Lecture 4b

Solutions to the master equation: method of generating functions and Gillespie algorithm

CONTENTS

I. INTRODUCTION

Consider a random experiment with discrete outcomes defined by the probability space $(\Omega \subset \mathbb{R}, \mathcal{A}, P)$ and the master equation

$$\frac{\partial}{\partial t} p(x,t) = \int_{\Omega} dx' \left[W(x,t|x',t)p(x',t) - W(x',t|x,t)p(x,t) \right] , \tag{1}$$

or the equivalent for processes defined on discrete sample spaces $\Omega \subset \mathbb{Z}$:

$$\frac{\partial}{\partial t} p(n,t) = \sum_{n'} \left[W(n,t|n',t)p(n',t) - W(n',t|n,t)p(n,t) \right] . \tag{2}$$

To solve the master equation there are several options, for example:

- If the rates $W(x,t|x',t)$ and $W(x',t|x,t)$ are linear: probability generating functions;
- Time-driven or event-driven (e.g. Gillespie algorithm) simulations;

Here we see the method of generating functions applied to the pure birth process, and the Gillespie algorithm for a generic system with N states and R reactions.

II. MASTER EQUATION OF A PURE BIRTH PROCESS

Consider a population of n individuals that can increase by one individual at a rate of μ:

$$n \xrightarrow{\mu} n + 1\,. \tag{3}$$

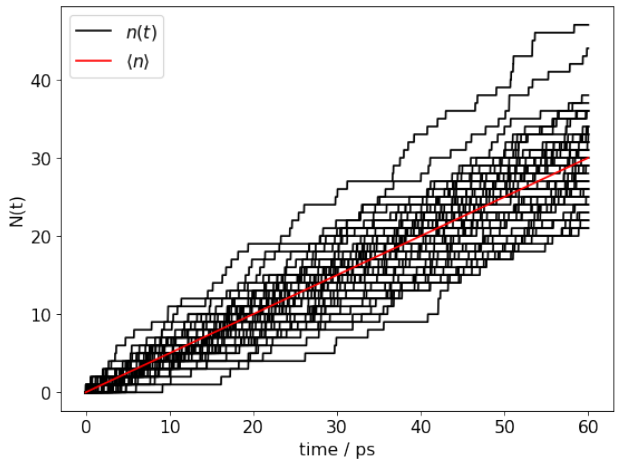

FIG. 1. Pure birth process.

The transition rate μ (units [time^{-1}]) represents the probability that an event occurs in an infinitesimal timestep, then the transition probability (unit less) in a timestep Δt is defined as

$$P(n+1,\, t+\Delta t|\, n,\, t) = \mu \Delta t\,. \tag{4}$$

The probability $P(n, t+\Delta t)$ to have n individuals at time $t+\Delta t$ is given by the sum of

- the probability there were $n-1$ individuals at time t, times the probability to increase the population by one in a time step Δt (eq. 4):

$$P(n-1, t) \cdot \mu \Delta t\,,$$

- the probability there were n individuals at time t, times the transition probability that no increase will occur:

$$P(n, t) \cdot (1 - \mu \Delta t)\,.$$

Then

$$P(n, t+\Delta t) = P(n-1, t) \cdot \mu \Delta t + P(n, t) \cdot (1 - \mu \Delta t)\,. \tag{5}$$

Rearranging eq. 5 and taking the limit $\Delta t \to 0$ yields the master equation

$$\boxed{\frac{\partial P(n,t)}{\partial t} = \mu P(n-1,t) - \mu P(n,t)\,.} \tag{6}$$

III. GENERATING FUNCTION METHOD

Consider the probability generating function

$$G(z,t) = \sum_{n=0}^{\infty} z^n \cdot P(n,t)\,, \tag{7}$$

where z is a complex number. Multiply the master equation defined in eq. 6 by z^n and sum over n:

$$\begin{aligned}
\frac{\partial \sum_n^\infty z^n P(n,t)}{\partial t} &= \mu \sum_n^\infty z^n P(n-1,t) - \mu \sum_n^\infty z^n P(n,t) \\
\frac{\partial G}{\partial t} &= \mu \sum_n^\infty z^n P(n-1,t) - \mu \sum_n^\infty z^n P(n,t) \\
&= \mu z \sum_n^\infty z^{n-1} P(n-1,t) - \mu \sum_n^\infty z^n P(n,t) \\
&= \mu z G(z,t) - \mu G(z,t) \\
&= \mu(z-1) G(z,t)\,,
\end{aligned} \tag{8}$$

