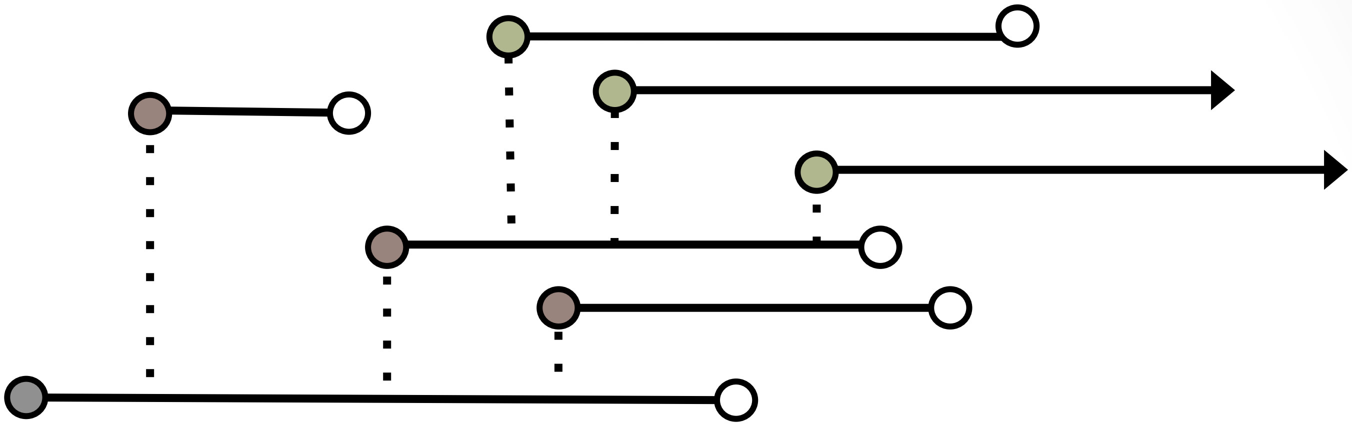


Basic stochastic simulation models

Clinic on Meaningful Modeling of Epidemiological Data
African Institute for the Mathematical Sciences
Muizenberg, South Africa
31 May 2016

Rebecca Borchering, BS, MS
Department of Mathematics and Emerging Pathogens Institute
University of Florida

Model taxonomy

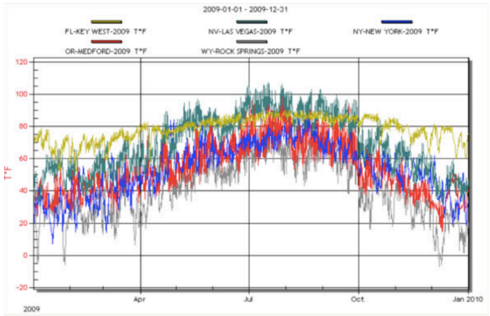
	CONTINUOUS TREATMENT OF INDIVIDUALS (averages, proportions, or population densities)	DISCRETE TREATMENT OF INDIVIDUALS
DETERMINISTIC	CONTINUOUS TIME • Ordinary differential equations • Partial differential equations DISCRETE TIME • Difference equations (eg, Reed-Frost type models)	
STOCHASTIC	CONTINUOUS TIME • Stochastic differential equations DISCRETE TIME • Stochastic difference equations	CONTINUOUS TIME • Gillespie algorithm DISCRETE TIME • Chain binomial type models (eg, Stochastic Reed-Frost models)

Why stochastic?

- Small populations, extinction

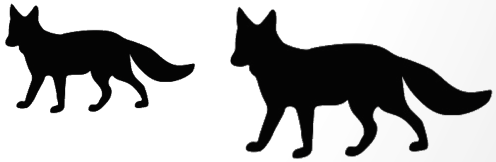
- Noisy data
  - imperfect observation
  - small samples

- Environmental stochasticity
  - long term variation in external drivers
  - changes in rates, including birth and death rates

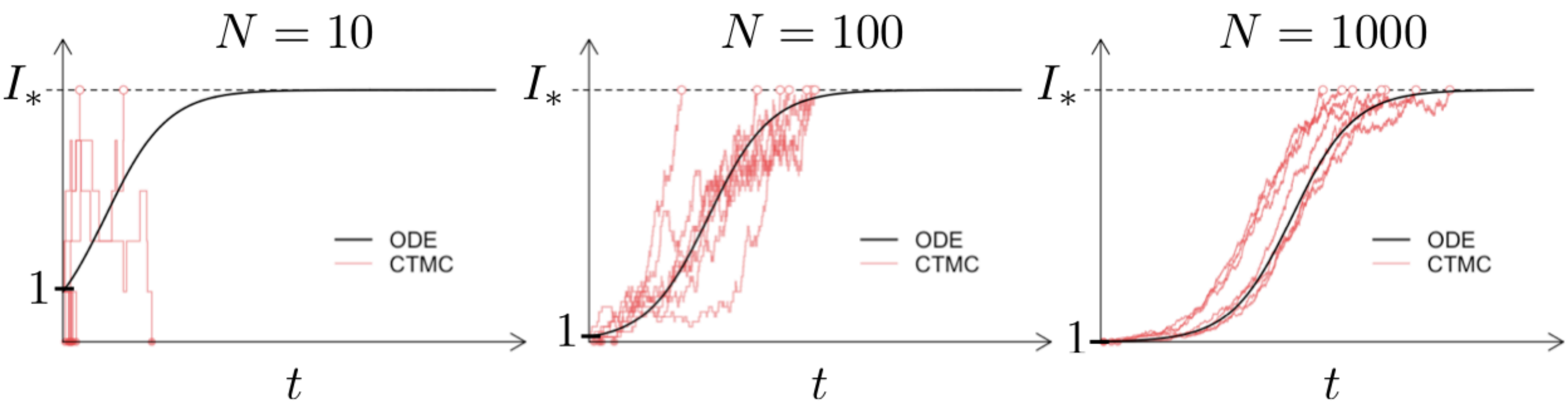


www.imagepermanenceinstitute.org

- Demographic stochasticity
  - comes out of having discrete individuals



Population size - N



Continuous Time Markov Chain (CTMC)

