Determination of Important Disease Spread Characteristics Using Markov Chain Models

Zeynep Gökçe Yıldız

Refik Güllü

Wolfgang Hörmann

Industrial Engineering Department
Boğaziçi University

July 10, 2018

Outline

▶ Introduction

▶ Markov Stochastic SIS Disease Spread Models

▶ Markov Stochastic SIR Disease Spread Models

▶ Models with Erlang Disease Time

▶ Numerical Results

▶ Conclusion

Markov Stochastic Disease Spread Models

▶ **The important epidemiological quantities to be analyzed**

1. Time to extinction
2. Quasi stationary distribution
3. Probability of an outbreak
4. Final outbreak size distribution
5. The distribution of maximum number of infected individuals during disease

Model Definition

▶ The disease duration for an infected $\sim$ Exponential distribution (μ)

▶ The number of contacts $\sim$ Poisson distribution (λ)

▶ $S + I = N$

▶ State space $\rightarrow \Omega = I(t)$

▶ Continuous time Markov Chain having property

$$Prob(I(t_{n+1})|I(t_0), I(t_1)..., I(t_n)) = Prob(I(t_{n+1}|I(t_n)).$$

Important epidemiological quantities

▶ Markov chain diagram

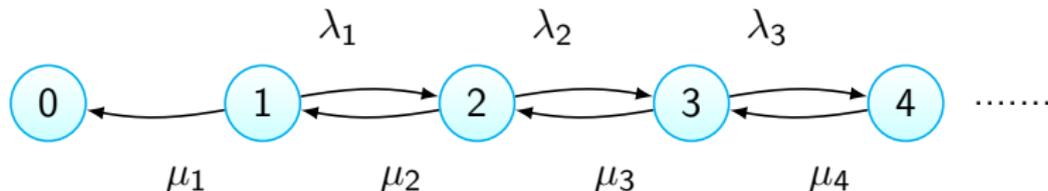

Figure: Markov chain diagram for SIS disease spread models

▶ Transition rates $\mu_i = i\mu$ and $\lambda_i = \lambda q_i$ where

$$q_i = ki(N - i).$$

▶ One absorbing state $\rightarrow$ (0)

▶ Interested in expected time to extinction

Expected time to extinction

▶ Possible to return to a state previously visited

▶ τ_k denoting the time to go from state k to state $k-1$

▶ Expected time to extinction starting with i infected

$$E[T_i] = \sum_{k=1}^{i} E[\tau_k].$$

Conditioning on next event

▶ Define

$$I_k = \begin{cases} 1, & \text{if first state change is due to a recovery.} \\ 0, & \text{if first state change is due to an infection.} \end{cases}$$

▶ $E[\tau_k]$ conditioning on state change

$$E[\tau_k | I_k = 1] = \frac{1}{\lambda_k + \mu_k},$$

$$E[\tau_k | I_k = 0] = \frac{1}{\lambda_k + \mu_k} + E[\tau_{k+1}] + E[\tau_k]$$

▶ where the probability

$$P[I_k = 1] = \frac{\mu_k}{\lambda_k + \mu_k} \quad \text{and} P[I_k = 0] = \frac{\lambda_k}{\lambda_k + \mu_k}.$$

$E[T_i]$

▶ $E[\tau_k]$

$$E[\tau_k] = \frac{1}{\mu_k} + \frac{\lambda_k}{\mu_k} E[\tau_{k+1}].$$

▶ For $k = N$, exact value of $E[\tau_k]$

$$E[\tau_N] = \frac{1}{\mu_N}$$

▶ Exact value of $E[\tau_k]$

$$E[\tau_k] = \frac{1}{\mu_k} + \sum_{j=1}^{N-k} \frac{1}{\mu_{k+j}} \prod_{l=0}^{j-1} \frac{\lambda_{l+k}}{\mu_{l+k}}.$$

$$E[T_i] = \sum_{k=1}^{i} E[\tau_k].$$

Model Definition

▶ The disease duration for an infected $\sim$ Exponential distribution (μ)

▶ The transition rate from infected to recovered $i\mu$

▶ The number of contacts for every infected individual $\sim$ Poisson distribution (λ)

▶ The total number of contacts $\sim$ Poisson distribution $(is\lambda/N)$ (Hernandez Suarez et al., 2010; Nasell, 2002)

