On the exact measure of the disease spread in stochastic epidemic models

Jesús Artalejo
Faculty of Mathematics
Complutense University of Madrid
Madrid, Spain

jesus_artalejo@mat.ucm.es

DSABNS 2013, Lisbon, February 13-15

Talk Schedule

1. Introduction
2. The SIS and SIR stochastic epidemic models
3. The exact reproduction number R_{e0}
4. The population transmission number R_p
5. Applications to the control of the infectious disease
6. Conclusions and References

This talk is based on the paper:

Artalejo, J.R. and Lopez-Herrero, M.J. (2013)

On the exact measure of the disease spread in stochastic epidemic models

Bulletin of Mathematical Biology (under revision).

1. Introduction

Stochastic epidemic models

• Andersson and Britton (2000) conclude that stochastic models are to be preferred when their analysis is possible.

• **Symplifying assumptions**: exponential distribution, homogeneous populations, random-mixing.

• **Small communities** include hospitals (Forrester and Pettitt 2005; Artalejo and Lopez-Herrero 2011; Wang et al. 2011), educational establishments (Stone et al. 2008; Artalejo et al. 2010), prisons (Hotta 2010) or small herds (de Keijer et al. 2008).

• **Stochastic techniques**: Markov chains, branching and diffusion processes.

The basic reproduction number R_0

It is a common practice to define R_0 as follows (Heesterbeek and Dietz 1996, Hethcote 2000):

R_0 is the expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual during its entire period of infectiousness.

Most remarkable features of R_0 (Roberts 2007; Li et al. 2011)

(i) the threshold value of R_0 that establishes that an infection persists only if $R_0 > 1$,

(ii) the usefulness of R_0 as a direct measure of the control effort required to eradicate the infection.

Methods to calculate R_0 include among others the survival function, the next-generation method, the eigen values of the Jacobian matrix and the constant term of the characteristic polynomial (Li et al. 2011, comparative analysis).

Flows of R_0

(i) Few of the calculation methods agree with each other.

(ii) Few of which produce the true average number of secondary infections.

(iii) In the case of seasonal changes, large epidemics can happen even if $R_0 < 1$ and the final size may not be an increasing function of R_0 (Bacaër and Gomes 2009).

Li, Blakeley and Smith? (2011) provide an excellent **overview** including exhaustive discussion of the problems with R_0, the comparative analysis of methods for its calculation and an examination of a number of alternatives to R_0.

Remark: Most of the existing works assume a *deterministic point of view* to calculate R_0. In particular, linearization is assumed to ignore that the infectious disease itself diminishes the availability of susceptibles (Diekmann and Heesterbeek 2000, Sect. 5.1). The consequence is that R_0 is clearly not the exact expected number of secondary cases.

Goal of this talk: We present two alternative measures, namely

R_{e0}: the exact reproduction number,

R_p: the population transmission number.

Both quantities do not count the number of contacts affecting to individuals which have been previously infected. As a result, R_{e0} and R_p correct the effect of the repeated contacts that R_0 overestimates.

2. The SIS stochastic model

- Closed population model of N individuals.
  - Classified either as susceptible or infective individual.

- Susceptible can be infected, then they recover and return to the susceptible pool.

- Evolution of the epidemic:
  - Birth and death process $\{I(t);\, t \geq 0\}$.
  - $I(t)$: number of infective individuals at time t.
  - $S = \{0, 1, ..., N\}$ (0 is an absorbing state).

- Classical SIS rates
  - Infection rate $\lambda_i = \frac{\beta}{N} i(N-i)$.
  - Recovery rate $\mu_i = \gamma i$.

- $R_0 = \frac{\beta}{\gamma}$ denotes the classical reproduction number.