- **finite** population size
- stochastic

Ordinary Differential Equation (ODE)

- large (**infinite**) population size
- deterministic

The Reed-Frost model

Cases/Infected

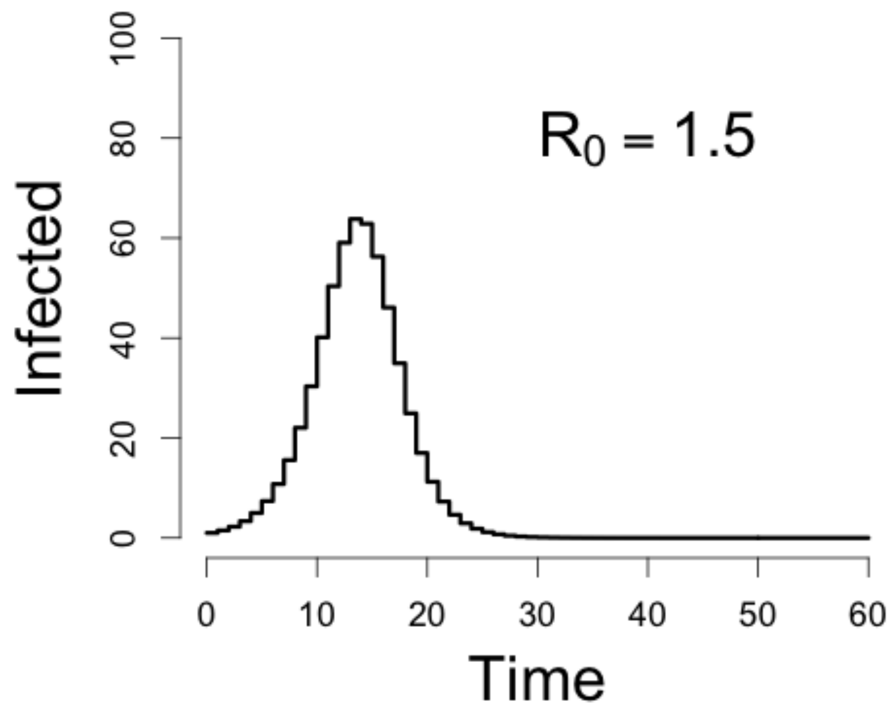
$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

Susceptible

$$S_{t+1} = S_t - C_{t+1}$$

Recovered

$$R_{t+1} = R_t + C_t$$



The Reed-Frost model

- The probability of getting infected by any infectious individual is

$$1 - q^{C_t}$$

- The expected number of cases in the next time unit is

$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

Reed-Frost (Chain Binomial)

- Fixed infectious period duration
- Generations of infectious individuals don't overlap
- Define

p probability of infection

C_t cases at time t

S_t susceptibles at time t

The Reed-Frost model

- For each susceptible individual, at time t:

prob. not getting infected by any infectious individual

$$1 - (1 - p)^{C_t}$$

prob. not getting infected by a particular infectious individual

prob. of getting infected by any infectious individual

- Let $q = 1 - p$
- So the probability of getting infected by any infectious individual is $1 - q^{C_t}$

