A probabilistic approach to calculate R0 in complex epidemic models

Carlos Hernandez (carlosmh@mac.com)
Universidad Francisco Gavidia  https://orcid.org/0000-0003-2216-0905

Osval Montesinos-López
Universidad de Colima  https://orcid.org/0000-0002-3973-6547

Research Article

Keywords: Epidemic models, Basic reproductive number, R0 Markov chains, Stochastic epidemic models

Posted Date: September 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2082391/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
A probabilistic approach to calculate $R_0$ in complex epidemic models

Carlos Hernandez-Suarez\textsuperscript{a,\,*}, Osval Montesinos López\textsuperscript{b}.

\textsuperscript{a}Instituto de Innovación y Desarrollo, Universidad Francisco Gavidia, San Salvador, El Salvador
\textsuperscript{b}Facultad de Telemática, Universidad de Colima, Colima 28040, Mexico

Abstract

Epidemic models are a valuable tool in the decision-making process. Once a mathematical model for an epidemic has been established, the very next step is calculating a mathematical expression for the basic reproductive number, $R_0$, which is the average number of infections caused by an individual that is introduced in a population of susceptible. Finding a mathematical expression for $R_0$ is important because it allows to analyze the effect of the different parameters in the model on $R_0$, so that we can act on some of them, in an attempt to keep $R_0 < 1$ to avoid a large epidemic. In this work we show how to calculate $R_0$ in complex epidemic models by using only basic concepts of Markov chains. We compare our approach with the traditional methods used in epidemiology and demography.

Keywords: Epidemic models, Basic reproductive number, $R_0$, Markov chains, Stochastic epidemic models

1. Introduction

In this note we will not review the history of $R_0$, also called basic reproductive number in epidemiology and demography, which is nicely presented in Heesterbeek (2002). The importance of $R_0$ as a parameter that determines the

\*Corresponding author

Email addresses: cmhs1@cornell.edu (Carlos Hernandez-Suarez), oamontes1@ucol.mx (Osval Montesinos López)
fate of a population of agents that reproduce, can be found in the theory of branching process in probability: in this process an individual gives birth to a random number of descendants and each of the descendants in turn gives birth to a random number of individuals and so on, and we are interested in calculating the probability that the population will become extinct. It is assumed that every individual of every generation, leaves $X$ descendants before dying, where $X$ is a random variable with some probability distribution. If $p_e$ is the probability of extinction of a population that starts with one individual and $P_i$ is the probability that an individual has $i$ descendants, then we can provide an expression for $p_e$ conditioning on the number of descendants of the first individual:

$$p_e = (p_e|X = 0)P_0 + (p_e|X = 1)P_1 + (p_e|X = 2)P_2 + (p_e|X = 3)P_3 \ldots$$

$$= P_0 + p_e P_1 + p_e^2 P_2 + p_e^3 P_3 \ldots \quad (1)$$

Observe that this expression assumes that if there are $k$ descendants at the first generation, then, the probability that the population will vanish is $p_e^k$, that is, the fate of each descendant is independent of the others but, most important, it assumes that each descendant has identical reproductive capabilities than the parent, which is frequently overlooked and will be relevant later. We leave for appendix A1 the proof of the following result: if the expectation of the random variable $X$ is $\theta$, then, if $\theta \leq 1$ the population will vanish eventually with probability 1. A corollary is that if $\theta > 1$, then there is a chance that the population will prevail, and the chances increase non-linearly with $\theta$.

The previous model assumes that all individuals are equally capable to reproduce, that is, that $X$ follows the same distribution for all individuals, which is not always true. Assume for instance that the resources to survive are limited, for instance, imagine a bacteria population in a petri dish, where the population will start competing for resources (besides perhaps the toxicity of chemicals produced by the same bacteria) and in this situation, the average number of descendants is not constant, because the first individual can produce on average
more descendants than those from later generations, due to saturation. But there is a fact we can use: if we find out that the first individual has an average number of descendants smaller or equal to one, then, descendants at later generations will have even worst conditions to reproduce than the first one and extinction is certain. This is the reason why in to the probability of extinction of a population we only need to calculate the reproductive potential of the first individual, because no other individual born at later generations can have a higher reproductive capability, when resources to reproduce becomes scarce when population grows. This also works in general for a population that starts with \( n_0 \) individuals: if the probability of extinction of a population that starts with a single individual is \( p_e \), then the probability of extinction of a population that starts with \( n_0 \) individuals is \( p_e^{n_0} \). If \( p_e < 1, p_e^{n_0} < 1 \)

The previous theory can be applied to contagion, where we now follow a population of infected individuals and, if the average number of infections per infected individual is larger than one, the disease may spread to become a large epidemics. Nevertheless, the first individual has always more chances to find susceptible individuals to infect, whereas subsequently infected individuals will find previously infected (and perhaps immune) individuals, so, it is harder for them to keep the same average number of infections than the first infected. But, again, if the average number of infections of the first individual, the one with more potential to produce more infections (when everybody else is susceptible) is smaller than one, then, all other subsequently infected, with less potential due to saturation, will have an average number of infections less than one and no large epidemics will occur.

In mathematical epidemiology the equivalent to the average number of descendants is thus the average number of infections, and the average number of descendants, \( \theta \), becomes \( R_0 \) in epidemiology. It is clear why the definition of \( R_0 \) has been established in terms of the potential of the first individual (Diekmann \textit{et al.}, 1990; Heesterbeek & Dietz, 1996).