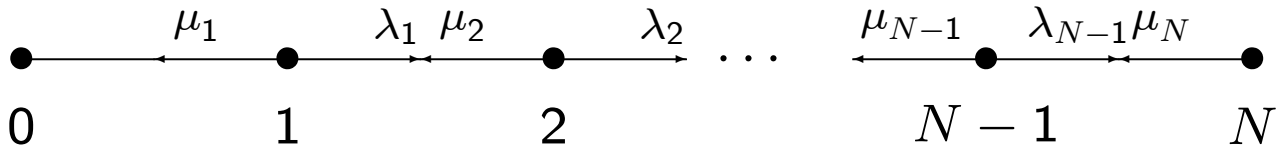


Figure 1. States and transitions of the birth and death model

The SIR stochastic model

- Closed population model of N individuals.
  - Classified as susceptible, infective or removed individuals.

- Susceptible can be infected, then they recover and become immune.

- Evolution of the epidemic:
  - Bidimensional CTMC $\{(I(t), S(t));\, t \geq 0\}$.
  - $I(t)$: number of infective individuals at time t ($I(0) = m \geq 1$).
  - $S(t)$: number of susceptibles at time t ($S(0) = n$).
  - $S = \{(i, j); 0 \leq j \leq n, 0 \leq i \leq m+n-j\}$ ($S_T = \{(0, j); 0 \leq j \leq n\}$).

- Classical SIR rates
  - Infection rate $\lambda_{ij} = \frac{\beta}{N} ij$.
  - Recovery rate $\mu_i = \gamma i$.

- $R_0 = \frac{\beta}{\gamma}$ is the classical reproduction number.

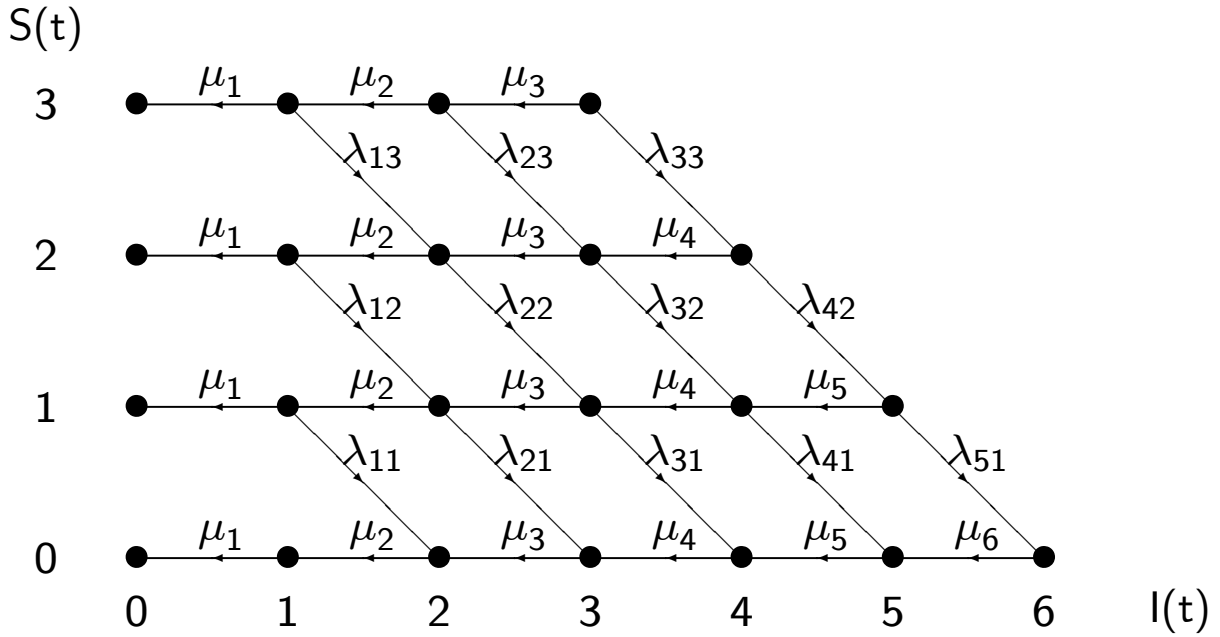


Figure 2. States and transitions of the SIR epidemic model

3. The exact reproduction number R_{e0}

R_{e0} is the exact number of secondary cases produced by a typical infective individual during its entire infectious period.

