t and show that as $t \to \infty$

$$p_k(t) = P(K(t) = k) \to 2^{-k}, \quad \text{for } k = 1, 2, \ldots \tag{1}$$

Thus at a suitable large time after the introduction of the infection, the distribution of the number of people infected by infected individuals follows a geometric distribution with parameter $1/2$ on $1, 2, \ldots$.

Chan *et al.* [5] also consider the distribution of the *ring number*, $R(t)$ (number of links back to the primal infection) over the whole tree, and show that R_t has a Poisson-like[1] distribution

$$P(R(t) = r) = \frac{(\lambda t)^{r+1}}{(e^{\lambda t} - 1)(r+1)!} \quad r = 1, \ldots \tag{2}$$

which has mean value $\lambda t(1 - e^{-\lambda t})^{-1} - 1 \sim \lambda t$ (as $t \to \infty$).

While this model may be suitable for some persistent diseases (*e.g.* oral herpes), it is probably not suitable for most STDs. Studies on networks

[1] In fact $R_t + 1$ has a zero-truncated Poisson distribution.

of sexual partners (*e.g.* [1]) seem to suggest that the degree (number of partners) distribution follows close to a power-law (scale-free) distribution. One would expect then that in a tree representing the spread of an STD, the degree distribution would similarly follow close to a power-law, rather than the geometric distribution of the Yule tree. A model which relaxes the homogeneous infectivity assumption and which yields such a distribution is

(ii) Reed-Hughes tree. In this model the infection rate functions $\lambda_i(t)$ are modelled as

$$\lambda_i(t) = \mu K_i(t)$$

where $K_i(t)$ is the degree of node i at time t. Thus individuals who have already infected many others are assumed to be more likely to infect new individuals than those who have infected few or no others. This assumption is meant to reflect the fact that some individuals are more promiscuous than others. The fact that an individual has already infected a large number of others is an indicator of that individual's promiscuity, and therefore it seems reasonable to assume a higher probability for further infections for such an individual. It would probably be more realistic to assume a distribution of the infectivity parameter λ_i over the population (*i.e.* λ_i varying with the i), or failing that, that $\lambda_i(t)$ depends on the previous number of sexual partners of individual i by time t. However neither of these assumptions lead to a model, which can be easily analyzed. In contrast the assumption $\lambda_i(t) = \mu K_i(t)$ does lead to a analytically tractable model and although it is a simplifiaction it does capture something of the essence of the variation in promiscuity over

the population.

Chan *et al.* [5] show that for this model with the clock started ($t = 0$) when a second person is infected ($N(0) = 2$), the number of individuals, $N(t)$, infected by time t follows a geometric distribution (on $2, 3, \ldots$) with

$$\mathrm{E}(N(t)) = 1 + e^{2\mu t}, \qquad \mathrm{var}(N(t)) = e^{2\mu t}(e^{2\mu t} - 1).$$

This is very similar to that for the Yule tree. However the degree distribution of a randomly selected node, asymptotically follows a Yule distribution [8, p. 276].

$$p_k(t) = P(K(t) = k) \to \frac{4}{k(k+1)(k+2)} \sim \frac{4}{k^3}, \quad \text{for } k = 1, 2, \ldots \tag{3}$$

So for this model, after the infection has spread within the community, the number of individuals infected by any one individual, in the upper-tail, follows a power-law (or scale-free) distribution with exponent -3. Note that this distribution does not have a finite variance, which has important implications concerning the epidemic threshold [7, 9, 2].

Chan *et al.* [5] use mean-field methods (using a deterministic approximation to the stochastic model in which probabilities are replaced by proportions) to obtain the following approximation to the distribution of the ring number

$$P(R(t) = r) \approx \frac{e^{\mu t}(\mu t)^{r-1}}{(e^{2\mu t} + 1)(r-1)!} \quad r = 1, 2, \ldots \tag{4}$$

with $P(R(t) = 0) \approx 1/(e^{2\mu t} + 1)$. By comparing this distribution with the results of simulations they show that the approximation is quite good. They

also obtain an approximate expression for the ring number distribution in terms of the size of the network.