The last line of eq. 8 is an ordinary differential equation, whose solution is

$$G(z,t) = Ae^{-\mu(z-1)t}\,, \tag{9}$$

where A is an arbitrary constant. Using $G(1,t) = \sum_{n=0}^\infty P(n,t) = 1$ (from eq. 7) and $G(1,t) = A$ (from eq. 9), we obtain $A = 1$. Thus we have

$$\begin{aligned}
G(z,t) &= e^{-\mu(z-1)t} \\
&= e^{\mu z t} e^{-\mu t} \\
&= e^{-\mu t} \sum_{n=0}^\infty \frac{1}{n!} (\mu z t)^n \\
&= \sum_{n=0}^\infty z^n P(n,t)
\end{aligned} \tag{10}$$

From eq. 10, we find the solution to the master equation

$$P(n,t) = \frac{1}{n!} (\mu t)^n e^{-\mu t}\,, \tag{11}$$

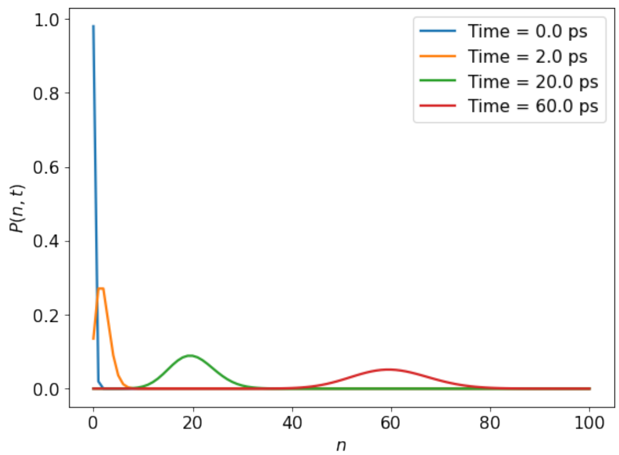

FIG. 2. Poisson distribution.

which is the Poisson distribution.

Using the generating functions, we find the moments. The mean is

$$\begin{aligned} \langle n \rangle &= \left.\frac{\partial G}{\partial z}\right|_{z=1} && (12)\\ &= \sum_{n=0}^{\infty} n\, z^{n-1}\Big|_{z=1} P(n,t) && (13)\\ &= \sum_{n=0}^{\infty} nP(n,t) && (14)\\ &= \mu t\,, && (15) \end{aligned}$$

where we used eq. 11 into eq. 14. Likewise, the variance is

$$\begin{aligned} \langle n^2 \rangle - \langle n \rangle^2 &= && (16)\\ \left.\frac{\partial^2 G}{\partial z^2}\right|_{z=1} + \langle n \rangle - \langle n \rangle^2 &= \langle n^2 \rangle - \langle n \rangle + \langle n \rangle - \langle n \rangle^2 && (17)\\ &= \mu^2 t^2 + \mu t - \mu^2 t^2 && (18)\\ &= \mu t\,, && (19) \end{aligned}$$

where we used

$$\left.\frac{\partial^2 G}{\partial z^2}\right|_{z=1} = \sum_{n=0}^{\infty} n(n-1)\left. z^{n-2}\right|_{z=1} P(n,t) \tag{20}$$

$$= \sum_{n=0}^{\infty} n(n-1)P(n,t) \tag{21}$$

$$= \sum_{n=0}^{\infty} n^2 P(n,t) - \sum_{n=0}^{\infty} nP(n,t) \tag{22}$$

$$= \langle n^2\rangle - \langle n\rangle \tag{23}$$

$$= \mu^2 t^2 \,. \tag{24}$$

IV. GILLESPIE'S ALGORITHM

Gillespie's algorithm is used to generate paths, whose time-dependent distribution is the solution of the master equation. Consider a discrete stochastic process characterized by N possible states and R reactions with rates $\mu(\mathbf{n},\, t) = \{\mu_1(\mathbf{n},\, t),\ \mu_2(\mathbf{n},\, t),\, \dots,\, \mu_R(\mathbf{n},\, t)\}$. For example the chemical reaction