The Reed-Frost model

- The probability of getting infected by any infectious individual is

$$1 - q^{C_t}$$

- The expected number of cases in the next time unit is

$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

- Susceptible individuals in the next time unit

$$S_{t+1} = S_t - C_{t+1}$$

- Recovered individuals in the next time unit

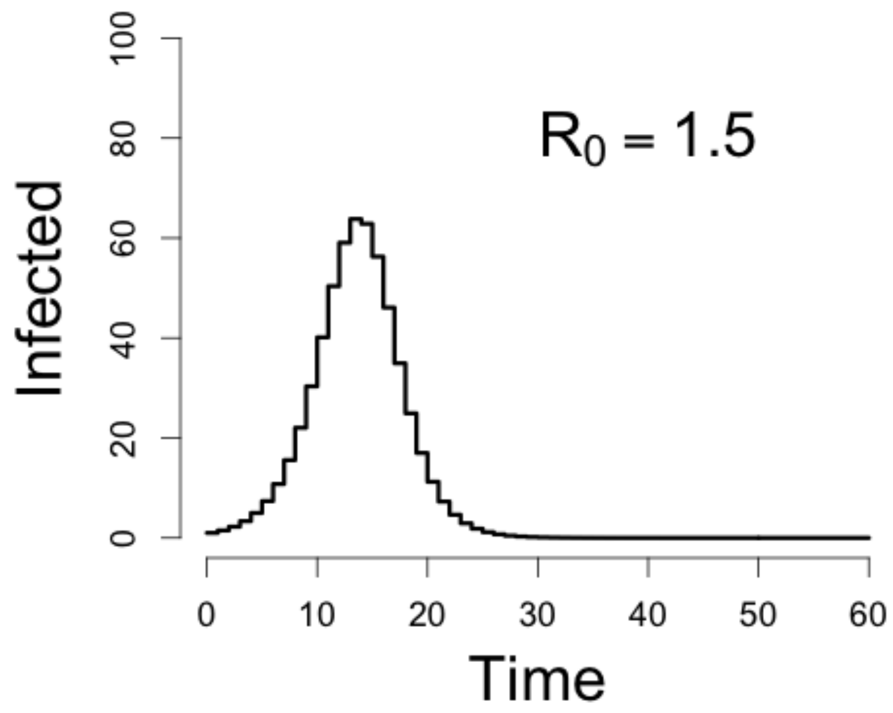
$$R_{t+1} = R_t + C_t$$

The Reed-Frost model

$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

$$S_{t+1} = S_t - C_{t+1}$$

$$R_{t+1} = R_t + C_t$$



The Reed-Frost model

- The full set of equations describing the deterministic population update is:

$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

$$S_{t+1} = S_t - C_{t+1}$$

$$R_{t+1} = R_t + C_t$$

- If $N = S + C + R$ is the total population size, the basic reproductive number for this model is

$$R_0 = (N-1)(1-q)$$

Building stochastic R-F model

- For each susceptible individual, at time t:

prob. not getting infected by any infectious individual

$$1 - (1 - p)^{C_t}$$

prob. not getting infected by a particular infectious individual

prob. of getting infected by any infectious individual

- Let $q = 1 - p$
- So the probability of getting infected by any infectious individual is $1 - q^{C_t}$

Building stochastic R-F model

- For each susceptible individual, at time t:

prob. of not getting infected by any infectious individual

$$q^{C_t}$$

prob. not getting infected by a particular infectious individual

- ($q = 1 - p$)

The stochastic R-F model

Putting it all together:

$$\mathbb{P}(C_{t+1} = x) = \binom{S_t}{x} \left(1 - q^{C_t}\right)^x \left(q^{C_t}\right)^{S_t - x}$$

- number of ways to choose x individuals (the binomial coefficient)
- prob. of x individuals getting infected by any infectious individual (the term $\left(1 - q^{C_t}\right)^x$)
- prob. of S_t - x individuals not getting infected by any infectious individual (the term $\left(q^{C_t}\right)^{S_t - x}$)

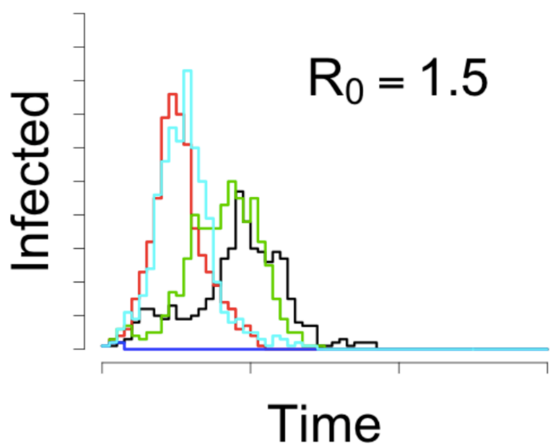
$$S_{t+1} = S_t - C_{t+1}$$
$$R_{t+1} = R_t + C_t$$

The Reed-Frost model

Stochastic:

$$\mathbb{P}(C_{t+1} = x) = \binom{S_t}{x} \left(1 - q^{C_t}\right)^x \left(q^{C_t}\right)^{S_t - x}$$

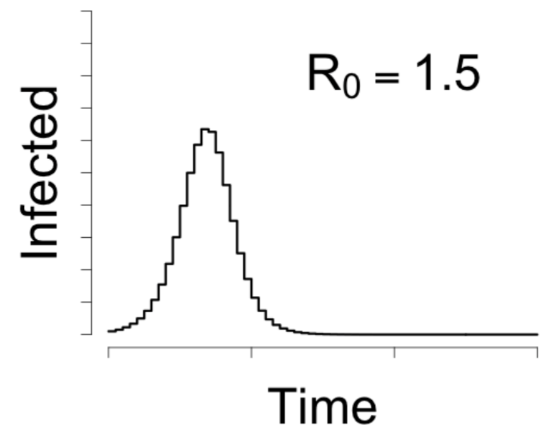
$$S_{t+1} = S_t - C_{t+1}$$
$$R_{t+1} = R_t + C_t$$



Deterministic:

$$C_{t+1} = S_t \left(1 - q^{C_t}\right)$$

$$S_{t+1} = S_t - C_{t+1}$$
$$R_{t+1} = R_t + C_t$$



Chain binomial models

- Chain binomial models can also be formulated based on the same parameters we used in the ODE models and with overlapping generations.
- Instantaneous hazard of infection for an individual susceptible individual is $\beta I/N$
- For a susceptible at time t, the probability of infection by time $t + \Delta t$ is
$$p = 1 - e^{-\frac{\beta I}{N}\Delta t}$$
- Similarly, for an infectious individual at time t, the probability of recovery by time $t + \Delta t$ is
$$r = 1 - e^{-\gamma \Delta t}$$

Chain binomial models

The stochastic population update can then be described as

X : new infectious individuals
Y : new recovered individuals

} random variables

$$S_{t+\Delta t} = S_t - X$$
$$I_{t+\Delta t} = I_t + X - Y$$
$$R_{t+\Delta t} = R_t + Y$$

$$\mathbb{P}(X = x) = \binom{S_t}{x} p^x (1-p)^{S_t - x}$$

$$\mathbb{P}(Y = y) = \binom{I_t}{y} r^y (1-r)^{I_t - y}$$

Chain binomial models

- For this model, if D is the average duration of infection, the basic reproductive number is:

$$R_0 = (N-1)\left(1 - e^{-\frac{\beta}{N}D}\right)$$

- Non-generation-based chain binomial models can be adapted to include many variations on the natural history of infection.

