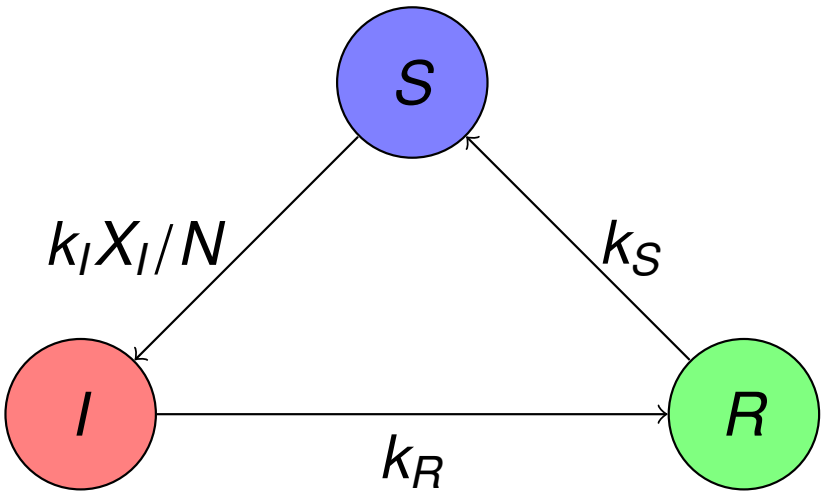
Outline

1. Preliminaries
   - Exponential Distribution
2. Continuous Time Markov Chains
   - Main concepts
   - Poisson Process
   - Time-inhomogeneous rates
3. Population Continuous Time Markov Chains
4. Simulation
   - SSA
   - Next Reaction Method
   - τ-leaping

POPULATION PROCESSES

SIR epidemics model
single individual



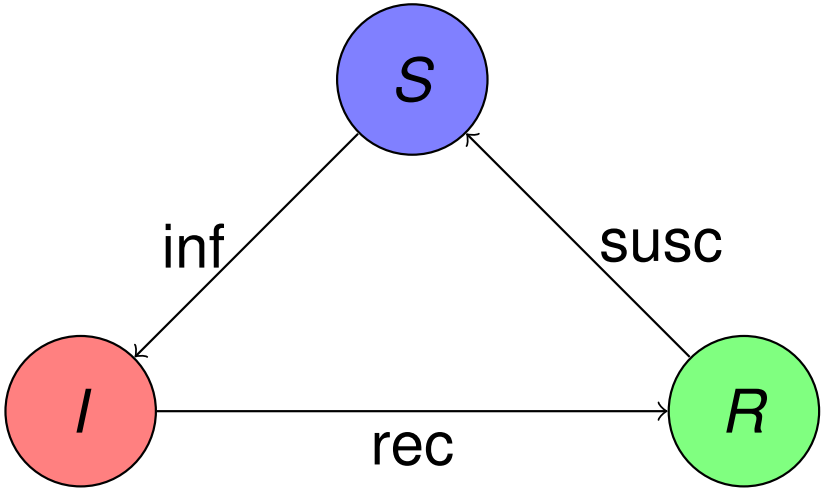
- Consider a CTMC model of a population epidemics in which each of N individuals can be in one of three states: susceptible (S), infected (I), and recovered (R);
- Infection rate depends on the density of infected individuals;
- The CTMC for N agents has 3^N states (if we distinguish the individuals) or $(N+1)^2$ states (if we just count them): *it's impossible to write down the Q matrix explicitly*.
- We need a better description of population CTMCs.

POPULATION CTMC

A population CTMC model is a tuple $\mathcal{X} = (\mathbf{X}, \mathcal{D}, \mathcal{T}, \mathbf{x_0})$, where:

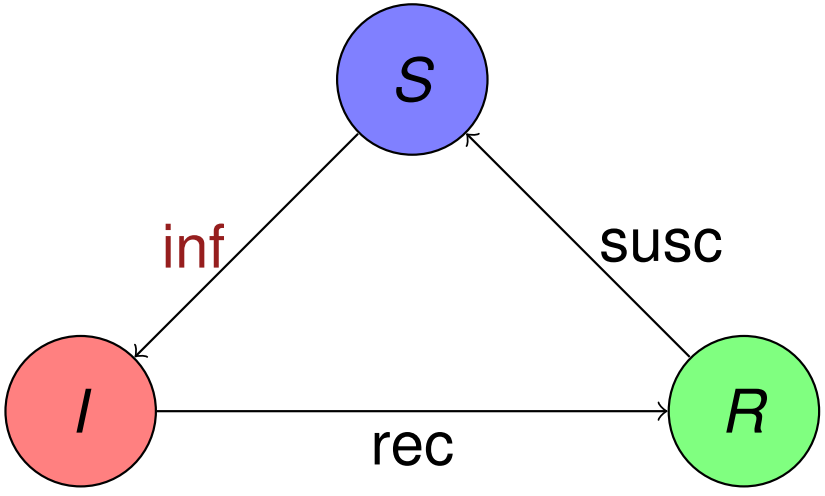
1. $\mathbf{X}$ — vector of *variables* counting how many individuals in each state.
2. $\mathcal{D} = \prod_i \mathcal{D}_i$ — (countable) state space.
3. $\mathbf{x_0} \in \mathcal{D}$ —*initial state*.
4. $\eta_i \in \mathcal{T}$ — *global transitions*, $\eta_i = (a, \phi(\mathbf{X}), \mathbf{v}, r(\mathbf{X}))$
    1. a — event name (optional).
    2. $\phi(\mathbf{X})$ — guard.
    3. $\mathbf{v} \in \mathbb{R}^n$ — *update vector* (from $\mathbf{X}$ to $\mathbf{X} + \mathbf{v}$)
    4. $r : \mathcal{D} \to \mathbb{R}_{\geq 0}$ — rate function.