The distribution (4) is similar to a Poisson distribution and has mean value $(\mu t+1)(1+e^{-2\mu t})^{-1} \sim \mu t$ (as $t \to \infty$). (Note that for this model, since it is assumed two individuals are infected at $t = 0$, the ring number is the distance to either of these two individuals).

The two models yield similar results concerning the growth of the infection in the community (exponential in expectation, with a geometric distribution for the number infected) and for the distribution of the ring number (both Poisson-like and asymptotically linear in t). However they differ considerably in the degree distribution – the Yule tree model yields a geometric distribution with parameter $1/2$, while the Reed-Hughes tree model yields a distribution following a power-law with exponent -3 in the upper tail. The second (Reed-Hughes) tree model seems more appropriate for an STD in that it recognizes non-homogeneity in infectivities and that it produces a scale-free network in agreement with the empirical results of Liljeros *et al.* [1] on a network of sexual partners. Indeed the predicted power-law exponent of 3 is quite close to the empirical values determined by Liljeros *et al.* for the Swedish network (3.1 ± 0.3 for females and 2.6 ± 0.3 for males). To obtain different distributions for males and females, one needs a two-sex model. Such a model, analogous to the Reed-Hughes tree model above, is developed in the next section.

3 A two-sex model.

Again consider an evolving tree with two individuals of opposite sex, A and B, considered connected if A infected B or *vice versa*. If homosexual infections are ignored, the tree will form a bi-partite graph, with the two parts corresponding to infected males and infected females. Let

$M(t)$ denote number of males infected by time t;

$N(t)$ denote number of females infected by time t;

$L(t)$ denote the number of links between males and females time t; and

$K(t)$ denote the degree of a specified infected male * at time t;

Assume now that in the infinitesimal interval $(t, t+h]$ a male i with degree k_i can infect a new female with probability $\lambda k_i h + o(h)$; and that a female j with degree f_j can infect a new male with probability $\rho f_j h + o(h)$ and that all infections are independent. Under these assumptions the following states of (M, N, L, K) at t could lead to state $(M = m, N = n, L = l, K = k)$ at $t + h$, with the given probabilities:

prior state	event	probability
$(m-1, n, l-1, k)$	female infects male	$\rho(l-1)h + o(h)$
$(m, n-1, l-1, k)$	male other than * infects female	$\lambda(l-1-k)h + o(h)$
$(m, n-1, l-1, k-1)$	male * infects female	$\lambda(k-1)h + o(h)$
(m, n, l, k)	no infection	$1-(\lambda+\rho)lh + o(h)$

Thus letting

$$p_{m,n,l,k}(t) = \mathrm{P}(M(t) = m, N(t) = n, L(t) = l, K(t) = k)$$

and using the Law of Total Probability one obtains

$$\begin{aligned} p_{m,n,l,k}(t+h) &= \rho(l-1)hp_{m-1,n,l-1,k}(t) + \lambda(l-1-k)hp_{m,n-1,l-1,k}(t) \\ &\quad +\lambda(k-1)hp_{m,n-1,l-1,k-1}(t) + [1-(\lambda+\rho)lh]p_{m,n,l,k}(t) + o(h) \end{aligned} \tag{5}$$

On subtracting $p_{m,n,l,k}(t)$ from both sides, dividing by h and letting $h \to 0$ one obtains the following Kolmogorov forwards equation

$$\begin{aligned} \frac{d}{dt}p_{m,n,l,k} &= \rho(l-1)p_{m-1,n,l-1,k} + \lambda(l-1-k)p_{m,n-1,l-1,k} \\ &\quad +\lambda(k-1)p_{m,n-1,l-1,k-1} - (\lambda+\rho)lp_{m,n,l,k} \end{aligned} \tag{6}$$

On summing out k and n one obtains the following differential equation for $p_{m,l} = \mathrm{P}(M(t) = m, L(t) = l)$

$$\frac{d}{dt}p_{m,l} = \rho(l-1)p_{m-1,l-1} + \lambda(l-1)p_{m,l-1} - (\lambda+\rho)lp_{m,l}. \tag{7}$$

Also on summing out m, n and l in (6) one obtains the following equation for $p_k(t) = \mathrm{P}(K(t) = k)$

$$\frac{d}{dt}p_k = \lambda(k-1)p_{k-1} - \lambda k p_k \tag{8}$$

This last equation is easily recognized as the Kolmogorov equation of Yule process [6]. If t^* is the time of infection of individual *, then $K(t^*) = 1$ and

from well-known results on the Yule process the solution to (8) with this initial condition is, for $t \geq t^*$

$$p_k(t) = e^{-\lambda(t-t^*)}(1 - e^{-\lambda(t-t^*)})^{k-1}, \qquad \text{for } k = 1, 2, \ldots \tag{9}$$

i.e $K(t)$ follows a geometric distribution with parameter $e^{\lambda(t-t^*)}$.