- Discrete-time simulation of chain binomials is far more computationally efficient than event-driven simulation in continuous time.

Chain binomial simulation

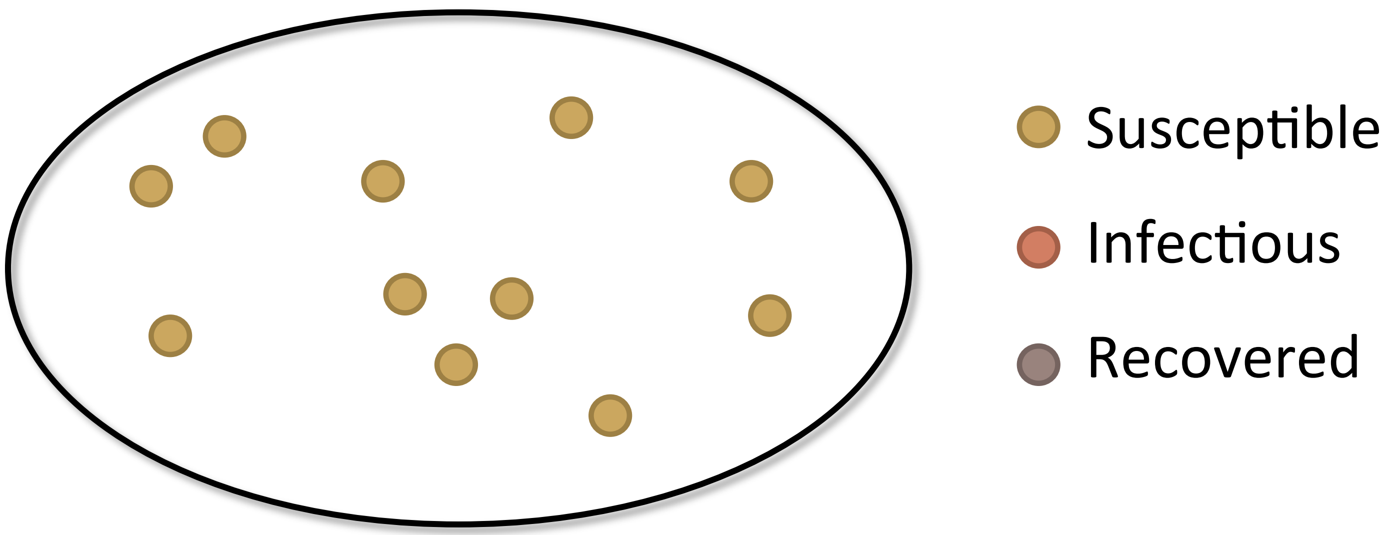
```
while (I > 0 and time < MAXTIME)
     Calculate transition probabilities
     Determine number of transitions for
     each type
     Update state variables
     Update time
end
```