EXAMPLE: SIR EPIDEMICS



Three variables: X_S, X_I, X_R.
State space:
$\mathcal{D} = \{(n_1, n_2, n_3) \mid n_1 + n_2 + n_3 = N\} \subset \{0, \ldots, N\}^3$.

EXAMPLE: SIR EPIDEMICS

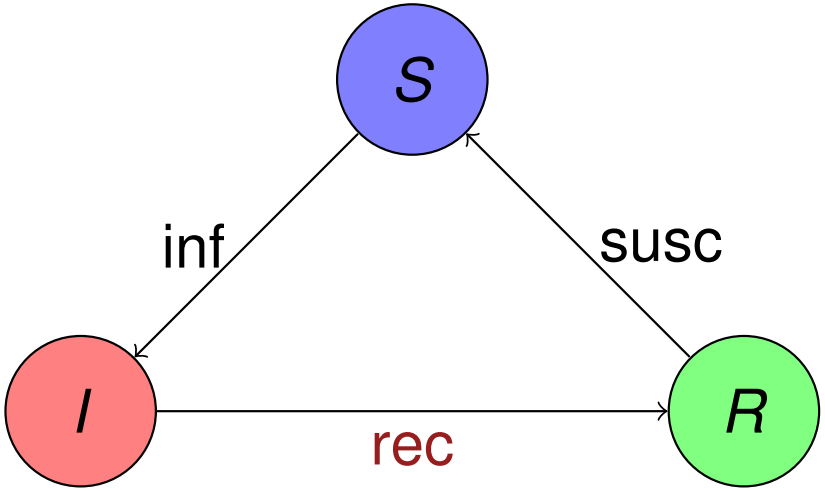


Three variables: X_S, X_I, X_R.
State space:
$\mathcal{D} = \{(n_1, n_2, n_3) \mid n_1 + n_2 + n_3 = N\} \subset \{0, \ldots, N\}^3$.

Transitions:

- $(inf, \top, (-1, 1, 0) k_I \frac{X_I}{N} X_S)$

EXAMPLE: SIR EPIDEMICS

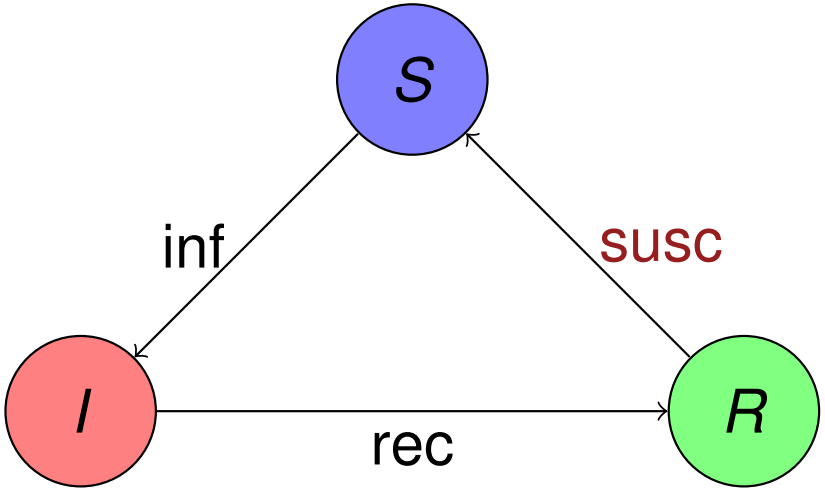


Three variables: X_S, X_I, X_R.
State space:
$\mathcal{D} = \{(n_1, n_2, n_3) \mid n_1 + n_2 + n_3 = N\} \subset \{0, \ldots, N\}^3$.

Transitions:

- $(inf, \top, (-1, 1, 0) k_I \frac{X_I}{N} X_S)$
- $(rec, \top, (0, -1, 1), k_R X_I)$

EXAMPLE: SIR EPIDEMICS

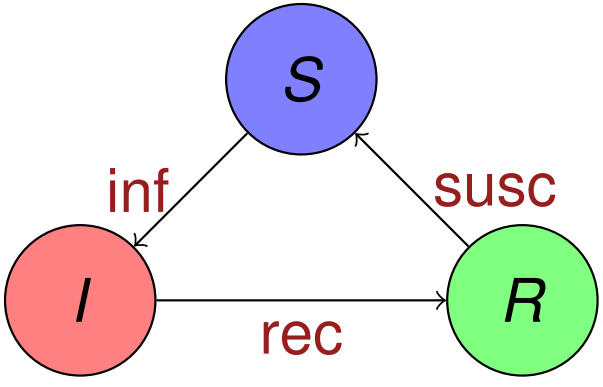


Three variables: X_S, X_I, X_R.
State space:
$\mathcal{D} = \{(n_1, n_2, n_3) \mid n_1 + n_2 + n_3 = N\} \subset \{0, \ldots, N\}^3$.