An alternative construction of an epidemic that suggests a broader meaning for \( R_0 \) was suggested by Hernández-Suárez & Mendoza-Cano (2009) in terms
of urn models, a set of probability distributions that arise from placing balls in urns at random (Johnson & Kotz, 1977). In the simplest urn model, one places \( b \) balls at random in \( N \) urns, where each urn can contain any number of balls. The expected number of occupied (non-empty) urns is:

\[
1 - e^{-b/N}.
\]

From this, Hernández-Suárez & Mendoza-Cano suggested to construct an epidemic by placing a single ball at random in one of the \( N \) urns and establishing the rule that, every time a ball falls in an empty urn, it throws away \( X \) balls at random to the \( N \) urns, where \( X \) follows some arbitrary distribution with \( E[X] = R_0 \). If the expected number of occupied urns is \( E[Z] \), then:

\[
E[Z] \approx 1 - e^{-E[Z]R_0/N},
\]

which is due to the fact that the total number of balls thrown has expectation \( E[Z]R_0 \). The final epidemic size of the epidemics can be obtained by solving numerically for \( E[Z] \). Conversely, if we observe a realization of \( Z \), \( R_0 \) can be estimated from the last expression. Interesting enough, expression (2) is the same result for the final epidemic size as the derived by Kermack & McKendrick (1991) using a deterministic SIR model.

Hernández-Suárez & Mendoza-Cano also showed that the distribution of \( Z \) is irrelevant, and we only need the first moment. Observe that in this construction of an epidemic, the first infected is the first ball placed in the \( N \) urns and each set of \( X \) balls thrown at random to the remaining \( N \) urns, are the contacts, that will result in a new infection if a ball falls in an empty urn, which is in essence, the basic mechanism of transmission in an epidemic. Observe also that this construction does not require to define \( R_0 \) as the average number of balls thrown by the first infected, but the average number of balls thrown by any infected. Extending this to the traditional definition of \( R_0 \) as the average number of infections of the first infected, we can see that \( R_0 \) is the average number of contacts made by any infected individual during its infectious life.
Whether these contacts result in infections or not is irrelevant.

In the Discussion section, we will analyze how the probabilistic approach presented here compares with the traditional mathematical method to calculate $R_0$ in epidemiology (by using mainly the approach of Diekmann et al., 2010) and with that used in demography based on matrix population growth models of Leslie (1945, 1948) and Caswell (2009).

1.1. The life circle

There is some care we need to take in the calculation of the average number of individuals produced by an individual. Figure 1 is a typical example of the problem at hand, related to *Eratyrs mucronatus*, the *kissing bug* that transmits Chagas’ disease (see Hernandez-Suarez et al., 2019). This figure shows that the life cycle of an individual has been decomposed in seven biological stages, and that individuals in the adult stage produce individuals in the egg stage, but the average number of eggs produced is not $R_0$ because there is no guarantee that an egg will become an adult and thus, the average number of eggs produced by an adult is not a measure of the ability of an individual to replace itself. A requirement from branching process (see eqn. 1) is that every descendant has the same fitness of the parent to reproduce, therefore, what we need to calculate is the average average number of eggs produced by an egg, or the average number of adults produced by an adult, or the average number of nymph 1 individuals produced by a nymph 1 individual and so on.

The previous rationale applies to epidemic models, where a particular individual moves along a series of states or compartments (susceptible, infected, isolated, vaccinated, dead, etc.) according to certain rules, and, for obvious reasons, we are interested in how many infected will be produced by an infected, and it has been clarified already that we need to analyze the infectious potential of the first infected only.

Define a *contact* as any act between two individuals that will cause the infectious of the susceptible if the contact involves an infectious and a susceptible. This could be sexual contact, sharing hands, sharing needles, talking, etc. With
Figure 1: The life cycle of *Eratyrus mucronatus*. In this example, the Adult stage is the reproductive one, and all newborns are eggs. The dashed arrow indicates that all newborns belong to the first stage, and $f_7$ is the fertility rate.

this in mind, observe that the number of contacts that an individual has per unit of time (day, hour, week, etc) is a random variable $Y$ with expectation $\theta$, and this is the number that matters. Everything we need to know is the number of contacts that an infectious individual has during the time s/he is infected, because, by assumptions in the model, $Y$ has the same distribution for every individual, regardless if it is the first infected or the one in the middle or the last one, because individuals have contacts whether they are infected or not, that is, contacts occur even before the epidemic starts. This means that to obtain $R_0$, we need to calculate the average number of contacts an individual has per unit of time and then multiply this number by the average time the individual spends infectious.

