Three Disguises of $1 - x = e^{-\lambda x}$

Chathuri Karunarathna Mudiyanselage
Rabi K.C.
Winfried Just
Department of Mathematics, Ohio University

Mathematical Biology and Dynamical Systems Seminar
Ohio University
November 7, 2017

Solutions of $1 - x = e^{-\lambda x}$

The equation $1 - x = e^{-\lambda x}$ is quite interesting.

It always has the solution $x = 0$.

When $\lambda > 1$, then the function $e^{-\lambda x}$ is decreasing, concave up, and its first derivative at 0 is less than -1, which is the first derivative of the function $1 - x$.

Thus in this case there also exists exactly one solution $x = x(\lambda) \in (0, 1)$.

Moreover, $x(\lambda)$ is a strictly increasing function such that

$$\lim_{\lambda \to 1^+} x(\lambda) = 0 \qquad \text{and} \qquad \lim_{\lambda \to \infty} x(\lambda) = 1.$$

This solution $x(\lambda)$ pops up in very different contexts, both in mathematical biology and in other areas of mathematics. We will now present three different topics where this happens: The spread of immunizing infections, Galton-Watson birth processes, and random graphs.

SIR-models of disease transmission

Chathuri: Consider an infectious disease that is caused by pathogens (such as bacteria or viruses) that spread through direct contact in a population of N people, aka hosts.

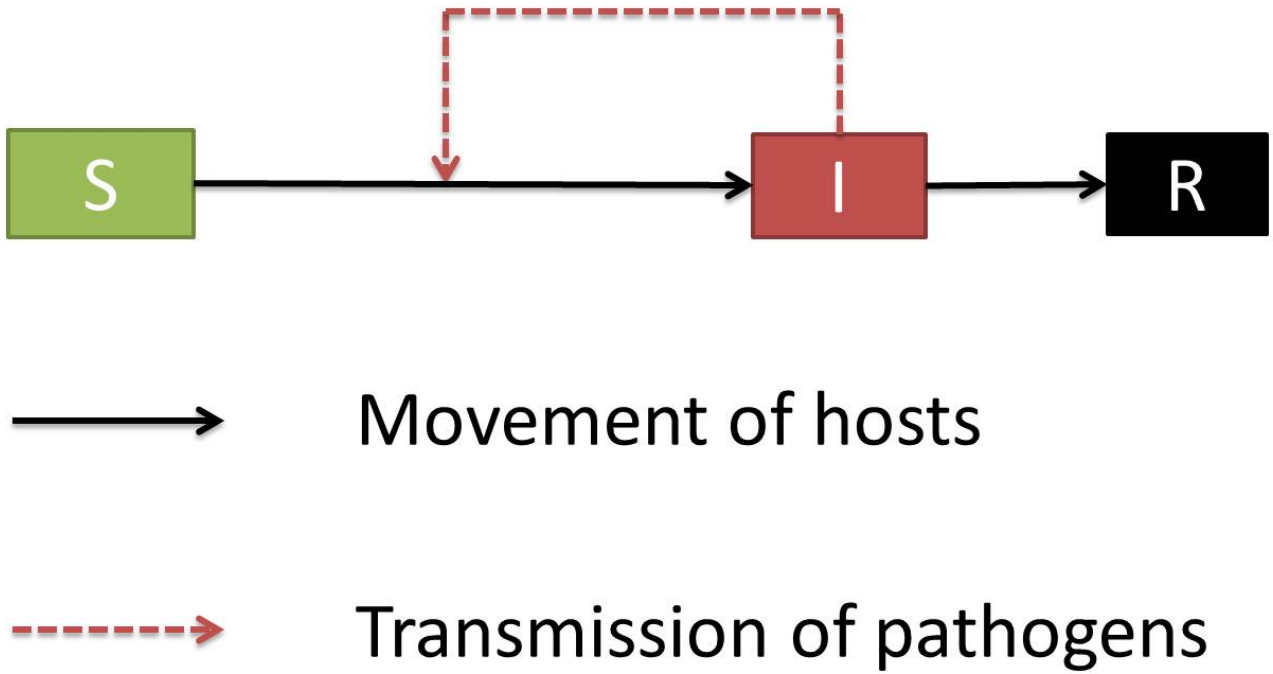
Assume, moreover, that the infection is immunizing so that nobody who recovers can ever be re-infected.