▶ The transition rate from susceptible to infected $is\lambda/N$

Model Definition

▶ $N = S + I + R$

▶ $R = N - I - S$

▶ $\Omega = (I(t), S(t))$

▶ Continuous time Markov Chain having property

$$Prob((I(t_{n+1}), S(t_{n+1}))|(I(t_0), S(t_0)), ..., (I(t_n), S(t_n))) =$$
$$Prob((I(t_{n+1}), S(t_{n+1}))|((I(t_n), S(t_n))).$$

Markov Chain Diagram for SIR
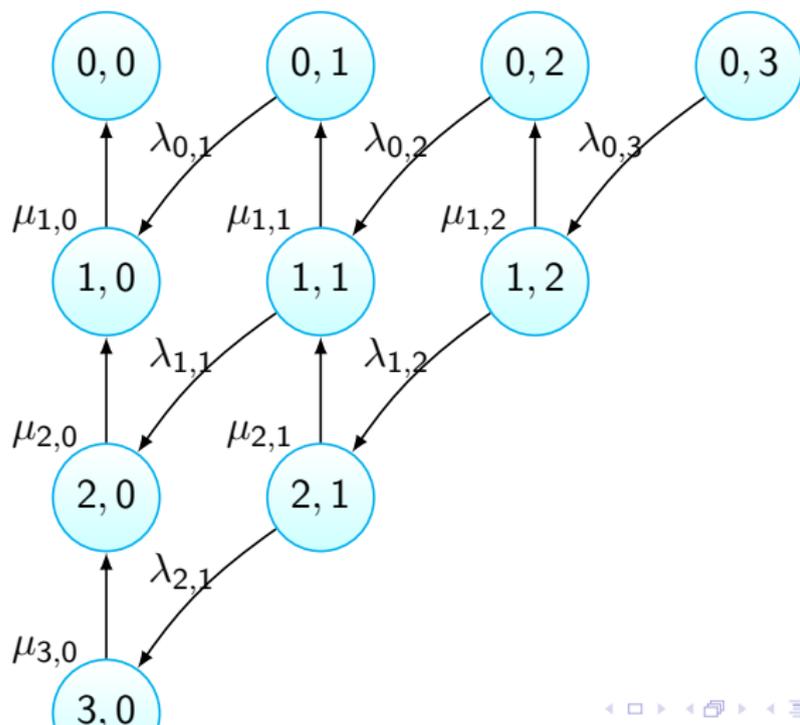
Expected time to extinction

▶ Impossible to return a state previously visited

▶ Expected time to extinction starting with i infected and s susceptible, $E[T_{i,s}]$ calculated directly

▶ Absorbing states $=> (0, s)$ for $s = 0, 1, ...N$

▶ Boundary Conditions

$$T_{0,s} = 0 \quad \text{for s=1....N.}$$

$E[T_{i,s}]$ for states with zero susceptible

$$E[T_{i,0}] = \sum_{k=1}^{i} \frac{1}{\mu_k}.$$

Implementation of First Step Analysis

▶ Start with calculate $E[T_{i,s}]$ for states with one susceptible and one infected

$$E[T_{1,1}] = \frac{1}{\lambda_{1,1} + \mu_i} + \frac{\lambda_{1,1}}{\lambda_{1,1} + \mu_i} E[T_{2,0}] + \frac{\mu_1}{\lambda_{1,1} + \mu_1} E[T_{0,1}].$$

▶ By increasing the number of infected individuals one by one, calculate $E[T_{i,1}]$ for all i values.

$$E[T_{i,1}] = \frac{1}{\lambda_{i,1} + \mu_i} + \frac{\lambda_{i,1}}{\lambda_{i,1} + \mu_i} E[T_{i+1,0}] + \frac{\mu_i}{\lambda_{i,1} + \mu_i} E[T_{i-1,1}].$$

$E[T_{i,s}]$ for s susceptible cases

▶ Calculate $E[T_{i,s}]$ by increasing the number of susceptible cases one by one

$$E[T_{i,s}] = \frac{1}{\lambda_{i,s} + \mu_i} + \frac{\lambda_{i,s}}{\lambda_{i,s} + \mu_i} E[T_{i+1,s-1}] + \frac{\mu_i}{\lambda_{i,s} + \mu_i} E[T_{i-1,s}]$$