$$A + B \underset{\mu_2}{\overset{\mu_1}{\rightleftharpoons}} AB\,, \tag{25}$$

has $N = 3$ possible states and $R = 2$ possible reactions. At time t, the system is in a state $\mathbf{n}(t) = \{n_1(t),\, n_2(t),\, \dots,\, n_N(t)\}$. Then the algorithm is used (i) to calculate the time $t+\tau$ at which the next reaction occurs, (ii) to select which reaction occurs.

To derive the precise steps of the algorithm, we introduce the next-jump probability density function [1], which represents the probability that, given the process is in state $\mathbf{n}$ at time t, its next jump $\mathbf{n} \to \mathbf{n}'$ will occur between $t+\tau$ and $t+\tau+d\tau$:

$$p(\mathbf{n}',\, t+\tau+d\tau|\mathbf{n},\, t)\,. \tag{26}$$

Eq. 26 is the product of three terms:

$$\begin{aligned} &p(\mathbf{n}',\, t+\tau+d\tau|\mathbf{n},\, t)\\ &= q(\mathbf{n}'\neq\mathbf{n},\, t+\tau+d\tau|\,\mathbf{n},\, t+\tau) \times (1 - q(\mathbf{n}'\neq\mathbf{n},\, t+\tau|\,\mathbf{n},\, t)) \times w(\mathbf{n}',\, t+\tau|\mathbf{n},\, t+\tau)\,, \end{aligned} \tag{27}$$

where

1. The first term is the probability that the system in state $\mathbf{n}$ at time $t+\tau$ will change state in the next infinitesimal timestep $d\tau$, independently on the arrival state $\mathbf{n}'$ (we just require that $\mathbf{n}' \neq \mathbf{n}$):

$$q(\mathbf{n}' \neq \mathbf{n},\, t+\tau+d\tau |\, \mathbf{n},\, t+\tau) = \sum_{i=1}^{R} \mu_i(\mathbf{n},\, t+\tau) d\tau = a(\mathbf{n},\, t+\tau) d\tau\,; \tag{28}$$

2. The second term is the probability that no system change will occur in the time interval $[t, t+\tau]$:

$$(1 - q(\mathbf{n}' \neq \mathbf{n},\, t+\tau |\, \mathbf{n},\, t)) = \exp\left(-a(\mathbf{n}, t)\tau\right) \tag{29}$$

3. The third term is by definition the probability to make the specific transition $\mathbf{n} \to \mathbf{n}'$ in a lag time τ.

Note that all three terms are unit-less.

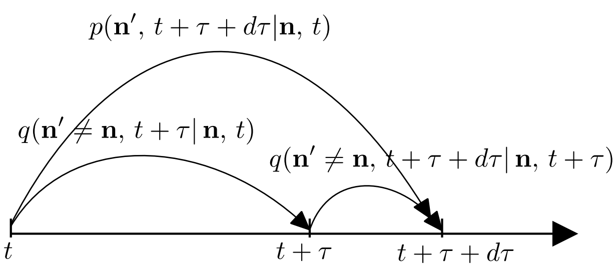

FIG. 3. Description of jump probabilities.

To derive eq. 29, we divide τ in $k \ll 1$ equal intervals of size $\varepsilon = {}^{\tau}/_{k}$. The probability that the system will not change state in a timestep ε is

$$\prod_{i=1}^{R} (1 - \mu_i(\mathbf{n},\, t)\varepsilon) \approx 1 - \sum_{i=1}^{R} \mu_i(\mathbf{n},\, t)\varepsilon + O(\varepsilon^2)\,. \tag{30}$$