Differences between R_0 and R_{e0}

• R_{e0} does not count the number of contacts taking place between the typical infective individual and any contacted individual which has previously been infected. Then, the real measurement of the disease spread is obtained.

• R_0 is only defined at the time of invasion, when the typical infective is introduced into a completely susceptible population. In contrast, R_{e0} can be defined at all times.

• R_{e0} is a random variable, rather than an expected number, so we can study its whole probability distribution (i.e., probability mass function, expectation and higher order moments).

By assuming that the invasion starts at $t = 0$ with $I(0) = 1$, the expectation $\overline{R}_{e0} = E[R_{e0} \,|I(0) = 1]$ amounts to the exact quantification of the disease transmission aimed in the definition of R_0.

The SIS model

We decompose the contact rate as $\beta_i = \beta_i^* + \widetilde{\beta}_{i-1}$, for $1 \leq i \leq N$, where

$\beta_i^* = \frac{\beta}{N}(N-i)$: the individual rate at which the typical infective contacts with the susceptible population,

$\widetilde{\beta}_{i-1} = \frac{\beta}{N}(i-1)(N-i)$: the superposition of the contact rates of the remaining $i-1$ infectives.

For a given i, $0 \leq i \leq N$, we define

$$\varphi_i(z) = E\left[z^{R_{e0}} \,|I(0)=i\right] = \sum_{k=0}^{\infty} z^k P\left\{R_{e0}=k \,|I(0)=i\right\}, \, |z| \leq 1.$$

$$m_i^k = E\left[R_{e0}(R_{e0}-1)\cdots(R_{e0}-k+1)\,|I(0)=i\right], \; 1 \leq i \leq N, \; k \geq 0$$
$(m_i^0 = 1)$.

Four possibilities for the next event:

i) recovery of the tagged (typical) infective,

ii) recovery of a non-tagged infective,

iii) effective contact between the tagged infective and any susceptible individual,

iv) effective contact between a non-tagged infective and any susceptible.

Conditioning on the first transition, we have

$$\varphi_i(z) = \frac{\gamma}{\beta_i + \gamma_i} + \frac{\gamma(i-1)}{\beta_i + \gamma_i}\varphi_{i-1}(z) + \frac{\beta_i^*}{\beta_i + \gamma_i} z\varphi_{i+1}(z) + \frac{\widetilde{\beta}_{i-1}}{\beta_i + \gamma_i}\varphi_{i+1}(z),\ 1 \leq i \leq N.$$

By differentiating equations for $\varphi_i(z)$ $k \geq 1$ times, and setting $z = 1$, we get

$$m_i^k = \frac{\gamma(i-1)}{\beta_i + \gamma_i} m_{i-1}^k + \frac{\beta_i}{\beta_i + \gamma_i} m_{i+1}^k + \frac{\beta_i^*}{\beta_i + \gamma_i} k m_{i+1}^{k-1},$$
$$1 \leq i \leq N,\ k \geq 1.$$

In particular, $\overline{R}_{e0} = m_1^1$.

We next summarize a recursive scheme for the computation of $\left\{m_i^1;\ 1 \leq i \leq N\right\}$. The proposed scheme only deals with algebraic operations involving positive terms, which guarantees that the computation is stable.

Minor modifications lead to the computation of higher order moments.

Theorem 1 *The expected values* $\left\{m_i^1;\ 1 \leq i \leq N\right\}$ *are computed by the equations*

$$m_N^1 = \frac{(N-1)D_{N-1}}{Na_{N-1} + \beta_{N-1}},$$
$$m_i^1 = \frac{D_i + \beta_i m_{i+1}^1}{a_i + \beta_i},\ i = N-1, ..., 1,$$

where the coefficients a_i *and* D_i*, for* $1 \leq i \leq N-1$*, are given by*

$$a_1 = \gamma,$$
$$a_i = \frac{\gamma(ia_{i-1} + \beta_{i-1})}{a_{i-1} + \beta_{i-1}},\ 2 \leq i \leq N-1,$$
$$D_1 = \beta_1^*,$$
$$D_i = \beta_i^* + \frac{\gamma(i-1)D_{i-1}}{a_{i-1} + \beta_{i-1}},\ 2 \leq i \leq N-1.$$