We can model the spread of such a disease by assuming that each host in the population is in one of three compartments at any given time:

- S: Susceptible to infection.
- I: Infected and able to infect others.
- R: Removed has recovered from the disease with permanent immunity or died from it.

SIR-models of disease transmission

We can now construct *SIR*-models of how membership in the compartments changes. Schematically, they look as follows:



On a quantitative level, we may consider variables s, i, r that represent the proportions of hosts in the three compartments at any given time, and study how these proportions change.

ODE-based *SIR*-models

These models take the form

$$\begin{aligned} s' &= -\beta si, \\ i' &= \beta si - \gamma i, \\ r' &= \gamma i, \end{aligned}$$

where $\beta, \gamma > 0$ are constants.

The equilibria are all vectors $(s^*, 0, r^*)$ with $s^* + r^* = 1$; $s^*, r^* \geq 0$.

Taking advantage of Leibniz notation for the derivative, we see that

$$\frac{di}{ds} = \frac{di/dt}{ds/dt} = \frac{\beta si - \gamma i}{-\beta si} = -1 + \frac{\gamma}{\beta s}.$$

which we can be solved by

$$\begin{aligned} \implies \int di &= \int \left(-1 + \frac{\gamma}{\beta s}\right) ds \\ \implies i &= -s + \frac{\gamma}{\beta} \ln s + C. \end{aligned}$$

The proportion s_∞ of hosts who escape the infection

Imposing initial conditions gives us that

$$i = -s + \frac{\gamma}{\beta} \ln s + i_0 + s_0 - \frac{\gamma}{\beta} \ln s_0.$$

Taking the limit as t goes to infinity, we have that

$$i_\infty = 0 = -s_\infty + \frac{\gamma}{\beta} \ln s_\infty + i_0 + s_0 - \frac{\gamma}{\beta} \ln s_0.$$

We will assume here that no one is recovered at the beginning of the epidemic (i.e., $r_0 = 0$, so that $s_0 + i_0 = 1$). Then we have

$$s_\infty - \frac{\gamma}{\beta} \ln s_\infty = 1 - \frac{\gamma}{\beta} \ln s_0.$$

Calculating s_∞ for epidemics caused by one index case

The equation

$$s_\infty - \frac{\gamma}{\beta} \ln s_\infty = 1 - \frac{\gamma}{\beta} \ln s_0$$

can be rewritten as

$$R_0 s_\infty - \ln s_\infty = R_0 - \ln s_0,$$

where $R_0 = \frac{\beta}{\gamma}$ is called the basic reproductive ratio.

When the epidemic is started by **one** infectious host called the index case, the initial population will be nearly completely susceptible (i.e., $s_0 \approx 1$, which implies that $\ln s_0 \approx 0$).
In this case it follows that

$$-\ln s_\infty = (1 - s_\infty) R_0.$$

Calculating the final size for *SIR*-models

The equation

$$-\ln s_\infty = (1 - s_\infty)R_0$$

can be written as

$$1 - (1 - s_\infty) = e^{-R_0(1-s_\infty)}.$$

When we let $x = 1 - s_\infty$ denote the final size of the epidemic, that is, the proportion of hosts who experienced infection, and let $\lambda = R_0$, we obtain the equation:

$$1 - x = e^{-\lambda x}.$$

Galton-Watson birth processes

Rabi: Consider an initial set of objects/individuals, which we call 0^{th} generation. We assume that each individual produces a random number of daughters at the end of every period producing successive generations. Also, the number of daughters born from a parent does not depend on the size of the population.

Let X_n denote the number of individuals in the n^{th} generation for $n = 0, 1, \ldots$ with $X_0 = 1$.