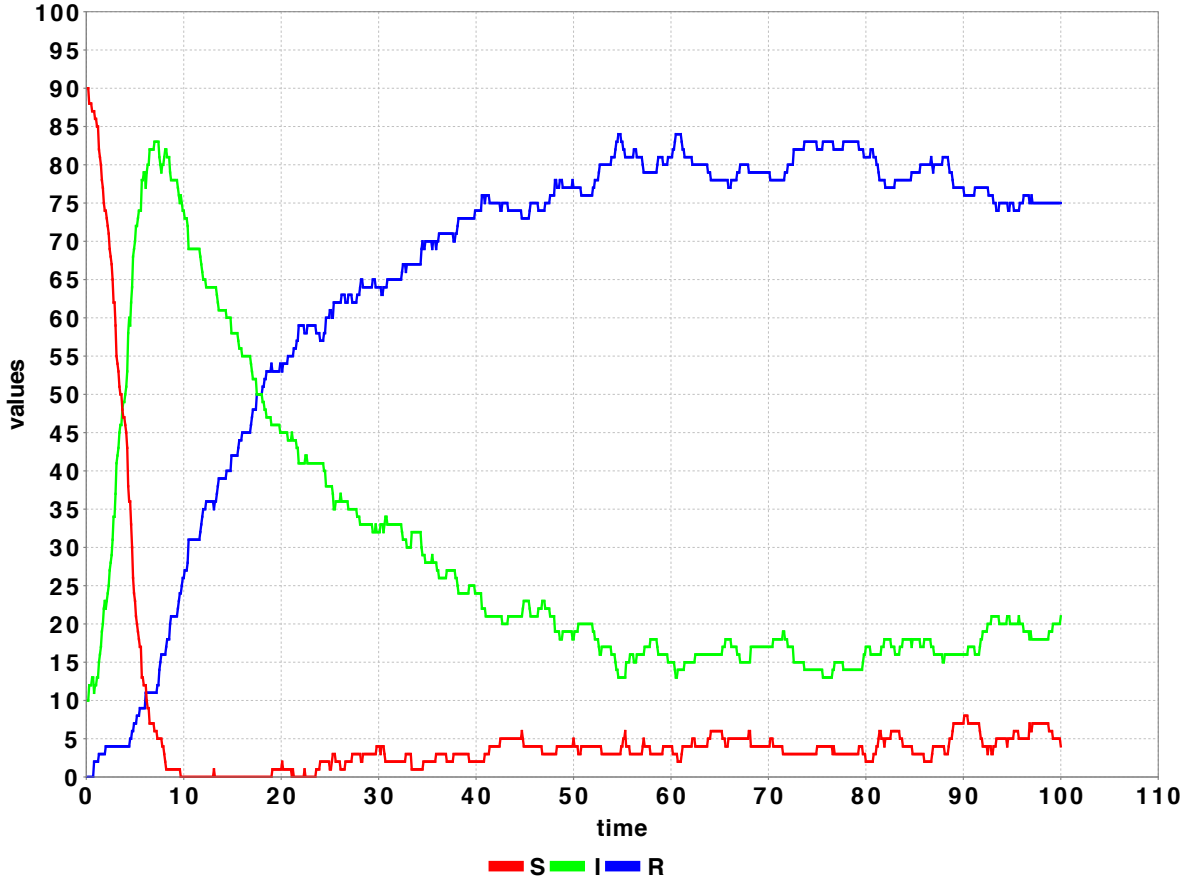
Transitions:

- $(inf, \top, (-1, 1, 0) k_I \frac{X_I}{N} X_S)$
- $(rec, \top, (0, -1, 1), k_R X_I)$
- $(susc, \top, (1, 0, -1), k_S X_R)$