2. Methodology

We have highlighted the key to calculate $R_0$, regardless how complicated the model is: if $a$ is the rate (number of events per unit of time) at which an individual has contacts, and $w$ is the average time an infected individual is infectious, then $R_0 = a \cdot w$. The problem is not $a$, since most of the times it is one of the parameters in the model. The problem is $w$, the average time infectious, since the model may be very complex and it is not clear how to calculate it. For that, we need some theory of Markov chains. Nevertheless, in many cases, $R_0$ can be calculated by inspection. For instance, consider the classical SIR model
in Figure 1. Above each arrow between any two compartments, there is the rate at which individuals move from one compartment to another. The word \textit{rate} needs to be clearly defined, and we use for this \textit{rate} in the sense of a stochastic model: if an individual leaves a box at a rate $\alpha$, it means that the average time an individual spends in that compartment is exponential with mean $1/\alpha$.

\[ S \xrightarrow{\lambda S/N} I \xrightarrow{\mu I} R \]

\textbf{Figure 2: The Susceptible-Infected-Removed epidemic model. $R_0 = \lambda/\mu$.}

Now, in the SIR model, what is the average time an individual lasts infectious? We can see that the individual exit rate from compartment $I$ is $\mu$, so on average, the individual remains infected for an average time $1/\mu$. But according to the model in Figure 2, every individual has contacts with another at a rate $\lambda$, therefore, the average number of contacts of an infectious individual is $\lambda/\mu$, which is $R_0$. Applying the same rationale, word by word, the $R_0$ value is the same for the SEIR model of Figure 2, where $E$ stands for an infected, noninfectious (“latent”) stage.

\[ S \xrightarrow{\beta S/N} E \xrightarrow{\sigma E} I \xrightarrow{\mu I} R \]

\textbf{Figure 3: The Susceptible-Latent-Infected-Removed epidemic model. $R_0$ is also $\lambda/\mu$.}

Now, consider a particular case of a modified SEIR model depicted in Figure 4, where an individual in the latent stage can be detected and removed before
going to the I state with probability $p$. This is a big change: in the SIR model, an infectious (I) individual produces infected individuals, whereas in the SEIR model, an infected individual produces latent (E) individuals. Since in an SEIR model every latent individual will become type I eventually with probability one (nothing in the model indicates otherwise) there is no difference in $R_0$ between SIR and SEIR, but in the model of Figure 4, an infectious individual produces latent individuals and now there is a chance that they may not become infectious. This is a similar problem to the “life cycle” problem mentioned in the previous section. Since what is needed is the average of infectious produced by an infectious, in the modified SEIR the chance of being removed before becoming infectious, must be considered. Thus, the average amount of time that a recently infected individual spends in the infectious compartment I changes from $1/\mu$ to $(1-p)/\mu$. Since the number of contacts is still $\lambda$, the $R_0$ for this modified SEIR is $\lambda(1-p)/\mu$.

Figure 4: The Susceptible- Latent -Infected -Removed epidemic model with detection and removal of latent. Now there is a chance that a newly infected will not become infectious. $R_0$ for this model is $\lambda(1-p)/\mu$.

Another example, a bit more complicated is the model of Figure 5, in which an individual infected may alternate between infectious (I) and non-infectious (C) states several times before being removed. That is, the individual may move through the states like S-I-C-I-C-I-R, or S-I-C-I-C-I-C-I-C-I-C-I-C-I-C-I-R or S-I-R, just to illustrate the problem. Once the expected number of times the individual visits stage I, say $k$, we multiply this value by $(\mu + \alpha)^{-1}$ which is the average time in each visit to I, to give the average time spent in the infectious stage, and then, this the result is multiplied by the contact rate to yield $R_0$. 

8
It requires some training and experience to calculate $R_0$ just by inspection, that is why we need the theory of Markov chains. The resulting theory can be applied to any model, simple or not.

![Diagram of a model in which an infected individual switches between a carrier (C), non-infectious stage and an infectious stage (I). The number of visits to state I is a geometric random variable.](image)

Figure 5: A model in which an infected individual switches between a carrier (C), non-infectious stage and an infectious stage (I). The number of visits to state I is a geometric random variable.

2.1. Markov chains

We will focus on discrete time, discrete space Markov chains. An excellent review of these processes at the level required in this paper is found in Ross (2007); Grinstead & Snell (1997) and a deeper treatment in Iosifescu (1980). Imagine an individual moving across a series of states (see Figure 6) in some random fashion: when it is in stage $i$, it moves to stage $j$ with some constant probability $p_{ij}$. The time spent in each stage is, by now, irrelevant. For now, assume that the time spent on each stage is unitary. This is a very useful model with applications in a wide range of areas. Observe that, regardless on where the individual starts, the process never ends, the individual keeps moving: $3,4,2,3,2,3,4,2,1,4,2,3,...$

There are many interesting questions related to this model, for instance, what is the probability that an individual starting in stage 3 will be on stage 1 after $k$
transitions? Or, what will the proportion of time spent on each stage when the number of transitions goes to infinity?

![Figure 6: A Markov chain with four states.](image)

But there is a class of Markov chains we will focus on: those that contain one or more absorbing states, where an absorbing state is a black hole that swallows the individual and this cannot leave this stage, and thus the process stops (see Figure 7). These Markov chains also raise many questions as, for instance: what is the probability that the individual ends trapped on some state \(a\) rather than in state \(b\)? (assuming that states \(a\) and \(b\) are absorbing states). Or, what is the average number of visits made to some state \(x\) before the process stops? The answers to these questions are simple and straightforward, although we refer to other texts for the proof.