Then

$$(1 - q(\mathbf{n}' \neq \mathbf{n},\, t+\tau |\, \mathbf{n},\, t)) = \left(1 - \sum_{i=1}^{R} \mu_i(\mathbf{n},\, t)\varepsilon\right)^k \tag{31}$$

$$= \exp\left(-\tau \sum_{i=1}^{R} \mu_i(\mathbf{n},\, t)\right) \tag{32}$$

$$= \exp\left(-\tau a(\mathbf{n},\, t)\right)\,, \tag{33}$$

where we used

$$\lim_{k\to\infty}\left(1+\frac{1}{k}\right)^k = e\,. \tag{34}$$

In conclusion, the next-jump probability density function (eq. 26) is written as

$$p(\mathbf{n}',\, t+\tau+d\tau|\mathbf{n},\, t) = a(\mathbf{n})d\tau\, \exp\left(-\tau a(\mathbf{n},\, t+\tau)\right)\, w(\mathbf{n}',\, t+\tau|\mathbf{n},\, t+\tau)\,. \tag{35}$$

In a event-driven simulation, the first two terms can be used to determine the time $t+\tau$ at which the next reaction occurs, while the last term is used to determine a specific reaction.

The lag time τ can be determined calculating the cumulative distribution function (CDF)

$$\int_0^\tau d\tau'\, a(\mathbf{n},\, t+\tau)\exp\left(-\tau a(\mathbf{n},\, t+\tau')\right) \tag{36}$$

Assuming that the process is temporally homogeneous, i.e. that the function a does not depend on time, we obtain

$$\int_0^\tau d\tau'\, a(\mathbf{n})\exp\left(-\tau a(\mathbf{n})\right) = -\exp\left(-\tau a(\mathbf{n})\right)\Big|_0^\tau = 1-\exp\left(-\tau a(\mathbf{n})\right)\,. \tag{37}$$

Because the CDF is a number between 0 and 1, we estimate τ applying the probability integral transform:

$$1-e^{a(\mathbf{n})\tau} = u_1 \in \mathcal{U}(0,1)\,, \tag{38}$$

where u_1 is a random number drawn from the uniform distribution. From the inverse of the CDF, we obtain

$$\tau = -\frac{\log u_1}{a(\mathbf{n})}\,. \tag{39}$$

After having randomly drawn τ, we select which reaction occurs. The ratio

$$\frac{\mu_i(\mathbf{n})}{a(\mathbf{n})}\,, \tag{40}$$

is a number between 0 and 1 and represents the probability that a certain reaction i occurs. Then we draw a random number u_2 from the uniform distribution and the next reaction is given by the first integer j for which

$$\frac{\sum_i^j \mu_i(\mathbf{n})}{a(\mathbf{n})} > u_2\,. \tag{41}$$

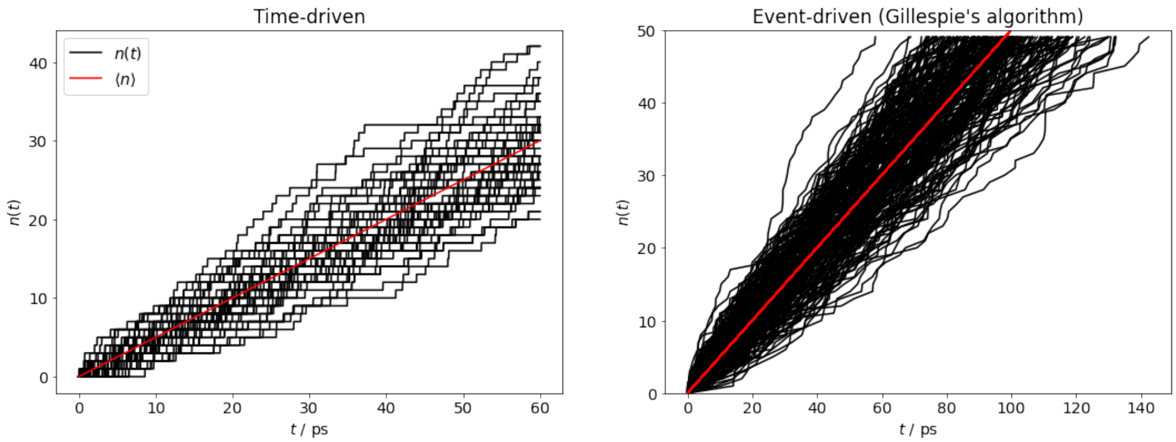

FIG. 4. Trajectories generated by time-driven and event-driven simulations.