A direct computation of the probabilities $x_i^k = P\{R_{e0} = k \,|I(0) = i\}$, for $1 \leq i \leq N$ and $k \geq 0$, can be done with the help of the following equations:

$$x_i^k = \delta_{k0}\frac{\gamma}{\beta_i + \gamma_i} + \frac{\gamma(i-1)}{\beta_i + \gamma_i}x_{i-1}^k + (1 - \delta_{k0})\frac{\beta_i^*}{\beta_i + \gamma_i}x_{i+1}^{k-1} + \frac{\widetilde{\beta}_{i-1}}{\beta_i + \gamma_i}x_{i+1}^k,\ 1 \leq i \leq N,\ k \geq 0.$$

The SIR model

We define the generating functions

$$\varphi_{ij}(z) = \sum_{k=0}^{j} z^k P\left\{R_{e0} = k \mid (I(0), S(0)) = (i,j)\right\},\ (i,j) \in S_T,\ |z| \leq 1,$$

which verify the following triangular system of equations:

$$\varphi_{ij}(z) = \frac{\gamma}{\beta_{ij}+\gamma_i} + \frac{\gamma(i-1)}{\beta_{ij}+\gamma_i}\varphi_{i-1,j}(z) + \frac{\beta^*_{ij}}{\beta_{ij}+\gamma_i} z\varphi_{i+1,j-1}(z) + \frac{\widetilde{\beta}_{i-1,j}}{\beta_{ij}+\gamma_i}\varphi_{i+1,j-1}(z),\ (i,j) \in S_T,$$

where $\beta^*_{ij} = \frac{\beta}{N} j$ and $\widetilde{\beta}_{i-1,j} = \frac{\beta}{N}(i-1)j$, for $(i,j) \in S_T$.

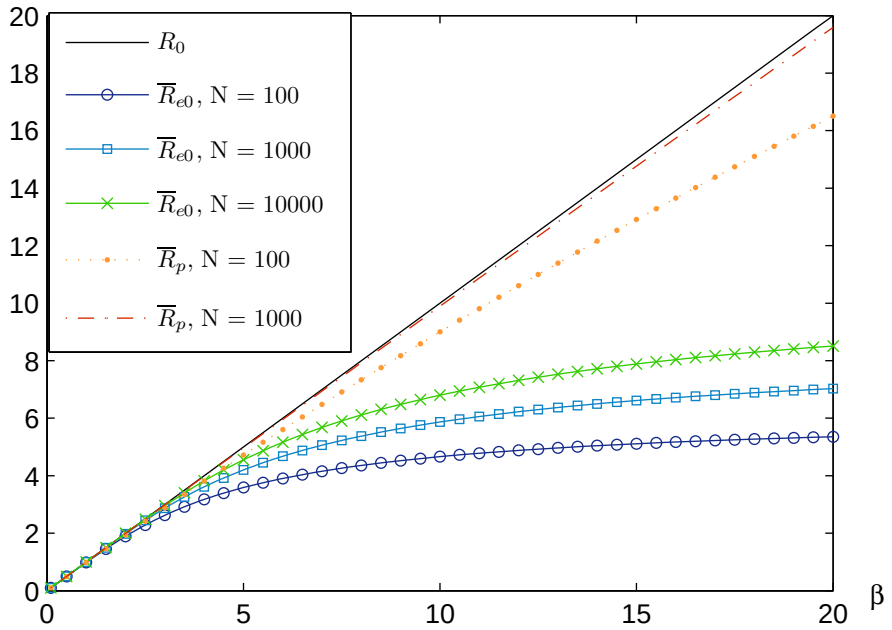
After appropriate differentiation, we derive recursive equations for the factorial moments m_{ij}^k

$$m_{ij}^k = \frac{\gamma(i-1)}{\beta_{ij}+\gamma_i} m_{i-1,j}^k + \frac{\beta_{ij}}{\beta_{ij}+\gamma_i} m_{i+1,j-1}^k$$
$$+ \frac{\beta_{ij}^*}{\beta_{ij}+\gamma_i} k m_{i+1,j-1}^{k-1},\ (i,j) \in S_T,\ k \geq 1.$$