Consider for instance the Markov chain of Figure 7. We first need to construct a matrix of transitions between states, where \(p_{ij}\) is the probability that once in state \(i\) the individual moves to state \(j\).
Figure 7: A Markov chain with five states, where states 3 and 5 are absorbing states. States 1, 2 and 4 are called transient states.

\[ P = \begin{bmatrix}
1 & 2 & 3 & 4 & 5 \\
1 & 0 & 0.2 & 0 & 0.8 & 0 \\
2 & 0.7 & 0 & 0.3 & 0 & 0 \\
3 & 0 & 0 & 1 & 0 & 0 \\
4 & 0 & 0.5 & 0 & 0 & 0.5 \\
5 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix} \]

Absorbing states are easily located and correspond to the 1’s in the diagonal (states 3 and 5). The rest of the states is called transient states. We first need to rearrange rows and columns of the matrix \( P \) so that the absorbing states are located at the bottom right corner, to the form:

\[ P = \begin{bmatrix}
U & R \\
Z & I \\
\end{bmatrix} \]
where $U$ is a $k \times k$ matrix containing the transitions between the $k$ transient states, $R$ is a $k \times r$ matrix containing the transitions between the $k$ transient states to the $r$ absorbing states, $Z$ is a $r \times k$ matrix of zeros and $I$ is the identity matrix of size $r$. Rearranging matrix $P$ yields:

$$
P = \begin{bmatrix}
1 & 2 & 4 & 3 & 5 \\
1 & 0 & 0.2 & 0 & 0.8 & 0 \\
2 & 0.7 & 0 & 0.3 & 0 & 0 \\
3 & 0 & 0 & 0 & 1 & 0 \\
4 & 0 & 0.5 & 0 & 0 & 0.5 \\
5 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
$$

Then:

$$
U = \begin{bmatrix}
0 & 0.2 & 0 \\
0.7 & 0 & 0.3 \\
0 & 0.5 & 0 \\
\end{bmatrix}, \quad R = \begin{bmatrix}
0.8 & 0 \\
0 & 0 \\
0 & 0.5 \\
\end{bmatrix}
$$

2.1.1. The fundamental matrix

Among the many useful results that can be derived from the previous decomposition of matrix $P$, the one we need is the following result:

The average number of visits to each transient state before absorption, is given by $N = (I - U)^{-1}$.

The proof is in Appendix A2. Matrix $N$ is called the fundamental matrix of $P$.
which is read as follows: the average number of visits to state 2, if the initial state is state 1, is 0.282. If the initial state is 4, the average number of visits to stage 4 is 1.211.

2.1.2. From rates to probabilities

A useful result from probability theory, in particular, from properties of exponential distributions is related to how likely is that an individual moves from one state to another, given that we do not know transition probabilities, but only exit rates. Figure 8 illustrate the problem. In this Figure, an individual moves from compartment A to B or C at rates $\alpha$ and $\beta$ respectively. But it can only move to one of them. The result is that the probability that it moves to A is $\frac{\alpha}{\alpha + \beta}$ and to B with the remaining probability $\frac{\beta}{\alpha + \beta}$. This result can be extended to any number of compartments.

2.1.3. Average time in a compartment

The last result we need is related to the average time in a compartment. Figure 8 is useful to illustrate the result: the average time in a compartment is equal to the inverse of the sum of all exit rates, regardless the number of compartments. In this case is $(\alpha + \beta)^{-1}$.

We are now in possession of the three elements that will be used to calculate $R_0$. Once a model has been established, we need to follow four steps:

(i) Convert all rates in the model to probabilities, to build a transition matrix $P$ between states.

(ii) Calculate the average number of visits to the infectious state using the fundamental matrix $N$.

(iii) Calculate the average time spent in a single visit to the infectious state and use the result in (ii) to obtain the average time total expected time in the infectious state.

(iv) Multiply the result in (iii) by the contact rate, to obtain the total number of contacts of an infectious individual. This is $R_0$. 

13
Figure 8: An individual moves from compartment A to B or C at rates $\alpha$ and $\beta$ respectively. The probability that it moves to A is $\alpha/(\alpha + \beta)$. The average time in compartment A is $(\alpha + \beta)^{-1}$.

3. Examples

3.1. Modified SEIR

The first example is the modified SEIR (Figure 4). As previously shown, it does not require the use of Markov chains because is simple enough to be calculated by inspection, but it will be used with the purpose of preparation of more complicated examples.

Step (i) is the calculation of the transition matrix $P$ in this case is:

$$
P = \begin{bmatrix}
S & E & I & R \\
S & 0 & 1 & 0 & | & 0 \\
E & 0 & 0 & 1-p & | & p \\
I & 0 & 0 & 0 & | & 1 \\
R & 0 & 0 & 0 & | & 1 
\end{bmatrix}
$$
where:

\[
\begin{bmatrix}
S & E & I \\
S & 0 & 1 & 0 \\
E & 0 & 0 & 1-p \\
I & 0 & 0 & 0 \\
\end{bmatrix}
\]

To calculate step (ii) we need the *fundamental matrix* \( N \) which is:

\[
\begin{bmatrix}
S & E & I \\
S & 1 & 1 & 1-p \\
E & 0 & 1 & 1-p \\
I & 0 & 0 & 1 \\
\end{bmatrix}
\]

where we can see the expected number of visits to the infectious state \( I \) starting from state \( S \) (we always start in this state) is \( 1-p \).

For step (iii), the average time spent in a single visit to the infectious state, this is the inverse of the sum of all exit rates from \( I \), which is \( 1/\mu \) (we need to consider only the *individual* exit rate, since the total exit rate is \( \mu I \), the result follows). Multiplying this by the average number of visits to the infectious state gives \( (1-p)/\mu \) as the average total amount of time that an individual stays infectious.

\( R_0 \) from step (iv) is just the product of the contact rate and the previous result, \( \lambda(1-p)/\mu \).

