# “Introduction to the Stochastic modelling of reaction networks”

## Day 1: The Gillespie algorithm (intro)

Manuel Lera Ramírez

manuel.lera@embl.de

CBB – François Nédélec Group

# Deterministic vs. stochastic

- Deterministic: every set of variable states is uniquely determined by parameters in the model and by sets of previous states of these variables; therefore, a deterministic model always performs the same way for a given set of initial conditions.
- Stochastic trajectory: randomness is present, and every trajectory can be different from the others.
- Stochastic: calculate deterministically some variable of the stochastic distribution

# Notation: Describing a chemical system over time

- The total volume of the system, *V*.
- A series of chemical species: $S_i$ $(i = 1,2, \dots, N)$, N total number of chemical species.
- The quantities of those elements at time *t*: $\mathrm{X}_i$ $(i = 1,2, \dots, N)$
- The possible reactions: $R_\mu (\mu = 1,2, \dots, M)$ M, total number of reactions.
- Parameter for the evolution of the system:
  - $k_\mu$ (deterministic)
  - $c_\mu$(stochastic)

# The deterministic approach to chemical reactions

- Back to high-school: A simple equilibrium

$S_i = (A, B, C); X_i = (150, 100, 0);\ R = (R_1, R_2); k = (k_1 = 0.5, k_2 = 0.05);\ V{=}1$