EXAMPLE: SIR EPIDEMICS

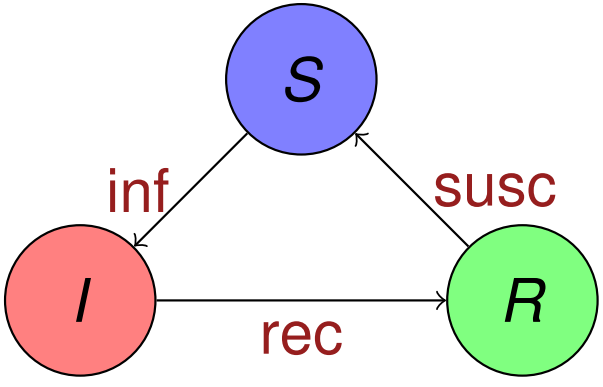


$N = 100,\ k_I = 1,\ k_R = 0.05,\ k_S = 0.01$

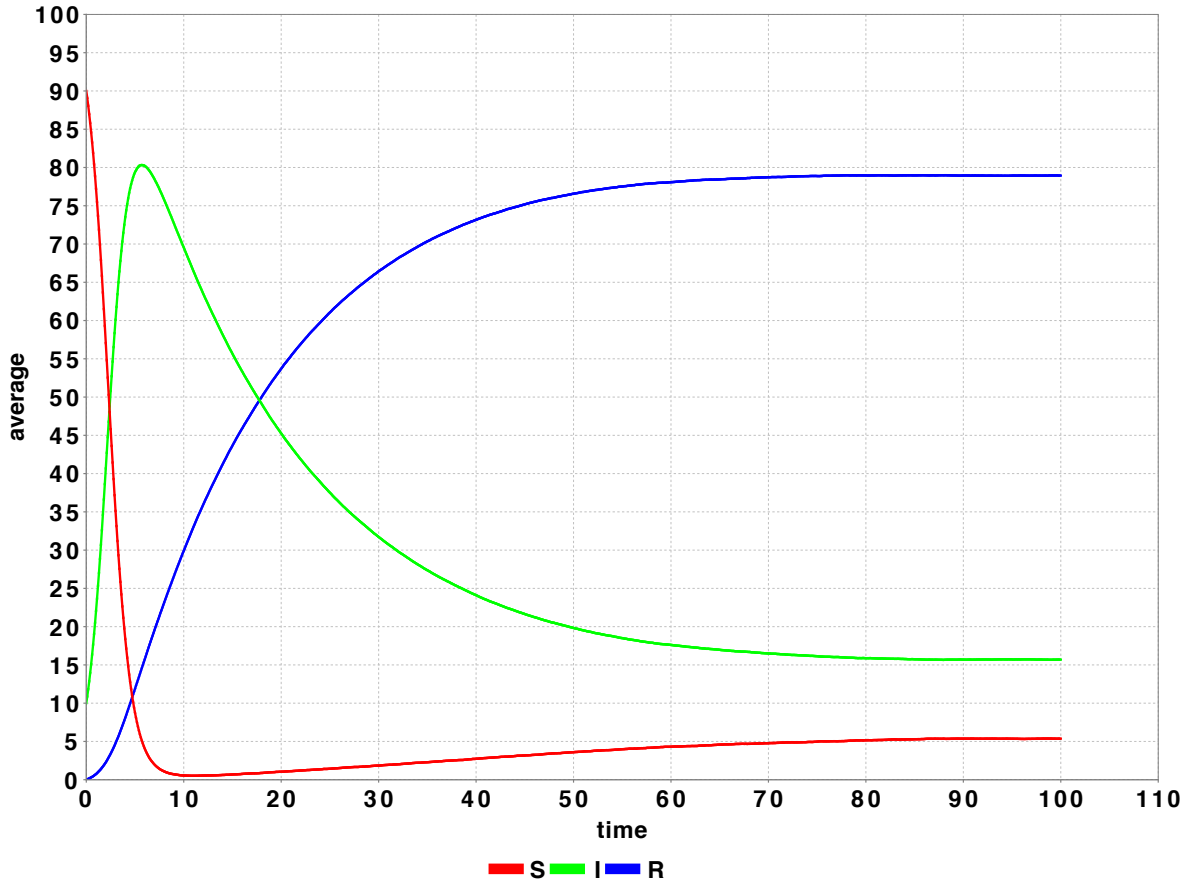


(1 run)

EXAMPLE: SIR EPIDEMICS



$N = 100,\ k_I = 1,\ k_R = 0.05,\ k_S = 0.01$



(average)

MASTER EQUATION

The Kolmogorov equation in the context of Population Processes is often know as master equation.

There is one equation per state $\mathbf{x} \in \mathcal{D}$, for the probability mass $P(\mathbf{x}, t)$, which considers the inflow and outflow of probability at time t.

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{\eta \in \mathcal{T}} r_\eta(\mathbf{x} - \mathbf{v}_\eta) P(\mathbf{x} - \mathbf{v}_\eta, t) - \sum_{\eta \in \mathcal{T}} r_\eta(\mathbf{x}) P(\mathbf{x}, t)$$

POISSON REPRESENTATION

Population CTMC admit a simple description in terms of Poisson processes.
Essentially, we introduce variables $R_\eta(t)$ counting how many times each transition η has fired up to time t. Hence we can write:

$$X(t) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta R_\eta(t).$$

It turns out that $R_\eta(t)$ is a time-inhomogeneous Poisson process with cumulative rate $\int_0^t r_\eta(X(s))ds$, independent from the other $R_{\eta'}$.
Hence, let $\mathcal{N}_\eta$ be independent Poisson processes. For each $t \geq 0$:

$$X(t) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta \mathcal{N}_\eta \left(\int_0^t r_\eta(X(s))ds \right).$$

Equivalently, let $\mathcal{Y}_\eta$ be independent Poisson r.v. It holds:

$$X(t) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta \mathcal{Y}_\eta \left(\int_0^t r_\eta(X(s))ds \right).$$

Outline

1. Preliminaries
   - Exponential Distribution
2. Continuous Time Markov Chains
   - Main concepts
   - Poisson Process
   - Time-inhomogeneous rates
3. Population Continuous Time Markov Chains
4. Simulation
   - SSA
   - Next Reaction Method
   - τ-leaping

SIMULATING A POPULATION CTMC

Population CTMC have generally a complex dynamics and state space which is too large for

1. Solving the CTMC analytically
2. Performing numerical computations like solution of the Kolmogorov equation, transient analysis by uniformization, or computation of steady state.

Therefore, one can resort to statistical tools.
One samples a (large) set of trajectories from the distribution induced by the CTMC in the space of traces (cadlag functions), and then uses statistical methods to extract information about the process from these samples.
We will review some simulation algorithms, exploiting the different characterizations of (population) CTMCs.

DIRECT METHOD

RACE CONDITION CHARACTERIZATION OF A PCTMC

In each state $\mathbf{x}$, the m transitions in $\mathcal{T}$ compete in a race condition: the fastest wins and is executed.

DIRECT METHOD

At each step, with current state $\mathbf{x}$ and current time t

1. sample m uniform r.v. U_η;
2. compute $T_\eta = -\frac{1}{r_\eta(\mathbf{x})} \log(U_\eta)$;
3. find $\bar{\eta} = \operatorname{argmin}_{\eta \in \mathcal{T}} T_\eta$;
4. execute transition $\bar{\eta}$ updating the current state from $\mathbf{x}$ to $\mathbf{x} + \mathbf{v}_\eta$ and current time to $t + T_\eta$.