To determine the degree distribution over the whole network, one needs to integrate this with respect to the distribution of t^*, the time of infection of any individual. We will return to this later, but first we consider how the size of the epidemic evolves. In the Appendix it is shown that, assuming that the time origin ($t = 0$) is set at the time when the original infective infects a person of the opposite sex (so that $M(0) = 1, N(0) = 1$), then $M(t)$ and $N(t)$ are both geometrically distributed with probability mass functions (pmfs)

$$p_m(t) = \mathrm{P}(M(t) = m) = \phi_m(t)(1 - \phi_m(t))^{m-1}, \qquad m = 1, 2, \ldots \tag{10}$$

and

$$p_n(t) = \mathrm{P}(N(t) = n) = \phi_n(t)(1 - \phi_n(t))^{n-1}, \qquad n = 1, 2, \ldots \tag{11}$$

where

$$\phi_m(t) = \frac{(\lambda + \rho)e^{-(\lambda+\rho)t}}{\rho + \lambda e^{-(\lambda+\rho)t}}, \qquad \phi_n(t) = \frac{(\lambda + \rho)e^{-(\lambda+\rho)t}}{\lambda + \rho e^{-(\lambda+\rho)t}}. \tag{12}$$

The total number of infectives $T(t) = M(t) + N(t)$ is thus the sum of two geometric random variables, with expected value

$$\mathrm{E}[T(t)] = \mathrm{E}[M(t)] + \mathrm{E}[N(t)] = \frac{1}{\phi_m(t)} + \frac{1}{\phi_n(t)} = e^{(\lambda+\rho)t} + 1. \tag{13}$$

Thus in expectation, the epidemic grows (almost) exponentially at rate $\lambda+\rho$. These results parallel those for the single-sex Reed-Hughes tree model of the previous section.

To determine the distribution of the times t^* at which males are infected the notion of an *order-statistic process* [10] can be employed. Such processes are point processes with the property that the joint distribution of the ordered event times in say $[0,T]$ have the same distribution as the order statistics of independent identically distributed random variables (iid rvs) with support on $[0,T]$. The best known example is the Poisson process whose ordered event times on $[0,T]$ have the same joint distribution as the order statistics of iid rvs uniformly distributed on $[0,T]$. Puri [11] provides necessary and sufficient conditions for a birth process to be an order-statistic process. The conditions are in terms of the birth rate function. For the process $\{M(t)\}$ (whose events are new infections of males) this is

$$\begin{aligned}\theta_m(t) &= \lim_{h\to 0}[\mathrm{P}(M(t+h)=m+1|M(t)=m)/h] \\ &= \lim_{h\to 0}\left[\sum_{l=1}^{\infty}\mathrm{P}\left(M(t+h)=m+1|M(t)=m,L(t)=l\right)\ \mathrm{P}(L(t)=l|M(t)=m)/h\right] \\ &= \rho\ \mathrm{E}[L(t)|M(t)=m]. \end{aligned} \tag{14}$$

In the Appendix it is shown that the distribution of $L(t)$ conditional on $M(t)=m$ is negative binomial with pmf

$$P(L(t)=l|M(t)=m)=\binom{l-1}{m-1}\pi(t)^m[1-\pi(t)]^{l-m}, \qquad l=m,m+1,\ldots \tag{15}$$

where

$$\pi(t) = \frac{\rho + \lambda e^{-(\lambda+\rho)t}}{\lambda + \rho}$$

so that

$$\theta_m(t) = \rho E[L(t)|M(t) = m] = \frac{\rho m}{\pi(t)} = \frac{\rho m(\lambda + \rho)}{\rho + \lambda e^{-(\lambda+\rho)t)}}. \tag{16}$$