Galton-Watson birth/branching processes (GWbps) are discrete-time Markov chains, where the r.v.s $\{X_n\}_{n=0}^{\infty}$, and the time $n = 0, 1, 2, \ldots$ are discrete variables.

Some more assumptions for GWbps

- All individuals in the population gives birth to their daughters independently.
- The same offspring distribution applies to all generations.
- In the n^{th} generation, each individual gives birth to Y daughters.The discrete random variable Y takes non-negative integer values with probabilities $(p_0, p_1, ...)$, where

$$p_r = P(Y = r), \qquad r = 0, 1, 2, \ldots$$

Thus the size of $(n+1)^{th}$ generation is

$$X_{n+1} = \sum_{i=1}^{X_n} Y_i,$$

where $Y_1, Y_2, \ldots$ are i.i.d. copies of Y.

Probability generating function

The probability generating function (pgf) for X_n is given by

$$g_{X_n}(s) = E(s^{X_n}) = \sum_{r=0}^{\infty} p_r s^r, \tag{1}$$

where s is a real number in the unit interval.

If X_n has the Poisson distribution with the parameter λ, then the pgf is given by

$$g_{X_n}(s) = \sum_{r=0}^{\infty} s^r e^{-\lambda} \frac{\lambda^r}{r!} \tag{2}$$

$$= e^{-\lambda(1-s)}. \tag{3}$$

Extinction probability of a GWbp

Let us consider a family of one arbitrarily chosen individual. Let d_n denote the probability that the family of an individual dies out by the n^{th} generation, i.e., $d_n = \ P(X_n = 0)$.

The family of such an individual will become extinct if the individual has 0 daughters, or 1 daughter whose family becomes extinct by the $(n-1)^{th}$ generation, or 2 daughters both of whose families become extinct by the $(n-1)^{th}$ generation, and so on. So the probabilities of the events described above are:

$$p_0,\ p_1 d_{n-1}^1,\ p_2 d_{n-1}^2,\ \ \ldots$$

Since these events are mutually disjoint, the probability of occurrence of one of these event is

$$\sum_{r=0}^{\infty} p_r d_{n-1}^r = d_n.$$

Extinction probability of a GWbp

Now using $g(s) = \sum_{r=0}^{\infty} p_r s^r$ we get,

$$\sum_{r=0}^{\infty} p_r d_{n-1}^r = g(d_{n-1}).$$

Hence

$$d_n = g(d_{n-1}). \tag{4}$$

Since $X_n = 0$ implies $X_{n+1} = 0$, we have

$$0 = d_0 \leq d_1 \leq d_2 \leq \cdots \leq 1,$$

i.e., $\{d_n\}$ is a bounded monotone sequence. Hence

$$d = \lim_{n \to \infty} d_n \tag{5}$$

exists. It is called the probability of extinction.

Extinction probability of a GWbp

Taking limit on both sides of the equation $d_n = g(d_{n-1})$,
as $n \to \infty$ we get

$$d = g(d) = e^{-\lambda(1-d)} \tag{6}$$

Let x be the survival probability of the process. Then $x = 1 - d$.

Now substituting value of d as $1 - x$ in Equation (6), we get

$$1 - x = e^{-\lambda x}. \tag{7}$$

Graphs, paths, connected components

WJ: A graph is a pair $G = (V(G), E(G))$, where $V(G)$ is a nonempty set of nodes or vertices, and $E(G)$ is a set of unordered pairs of nodes that represent the edges.

The size of a graph is the number of nodes N.

The degree $deg(i)$ of node i is the number of edges that contain it.

A path in a graph G is a sequence of nodes $P = (i_1, i_2, \ldots, i_m)$ such that each of the pairs $\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_{m-1}, i_m\}$ is an edge of G.

The connected component of a node i in a graph G as the set of all nodes that can be reached by a path in G that starts at i.

Disclaimer: The terminology in graph theory is not as well-established as in other areas of mathematics. For example, some authors would call P as above a "walk" and/or define a connected component as a subgraph rather than a set of nodes.