Outline

1. Preliminaries
   - Exponential Distribution
2. Continuous Time Markov Chains
   - Main concepts
   - Poisson Process
   - Time-inhomogeneous rates
3. Population Continuous Time Markov Chains
4. Simulation
   - SSA
   - Next Reaction Method
   - τ-leaping

STOCHASTIC SIMULATION ALGORITHM

JUMP CHAIN AND HOLDING TIMES

We can improve the previous simulation by using the characterization with Jump Chain and Holding Times, which for population CTMC reads:

HOLDING TIME $r(\mathbf{x}) = \sum_{\eta \in \mathcal{T}} r_\eta(\mathbf{x})$

JUMP CHAIN $P(\eta \mid \mathbf{x}) = \frac{r_\eta(\mathbf{x})}{r(\mathbf{x})}$

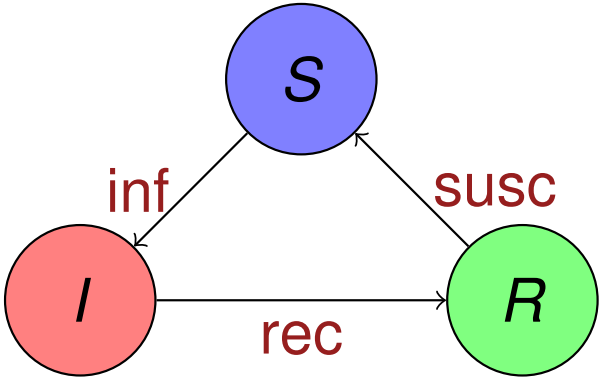
SSA

At each step, with current state $\mathbf{x}$ and current time t

1. sample the next transition η from the jump chain;
2. sample the holding time from an $Exp(r(\mathbf{x}))$;
3. update current state and current time.

This method in biochemistry and system biology is also known as Gillespie Algorithm.

EXAMPLE: SIR EPIDEMICS



$N = 10$, $k_I = 1$, $k_R = 0.05$, $k_S = 0.01$
$X_S(0) = 8$, $X_I(0) = 2$, $X_R(0) = 0$.

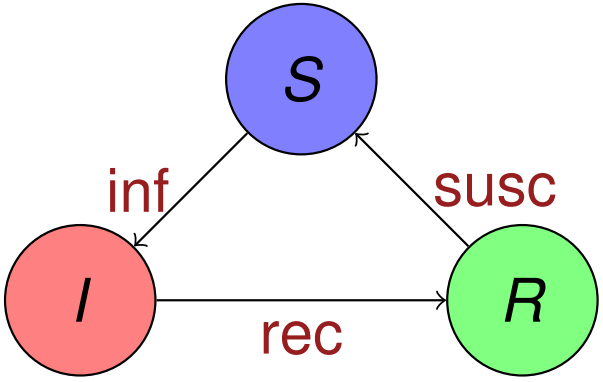
STEP 0: RATES OF TRANSITIONS

INFECTION: $\frac{1}{10} \cdot 8 \cdot 2 = 1.6$

RECOVERY: $0.05 \cdot 2 = 0.1$

IMMUNITY LOSS: 0

EXAMPLE: SIR EPIDEMICS



$N = 10$, $k_I = 1$, $k_R = 0.05$, $k_S = 0.01$
$X_S(0) = 8$, $X_I(0) = 2$, $X_R(0) = 0$.

STEP 0: RATES OF TRANSITIONS

INFECTION: $\frac{1}{10} \cdot 8 \cdot 2 = 1.6$

RECOVERY: $0.05 \cdot 2 = 0.1$

IMMUNITY LOSS: 0

TIME DELAY

Exponential with rate $1.6 + 0.1 = 1.7$.

NEXT STATE

- $X_S(0) = 7$, $X_I(0) = 3$, $X_R(0) = 0$ with prob. $\frac{1.6}{1.6+0.1} = 0.9412$
- $X_S(0) = 8$, $X_I(0) = 1$, $X_R(0) = 1$ with prob. $\frac{1.6}{1.6+0.1} = 0.0588$

Outline

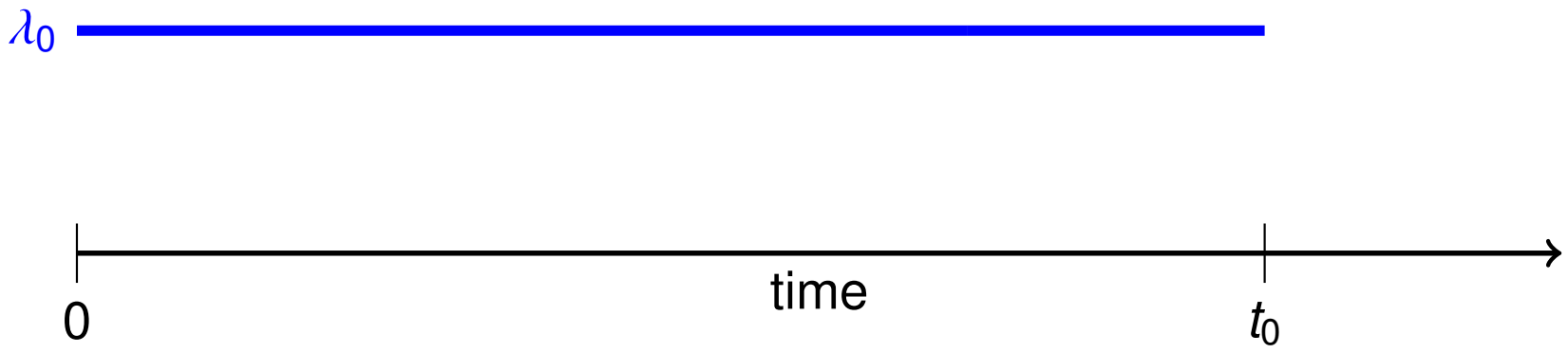
1. Preliminaries
   - Exponential Distribution
2. Continuous Time Markov Chains
   - Main concepts
   - Poisson Process
   - Time-inhomogeneous rates
3. Population Continuous Time Markov Chains
4. Simulation
   - SSA
   - Next Reaction Method
   - τ-leaping

NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)

- Consider a single η transition in a time interval $[0, t]$ in which it never fires.
- As other transitions may fire, its rate $r_\eta(\mathbf{X}(s))$ is a time-dependent function.
- Therefore, we can sample the firing time of η using the inversion method for time-inhomogeneous exponential distribution, solving for t

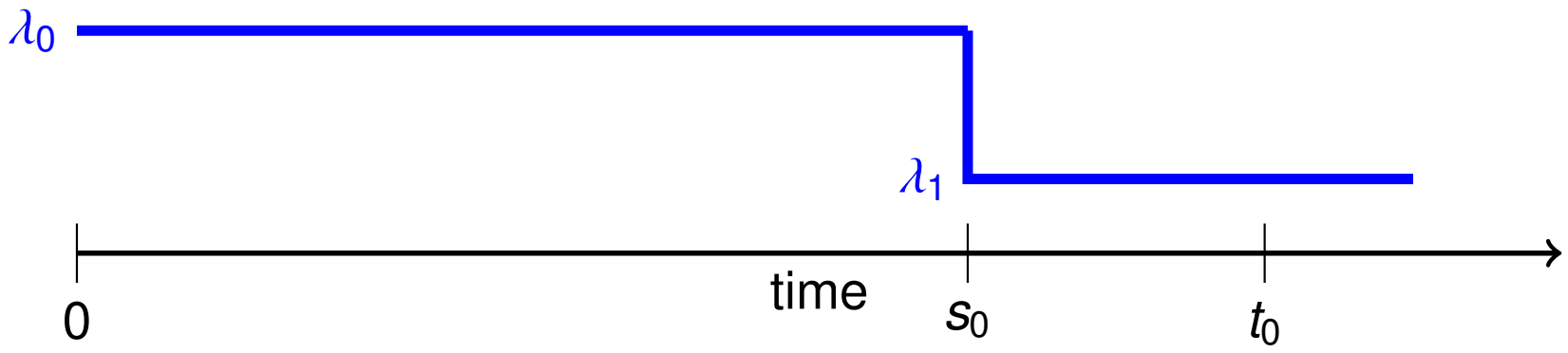
$$\Lambda_\eta(t) = \int_0^t r_\eta(\mathbf{X}(s))ds = \xi \sim Exp(1).$$

NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)



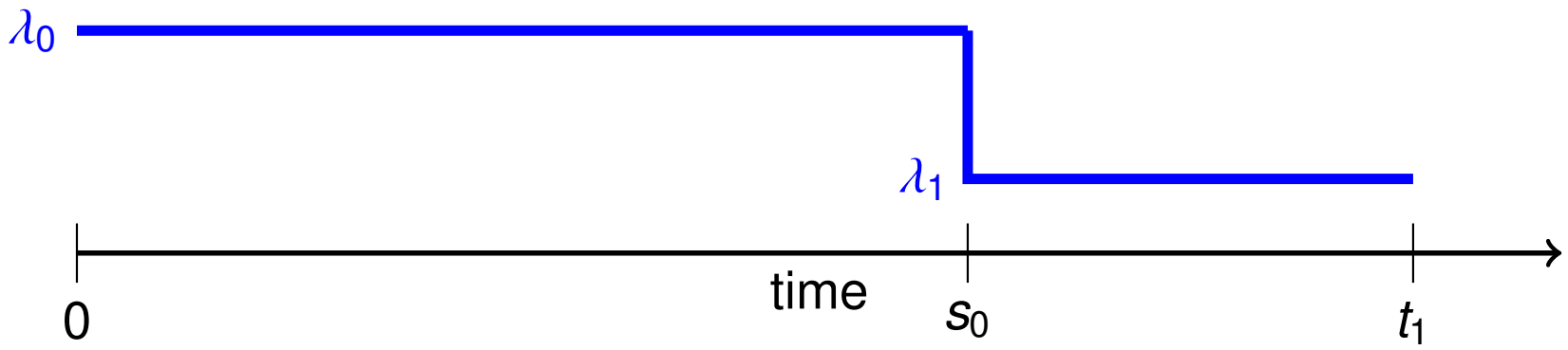
- Start at time 0, and suppose the rate of η is λ_0. Assuming it does not change in time, the firing time would be $t_0 = \frac{1}{\lambda_0}\xi \sim Exp(\lambda_0)$.

NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)



- Start at time 0, and suppose the rate of η is λ_0. Assuming it does not change in time, the firing time would be $t_0 = \frac{1}{\lambda_0}\xi \sim Exp(\lambda_0)$.
- Now, suppose at time s_0 another event η' fires, and this changes the rate of η to λ_1.