This is easily confirmed to satisfy Puri's [11] necessary and sufficient conditions (with Puri's $L(i) = \Gamma(i)$ and $h(t) = \rho e^{(\lambda+\rho)t}$) so that $\{M(t)\}$ is confirmed to be an order statistic process. From this it is easy to show [10, Thm.2] that the joint distribution of the times of new male infections in $(0, t]$ are iid random variables with probability density function (pdf)

$$f(t^*) = \frac{(\lambda + \rho)e^{-(\lambda+\rho)(t-t^*)}}{1 - e^{-(\lambda+\rho)t}}, \quad 0 < t^* \leq t. \tag{17}$$

The distribution of infection times over *all* nodes in the network is obtained by a mixture of this distribution with an atomic distribution at 0 (corresponding to the original male infective) with mixing weights $\alpha_1 = 1 - 1/\mathrm{E}(M(t))$ and $\alpha_0 = 1/\mathrm{E}(M(t))$. Integrating the pmf (9) of the degree of node * with repect to this mixed distribution yields the pmf $\tilde{p}_k$ of the degree disribution of a randomly selected node in the network at time t

$$\tilde{p}_k = \alpha_0 e^{-\lambda(t)}(1-e^{-\lambda(t)})^{k-1} + \alpha_1 \int_0^t e^{-\lambda(t-t^*)}(1-e^{-\lambda(t-t^*)})^{k-1} \frac{(\lambda + \rho)e^{-(\lambda+\rho)(t-t^*)}}{1 - e^{-(\lambda+\rho)t}} dt^*$$

Using the fact that $M(t)$ is geometrically distributed it follows that

$$\alpha_0 = \mathrm{E}\left(\frac{1}{M(t)}\right) = -\frac{\phi_m(t)}{1 - \phi_m(t)} \log \phi_m(t).$$

As $t \to \infty$, $\alpha_0 \to 0$ and $\alpha_1 \to 1$; hence letting $\tau = t - t^*$ it follows that

$$\begin{aligned} \tilde{p}_k &\to \int_0^\infty e^{-\lambda\tau}(1-e^{-\lambda\tau})^{k-1}(\lambda+\rho)e^{-(\lambda+\rho)\tau}d\tau \\ &= \left(\frac{\lambda+\rho}{\lambda}\right)\frac{\Gamma(k)\Gamma(\rho/\lambda+2)}{\Gamma(k+2+\rho/\lambda)}, \end{aligned} \tag{18}$$

a Yule distribution (Johnson *et al.*, 1993, p.276) for which

$$\tilde{p}_k \sim \left[\frac{(\lambda+\rho)\Gamma(2+\rho/\lambda)}{\lambda}\right] k^{-(2+\rho/\lambda)}, \qquad (k \to \infty). \tag{19}$$

So at a suitably long time after the introduction of the infection the degree distribution of males in the network will be scale-free with exponent $2+\rho/\lambda$. In an exactly similar way one can establish that the degree distribution of females in the network will be scale-free with parameter $2+\lambda/\rho$. Typically one would expect $\lambda > \rho$ (males more promiscuous on average than females) which would lead to a smaller power-law exponent (*i.e* a degree distribution with a longer tail) for males than females. This is in agreement with the results of Liljeros [1] in their study of a network of sexual partners. They obtained 95% confidence intervals of 2.6 ± 0.3 for males and 3.1 ± 0.3 for females for the power-law exponents. Values of the ratio λ/ρ in the range 1.11 – 1.40 yield exponents $2+\rho/\lambda$ and $2+\lambda/\rho$ lying within these two confidence intervals.

It is possible to use mean-field theory to obtain an expression for the generating function of the approximate ring number distribution, using methods analogous to those of Chan *et al.* [5]. However it is very complicated and no simple closed-form expressions for the corresponding approximate probabilities are available. Nonetheless it is possible to use the generating function

to obtain numerically the approximate ring number probabilities, for given parameter values, by expanding the generating function using a symbolic mathematics software such as Maple IX[12]. The derivation of the generating function for the mean-field approximation to ring number distribution (for males and females) is outlined in the Appendix.

Fig. 1 shows the mean-field approximation to the ring number distribution for males and females when $\lambda = 0.2, \rho = 0.1$ and $t = 10$. Numerical results using the mean field approximation suggest that, as in the single sex mode, the expected ring number asymptotically grows linearly with time (both for males and females).