For $k = 0$, we observe that $m_{ij}^0 = P\{R_{e0} < \infty \,|(I(0), S(0)) = (i,j)\} = 1$, for $(i,j) \in S_T$. Then, the above equations can be efficiently calculated in the recursive order $k \geq 1$, $0 \leq j \leq n$ and $1 \leq i \leq m+n-j$.

A direct method for the recursive computation of the probability mass function of R_{e0} (i.e., $x_{ij}^k = P\{R_{e0} = k \,|(I(0), S(0)) = (i,j)\}$, for $0 \leq k \leq j$) can also be developed.

Figure 3 R_0, $\overline{R}_{e0}$ and $\overline{R}_p$ versus the contact rate β. We assume an SIS model with $\gamma = 1$ so $R_0 = \beta$, and $N = 100, 1000, 10000$.



As far as β increases and N decreases, we observe higher differences between R_0 and $\overline{R}_{e0}$. When $\beta = 20$, we find that $\overline{R}_{e0} = 5.35$, for $N = 100$, and $\overline{R}_{e0} = 8.50$, for $N = 10000$.

4. The population transmission number R_p

R_p is the exact number of secondary cases produced by all currently infective individuals prior to the first recovery.

Differences between R_0 and R_p

- R_p aims to count the exact number of effective contacts, irrespectively of the identity of the infective involved in each contact, during the elapsed time until the first recovery occurs.

- R_p can be defined either at the invasion time or at any time after the invasion.

If $t = 0$ and $I(0) = 1$, then we denote its expected value as $\overline{R}_p = E\left[R_p \,|I(0) = 1\right]$.

The general birth-and-death process in $S = \{0, 1, ..., N\}$

Dynamics of the generating functions

$$\Psi_i(z) = \frac{\mu_i}{\lambda_i + \mu_i} + (1 - \delta_{iN})\frac{\lambda_i}{\lambda_i + \mu_i} z \Psi_{i+1}(z), \ 1 \leq i \leq N.$$

Calculation of the kth factorial moments

$$M_i^k = (1 - \delta_{i,N-1}) k \sum_{l=i}^{N-2} M_{l+1}^{k-1} \prod_{j=i}^{l} \frac{\lambda_j}{\lambda_j + \mu_j} + \delta_{k1} \prod_{j=i}^{N-1} \frac{\lambda_j}{\lambda_j + \mu_j},$$
$$1 \leq i \leq N-1, \ k \geq 1,$$

The boundary equations $M_i^0 = 1$, $1 \leq i \leq N$, and $M_N^k = 0$, $k \geq 1$, respectively reflect that $P\{R_p < \infty \,|I(0) = i\} = 1$, for $1 \leq i \leq N$, and $R_p = 0$, when $I(0) = N$.

The mass probability function

$$y_i^k = P\{R_p = k \,|I(0) = i\} = \frac{\mu_{i+k}}{\lambda_{i+k} + \mu_{i+k}} \prod_{l=i}^{i+k-1} \frac{\lambda_l}{\lambda_l + \mu_l},\ 0 \le k \le N - i.$$

Moreover, for $i = N$, we have $y_N^0 = 1$.