▶ Replace $\lambda_{i,s}$ by $is\lambda/N$ and μ_i by $i\mu$

$$E[T_{i,s}] = \frac{N}{is\lambda + Ni\mu} + \frac{s\lambda}{s\lambda + N\mu} E[T_{i+1,s-1}] + \frac{N\mu}{s\lambda + N\mu} E[T_{i-1,s}]$$

Calculation of Final Outbreak Size

▶ Let's define τ

$$\tau = \inf\{t > 0 : I(t) = 0\}.$$

▶ $m \Longrightarrow$ total number of recovered individuals at τ

▶ To determine the final outbreak size distribution, calculate

$$P_m(i, s) = \Pr\{R(\tau) = m \quad | \quad (I(0), S(0)) = (i, s)\}.$$

Implementation of First Step Analysis

▶

$$P_m(i, s) = \frac{\lambda_{is}}{\lambda_{is} + \mu_i} P_m(i+1, s-1) + \frac{\mu_i}{\lambda_{is} + \mu_i} P_m(i-1, s).$$

$$P_m(i, s) = \frac{\lambda s}{\lambda s + \mu N} P_m(i+1, s-1) + \frac{\mu N}{\lambda s + \mu N} P_m(i-1, s).$$

▶ **Absorbing states** $=> (0, s)$ for $s = 0, 1, ... N$

▶ **Boundary Conditions**

$$P_m(0, s) = 1 \quad \text{for s=N-m} \quad \text{and} \quad P_m(0, s) = 0 \quad \text{for } s < N - m.$$

Calculation of Final Outbreak Size

▶ First calculate $P_m(i, s)$ for $s = N - m$

$$
\begin{aligned}
P_m(i, N - m) &= \frac{\mu N}{\lambda(N - m) + \mu N} P_m(i - 1, N - m) \\
&= \left\{ \frac{\mu N}{\lambda(N - m) + \mu N} \right\}^i
\end{aligned}
$$

▶ Then calculate $P_m(i, s)$ recursively by increasing the number of susceptible

$$
P_m(i, s) = \frac{s\lambda}{\lambda s + \mu N} P_m(i + 1, s - 1) + \frac{\mu N}{\lambda s + \mu N} P_m(i - 1, s).
$$

Algorithm 1 Final Size Distribution for Exponential Disease Time

1: Set m=1
2: *Set* $P_m(i,s) = 0$ *for* $s < N - m$ *and* $i = 0, 1, ...N - s$
3: *Set* $s = N - m$
4: *Set* $P_m(0,s) = 1$
5: **for** i=1,..N-s **do**
6: *Compute* $P_m(i,s)$ *from Equation* 1
7: **end for**
8: *Set* $s = s + 1$
9: *Set* $P_m(0,s) = 0$
10: **for** i=1,..N-s **do**
11: *Compute* $P_m(i,s)$ *from Equation* 1
12: **end for**
13: *Set* $s = s+1$. If s ≤N-1 go to step 9. Otherwise, go to step 14.
14: *Set* $m = m + 1$. If m≤N go to step 2. Otherwise, stop.

Distribution of maximum number of infected individuals

- $m =>$ maximum number of infected individuals until disease disappears
- $Q_m(i, s)$ denotes the probability that the maximum number of infected individuals becomes m starting with i infected and s susceptible
- **Absorbing states** $=> (0, s)$ for $s = 0, 1, ... N$
- **Boundary Conditions**

$$Q_m i, s = 0 \quad \text{for m=0,...i-1.}$$

Implementation of First Step Analysis

▶ First calculate $Q_m(i, s)$ for $m = N - s$

$$Q_{N-s}(i, s) = \frac{\lambda s}{\lambda s + \mu N} Q_m(i + 1, s) = \frac{\lambda s}{\lambda s + \mu N}^{N-s-i}.$$

▶ Then calculate $Q_m(i, s)$ recursively by decreasing m one by one until it becomes i+1

$$Q_m(i, s) = \frac{s\lambda}{\lambda s + \mu N} Q_m(i + 1, s - 1) + \frac{\mu N}{\lambda s + \mu N} Q_m(i - 1, s).$$

▶ Finally, calculate $Q_m(i, s)$ for $m = i$

$$Q_i(i, s) = 1 - \sum_{n=i+1}^{N-s} Q_n(i, s)$$

Problems with exponential disease time

- ▶ Lack of versatility with single parameter value
- ▶ Memoryless property
- ▶ To overestimate variance

Why Erlang disease time?