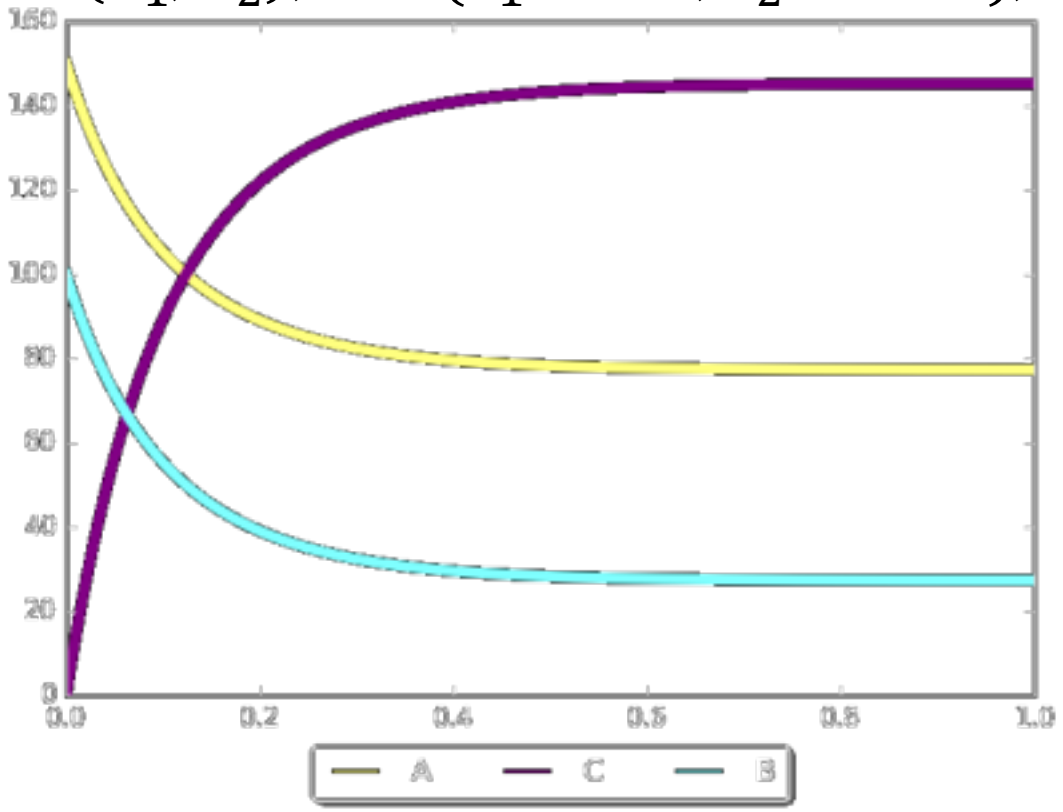
$R_1: A + B \xrightarrow{k_1} 2C$

$R_2: 2C \xrightarrow{k_2} A + B$

$$v_1 = \frac{dA}{dt} = \frac{dB}{dt} = -k_1[A][B]$$

$$v_2 = \frac{1}{2}\frac{dC}{dt} = -k_2[C]^2$$

$Only\ in\ the\ equilibrium: v_1 = v_2$

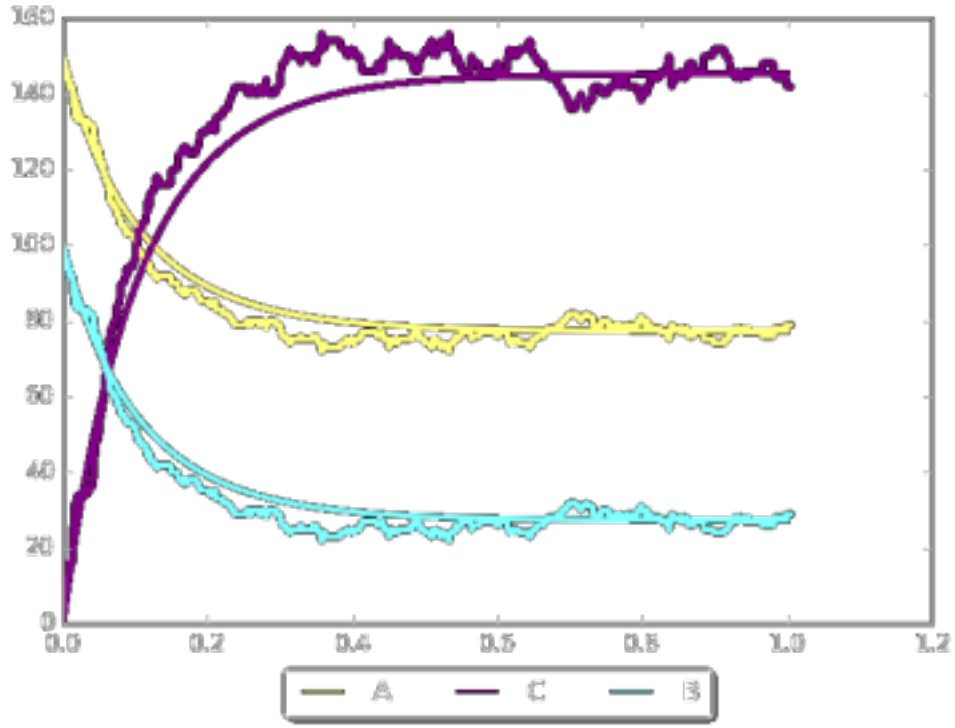


# A stochastic approach to the chemical system

- Time evolution of a chemical system is **not continuous**, chemical species quantities can only change by discrete integer amounts.
- It is **not deterministic** if we take into account only the quantities of the chemical species. It is impossible to predict the exact molecular population levels at a future time.
- A better approach is to consider the process as a random walk

# Stochastic model assumptions

- The volume *V* is fixed
- The molecules are in thermal equilibrium, collisions occur in a random manner. The mixture is spatially homogeneous.
- A reaction happens when two interacting molecules collide in an "appropriate way”.