### 3.2. Modified SEIR with natural mortality

The second example is another version of the SEIR in which individuals in states \( E \) or \( I \) may die (Figure 9). We can start directly writing matrix \( U \) the matrix that excludes the absorbing states (in this case state \( R \) and the death states). Observe that the death state is not indicated in the figure but is where all the arrows pointing down are directed to. The remaining states (transient) are \( S, E \) and \( I \). Matrix \( U \) is:
Figure 9: An SEIR model with disease induced mortality. In this model \( N = S + E + I \)

\[
\begin{pmatrix}
S & E & I \\
\end{pmatrix}
\]

\[
U = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & \frac{\alpha}{\alpha + \omega} \\
0 & 0 & 0
\end{pmatrix}
\]

The reason why the probability that the individual moves from state E to I is \( \frac{\alpha}{(\alpha + \omega)} \) has been explained in Figure 8.

The fundamental matrix \( N \) is:

\[
N = \begin{pmatrix}
1 & 1 & \frac{\alpha}{\alpha + \omega} \\
0 & 1 & \frac{\alpha}{\alpha + \omega} \\
0 & 0 & 1
\end{pmatrix}
\]

where we can see the expected number of visits to the infectious state I starting in state \( S \) is \( \frac{\alpha}{(\alpha + \omega)} \).

At this point and in preparation to more complex models of disease transmission, we will introduce two vectors that, together with \( N \) will allow for a quick calculation of \( R_0 \). This are vectors \( w \) and \( b \). Vector \( w \) contains elements \( w_j \) as the average time in stage \( j \) once in that stage, if stage \( j \) is infectious and 0 otherwise. Vector \( b \) contains elements \( b_j \) as the contact rate in stage \( j \) if stage \( j \) is infectious and 0 otherwise. Observe: \( w \circ b \) contains the number of contacts per visit to the infectious stages (with ‘\( \circ \)’ as the element-by-element product).
Now, since all infected come from the susceptible stage $S$, we only need the number of visits to every infectious stage by an individual that starts in stage $S$, which is contained in the first row of $N$. Let $N_1$ be this vector. Then:

$$R_0 = N_1(w \circ b)$$

(3)

Continuing with our example of a modified SEIR with natural mortality, we have $w' = \{0, 0, (\mu + \omega)^{-1}\}$ and $b' = \{0, 0, \beta\}$. Thus we have: $(w \circ b)' = \{0, 0, \beta/(\mu + \omega)\}$. Finally:

$$R_0 = N_1(w \circ b) = \frac{\beta \alpha}{(\alpha + \omega)(\mu + \omega)}$$

(4)

3.3. An Ebola transmission model

We will analyze another example where individuals are infectious at different stages and with a different degree of infectiousness. This is an Ebola model in infectious individuals infect susceptible ones while alive outside hospitals, in hospitals and dead but not buried (Legrand et al., 2007). The infectious states are $I$, $H$ and $F$. State $E$ is included since it communicates with one of the infectious states. The final set is $\{E, I, H, F\}$ (see Figure 10).

This is an interesting example because the answer allows to see the definition of $R_0$ when there are several types of infectious individuals.

We can skip directly to the transition matrix $U$, that excludes the absorbing states. Thus:

\[
\begin{bmatrix}
S & E & I & H & F \\
S & 0 & 1 & 0 & 0 & 0 \\
E & 0 & 0 & 1 & 0 & 0 \\
I & 0 & 0 & 0 & \theta_1 & \delta_1 (1 - \theta_1) \\
H & 0 & 0 & 0 & 0 & \delta_2 \\
F & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

The fundamental matrix $N$ is:
Figure 10: An Ebola transmission model in which hospitalized and unburied are a source of infection. S, Susceptible individuals; E, Exposed individuals; I, Infectious; H, hospitalized; F, dead but not yet buried; R, removed. The parameters are: $\beta_I$, transmission coefficient in the community; $\beta_H$, transmission coefficient at the hospital; $\beta_F$, transmission coefficient during funerals. $\theta_1$ fraction of infectious cases that are hospitalized. $\delta_1, \delta_2$ are computed in order that the overall case-fatality ratio is $\delta$. The incubation rate is $\alpha$. The mean duration from symptom onset to hospitalization is $\gamma_{ih}^{-1}$, $\gamma_{ih}^{-1}$ is the mean duration from hospitalization to death, and $\gamma_{ih}^{-1}$ denotes the mean duration of the infectious period for survivors. The mean duration from hospitalization to end of infectiousness for survivors is $\gamma_{ih}^{-1}$ and $\gamma_{ih}^{-1}$ is the mean duration from death to burial. Transmission coefficients are expressed in weeks$^{-1}$ (Legrand et al., 2007).
The infectious stages are $I$, $H$ and $F$. Thus $w' = [0,0,1/\gamma_I,1/\gamma_H,1/\gamma_F]$, and $b' = [0,0,\beta_I,\beta_H,\beta_F]$. The number of contacts in the infectious stages per visit to an infectious stage are $(w \circ b)' = [0,0,\beta_I/\gamma_I,\beta_H/\gamma_H,\beta_F/\gamma_F]$. Thus:

$$R_0 = N_1(w \circ b) = \beta_I/\gamma_I + \beta_H\theta_1/\gamma_H + \beta_F(\delta_2\theta_1 + \delta_1(1 - \theta_1))/\gamma_F).$$

This expression allows to see the contribution to $R_0$ from each type of infected and allows to see that the definition of $R_0$ is the expected number of individuals of type $I$ produced by an individual of type $I$, plus the expected number of individuals of type $H$ produced by an individual of type $H$ plus the expected number of individuals of type $F$ produced by an individual of type $F$.