The SIS model (particular case)

The expected values, M_i^1, and the variances, $\sigma_i^2 = M_i^2 + M_i^1 - (M_i^1)^2$, are reduced to the simple formulas

$$M_i^1 = \frac{(N-i)R_0}{R_0 + N},\ 1 \le i \le N,$$
$$\sigma_i^2 = \frac{N(N-i)R_0((N-i+1)R_0 + N)}{(2R_0 + N)(R_0 + N)^2},\ 1 \le i \le N.$$

The SIR model

The mass probabilities $y_{ij}^k = P\{R_p = k \,|I(0) = i, S(0) = j\}$, for $i \geq 1$ and $j \geq 1$, are given by

$$y_{ij}^k = \frac{\mu_{i+k}}{\lambda_{i+k,j-k} + \mu_{i+k}} \prod_{l=i}^{i+k-1} \frac{\lambda_{l,i+j-l}}{\lambda_{l,i+j-l} + \mu_l}, \; 0 \leq k \leq j.$$

The expected values and the variances

$$M_{ij}^1 = \frac{jR_0}{R_0 + N}, \; (i,j) \in S_T,$$
$$\sigma_{ij}^2 = \frac{NjR_0((j+1)R_0 + N)}{(2R_0 + N)(R_0 + N)^2}, \; (i,j) \in S_T.$$

5. Applications to the control of the infectious disease

Control of infectious diseases includes vaccination, isolation and culling (Keeling and Rohani 2008).

Vaccination based on R_0 reduces $R_0 > 1$ to $1 - v$ (fraction of susceptibles) such that

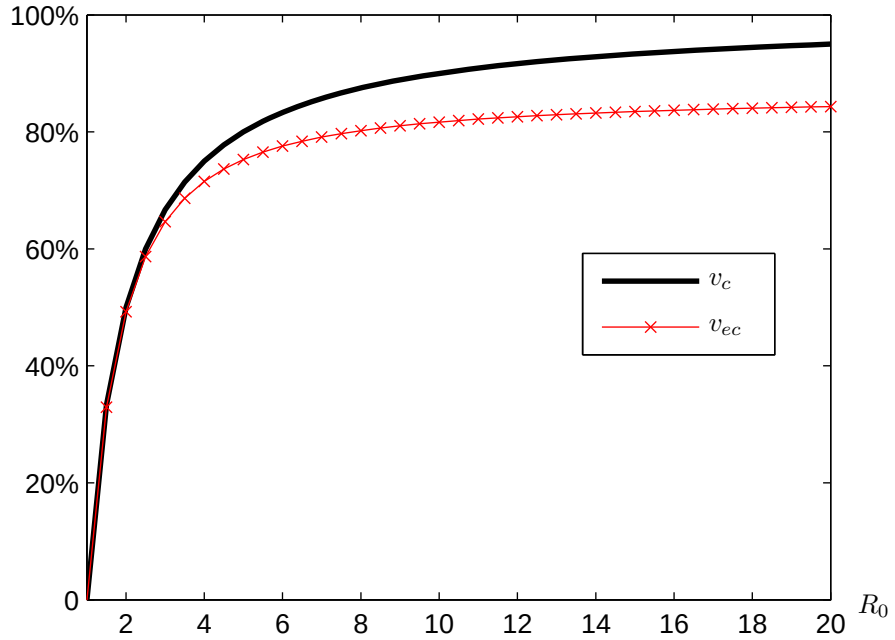
$$R_0(1 - v) < 1 \longrightarrow v_c = 1 - \frac{1}{R_0}.$$

The *critical vaccination coverage* v_c is interpreted in terms of the critical proportion of the population that should be vaccinated to prevent the spread of the epidemic.

Replacing R_0 by the expected value $\overline{R}_{e0}$, we get

$$v_{ec} = \begin{cases} 1 - \frac{1}{\overline{R}_{e0}}, & \text{if } \overline{R}_{e0} > 1, \\ 0, & \text{otherwise.} \end{cases}$$

Figure 4a Critical vaccination levels v_c and v_{ec} versus R_0. We assume an SIR model with $(m, n) = (1, 999)$ and $\gamma = 1$ so $R_0 = \beta$.



Example of the reduction obtained. $R_0 = 10$ (e.g. chickenpox) then $\overline{R}_{e0} = 5.45712$. As a result, we obtain $v_c = 0.9$ (i.e., 90% of the susceptible population should be vaccinated) and $v_{ec} = 0.81$.