Possible combinations

$$R_1: A + B \xrightarrow{k_1} 2C$$

$$R_2: 2C \xrightarrow{k_2} A + B$$

$$\frac{dP_{R_1}}{dt} = \frac{a_1}{dt} = c_1 n_A n_B$$

$$\frac{dP_{R_2}}{dt} = \frac{a_2}{dt} = \frac{c_2 n_C (n_c - 1)}{2}$$

$$v_1 = -k_1 \frac{n_A n_B}{V^2}$$

$$v_2 = -k_2 \frac{n_c^2}{V^2}$$

$$c1 = \frac{k_1}{V^2}$$

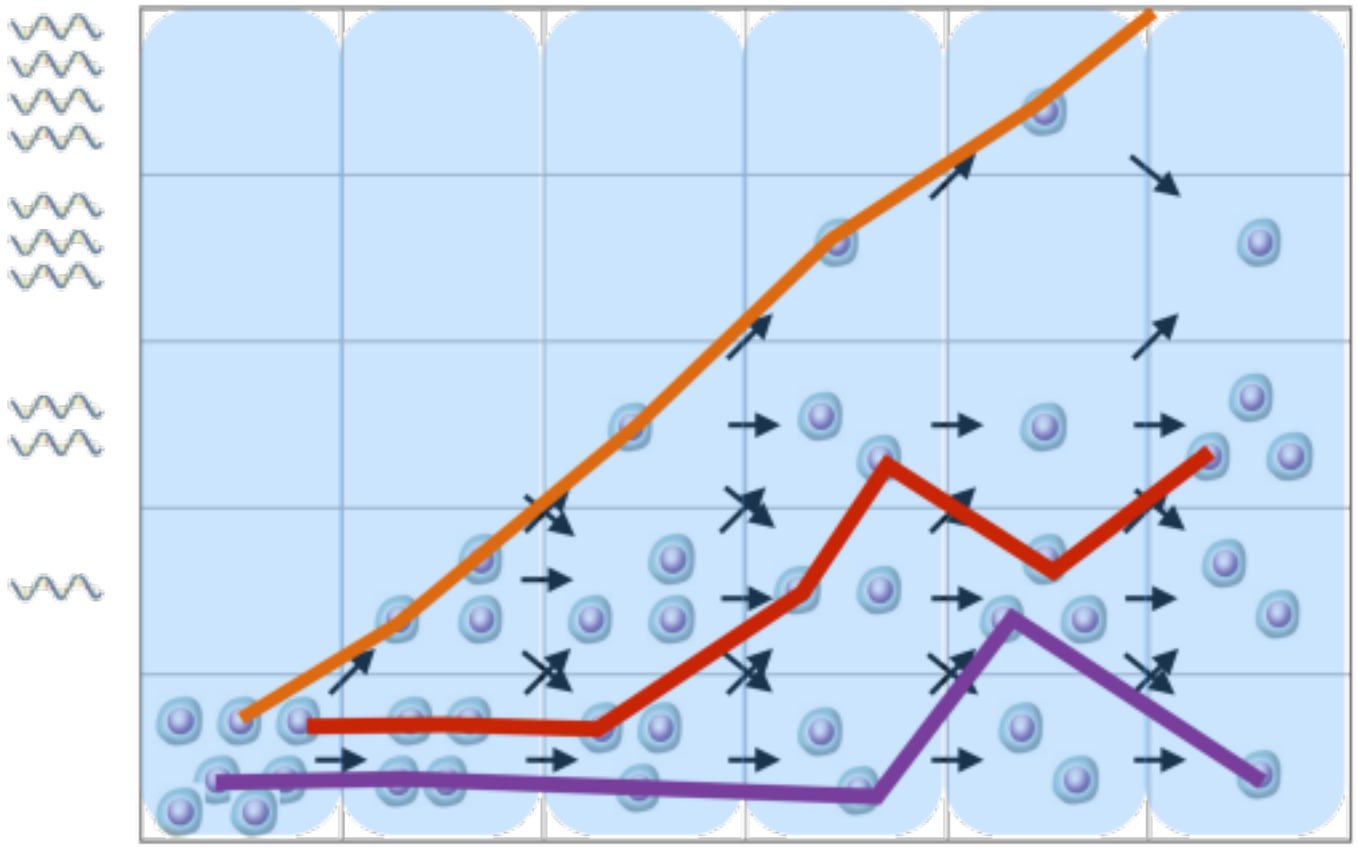
$$c2 = 2\frac{k_2}{V^2}$$

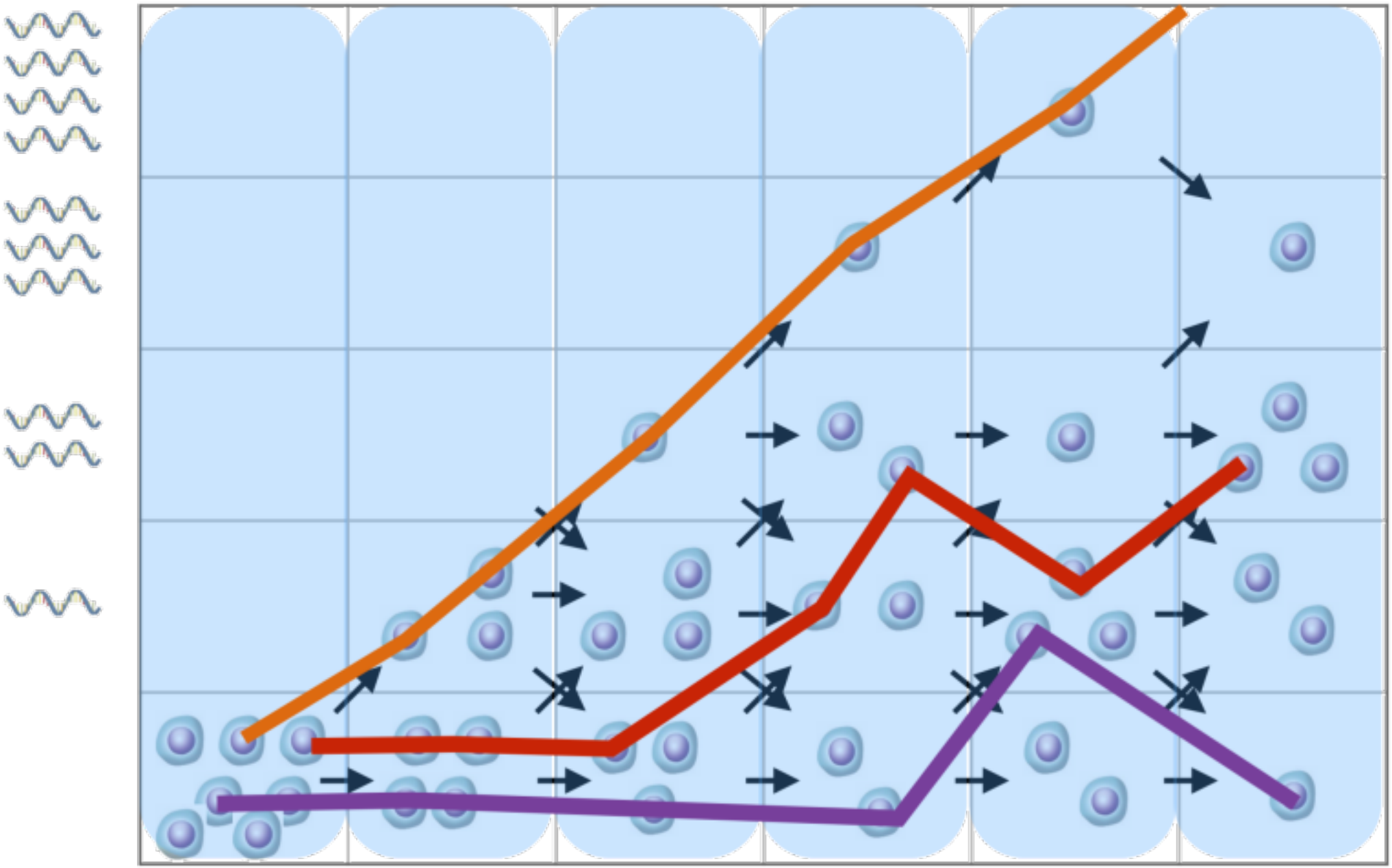
Units?