4 Conclusions.

The main results of this article concern a simple two-sex stochastic model for the spread of an STD. To reflect varying degrees of promiscuity among individuals the model assumes that the infectivity of an infected individual is proportional to the number of previous contacts the individual has had with other infected individuals. The proportionality coefficients are assumed different for males and females. The main results are that the epidemic grows exponentially in expectation, and that in the evolving bipartite tree graph of the epidemic the degree distributions in each part of the bipartite graph evolve towards ones with power-law (scale-free) tails, with different power-law exponents. Also a mean field approximation the distribution of the ring number (distance from the initial infective) is determined and numerical

results suggest that asymptotically the mean ring number grows linearly with time. The resulting scale-free network is consistent both qualitativley and quantitatively with the empirical results established for a network of Swedish sexual partners [1].

One limitation of the model is that it allows only heterosexual infection. A more realistic model would include the possibility of homosexual infection. While it is possible in principle to extend the model and methods to include, for example, four categories – heterosexual, homosexual and bisexual males and heterosexual females – the resulting process would have many state variables and the analysis become much more complicated. One might expect power-law degree distributions to occur in such a model. What would be of most interest, but remains to be determined, is the way the power-law exponents would depend on the infectivity parameters of each of the four groups.

References

[1] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley and Y. Åberg, The web of human sexual contacts. Nature, 411 (2001) 907-908.

[2] H. J. Jones and M. S. Handcock, Sexual contacts and epidemic thresholds. Nature, 423 (2003) 605–606.

[3] Pastor-Satorras, R. and A. Vespignani, Epidemic dynamics in finite size scale-free networks. Phys. Rev. E, 65 (2001) 035108

[4] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks. Rev. Modern Phys., 74 (2002) 47–97.

[5] D. Y. C. Chan, B. D. Hughes, A. S. Leong and W. J. Reed, Stochastically evolving networks. Phys. Rev. E, 68 (2003) 066124.

[6] S. Karlin and H. M. Taylor, A First Course in Stochastic Processes, Second Edition. Academic Press, New York, 1975.

[7] Pastor-Satorras, R. and A. Vespignani, Epidemics and immunization in scale-free networks, in S. Bornholdt and H. G. Schuster (Eds.), Handbook of Graphs and Networks: From the Genome to the Internet, Wiley-VCH, Berlin, 2002, pp. 113-132 .

[8] N. L. Johnson, S. Kotz and A. W. Kemp, Univariate Discrete Distributions, Second Edition. John Wiley and Sons, New York, 1993.

[9] R. M. May and A. L. Lloyd, Infection dynamics on scale-free networks. Phys. Rev. E. 64 (2001) 066112.

[10] P. D. Feigin, On the characterization of point processes with the order statistic property. J. Appl. Prob. 16 (1979) 297–304.

[11] P. S. Puri, On the characterization of point processes with the order statistic property without moment condition. J. Appl. Prob. 19 (1982) 39-51. X

[12] Maple IX, Maplesoft, Waterloo Maple Inc., 2003.

[13] R. Haberman, Elementary Partial Differential Equations, Second Edition. Prentice-Hall, Engelwood Cliffs, N.J, 1987.

Appendix

Solving the Kolmogorov equation.

Consider the Kolmogorov equation (7) for $p_{m,l} = \mathrm{P}(M(t) = m, L(t) = l)$

$$\frac{d}{dt}p_{m,l} = \rho(l-1)p_{m-1,l-1} + \lambda(l-1)p_{m,l-1} - (\lambda+\rho)lp_{m,l}. \tag{i}$$

To solve this consider the generating function

$$\Phi(x,y,t) = \sum_m \sum_n \sum_l p_{m,n,l}(t)x^m y^l = \mathrm{E}\left(x^{M(t)}y^{L(t)}\right). \tag{ii}$$

From (7) it follows that Φ satisfies the partial differential equation

$$\Phi_t = y[\lambda y + \rho xy - (\lambda+\rho)]\Phi_y. \tag{iii}$$

This can be solved by the method of characteristics (see *e.g.* [13, p.419]). With initial condition $M(0) = 1, L(0) = 1$ the solution is

$$\Phi(x,y,t) = \frac{xye^{-(\lambda+\rho)t}}{1 - y\left(\frac{\rho x+\lambda}{\rho+\lambda}\right)\left(1-e^{-(\lambda+\rho)t}\right)}. \tag{iv}$$