An example

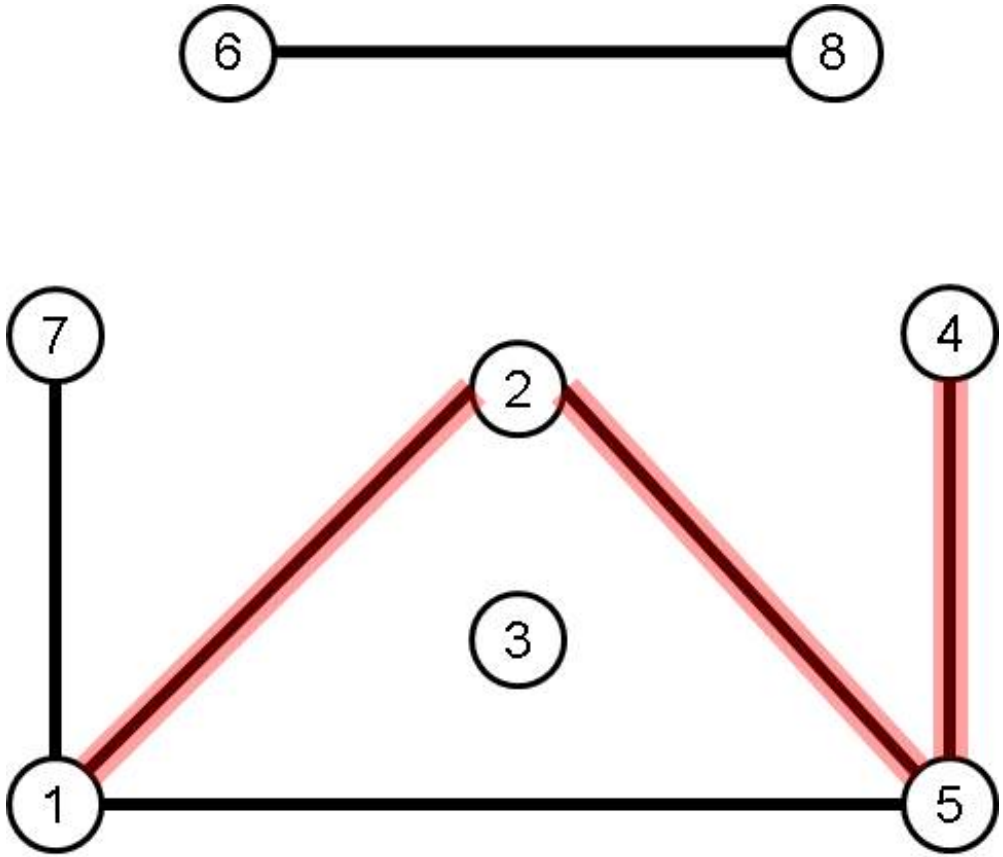


Figure: A graph $G = (V, E)$. Here $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$, while $E = \{\{1, 2\}, \{1, 5\}, \{1, 7\}, \{2, 5\}, \{4, 5\}, \{6, 8\}\}$. The edges traversed by the path $P = (1, 2, 5, 4)$ are highlighted. The connected components are $\{1, 2, 4, 5, 7\}$, together with $\{3\}$ and $\{6, 8\}$. The degrees are $deg(3) = 0$, $deg(i) = 1$ for $i \in \{4, 6, 7, 8\}$, $deg(2) = 2$, $deg(1) = deg(5) = 3$.

Random graphs

Graphs can be used to model various structures of interest in applied mathematics, for example contact networks.

Unfortunately, for large populations of people, we would not know the entire contact network. However, we might be able to estimate some of its statistical properties, such as the mean degree.

This would allow us to study other expected properties of the contact network if we assume that it is random graph, that is, randomly drawn from a probability distribution of graphs that tend to have these empirically verified statistical properties.

Note: In most cases, any graph of a specified size N could in principle be drawn from the distribution. Thus we will not be able to assert that a random graph has a given property for sure, but only that it will have this property asymptotically almost surely (a.a.s.), that is, with probability approaching 1 as $N \to \infty$.