Volume effect?

# Master equation vs. stochastic simulation algorithm

- Master equation: Probability distribution over time
- Stochastic simulation: Calculates a particular trajectory





# The Gillespie algorithm

- Basically, we can simulate the evolution of a chemical system by repeatedly asking this two following questions:
  - *Which reaction will happen next?*
  - *When will it happen?*
- How to find the response to them?
  - This is the main goal of this session

# The Gillespie algorithm: First, Refreshment

- The total volume of the system, *V*.
- A series of chemical species: $S_i$ $(i = 1,2, \dots, N)$, N total number of chemical species.
- The quantities of those elements at time $t$: $\mathrm{X}_i$ $(i = 1,2, \dots, N)$
- The possible reactions: $R_\mu (\mu = 1,2, \dots, M)$ M, total number of reactions.
- Parameter for the evolution of the system: $c_\mu$(stochastic)

$$\mathrm{R}_2\colon\ 2C \xrightarrow{k_2} A + B\ ; \frac{a_2}{dt} = \frac{c_2 n_C (n_c - 1)}{2} \rightarrow R_\mu\colon\ a_\mu = c_\mu h_\mu dt$$

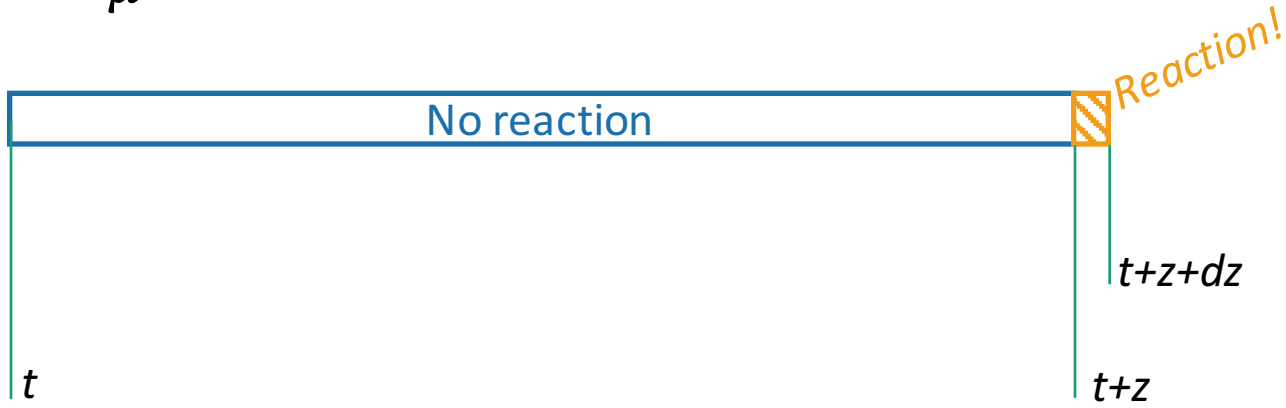
# The Gillespie algorithm:

$$P(z,\mu)dz = P_0(z)a_\mu dz$$

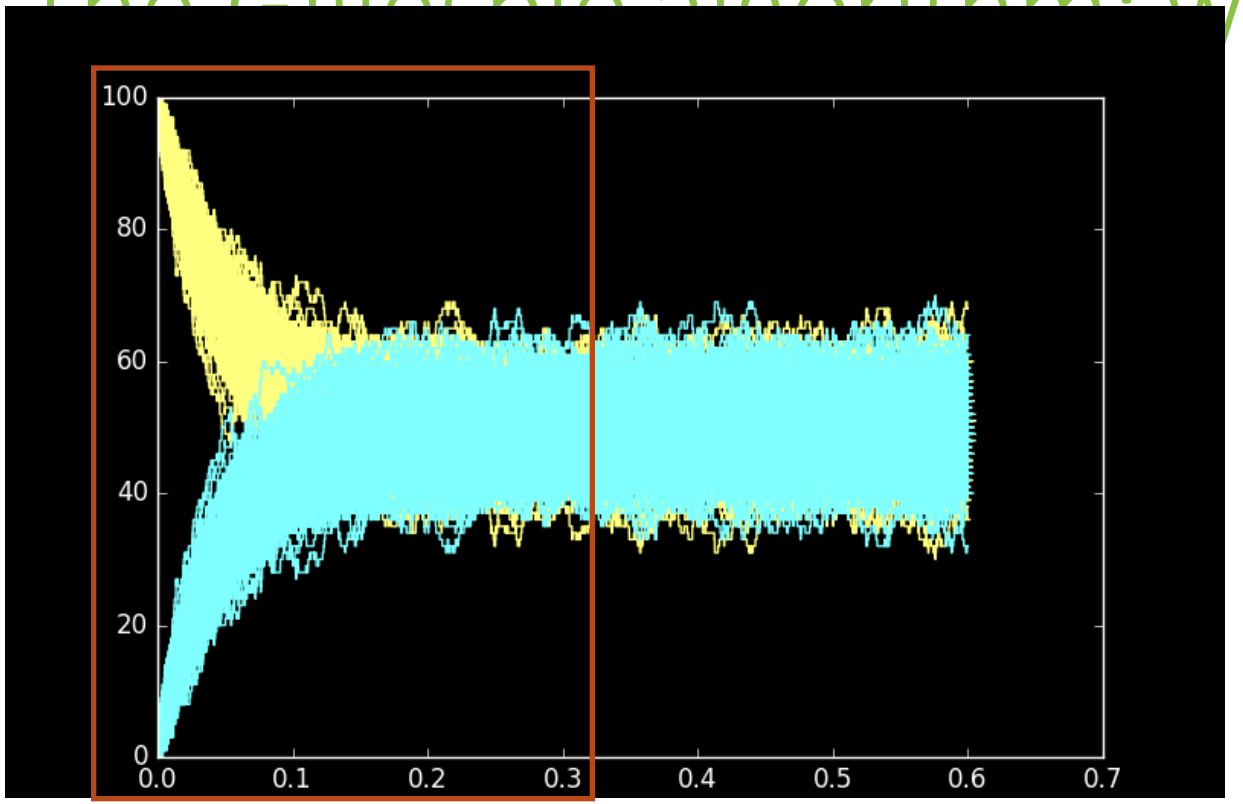
$P(z,\mu)dz$ : probability that no reaction happens in the interval *(t, t+z)*, and THE NEXT reaction happens in the interval *(t+z,t+z+dz) and it's* $R_\mu$.

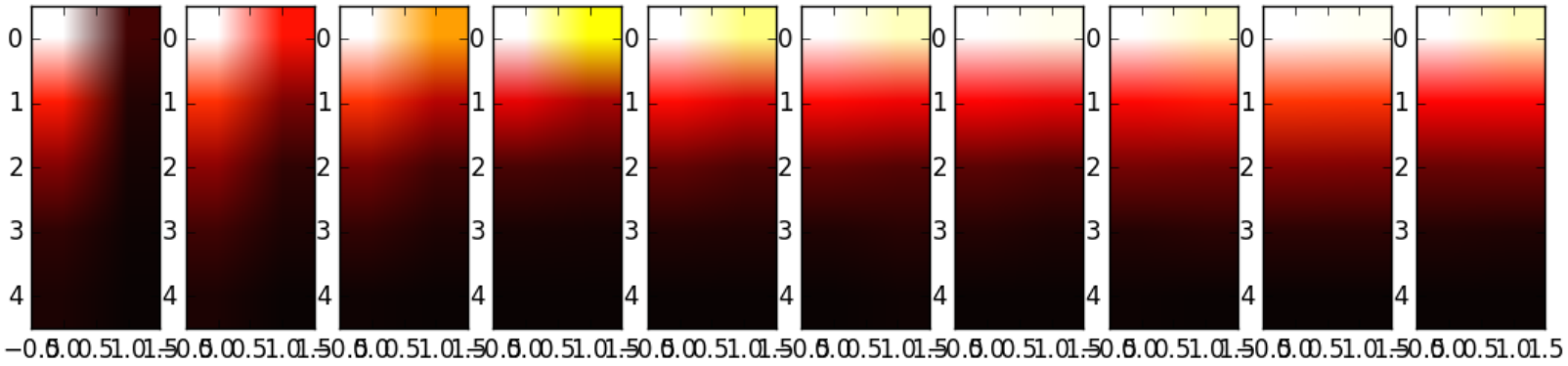
$P_0(z)$ : probability that no reaction happens in interval *(t, t+z)*

$a_\mu dz$ : probability that THE NEXT reaction happens in the interval *(t+z,t+z+dz) and it's* $R_\mu$.



# The Gillespie algorithm: what is $P(z, \mu)dz$?



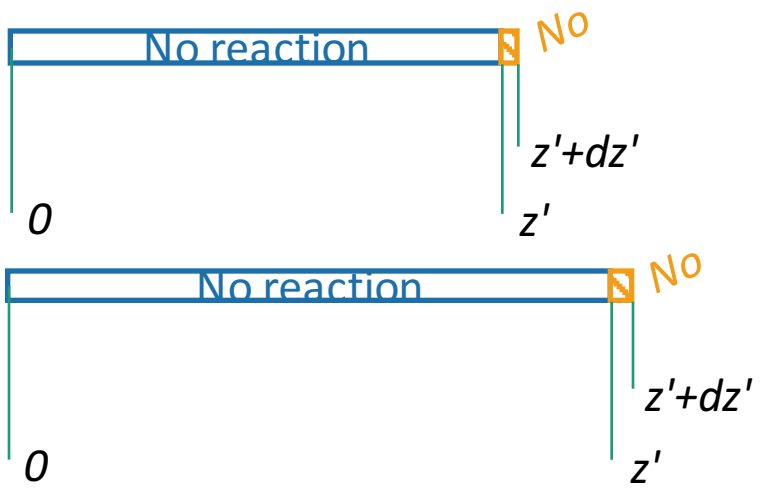
# The Gillespie algorithm

$$P(z,\mu)dz = P_0(z)a_\mu dz$$

$$P_0(z' + dz') = P_0(z')\left[1 - \sum_{m=1}^{M} a_m dz'\right] = P_0(z') - P_0(z')a_0 d_z$$