Another way to simulate stochastic epidemics…

event-driven simulation

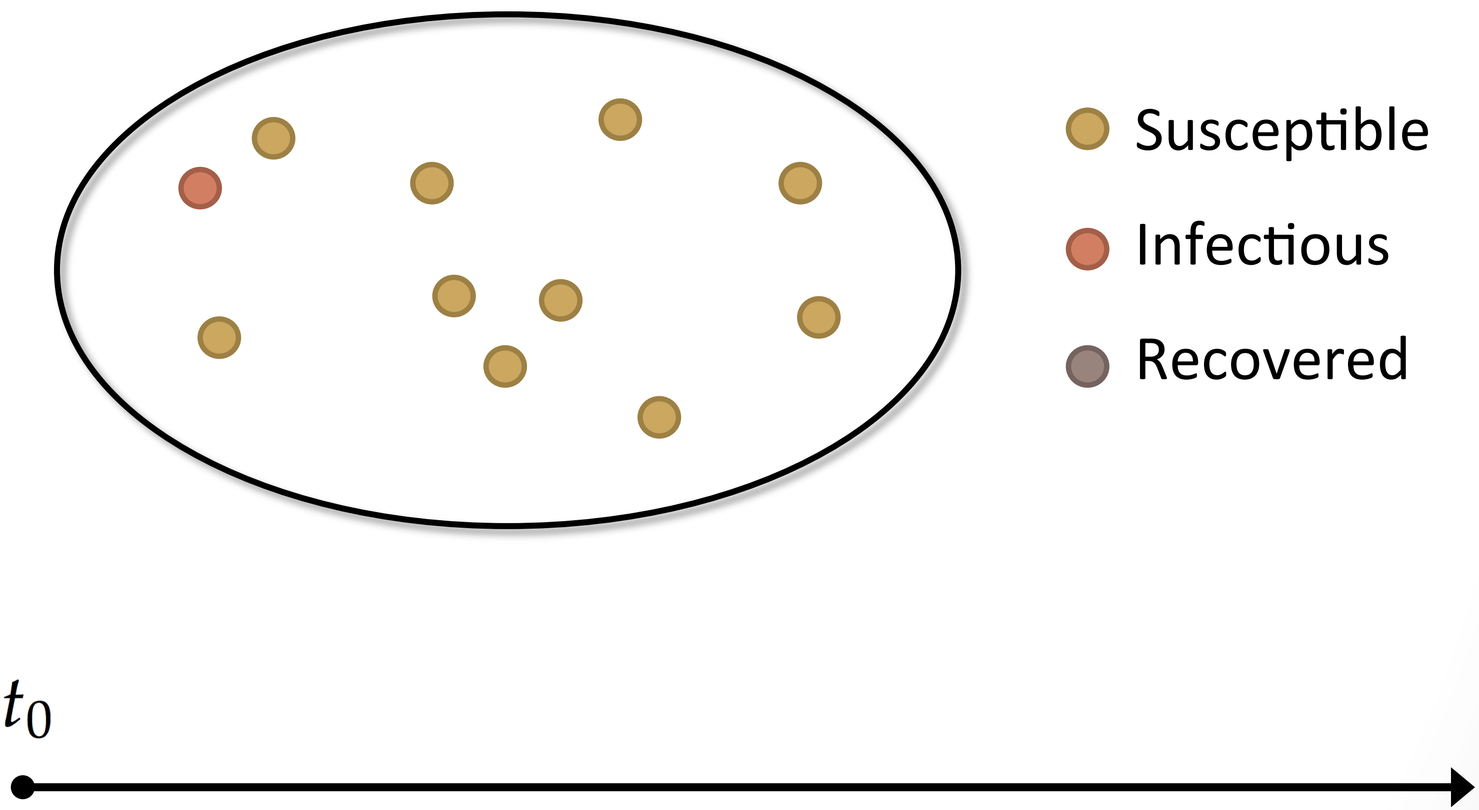
Stochastic SIR dynamics

Small population



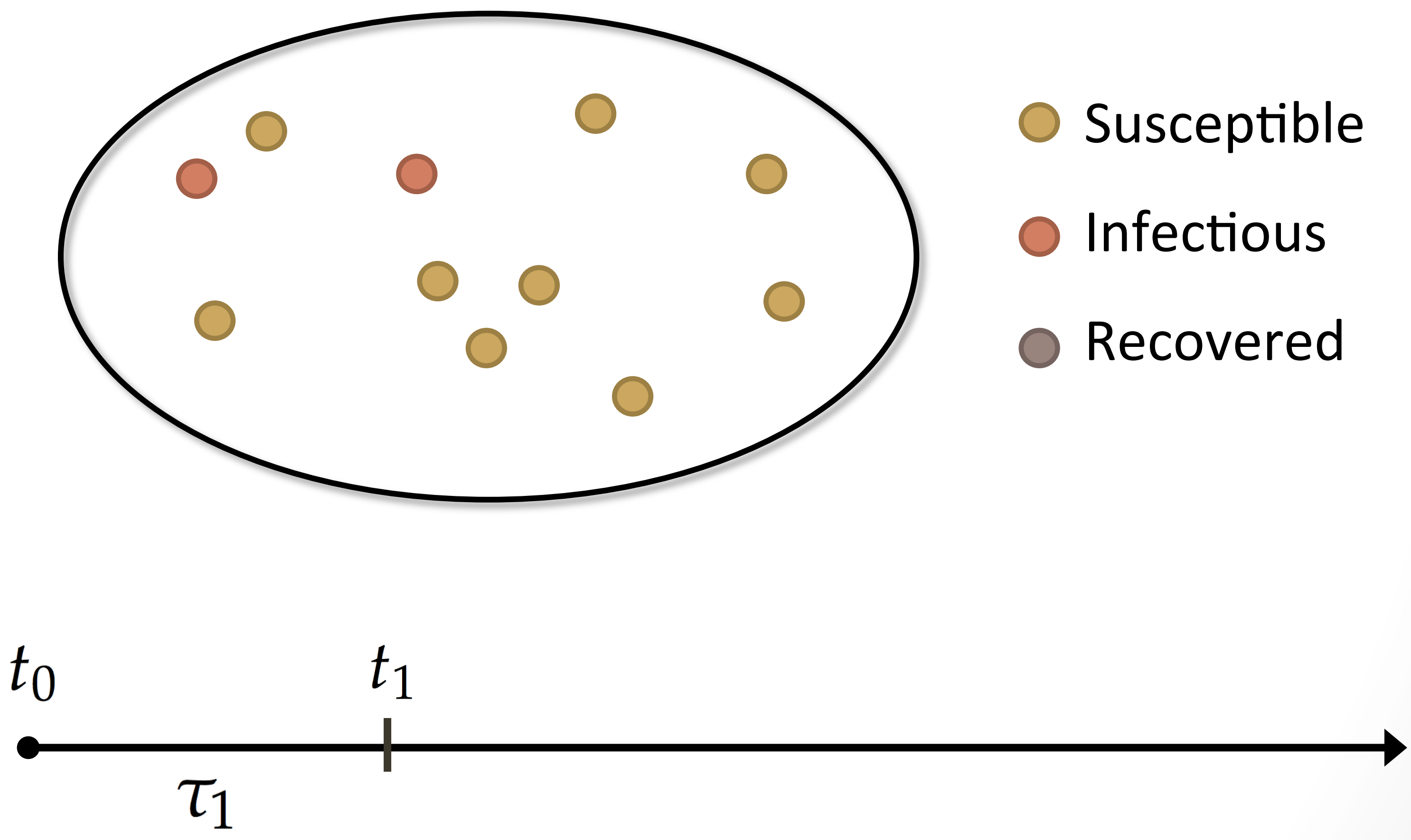
Stochastic SIR dynamics

Small population



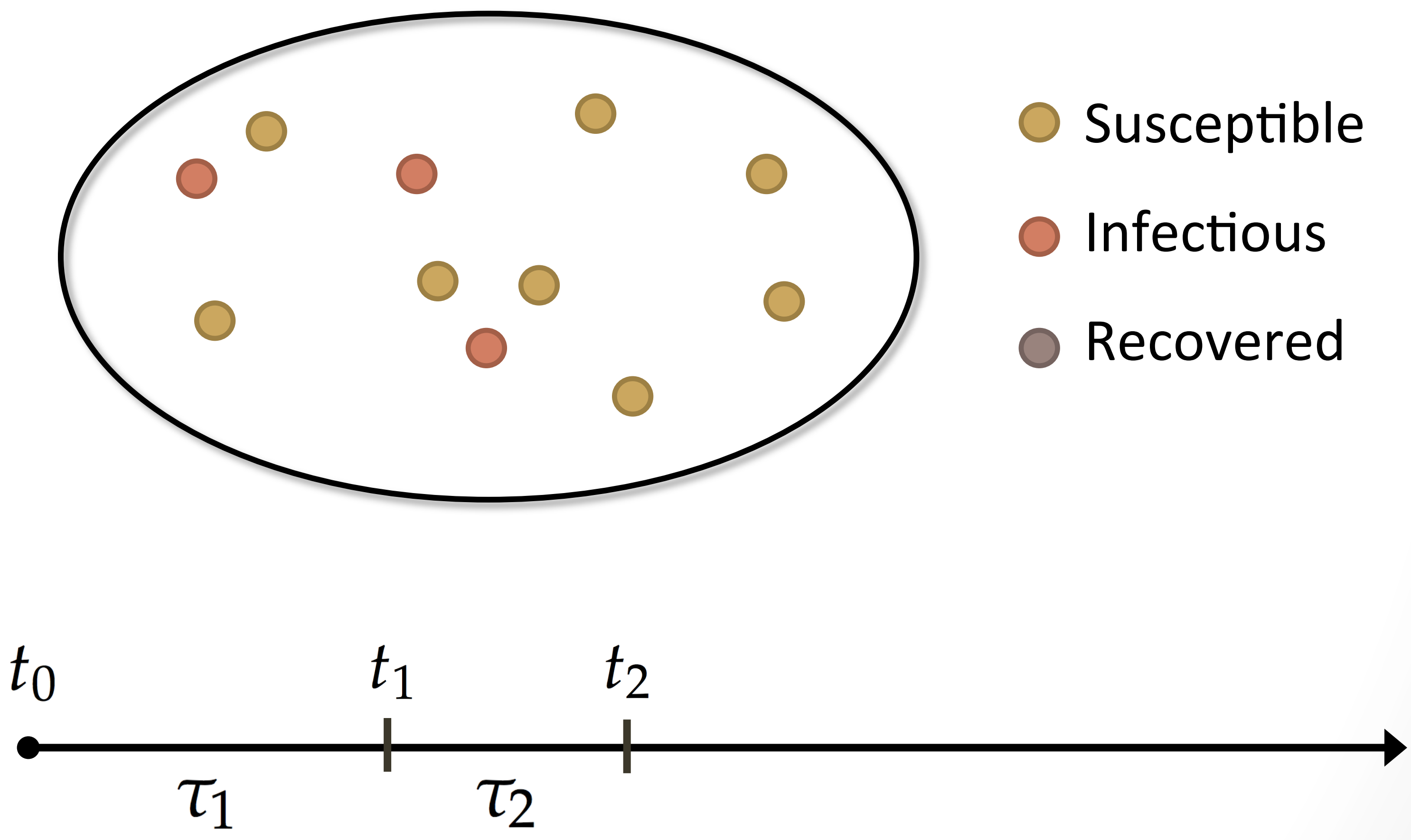
Stochastic SIR dynamics

Small population



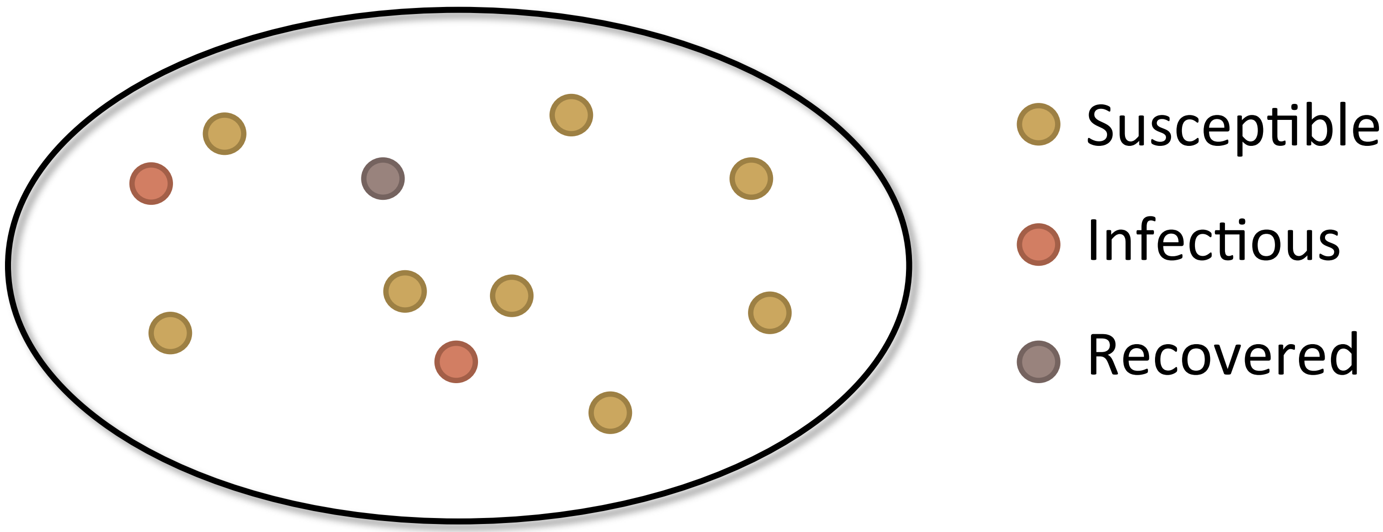
Stochastic SIR dynamics

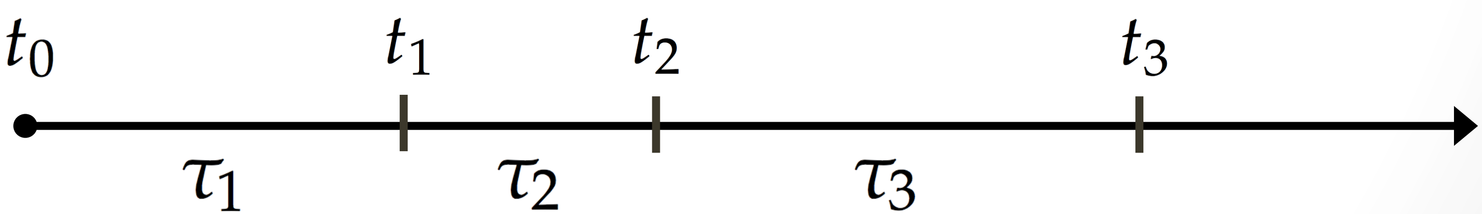
Small population



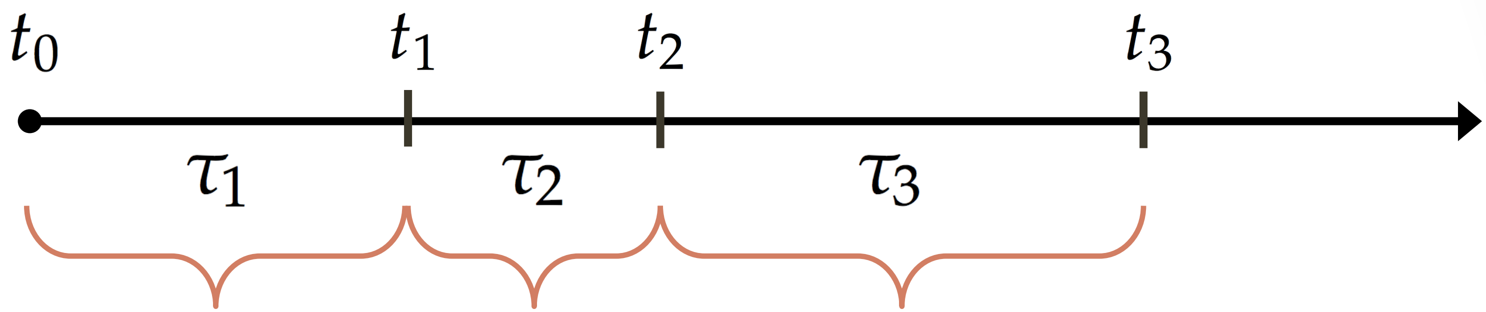
Stochastic SIR dynamics