Erdős-Rényi random graphs

To construct an Erdős-Rényi graph with N nodes, start by listing all possible edges $e_1, \ldots, e_{\frac{N(N-1)}{2}}$ between these nodes.

Then repeatedly toss a biased coin that comes up heads with probability p. We include edge e_ℓ as an actual edge if, and only if, the coin comes up heads in toss number ℓ.

The mean degree will be approximately $\lambda = p(N-1)$.

It will be more convenient to use the parameter λ instead of the parameter $p = \frac{\lambda}{N-1}$. The symbol $G_{ER}(N, \lambda)$ will denote an Erdős-Rényi random graph that is constructed with parameters N and λ.

For large N, the degree distribution in $G_{ER}(N, \lambda)$ tends to be approximately Poisson with parameter λ.

An example of a.a.s.: The maximal degree

Let us consider the probability of the event *Ldeg* that $G_{ER}(N, \lambda)$ will contain a node with degree $\geq \ln N$.

Since there are N nodes and the degree distribution is approximately Poisson with parameter λ, the probability of *Ldeg* can be estimated as follows:

$$P(Ldeg) \leq N\, P(deg(i) \geq \ln N) \approx N \sum_{k=\lceil \ln N \rceil}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!}$$

From a well-known estimate of the upper tail of the Poisson distribution we now get:

$$N \sum_{k=\lceil \ln N \rceil}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \leq N e^{-\lambda} \left(\frac{e\lambda}{\ln N}\right)^{\ln N} = \frac{e^{1+\ln N} \lambda^{-1+\ln N}}{N^{\ln N}}$$

The right-hand side approaches zero as $N \to \infty$, and it follows that $G_{ER}(N, \lambda)$ will a.a.s. contain no nodes with degree $\geq \ln N$.

This makes Erdős-Rényi random graphs rather unsuitable as models of contact networks, but they are still important.

Connected components in Erdős-Rényi random graphs

If the degree distribution in $G_{ER}(N, \lambda)$ is approximately Poisson, then a fraction of $e^{-\lambda}$ of all nodes will have degree $deg(i) = 0$, that is, will be in connected components of size 1.

How about the largest component, aka giant component?

Theorem (P. Erdős and A. Rényi, 1960)

Suppose $\lambda > 1$ and $\varepsilon > 0$. Let $0 < x < 1$ be such that $1 - x = e^{-\lambda x}$.
Then there exists a constant $c > 0$ such that in the class of Erdős-Rényi graphs $G_{ER}(N, \lambda)$ the following holds a.a.s.:

- *There exists one connected component GC of $G_{ER}(N, \lambda)$ of relative size $x - \varepsilon < \frac{|GC|}{N} < x + \varepsilon$.*
- *All other connected components of $G_{ER}(N, \lambda)$ have size $< c \ln N$.*

What is the connection?

Chathuri: The size of this giant component is exactly the same as the final size of the epidemic that I derived for the *SIR*-model! Is there a connection between random graphs and disease transmission?

Rabi: And it is exactly the same as the survival probability that I derived for the Galton-Watson birth process! Is there a connection between random graphs and these birth processes?

WJ: In fact there is. But first let's think about this: Would there be a connection between disease transmission and Galton-Watson birth-processes? How, exactly, are diseases transmitted?

Chathuri: In the models that I presented, by direct contact. Such a contact is called effective if a sufficient number of pathogens is transmitted to cause an infection. But whether or not a contact between any two persons occurs, and whether or not it is effective, these would be random events.

Generations of the infection

Rabi: So the spread of an outbreak with one index case (initially infectious host) really works like the Galton-Watson process that I described:

- The index case would be the individual in generation 0.
- The hosts that get infected by the infected case can be considered the "daughters" of the index case. They comprise generation 1.
- The hosts that become infected by an individual in generation 1 can be considered "daughters" of that individual. They comprise generation 2. And so on.
- If the population is large and each host has relatively few effective contacts during the time until they recover, the number of "daughters," that is, number of secondary infections caused by this host will have approximately a Poisson distribution with parameter $\lambda = R_0$.