Setting $y = 1$ yields the generating function of $M(t)$. This is easily recognized as the generating function of a geometrically distributed random variable with parameter

$$\phi_m(t) = \frac{(\lambda+\rho)e^{-(\lambda+\rho)t}}{\rho + \lambda e^{-(\lambda+\rho)t}}.$$

Thus

$$p_m(t) = \mathrm{P}(M(t) = m) = \phi_m(t)(1-\phi_m(t))^{m-1}, \qquad m = 1, 2, \ldots \tag{v}$$

In an exactly similar way one can establish the pmf of $N(t)$ as

$$p_n(t) = \mathrm{P}(N(t) = n) = \phi_n(t)(1 - \phi_n(t))^{n-1}, \qquad n = 1, 2, \ldots \tag{vi}$$

where

$$\phi_n(t) = \frac{(\lambda + \rho)e^{-(\lambda+\rho)t}}{\lambda + \rho e^{-(\lambda+\rho)t}}.$$

Conditional distribution of $L(t)$ given $M(t)$.

To find the conditional distribution of $L(t)|M(t) = m$ one can first find, by expanding (iv), the joint pmf of $M(t), L(t)$

$$p_{m,l} = e^{-(\lambda+\rho)t}\binom{l-1}{m-1}\left(\frac{\rho}{\lambda}\right)^{m-1}\left[\frac{\lambda(1 - e^{-(\lambda+\rho)t})}{\lambda + \rho}\right]^{l-1}. \tag{vii}$$

Dividing this by $p_m(t)$ above one obtains the conditional pmf as

$$\begin{aligned} p_{l|m} &= \binom{l-1}{m-1}\left[\frac{\lambda(1 - e^{-(\lambda+\rho)t})}{\lambda + \rho}\right]^{l-m}\left[\frac{\rho + \lambda e^{-(\lambda+\rho)t}}{\lambda + \rho}\right]^{m} \\ &= \binom{l-1}{m-1}[1 - \pi(t)]^{l-m}\pi(t)^m \end{aligned} \tag{viii}$$

where

$$\pi(t) = \frac{\rho + \lambda e^{-(\lambda+\rho)t}}{\lambda + \rho}.$$

The pmf (viii) can be recognized as that of a negative binomial distribution with index m and parameter $\pi(t)$. It follows from standard results that

$$\mathrm{E}[L(t)|M(t) = m] = \frac{m}{\pi(t)} = \frac{m(\lambda + \rho)}{\rho + \lambda e^{-(\lambda+\rho)t}}$$

The mean-field approximation for the ring number distribution.

The mean-field approximation technique involves replacing the stochastic system with the corresponding deterministic system in which probabilities are replaced by proportions. Thus in the two-sex model of Sec.3, it is asssumed that when $M(t) = m, N(t) = n$ and $L(t) = l$, the number of new females and males infected in the $(t, t+h]$ is $\lambda lh + o(h)$ and $\rho lh + o(h)$ respectively.

To apply this technique to determine an approximation for the ring number distribution, let

$m(k, r; t)$ denote the number of males with degree k and ring number r; and

$n(l, r; t)$ denote the number of females with degree l and ring number r at time t.

Then it follows that

$$\frac{d}{dt}m(k,r;t) = \lambda(k-1)m(k-1,r;t) - \lambda k m(k,r;t) + \delta_{k,1}\sum_{l\geq 1}\rho n(l,r-1;t) \quad \text{(ix)}$$

and

$$\frac{d}{dt}n(l,r;t) = \rho(l-1)n(k-1,r;t) - \rho k l n(l,r;t) + \delta_{l,1}\sum_{k\geq 1}\lambda m(k,r-1;t) \quad \text{(x)}$$

where $\delta_{j,1} = 1$ if $j = 1$ and is zero otherwise. The last terms on r.h.s of both equations correspond to new infectees, while the first terms correspond to the effect on the degree of the infecters.