Small population



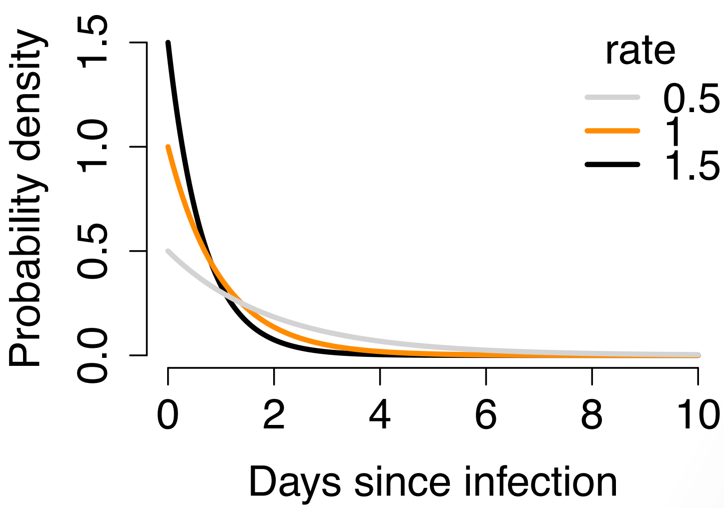


Exponential waiting times



time between events

waiting time distribution: distribution of times until an event occurs



Summary: Gillespie algorithm

Assumptions:

- finite, countable populations
- well-mixed contacts
- exponential waiting times (memory-less)

Summary: Gillespie algorithm

Assumptions:

- finite, countable populations
- well-mixed contacts
- exponential waiting times (memory-less)

Summary: Gillespie algorithm

Assumptions:

- finite, countable populations
- well-mixed contacts
- exponential waiting times (memory-less)

Summary: Gillespie algorithm

Assumptions:

- finite, countable populations
- well-mixed contacts
- exponential waiting times (memory-less)

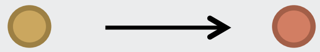
Notes:

- noise (stochasticity) is introduced by the discrete nature of individuals
- event-driven simulation
- computationally slow
  - especially for large populations

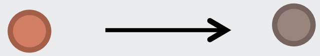
Need to know

- **What happened ?**
- When did it happen?