- ▶ Member of a versatile class of distribution
- ▶ Approximate to normal distribution
- ▶ Different coefficient of variation
- ▶ To allow Markov chains model

Model Definition

▶ $N = S + I + R$

▶ Not possible to use a single I for the number of infected individuals

▶ Representation for Erlang distribution

$$\alpha = (1, 0, ..0)$$

$$T = \begin{bmatrix} -\mu & \mu & 0 & 0 & \dots & 0 & 0 \\ 0 & -\mu & \mu & 0 & \dots & 0 & 0 \\ 0 & 0 & -\mu & \mu & \dots & 0 & 0 \\ \vdots & \ddots & & & & & \\ 0 & 0 & 0 & 0 & \dots & -\mu & \mu \\ 0 & 0 & 0 & 0 & \dots & 0 & -\mu \end{bmatrix}$$

.

Model Definition

- $\Omega = (\tilde{I}(t), S(t))$
- $\tilde{I}(t) => \{(i_1(t), i_2(t), i_3(t), ... i_k(t)) : \sum i_n \leq N \quad for \quad n \leq k\}$
- The number of states required to define model
  $=> (N + k + 1)!/(k + 1)!N!$
- Define $I(t)$ as

$$I(t) = \sum_{n=1}^{k} i_k$$

Calculation of Final Outbreak Size

To determine the final outbreak size distribution, calculate

$$P_m(\tilde{i}, s) = P_m(\{i_1, i_2, i_3, ...i_k\}, s) = \Pi_m(v, s).$$

Implementation of First Step Analysis

▶ $1_n = 0, 0, ..1, ...0$

$$P_m(\tilde{i}, s) = \frac{\lambda_{is}}{\lambda_{is} + \mu_i} P_m(\tilde{i}+1_k, s-1) + \sum_{n=1}^{k} \frac{\mu_{i_n}}{\lambda_{is} + \mu_i} P_m(\tilde{i}-1_n, s).$$

▶ **Absorbing states** $=> (0, s)$ for $s = 0, 1..N$

▶ **Boundary Conditions**

$$P_m(0, s) = 1 \quad \text{for s=N-m} \quad \text{and} \quad P_m(0, s) = 0 \quad \text{for s<N-m}.$$

Zeynep Gökce Yıldız

Determination of Important Disease Spread Characteristics Using Markov Chain Models   26 / 33

Computational Efficiency

▶ Time to find final outbreak size distribution increases like $kN^{(k+1)}$

▶ However, use of different modeling approach increases time to find final outbreak size like kN^2

Table: Time Required to Calculate Exact Final Outbreak Size Distribution

	time to perform calculation with		
	population size N		
k	100	500	1000
2	2.02 s	240.52 s	1886.14 s
5	5.10 s	614.49 s	4828.98 s
10	10.45 s	1271.74 s	10228.08 s

Zeynep Gökce Yıldız

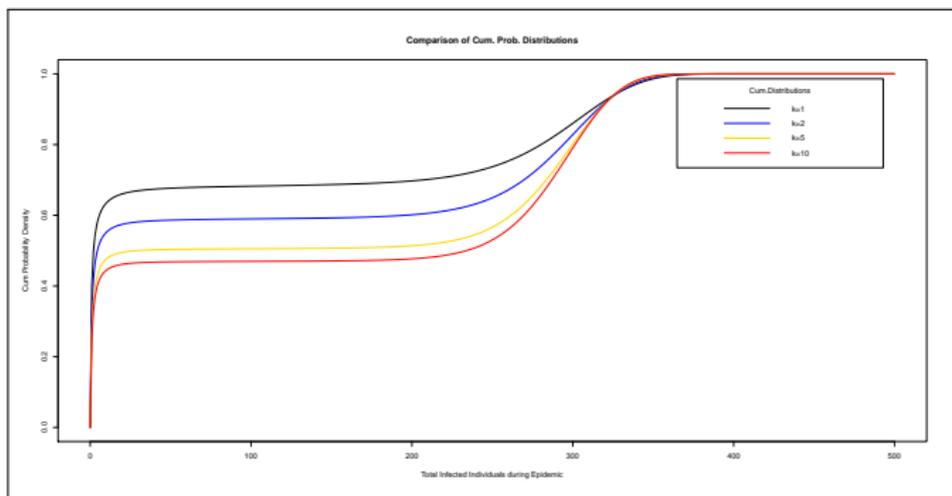Determination of Important Disease Spread Characteristics Using Markov Chain Models