Probability of no reaction

$$P_0(z' + dz') - P_0(z') = dP_0(z') = -P_0(z')a_0 d_z$$




$$\frac{dP_0(z')}{P_0(z')} = -a_0 dz' \rightarrow \ln(P_0(z')) = a_0 z' \rightarrow P_0(z) = e^{-a_0 z}$$

$$Substitute\ in\ the\ first\ equation:\ P(z,\mu)\cancel{dz} = e^{-a_0 z}a_\mu \cancel{dz}$$

$$P(z,\mu) = e^{-a_0 z}a_\mu$$

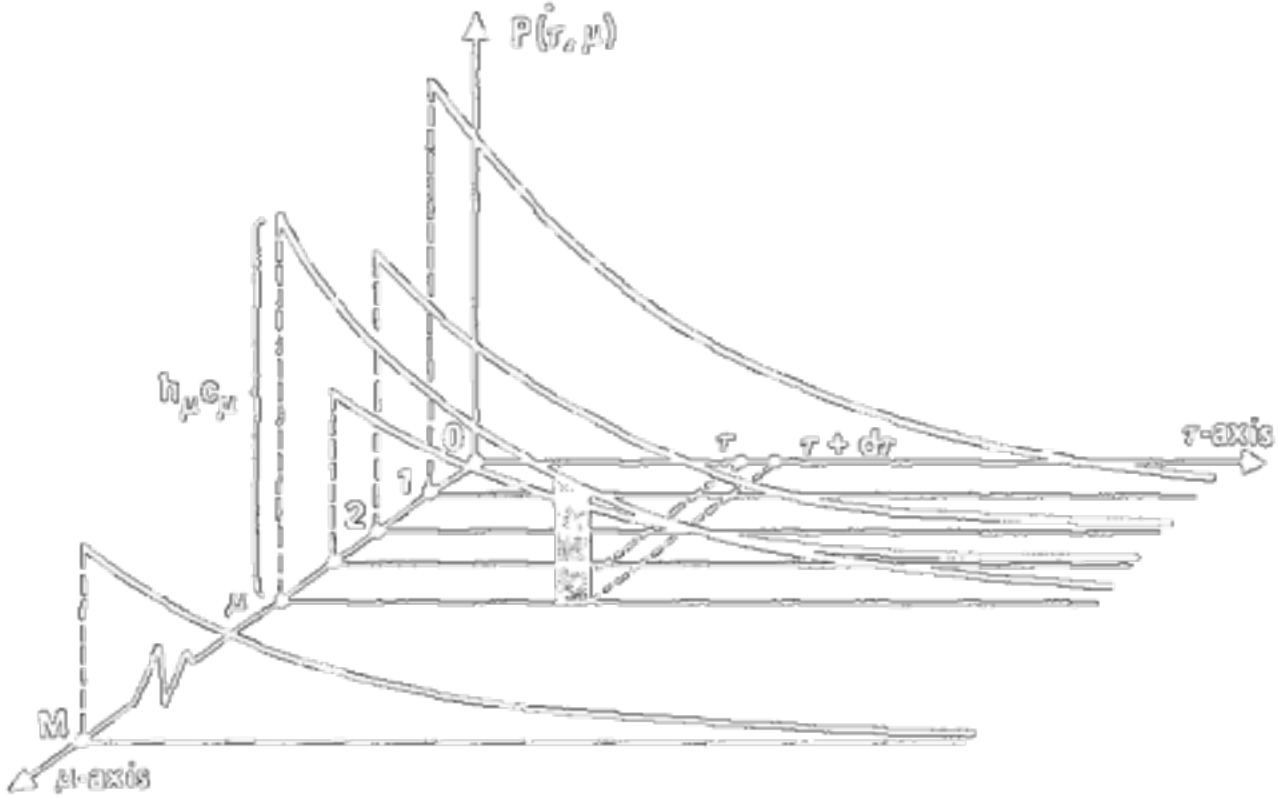
# The Gillespie algorithm

$P(z,\mu) = e^{-a_0 z} a_\mu$ is a probability density function

What does it describe?

Why is it an exponential?

Is it useful for our purpose?