The eradication of the disease requires a random time which depends not only on the transmission speed but also on its variability. We notice that the variance of R_p increases as a function of j. In fact, the variance is 0, if $j = 0$. Mass vaccination of the entire susceptible population seems unfeasible due to logistical difficulties and cost constraints.

In what follows, we use the probability mass function of R_p of the SIR model to determine a vaccination strategy going beyond the levels v_{ec} and v_c.

R_p has a decreasing probability mass function if and only if $R_0 < N$. The idea now is to guarantee that the decay of the disease spread is controlled by a fixed level $\rho \in (0, 1)$.

To reach this goal, for a fixed (i, j), we look for the large number of susceptibles j such that

$$y_{ij}^{k+1} < \rho y_{ij}^{k}, \ (\rho\text{-control decay of } R_p).$$

But it amounts to

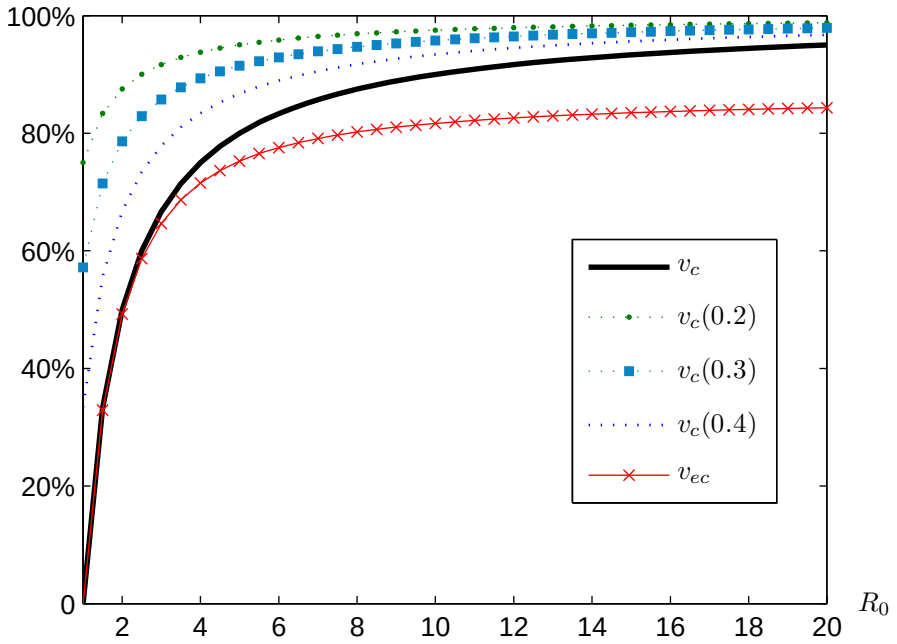
$$v_c(\rho) = 1 - \frac{\rho}{1-\rho}\left(\frac{1}{R_0} - \frac{1}{N}\right).$$

It is now easy to check that

$$v_c(\rho) > v_c \Leftrightarrow \rho < \left(2 - \frac{R_0}{N}\right)^{-1},$$

so we propose to employ the vaccination policy based on the ρ-*control decay* of R_p only if $R_0 \in (1, N)$ and $\rho < (2 - R_0/N)^{-1}$. In such a case, the critical level $v_c(\rho)$ provides not only the coverage required to prevent a major outbreak but also a control on the decay of the distribution of the random variable R_p.

Figure 4b The critical vaccination coverage $v_c(\rho)$ for the levels $\rho = 0.2, 0.3, 0.4$.



For a fixed R_0, $v_c(\rho)$ is decreasing with ρ, reflecting that a smaller level ρ implies a stronger control of the decay.

For $R_0 = 5$, we notice that $v_c(0.2) = 0.95$, $v_c(0.3) = 0.91$ and $v_c(0.4) = 0.86$, while $v_c = 0.8$ and $v_{ec} = 0.75$.

6. Conclusions and References

• The basic reproduction number is the most useful concept of the mathematical epidemiology. Despite its fundamental role and the apparent simplicity of its definition, R_0 is an intricate concept. When a deterministic approach is adopted, the impact of the depletion of susceptibles due to the infection process is neglected. As a result, R_0 counts for excess the reproductive potential.