V. PROBABILITY INTEGRAL TRANSFORM

Consider a random variable X defined on the probability space $(\Omega, \mathcal{A}, P)$, where $P : \mathcal{A} \to [0,1]$ is a probability measure with probability density function such that

$$P(X \in A) = \int_A dx\, p(x)\,, \tag{42}$$

with $A \in \mathcal{A}$ and $\int_\Omega dx\, p(x) = 1$. The cumulative density function (CFD) $F_X : \Omega \to [0,1]$ is defined as

$$F_X(x) = \int_{-\infty}^{x} dx\, p(x) = P(X \leq x)\,. \tag{43}$$

Theorem. Consider a random variable X with a continuous distribution P and CDF F_X strictly increasing, then the random variable $Y = F_X(X)$ has a uniform distribution $\mathcal{U}(0,1)$.

Proof.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(F_X(X) \leq y) \\ &= P(X \leq F_X^{-1}(y)) \\ &= F_X(F_X^{-1}(y)) \\ &= y\,. \end{aligned}$$

The CDF that satisfies $F_Y(y) = y$ is the CDF of the uniform distribution with probability density

$$p(X) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{else} \end{cases} , \tag{44}$$

indeed

$$F_X(x) = \int_{-\infty}^{x} dx\, p(x) = \begin{cases} x \in [0,1] & \\ 0 & \text{if x} < 0 \\ 1 & \text{if x} \leq 1 \end{cases} , \tag{45}$$

■

From the theorem, it follows that if we need to generate a random variable X from the distribution P, then we can draw a random number u from the uniform distribution, and take the inverse of the CDF:

$$X = F_X^{-1}(u) , \tag{46}$$

where $u \in \mathcal{U}(0,1)$.

We have required that the CDF F_X is strictly increasing, thus the inverse F_X^{-1} is well defined. The theorem can be generalized by introducing the quantile function, which is the generalization of the inverse of F_X:

$$F_X^{-1}(y) = \inf \{x : F_X(x) = y\} . \tag{47}$$

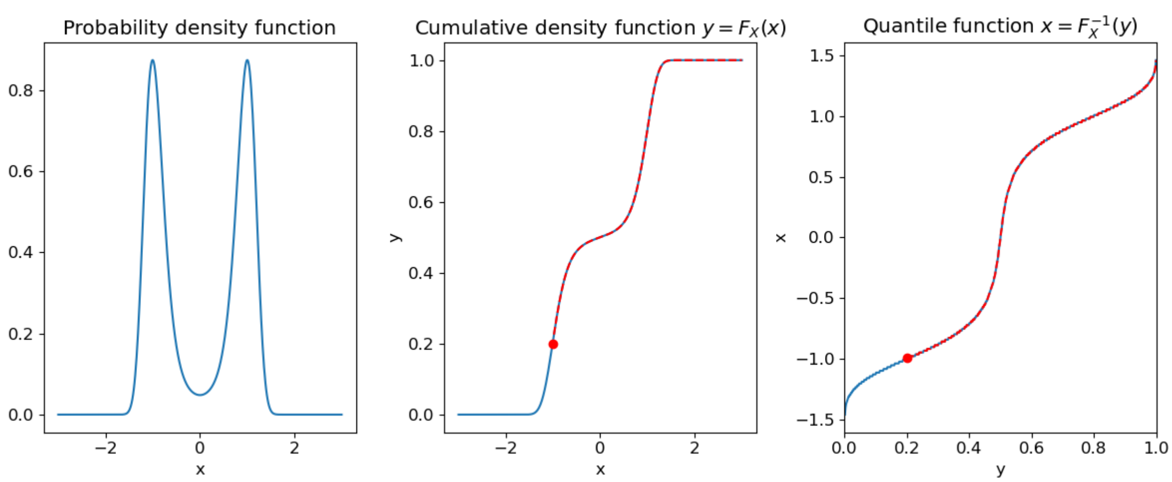

FIG. 5. Probability distribution, CDF and inverse CDF.