### 3.4. A model for COVID-19 transmission

Another example is a model for Covid-19 transmission (Figure 11). This model was chosen because it seems complex enough, with eight states and twelve possible transitions (Ndaïrou et al., 2020). The model has the following states: susceptible class ($S$), exposed class ($E$), symptomatic and infectious class ($I$), super-spreaders class ($P$), infectious but asymptomatic class ($A$), hospitalized ($H$), recovery class ($R$), and fatality class ($F$). Absorbing states are $A, R$ and $F$. Leaving the transient states $S, E, I, H$ and $P$. Observe the transition probabilities between the transient states: $P(S \rightarrow E) = 1, P(E \rightarrow I) = p_1, P(E \rightarrow P) = p_2, P(I \rightarrow H) = \gamma_a/(\gamma_a + \gamma_i + \delta_i), P(P \rightarrow H) = \gamma_a/(\gamma_a + \gamma_i)$. We can skip directly to writing the matrix $U$:
Figure 11: A Covid-19 mode. Susceptible class (S), exposed class (E), symptomatic and infectious class (I), super-spreaders class (P), infectious but asymptomatic class (A), hospitalized (H), recovery class (R), and fatality class (F). \( \beta_b, \beta_H, \beta_P \) are the contact rates of individuals in each of those states.
and \( N_1 \), the first row of fundamental matrix \( N \) is:

\[
N_1 = \begin{bmatrix} 1, 1, p_1, \gamma_p p_1/\gamma_a p_1/(\delta_i + \gamma_a + \gamma_i), \gamma_p p_2/(\gamma_a + \gamma_i), p_2 \end{bmatrix}
\]

Vectors \( w \) and \( b \) are respectively:

\[
w' = \begin{bmatrix} 0, 0, (\delta_i + \gamma_a + \gamma_i)^{-1}, (\delta_h + \gamma_r)^{-1}, (\gamma_p + \gamma_a + \gamma_i)^{-1} \end{bmatrix}
\]

\[
b' = \begin{bmatrix} 0, 0, \beta_I, \beta_H, \beta_P \end{bmatrix}
\]

Therefore \((w \circ b)' = [0, 0, \beta_I/(\delta_i + \gamma_a + \gamma_i), \beta_I/(\delta_h + \gamma_r), \beta_P/(\gamma_p + \gamma_a + \gamma_i)]\)

again, \( R_0 = N_1(w \circ b) \)

3.5. A Tuberculosis transmission model

This example is a model for Tuberculosis Blower et al. (1995), and it is interesting because the infectious individuals can recover temporarily and become infectious again, alternating between the two states (see Figure 12). In all states there is natural mortality and in some there is additional disease-induced mortality. The states are: \( S \), susceptible; \( L \), latent; \( T_i \), infected and infectious; \( T_n \), infected but non-infectious; \( R_i \), infectious individuals that become recovered and non-infectious for a while; \( R_n \), non-infectious individuals that become recovered for a while. All these states are transient, and the transition matrix \( U \) is:

\[
U = \begin{bmatrix}
S & E & I & H & P \\
S & 0 & 1 & 0 & 0 & 0 \\
E & 0 & 0 & p_1 & 0 & p_2 \\
I & 0 & 0 & 0 & \gamma_a \gamma_i + \delta_i & 0 \\
H & 0 & 0 & 0 & 0 & 0 \\
P & 0 & 0 & 0 & \gamma_a \gamma_i + \delta_i & 0
\end{bmatrix}
\]
Figure 12: Blower’s model for Tuberculosis transmission (Blower et al., 1995). Only individuals in $T_i$ are infectious. This state communicates with state $R_i$ and individuals can alternate between states $T_i$ and $R_i$ until removal.
The expression for $N_1$, the first row of the fundamental matrix $N$ is cumbersome and not shown here. On the other hand, since only $T_i$ is infectious:

$$w' = [0, 0, (c + \mu + \mu_t)^{-1}, 0, 0, 0]$$

$$b' = [0, 0, \lambda, 0, 0]$$

The result follows.

4. Discussion

We have presented a procedure that, given today’s computational capabilities, has as its most complicated step the construction of the transition matrix $U$. Most epidemic models start with a flux diagram between states with a differential equation for each arrow. In our approach, we substitute differential equations by probabilities. It is also very intuitive.

The basic requirement for this approach is that the transition probabilities are constant, so that $U$ does not contain the number of individuals in any stage, only parameters. There are models in which $U$ is not constant and thus, this method can not be applied, for instance Feng et al. (2000) suggested an SEIT model in which individuals in state $E$ can die at a constant rate $\mu$ or can move to state $I$ at a rate $cI/N$, therefore, the probability of moving to state $I$ is:

$$\frac{cI}{cI + N\mu}.$$
Therefore, we can not use a Markov chain approach to this model. This sort of models are not very common, and as it was shown in an earlier, less efficient version of the approach presented here (see Hernandez-Suarez, 2002), $R_0$ in Feng’s model is not maximized at the onset of the epidemic, that is, the first individual infected does not have the largest infectious potential.