Susceptible to Infectious



Infectious to Recovered



Two event types:

Transmission

$$(S, I, R) \longrightarrow (S-1, I+1, R) \text{ at rate } \frac{\beta S I}{N}$$

Recovery

$$(S, I, R) \longrightarrow (S, I-1, R+1) \text{ at rate } \gamma I$$

Need to know

- What happened ?
- **When did it happen?**

ODE analogue:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{\beta SI}{N}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta SI}{N} - \gamma I$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I$$

Two event types:

Transmission

$$(S, I, R) \longrightarrow (S-1, I+1, R) \text{ at rate } \frac{\beta SI}{N} \quad \leftarrow$$

Recovery

$$(S, I, R) \longrightarrow (S, I-1, R+1) \text{ at rate } \gamma I \quad \leftarrow$$

Need to know

- What happened ? EventType
- When did it happen? EventTime

Two event types:

Transmission

$$(S, I, R) \longrightarrow (S-1, I+1, R) \text{ at rate } \frac{\beta S I}{N}$$

Recovery

$$(S, I, R) \longrightarrow (S, I-1, R+1) \text{ at rate } \gamma I$$

The Gillespie algorithm

Two event types:

Transmission

$$(S, I, R) \longrightarrow (S-1, I+1, R) \text{ at rate } \frac{\beta S I}{N} = \lambda_1$$

Recovery

$$(S, I, R) \longrightarrow (S, I-1, R+1) \text{ at rate } \gamma I = \lambda_2$$

Time to the next event: $\tau \sim \mathrm{Exp}\left(\lambda = \sum_i \lambda_i\right)$

Probability the event is type i: $p_i = \frac{\lambda_i}{\sum_i \lambda_i}$

Simulating the Gillespie model

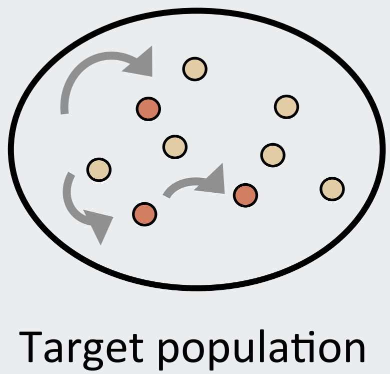
```
while (I > 0 and time < MAXTIME)
     Calculate rates
     Determine time to next event
     Determine event type
     Update state variables
     Update time
end
```

R code example

SIR model with *spillover*…

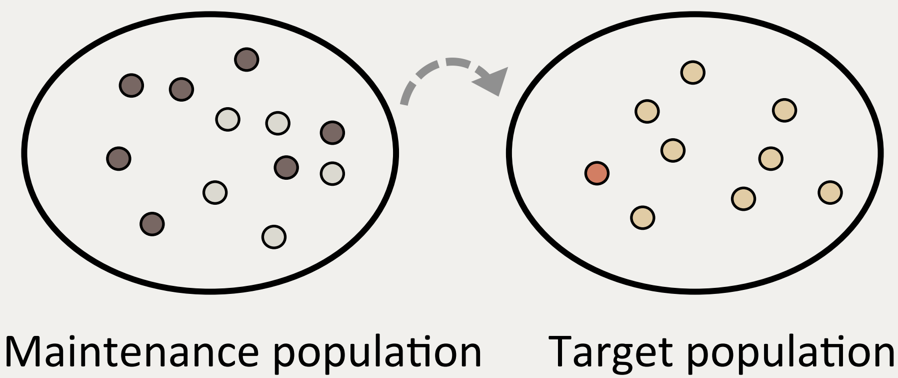
Two types of transmission

Within-population transmission



$$I \rightarrow I + 1 \text{ at rate } \frac{\beta S I}{N}$$

Spillover introductions



$$I \rightarrow I + 1 \text{ at rate } \lambda$$

R code example

SIR model with spillover

Download the associated file from ICI3D tutorial repository

Try changing:

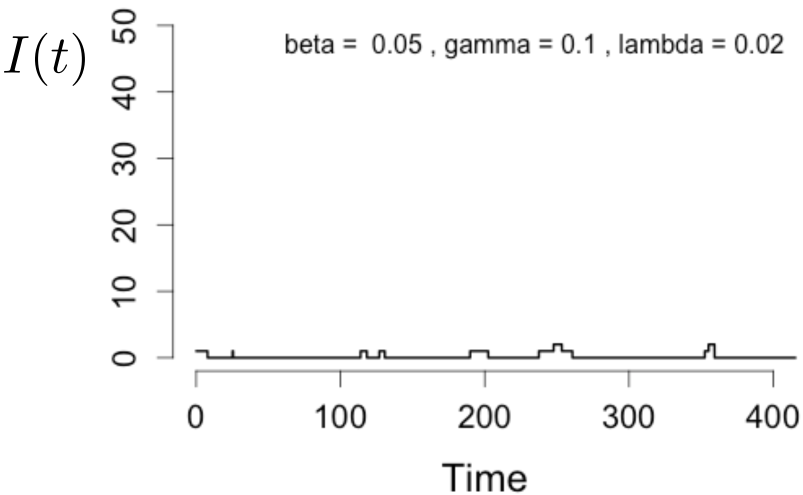
- population size
- spillover rate
- transmission rate
- recovery rate

Sub-critical or super-critical?

Basic reproduction number for SIR model:

$$R_0 = \frac{\beta}{\gamma}$$

Sub-critical



Super-critical