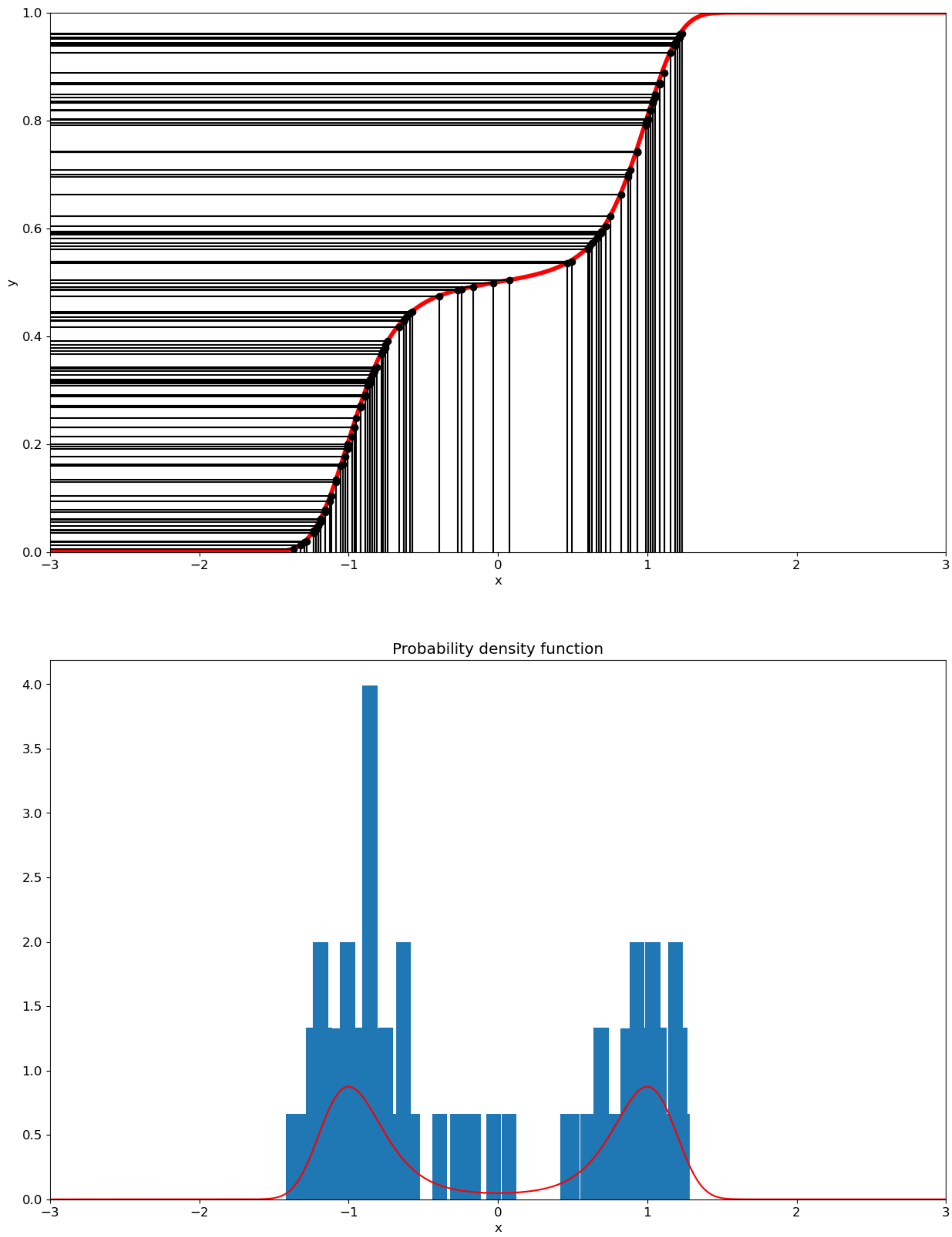

FIG. 6. Given a set of random numbers extracted from the uniform distribution (y-axis), the inverse of the CDF F_X^{-1} makes it possible to generate a sample of points distributed according to the P distribution (x-axis).

[1] D. Gillespie, Markov Processes: An Introduction for Physical Scientists (Elsevier Science, 1992).