4.1. Comparison with matrix population growth models

There is a correspondence between this approach and the Matrix Population Models theory for population growth (see Caswell, 2009), a stochastic version of deterministic modeling developed by Leslie (1945, 1948). If infections are considered births, and every individual remains in a stage one unit of time per visit, then, we can build a matrix of fertility $F$ containing values $f_{jk}$ where $f_{jk}$ is the number of individuals of class $j$ produced by an individual in class $k$ per unit of time. Observe that in general $f_{jk}$ is the contact rate of individuals in infectious compartment $k$, that produces individuals that move to compartment $j$ after infection.

Since the Leslie model assumes that every individual spends one unit of time in a compartment, we need to take this into account so that $f_{jk} = \lambda_k / \mu_k$, where $\lambda_k$ is the infectious rate of an individual in stage $k$ and $1/\mu_k$ is the average time in stage $k$ per visit. Then:

$$A = FN'$$

is called the next generation operator, since $x(k + 1) = A x(k)$, where $x(k)$ is the vector containing the number of individuals in each stage at generation $k$. Therefore, the dominant eigenvalue of $A$ is the production of infections of an infected individual during its lifetime. For instance, in the Ebola model in Figure 10, the fertility matrix would be:
\[
\begin{pmatrix}
        S & E & I & H & F \\
        S & 0 & 0 & 0 & 0 \\
        E & 0 & 0 & \frac{\beta_I}{\gamma_i} & \frac{\beta_H}{\gamma_H} & \frac{\beta_F}{\gamma_F} \\
        I & 0 & 0 & 0 & 0 \\
        H & 0 & 0 & 0 & 0 \\
        F & 0 & 0 & 0 & 0
\end{pmatrix}
\]

Observe that the row corresponding to stage \(E\) is the one that contains all births, since all newly infected are born to that compartment.

After obtaining Matrix \(\mathbf{A} = \mathbf{FN}'\) with \(\mathbf{N}\) already calculated in (5), we get that the dominant eigenvalue of \(\mathbf{A}\) is:

\[
\frac{\beta_I}{\gamma_i} + \frac{\beta_H}{\gamma_H} + \frac{\beta_F}{\gamma_F} \left( \frac{\theta_2 \theta_1 + \delta_1 (1 - \theta_1)}{\gamma_F} \right).
\]

which is the same as the \(R_0\) in (6).

### 4.2. Comparison with traditional mathematical methods

Diekmann et al. (1990) introduced a method to calculate \(R_0\) based on the construction of a matrix called the next generation matrix, NGM. Details to construct the NGM are given in more detail in Diekmann et al. (2010). The idea is that \(R_0\) is the dominant eigenvalue of the product \(-\mathbf{T}\Sigma^{-1}\), where \(\mathbf{T}\) is equivalent to our matrix of fertility \(\mathbf{F}\) above, and \(\Sigma\) is a matrix that contains elements \(s_{jk}\) with the transition rates from state \(k\) to state \(j\) if \(j \neq k\), and the sum of the exit rates (as negative values) if \(j = k\).

The equivalence of this approach with the matrix population growth above and the probabilistic approach presented here is straightforward: there is a direct relationship between \(\Sigma\) and \(\mathbf{U}\), our transition probability matrix. Let \(\mathbf{W}\) a diagonal matrix containing the average time in each stage. Thus, \(\mathbf{W}\) is equal to the inverse of the diagonal elements of \(\Sigma\) with positive sign. Then:

\[
\Sigma \mathbf{W} - \mathbf{I} = \mathbf{U}'
\]
that is:

\[ \Sigma = (U' + I)W^{-1} \]  \hspace{1cm} (7)

where \( U' \) is the transpose of our transition matrix described before. After a little algebra we arrive to:

\[ -\Sigma^{-1} = (I - U')^{-1}W = N'W \]

Recall \( N' \) contains the expected number of visits to each stage and when multiplied by \( W \) gives the expected times. Thus, \(-\Sigma^{-1}\) effectively contains the expected time on each stage for every initial stage. Therefore, \(-T\Sigma^{-1}\) is similar to \( A \), the matrix population growth model of the previous section and it is called the next generation matrix.

For instance, using again the Ebola model in Figure 10, following Diekmann et al. (2010) we have:

\[
\Sigma = \begin{bmatrix}
E & I & H & F \\
E & -\alpha & 0 & 0 & 0 \\
I & \alpha & -\gamma_i & 0 & 0 \\
H & 0 & \gamma_i\theta_1 & -\gamma_H & 0 \\
F & 0 & \gamma_i\delta_1(1 - \theta_1) & \delta_2\gamma_H & -\gamma_F \\
\end{bmatrix}
\]

and

\[
T = \begin{bmatrix}
E & I & H & F \\
E & 0 & \beta_i & \beta_H & \beta_F \\
I & 0 & 0 & 0 & 0 \\
H & 0 & 0 & 0 & 0 \\
F & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

The dominant eigenvalue of \(-T\Sigma^{-1}\) yields the same \( R_0 \) as in (6). Notice this method does avoid the susceptible stage, therefore, the matrices involved with this method are one dimension shorter.
It is remarkable the relationship between $\Sigma$ and $U$ given in (7). Clearly, $\Sigma$ is easier to construct and would allow to obtain the expected times on each stage $N'W$ more quickly, but lacks of a simple probabilistic intuition that we adopt here.