Now consider the generating functions

$$\mathcal{M}(x,z;t)=\sum_{k=1}^{\infty}\sum_{r=0}^{\infty}m(k,r;t)x^k z^r \quad \text{and} \quad \mathcal{N}(y,z;t)=\sum_{l=1}^{\infty}\sum_{r=0}^{\infty}n(l,r;t)y^l z^r$$

which satisfy the partial differential equations (subscripts denoting partial derivatives)

$$\mathcal{M}_t(x,z;t)=\lambda x(x-1)\mathcal{M}_x(x,z;t)+\rho xz\mathcal{N}_y(1,z;t) \tag{xi}$$

$$\mathcal{N}_t(y,z;t)=\rho y(y-1)\mathcal{N}_y(y,z;t)+\lambda yz\mathcal{M}_x(1,z;t). \tag{xii}$$

Using the method of characteristics and the initial conditions $M(0)=N(0)=K(0)=L(0)=1, R(0)=0$ one can show that

$$\mathcal{M}(x,z;t)=\int_0^t \frac{\rho zx\psi(\tau)d\tau}{x+(1-x)e^{\lambda(t-\tau)}}+\frac{x}{x+(1-x)e^{\lambda\tau}} \tag{xiii}$$

$$\mathcal{N}(y,z;t)=\int_0^t \frac{\lambda zy\phi(\tau)d\tau}{y+(1-y)e^{\rho(t-\tau)}}+\frac{y}{y+(1-y)e^{\rho\tau}} \tag{xiv}$$

where

$$\psi(t)=\mathcal{N}_y(1,z;t) \quad \text{and} \quad \phi(t)=\mathcal{M}_x(1,z;t); \tag{xv}$$

or using (xiii) and (xiv)

$$\psi(t)=\lambda z e^{\rho t}\int_0^t e^{-\rho\tau}\phi(\tau)d\tau+e^{\rho t} \tag{xvi}$$

$$\phi(t)=\rho z e^{\lambda t}\int_0^t e^{-\lambda\tau}\psi(\tau)d\tau+e^{\lambda t}. \tag{xvii}$$

Differentiating this pair of equations leads to the linear system

$$\frac{d}{dt}\begin{pmatrix}\psi\\ \phi\end{pmatrix}=\begin{pmatrix}\rho, & \lambda z\\ \rho z, & \lambda\end{pmatrix}\begin{pmatrix}\psi\\ \phi\end{pmatrix}+\begin{pmatrix}\rho e^{\rho t}\\ \lambda e^{\lambda t}\end{pmatrix} \tag{xviii}$$

with initial condition $(\psi(0), \phi(0))^T = (1, 1)$, since $\mathcal{M}(x, z; 0) = x$ and $\mathcal{N}(y, z; 0) = y$. This can be readily solved:

$$\begin{pmatrix} \psi \\ \phi \end{pmatrix} = e^{\boldsymbol{A}t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \int_0^t e^{\boldsymbol{A}(t-s)} \begin{pmatrix} \rho e^{\rho s} \\ \lambda e^{\lambda s} \end{pmatrix} ds \qquad \text{(xix)}$$

where

$$\boldsymbol{A} = \begin{pmatrix} \rho, & \lambda z \\ \rho z, & \lambda \end{pmatrix}$$

Now the generating functions $\mathcal{R}_m(z)$ and $\mathcal{R}_f(z)$ for the distribution of ring numbers for males and females are given by:

$$\mathcal{R}_m(z) = \frac{\mathcal{M}(1, z; t)}{\mathcal{M}(1, 1; t)} = \frac{\rho z \int_0^t \psi(\tau) d\tau + 1}{\lim_{z \to 1} \left[\rho z \int_0^t \psi(\tau) d\tau + 1 \right]} \qquad \text{(xx)}$$