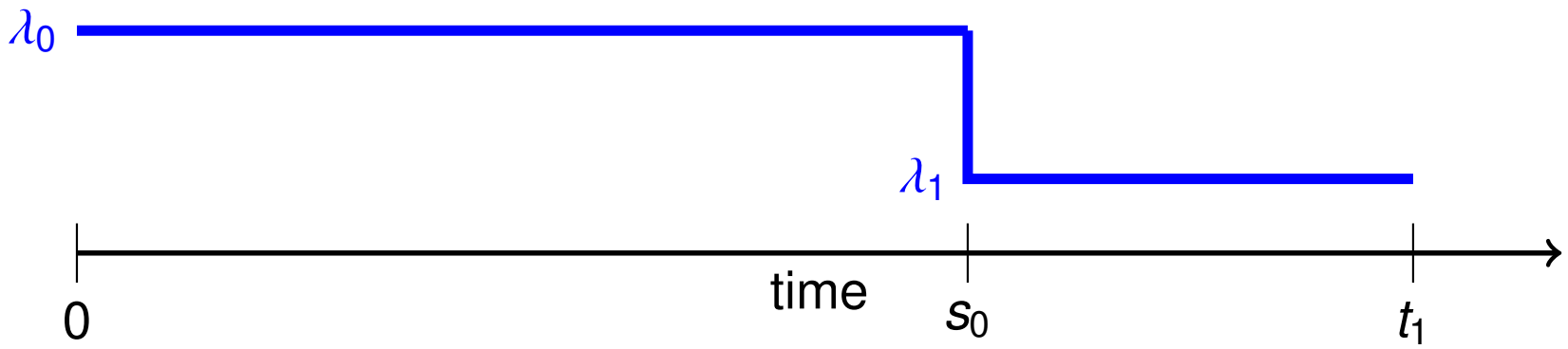
NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)



- Start at time 0, and suppose the rate of η is λ_0. Assuming it does not change in time, the firing time would be $t_0 = \frac{1}{\lambda_0}\xi \sim Exp(\lambda_0)$.
- Now, suppose at time s_0 another event η' fires, and this changes the rate of η to λ_1.
- Then the firing time of η would be found by solving $\lambda_0 s_0 + \lambda_1(t_1 - s_0) = \xi$, from which

$$t_1 = s_0 + \frac{\lambda_0}{\lambda_1}\left(\frac{1}{\lambda_0}\xi - s_0\right) = s_0 + \frac{\lambda_0}{\lambda_1}(t_0 - s_0).$$

NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)

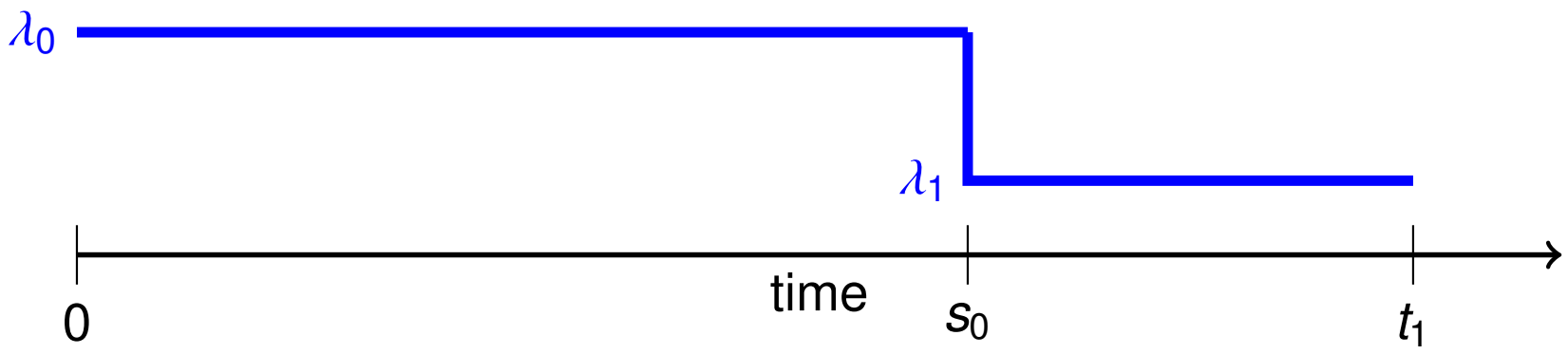


- Start at time 0, and suppose the rate of η is λ_0. Assuming it does not change in time, the firing time would be $t_0 = \frac{1}{\lambda_0}\xi \sim Exp(\lambda_0)$.
- Now, suppose at time s_0 another event η' fires, and this changes the rate of η to λ_1.
- Then the firing time of η would be found by solving $\lambda_0 s_0 + \lambda_1(t_1 - s_0) = \xi$, from which

$$t_1 = s_0 + \frac{\lambda_0}{\lambda_1}\left(\frac{1}{\lambda_0}\xi - s_0\right) = s_0 + \frac{\lambda_0}{\lambda_1}(t_0 - s_0).$$

- This is the update formula of Gibson-Bruck algorithm (can be easily generalized to n intermediate events by induction).