Figure: The cumulative probability function of total number of infected individuals during an epidemic for $R_0 = 1.5$.
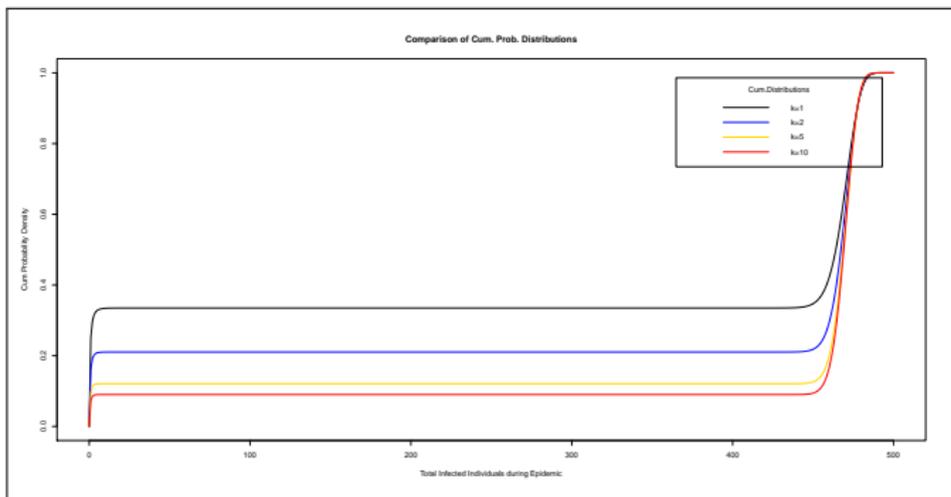
Figure: The cumulative probability function of total number of infected individuals during an epidemic for $R_0 = 3$.

Table: Probability that Less Than Fifty Individuals Recovered for Different R_0 and k Values

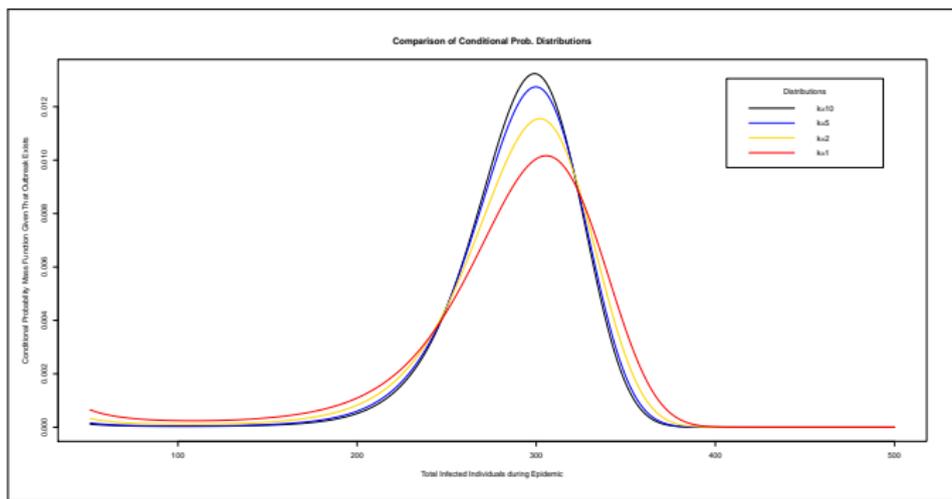|	k			
R_0	1	2	5	10
1	0.9353716	0.9278719	0.921807	0.9194013
1.5	0.6765616	0.5862919	0.5033751	0.4681467
2	0.5042293	0.3856372	0.2868653	0.2479226
2.5	0.4022905	0.2771721	0.1806280	0.1450215
3	0.3348617	0.2104195	0.1209246	0.0900515

Zeynep Gökce Yıldız

Figure: The Conditional probability function of total number of infected individuals during an epidemic given that total number of infected individuals greater than fifty for $R_0 = 1.5$.

Figure: The Conditional probability function of total number of infected individuals during an epidemic given that total number of infected individuals greater than fifty for $R_0 = 3$.

THANK YOU..