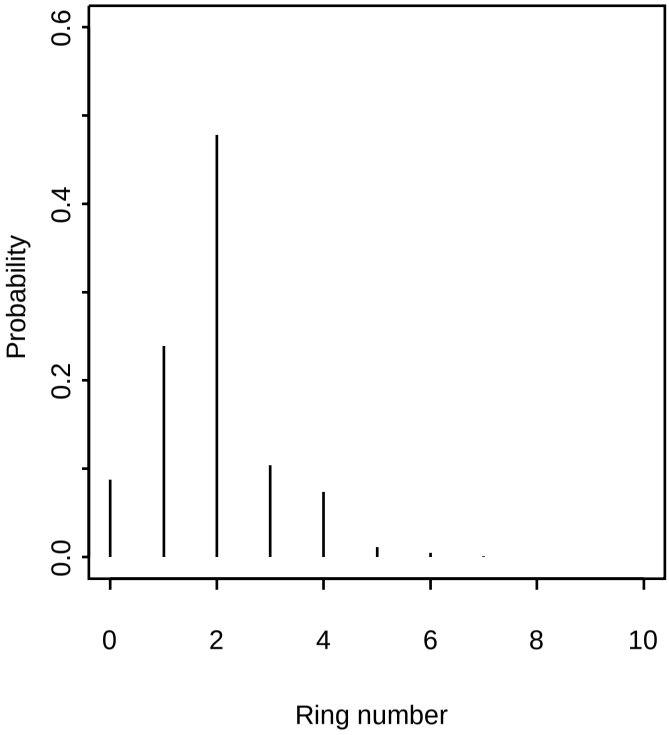
$$\mathcal{R}_f(z) = \frac{\mathcal{N}(1, z; t)}{\mathcal{N}(1, 1; t)} = \frac{\lambda z \int_0^t \phi(\tau) d\tau + 1}{\lim_{z \to 1} \left[\lambda z \int_0^t \phi(\tau) d\tau + 1 \right]} \qquad \text{(xxi)}$$

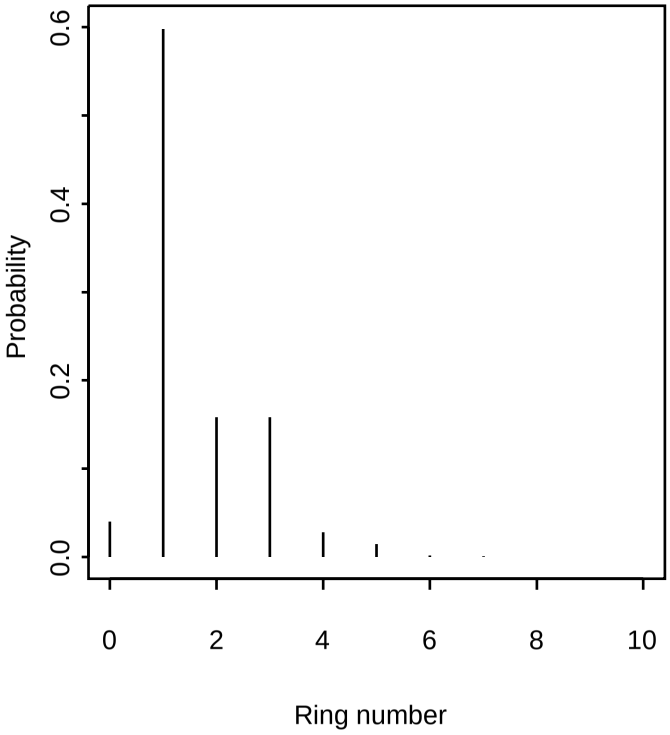
using (xiii) and (xiv). To obtain the mean-field approximation to the distribution of ring numbers, one can expand the generating functions as Taylor series about $z = 0$. This does not give rise to simple closed form expressions – the matrix exponentials in (xix) involve the eigen- values and vectors of the matrix $\boldsymbol{A}$; integrating the solution and repeatedly differentiating to obtain the Taylor expansion, yields expressions which rapidly become very complex and cumbersome. However it is possible, for given numerical values of the parameters, to use symbolic mathematics software (such as Maple IX[12]) to obtain numerical values of the mean-field approximation to the ring number probabilities.

The mean-field approximation of the expected ring number is given by $\mathcal{R}'_m(1)$ for males and $\mathcal{R}'_f(1)$.

Figure caption.

Fig. 1 Approximate ring number distributions for the two-sexx model obtained using the mean-field approximation and parameter values $\lambda = 0.1, \rho = 0.1$ and $t = 10$. The left hand panel is for males and the right-hand panel for females.


Probability
0.0
0.2
0.4
0.6
0
2
4
6
8
10
Ring number


Probability
0.0
0.2
0.4
0.6
0
2
4
6
8
10
Ring number