# The Gillespie algorithm

Another way to think about it… $P(z,\mu) = a_\mu e^{-a_0 z}$

$$P(z,\mu) = P_{any}(z)P(\mu|z)$$

Probability of any reaction to happen during z * that reaction being $\mu$

$$P_{any}(z) = \sum_{\mu=1}^{M} a_\mu e^{-a_0 z} = a_0 e^{-a_0 z}$$

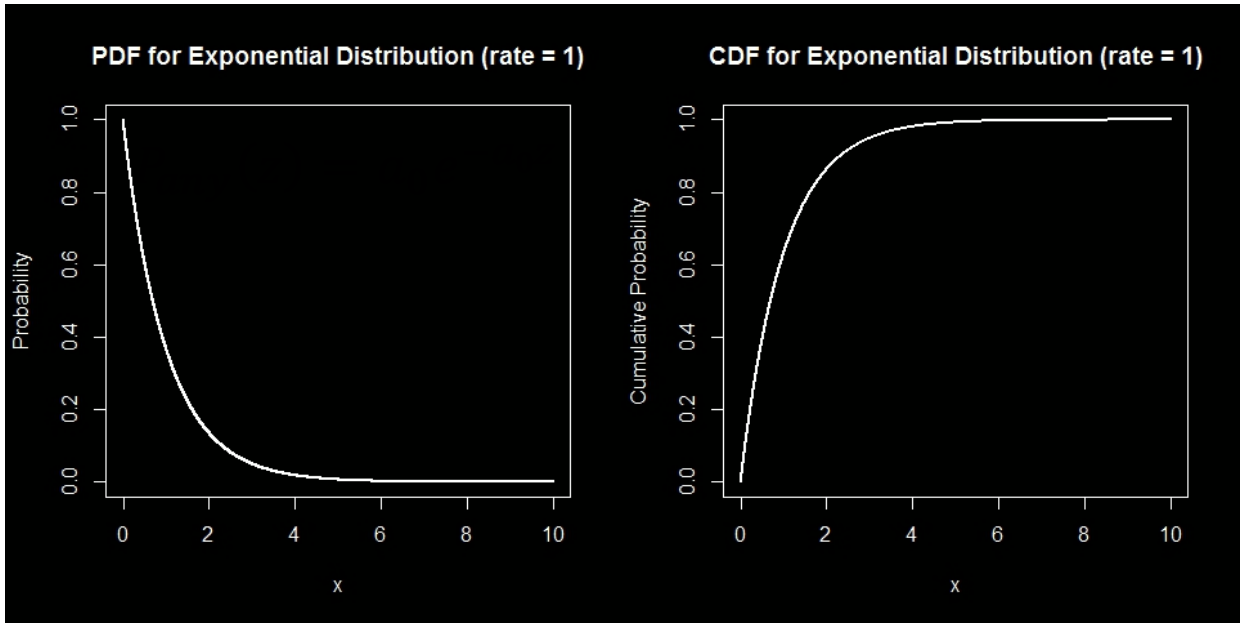
$$P(\mu|z) = \frac{a_\mu}{a_0}$$

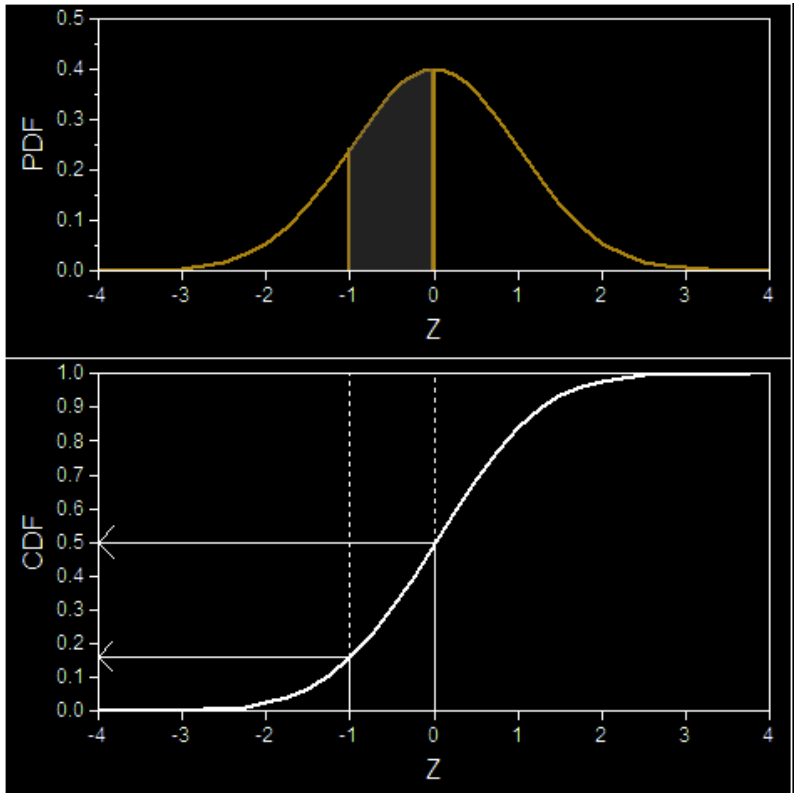
Two independent functions describe the probabilities of our questions

How do these pdfs look? What information can we extract from them?