5. Bibliography


If an individual produces $X$ descendants with probability $P_X$ during its life, then $P_e$, the probability of extinction of the population that starts with one individual, is:

$$P_e = (P_e | X = 0)P_0 + (P_e | X = 1)P_1 + (P_e | X = 2)P_2 + \cdots$$

But by independence of fates of individuals, $(P_e | X = k) = P_k e$, thus,

$$P_e = P_0 + P_e P_1 + P_e^2 P_2 + \cdots = \sum_{i=0}^{\infty} P_e^i P_i = M_X(P_e)$$

that is, the probability of extinction equates the Moment Generating Function of $X$, the offspring production. Thus, the probability of extinction $P_e$ is the solution of $z = M_X(z)$. This system has always the solution $z = 1$ and it may exist a second solution in $0 < z < 1$. The probability of extinction is the minimum value of all the solutions in $[0, 1]$. Branching process theory asserts that if the average offspring size is smaller than one, the probability of extinction is one since on average, at generation 1 there will be $R_0$ individuals; at generation 2, $R_0^2$; at generation 3, $R_0^3$ and in general, at generation $n$, $R_0^n$, which tends to 0 when $n \to \infty$ and $R_0 < 1$. To prove that the mean offspring size is the relevant factor to determine whether extinction is certain or not, is straightforward: observe that we are looking or the solutions of $z$ such that $M_X(z) = z$ in the interval $0 \leq z \leq 1$, and observe that in this interval, the second derivative of $M_X(z)$ is

$$\frac{d^2}{dz^2} M_X(z) = \sum_{i=2}^{\infty} i(i-1)z^{i-2}P(X = i) > 0$$


6. Appendix

6.1. Appendix A1

If an individual produces $X$ descendants with probability $P_X$ during its life, then $P_e$, the probability of extinction of the population that starts with one individual, is:

$$P_e = (P_e | X = 0)P_0 + (P_e | X = 1)P_1 + (P_e | X = 2)P_2 + \cdots$$

But by independence of fates of individuals, $(P_e | X = k) = P_k e$, thus,

$$P_e = P_0 + P_e P_1 + P_e^2 P_2 + \cdots = \sum_{i=0}^{\infty} P_e^i P_i = M_X(P_e)$$

that is, the probability of extinction equates the Moment Generating Function of $X$, the offspring production. Thus, the probability of extinction $P_e$ is the solution of $z = M_X(z)$. This system has always the solution $z = 1$ and it may exist a second solution in $0 < z < 1$. The probability of extinction is the minimum value of all the solutions in $[0, 1]$. Branching process theory asserts that if the average offspring size is smaller than one, the probability of extinction is one since on average, at generation 1 there will be $R_0$ individuals; at generation 2, $R_0^2$; at generation 3, $R_0^3$ and in general, at generation $n$, $R_0^n$, which tends to 0 when $n \to \infty$ and $R_0 < 1$. To prove that the mean offspring size is the relevant factor to determine whether extinction is certain or not, is straightforward: observe that we are looking or the solutions of $z$ such that $M_X(z) = z$ in the interval $0 \leq z \leq 1$, and observe that in this interval, the second derivative of $M_X(z)$ is

$$\frac{d^2}{dz^2} M_X(z) = \sum_{i=2}^{\infty} i(i-1)z^{i-2}P(X = i) > 0$$
that is, \( M_X(z) \) is a concave function. Since \( M_X(0) = P_0 \) and \( M_X(1) = 1 \), the shape of \( M_X(z) \) is one of the dotted line depicted in Figure 13. Observe also that

\[
\frac{d}{dz} M_X(z) \big|_{z=1} = \sum_{i=2}^{\infty} i P(X = i) = E[X]
\]

that is, the slope of \( M_X(z) \) at \( z = 1 \) is the average offspring size, therefore, to intersect the \( z \) line in another point in \( 0 < z < 1 \), it is necessary the slope at \( z = 1 \) (average offspring size) be larger than 1.

Figure 13: Plot of the Moment Generating Function \( M_X(z) \). We are interested in the minimum of the intersection points with \( z \), that is, the probability of extinction. Since \( M_X(z) \) is concave and goes always through \((0, 0)\) and \((1, P_0)\), if the slope of \( M_X(z) \) at \( z = 1 \) is larger than one, there is always another solution of \( M_X(z) = z \) in \((0, 1)\). Since \( M_X'(1) \) is by definition the expected value of \( X \), then, if \( E(X) > 1 \), the probability of extinction is less than 1.
6.2. Appendix A2

Let \((I - U)x = 0\); that is \(x = Ux\). Iterating this we can see that \(x = U^nx\).

Since \(\lim_{n \to \infty} U^n \to 0\) we have \(U^n x \to 0\), so \(x = 0\). Thus \((I - U)^{-1}\) exists.

The probability that the process will be in every state after \(n\) steps is \(U^n\), thus, the expected number of visits to each state after \(n\) transitions is:

\[
I + U + U^2 + U^3 + \ldots + U^n
\]

observe that:

\[
(I - U)(I + U + U^2 + U^3 + \ldots + U^n) = I - U^{n+1}
\]

multiplying both sides by \((I - U)^{-1}\) yields:

\[
I + U + U^2 + U^3 + \ldots + U^n = (I - U)^{-1}(I - U^{n+1})
\]

taking the limit when \(n \to \infty\) yields:

\[
\sum_{i=0}^{\infty} U^i = (I - U)^{-1} \equiv N
\]