NEXT REACTION METHOD/GIBSON-BRUCK (SKETCH)



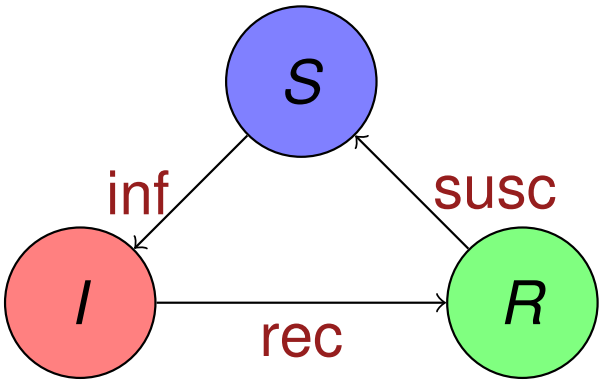
NEXT REACTION METHOD

At each step, with current state **x** and current time t

1. execute transition η with smallest time;
2. update rates and firing times of other transitions;
3. sample a new firing time for η.

the algorithm uses a priority queue and a dependency graph to speed up operations.

EXAMPLE: SIR EPIDEMICS



$N = 10$, $k_I = 1$, $k_R = 0.05$, $k_S = 0.01$
$X_S(0) = 8$, $X_I(0) = 2$, $X_R(0) = 0$.

STEP 1: RATES OF TRANSITIONS

INFECTION: $\frac{1}{10} \cdot 8 \cdot 2 = 1.6$

RECOVERY: $0.05 \cdot 2 = 0.1$

IMMUNITY LOSS: 0

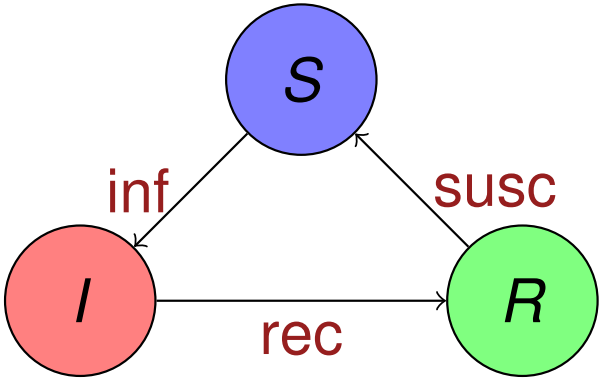
STEP 2: COMPUTE FIRING TIMES

INFECTION: $\frac{1}{1.6} \cdot 0.2228 = 0.1392$

RECOVERY: $\frac{1}{0.1} \cdot 1.9527 = 19.5273$

IMMUNITY LOSS: $\frac{1}{0} \cdot 0 = \infty$

EXAMPLE: SIR EPIDEMICS



$N = 10$, $k_I = 1$, $k_R = 0.05$, $k_S = 0.01$
$X_S(0.1392) = 7$, $X_I(0.1392) = 3$,
$X_R(0.1392) = 0$.

STEP 1: RATES OF TRANSITIONS

INFECTION: $\frac{1}{10} \cdot 7 \cdot 3 = 2.1$

RECOVERY: $0.05 \cdot 3 = 0.15$

IMMUNITY LOSS: 0

STEP 2: REEVALUATE FIRING TIMES

INFECTION: $\frac{1}{2.1} \cdot 3.3323 = 1.5868$

RECOVERY: $0.1392 + \frac{0.1}{0.15} \cdot (19.5273 - 0.1392) = 13.0646$

IMMUNITY LOSS: ∞

Outline

1. Preliminaries
   - Exponential Distribution
2. Continuous Time Markov Chains
   - Main concepts
   - Poisson Process
   - Time-inhomogeneous rates
3. Population Continuous Time Markov Chains
4. Simulation
   - SSA
   - Next Reaction Method
   - τ-leaping

τ-LEAPING (SKETCH)

Consider the Poisson representation of a population CTMC at time τ

$$X(\tau) = X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta \mathcal{Y}_\eta \left(\int_0^\tau r_\eta(X(s)) ds \right).$$

If τ is sufficiently small, we may assume that the rates $r_\eta(X(s))$ are approximately constant in $[0, \tau]$ and equal to a_η.
Then $\int_0^t r_\eta(X(s)) ds \approx a_\eta \tau$, hence

$$X(\tau) \approx X(0) + \sum_{\eta \in \mathcal{T}} \mathbf{v}_\eta \mathcal{Y}_\eta \left(a_\eta \tau \right).$$

τ-LEAPING (SKETCH)

τ-LEAPING

At each step, with current state $\mathbf{x}$ and current time t

1. choose τ;
2. for each η, sample n_η from the Poisson r.v. $\mathcal{Y}_\eta\left(a_\eta \tau\right)$;
3. update $\mathbf{x}$ to $\mathbf{x} + \sum_\eta \mathbf{v}_\eta n_\eta$ and time to $t + \tau$.

CHOICE OF τ: LEAPING CONDITION

The choice of τ is an art:

- it has to be small for rates to be approximately constant in $[t, t+\tau]$;
- it has to be as large as possible to make $\mathcal{Y}_\eta(a_\eta \tau)$ large to gain in computational efficiency;
- one has to avoid the generation of negative populations.

REFERENCES

- J.R. Norris. Markov Chains, Cambridge University Press, 1998.
- R. Durrett, Essentials of Stochastic Processes, Springer-Verlag, 1998.
- D.T. Gillespie (1976). A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. Journal of Computational Physics, 22(4): 403-434
- M.A. Gibson and J. Bruck (2000). Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. Journal of Physical Chemistry A, 104(9).
- Y Cao, DT Gillespie, LR Petzold (2006). Efficient step size selection for the tau-leaping simulation method. J Chem Phys, 28;124(4).