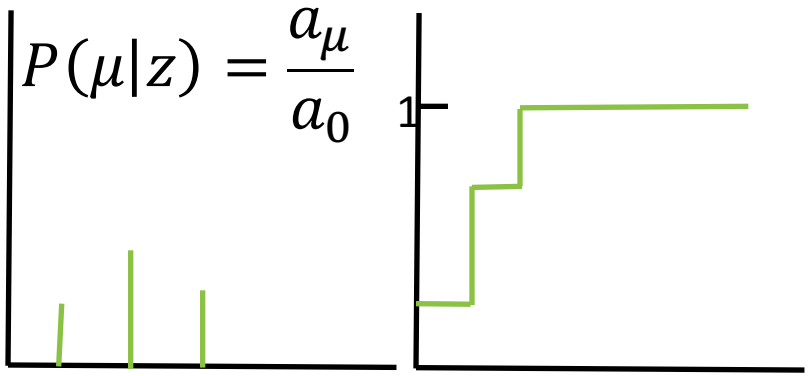
How can we use pairs of $(z,\mu)$ that satisfy those pdfs?

# How to sample random values from a PDF




$$P(\mu|z) = \frac{a_\mu}{a_0}$$




Sampling doing F(x)=random number

# How to sample values from a PDF

$$F(z) = \int_0^z P_{any}(z)\mathrm{dz} = \int_0^z a_0 e^{-a_0 z}\mathrm{dz} = 1 - e^{-a_0 z}$$

$$F(z) = 1 - e^{-a_0 z} \rightarrow \ln\big(1 - F(z)\big) = -a_0 z$$

$$z = \frac{1}{A}\ln\left(\frac{1}{1 - F(z)}\right)$$

$$z(r) = \frac{-1}{A}\ln(r)$$

Answer 1: When will the next reaction occur?

# How to sample values from a PDF

$$F(z) = \int_0^z P_{any}(z)\mathrm{dz} = \int_0^z a_0 e^{-a_0 z}\mathrm{dz} = 1 - e^{-a_0 z}$$

$$F(z) = 1 - e^{-a_0 z} \rightarrow \ln\big(1 - F(z)\big) = -a_0 z$$

$$z = \frac{1}{A}\ln\left(\frac{1}{1 - F(z)}\right)$$

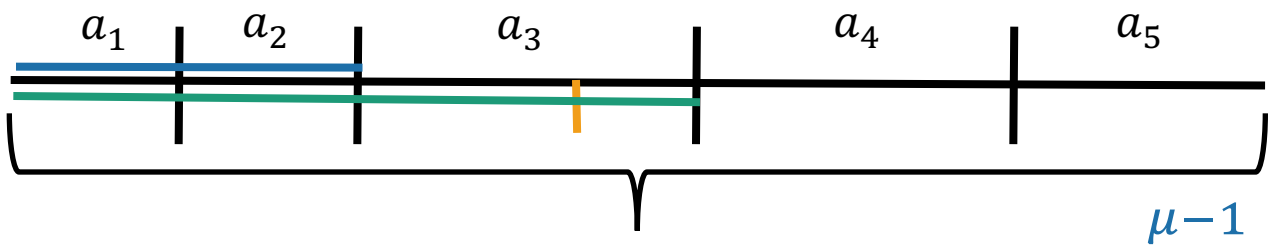
$$z(r) = \frac{-1}{A}\ln(r)$$

Answer 1: When will the next reaction occur?

# How to sample values from a PDF

$$P(\mu|z) = \frac{a_\mu}{a_0}$$

- How to think about it in this discrete case?



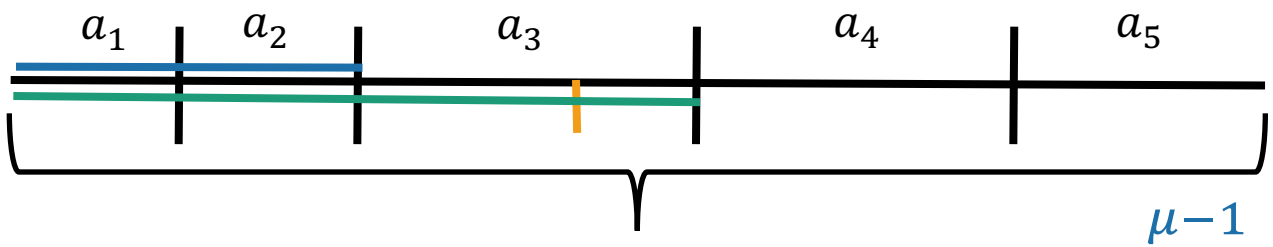

$$\mu \text{ is the integer for which: } \sum_{m=1}^{\mu-1} a_m < a_0 r \leq \sum_{m=1}^{\mu} a_m$$

Answer 2: What reaction will it be?

# How to sample values from a PDF

$$P(\mu|z) = \frac{a_\mu}{a_0}$$

- How to think about it in this discrete case?




$$\mu \text{ is the integer for which: } \sum_{m=1}^{\mu-1} a_m < a_0 r \leq \sum_{m=1}^{\mu} a_m$$

Answer 2: What reaction will it be?

# Possible projects

- Try it on your own data
- Build master equation
- Think of an example and compare stochastic and deterministic results
- How does tau change with time in different situations?
- Suggestions?