Chathuri: Right! The parameter R_0 is usually defined as the average number of secondary infections cause by an average index case in an otherwise susceptible population.

Are the assumptions of the Galton-Watson process satisfied?

WJ: But we need to check the assumptions of the Galton-Watson birth process.

Chathuri: All models are only approximations of the real world. In the study of disease dynamics we usually make the simplifying assumption that the effective contacts that different pairs of people have are independent. Under this assumption, the numbers of offspring will be independent random variables, as in Rabi's model.

WJ: Will they have approximately Poisson distributions with parameter $\lambda = R_0$?

Chathuri: Yes, we can make this simplifying assumptions, as long as there is not much variability in the time it takes until recovery from the disease.

Minor vs. major outbreaks

Rabi: Now we can see that one of two scenarios must occur:

- Either the birth process dies out and the spread of the infection will very quickly stop.
- Or the process will survive, with more and more people getting infected.

WJ: In the former case we will observe a minor outbreak with only a very, very small fraction of the population experiencing infection. In the latter case we will observe a major outbreak or epidemic with a significant fraction $x = 1 - s_\infty$ of hosts experiencing infection.

Chathuri: The ODE model will predict the former scenario for initial condition $(s_0, i_0, r_0) = (1, 0, 0)$, and the latter scenario for an initial condition $(s_0, i_0, r_0) = (1 - \varepsilon, \varepsilon, 0)$, where $\varepsilon > 0$ is very small.

WJ: Right! This ODE model is really just an approximation of a stochastic process that is very similar to the one that Rabi described.

Wait a minute ...

Rabi: But hold on: When the Galton-Watson process survives, new daughters will be produced forever. How could this happen in a finite population?

WJ: It can't. We still have not checked the assumption that the number of "daughters," or secondary infections, is identically distributed in all generations.

Chathuri: This one will actually be false. Only the number of effective contacts can be assumed identically distributed. When an infectious hosts makes contact with a host who is already removed or infectious, no new infection will result. Thus when there is already a significant proportion of such hosts, this will decrease the parameter λ in the distribution for your process.

WJ: So we can see that Rabi's Galton-Watson process will be a good approximation for the next-generation process in the beginning of an outbreak, but not in later stages of an epidemic.

Chathuri: But why do we get the same value for the survival probability of this process and the final size?

Let's construct a random graph

Let's forget about actual outbreaks for a moment and consider hypothetical ones in a fixed population of size N. Let's construct an Erdős-Rényi random graph $G_{ER}(N, \lambda)$ by making the hosts in this population the nodes. We will include an edge $\{i, j\}$ in this graph with probability $p = \frac{R_0}{N-1}$, so that $\lambda = R_0$.

We can interpret p as the conditional probability of the event that an effective contact between i and j during the time interval between infection of the first of these hosts and removal of that host who gets first infected, **if** in fact at least one of these hosts gets infected during the outbreak.

There is a close correspondence between such graphs and simulated outbreaks in an *SIR*-model with the given R_0. Now let's look at an outbreak that would correspond to the Erdős-Rényi graph that we have produced.

The connected component of a node i will be the set of all hosts who experience infection if the outbreak was caused by index case i.

The final size of a major outbreak and survival probabilites

The final size in the *SIR* model is the relative size of the connected component of the index case i.

If the index case is in the giant component, then the outbreak is a major one, and the estimate of the relative size of the giant component will be an estimate of the final size of the outbreak.

This explains why we get the same number for the relative size of the giant component and for the final size of a major outbreak in the ODE model.

If the index case is in a small component, then the next-generation process will die out before it reaches a significant proportion of the hosts, that is, while the Galton-Watson process is still a good approximation for the actual next-generation process. This will happen with probability $d = 1 - x$. Now if we pick the index case **randomly** for our given instance of $G_{ER}(N, \lambda)$, then this probability must also be equal to $1 - x$, where x is the relative size of the giant component. This explains why x must be equal to the survival probability of the Galton-Watson process.