• This talk presents two alternative measures called the exact reproduction number, R_{e0}, and the population transmission number, R_p, which provide a real measurement of the spread of a disease represented by a stochastic epidemic model.

• Since we deal with Markov chains with a finite population size, the epidemic dies out in a finite time with probability one. This fact is indeed an important difference between the deterministic and the stochastic approaches. Despite that R_0, R_{e0} and R_p are not actually a threshold, the three measures are valuable to quantify the spread and severity of the infectious disease.

• The classical reproduction number, R_0, aims to measure the expected number of secondary cases. The alternative measures, R_{e0} and R_p, are random variables so we can calculate not only their expected values but also their complete distributions (i.e., probability mass functions, higher order moments).

• R_0 is only defined at the time of invasion, when a typical infective is introduced into a virgin population. In contrast, R_{e0} and R_p can be defined at the initial invasion time or at any later time.

• When control measures are possible, R_0 provides a measure of the effort needed to prevent a major outbreak. Once the repeated contacts are discounted, the expected value $\overline{R}_{e0}$ contributes to the vaccination policy by providing the exact vaccination coverages v_{ec}. The use of the probability mass function of R_p leads to a coverage level $v_c(\rho)$, which provides not only the preventive coverage required but also a control on the decay of the distribution of the number of secondary infections.

References

Andersson, H., & Britton, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis.* Springer Lecture Notes in Statistics 151. Springer Verlag: New York.

Artalejo J.R., Economou, A., & Lopez-Herrero, M.J. (2010). On the number of recovered individuals in the SIS and SIR stochastic epidemic models. *Math. Biosci., 228,* 45-55.

Artalejo, J.R., & Lopez-Herrero, M.J. (2011). The SIS and SIR stochastic epidemic models: A maximum entropy approach. *Theor. Popul. Biol., 80,* 256-264.

Bacaër, N., & Gomes, M.G.M. (2009). On the final size of epidemics with seasonality. *Bull. Math. Biol., 71,* 1954-1966.

de Koejier, A.A., Diekmann, O., & de Jong, M.C.M. (2008). Calculating the extinction of a reactivating virus, in particular bovine herpes virus. *Math. Biosci., 212,* 111-131.

Diekmann, O., & Heesterbeek, J.A.P. (2000). *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation.* Wiley Series in Mathematical and Computational Biology. Chichester: Wiley.

Forrester, M., & Pettitt, A.N. (2005). Use of stochastic epidemic modeling to quantify transmission rates of colonization with methicillin-resistant *Staphylococcus aureus* in an intensive care unit. *Infect. Control Hosp. Epidemiol., 26,* 598-606.

Heesterbeek, J.A.P., & Dietz, K. (1996). The concept of R_0 in epidemic theory. *Stat. Neerl., 50,* 89-110.

Hethcote, H.W. (2000). The mathematics of infectious diseases. *SIAM Rev., 42,* 599-653.

Hotta, L.K. (2010). Bayesian melding estimation of a stochastic $SEIR$ model. *Math. Popul. Stud., 17,* 101-111.

Keeling, M.J., & Rohani, P. (2008). *Modeling Infectious Diseases in Humans and Animals.* Princeton University Press: Princeton.

Li, J., Blakeley, D., & Smith?, R.J. (2011). The failure of R_0. *Comput. Math. Methods Med.* Vol. 2011, Article ID 527610.

Roberts, M.G. (2007). The pluses and minuses of R_0. J*.R. Soc. Interface, 4,* 946–961.

Stone, P., Wilkinson-Herbots, H., & Isham, V. (2008). A stochastic model for head lice infections. *J. Math. Biol., 56,* 743-763.

Wang, J., Wang, L., Magal, P., Wang, Y., Zhuo, J., Lu, X., & Ruan, S. (2011). Modelling the transmission dynamics of methicillin-resistant *Staphylococcus aureus* in Beijing Tongren hospital. *J. Hosp. Infect., 79,* 302-308.

THANKS!