Nonequilibrium processes 2018
Notes 1

Ingrid Strandberg

March 28, 2018

The words *random* and *stochastic* are used interchangeably.

1 Stochastic processes

A stochastic process is a set of random variables indexed by some parameter t, that can be for example time. Let $\mathcal{S}$ be the sample space, and $\zeta \in \mathcal{S}$ be a random variable. A continuous time stochastic process is denoted by $X(t, \zeta)$, where for each value of t we associate a different random variable $\zeta \in \mathcal{S}$. Random processes can be continuous or discrete in either time or amplitude. Usually in the literature the dependence on ζ is not written explicitly, and we simply write $X(t) \equiv X_t$ or X_i.

1.1 Sample realizations

If one generates a signal X_t it is a *sample realization* of the process. In Fig. 1 you see three different

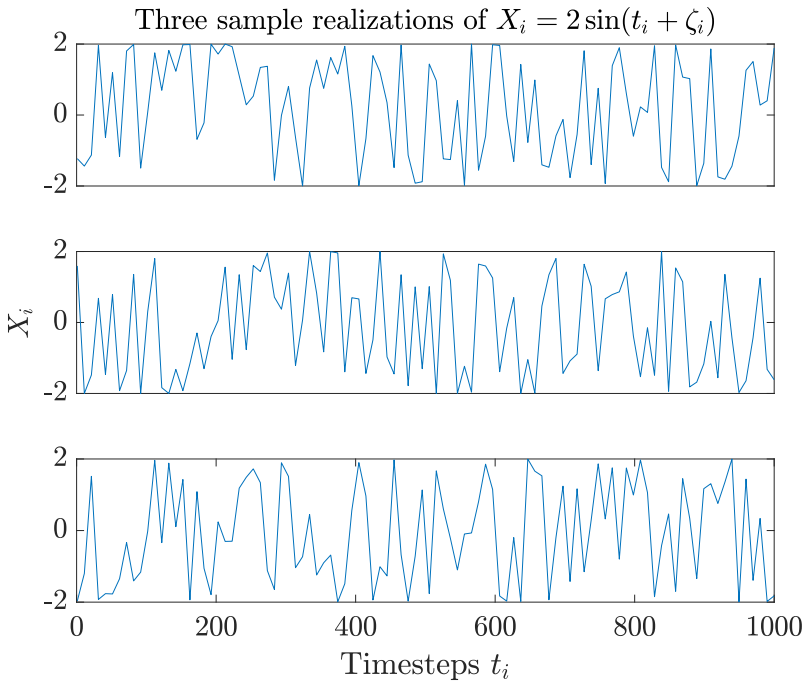

Figure 1: Caption

realizations of the stochastic process $X_t = 2\sin(\omega_0 t + \zeta)$, where ζ is uniformly distributed in the

interval $[0, 2\pi]$. The graphs are computer generated, so it is perhaps more accurate to label the process with a discrete index i than a continuous t: $X_i = 2\sin(t_i + \zeta_i)$.

Another example is the following:

A card player A is playing with a deck of standard cards numbered from 1 (ace) to 13 (king). The deck is shuffled and the player picks a card at random and observes its number. The card is then replaced, the deck reshuffled, and another card observed. This process is then repeated at unit intervals of time. The generated signal is discrete in both time and "amplitude", provided that the observed number only applies at the precise instant it is observed. This generates one sample realization of the process. In order to obtain an ensemble of sample signals, we need an ensemble of card players, each generating a different sample realization of the process.

1.2 Ensemble average vs. time average

1.2.1 Ensemble average

Eq. (1.2) in the compendium states the definition of the ensemble average. To calculate the ensemble average of a stochastic process using our notation above, we recognize that A in the compendium corresponds to the stochastic process X_t, and X in the compendium is our random variable ζ. This is an average over many realizations (ensembles) of the process.

We denote the ensemble average by a bracket $\langle\cdot\rangle$. In mathematics literature it is commonly denoted by $E[\cdot]$.

1.2.2 Time average

The time average is defined in Eq. (1.17) in the compendium. This is an average over time t for *a single realization* of the process X_t.

Fig. 2 illustrates the difference between these two types of averaging.

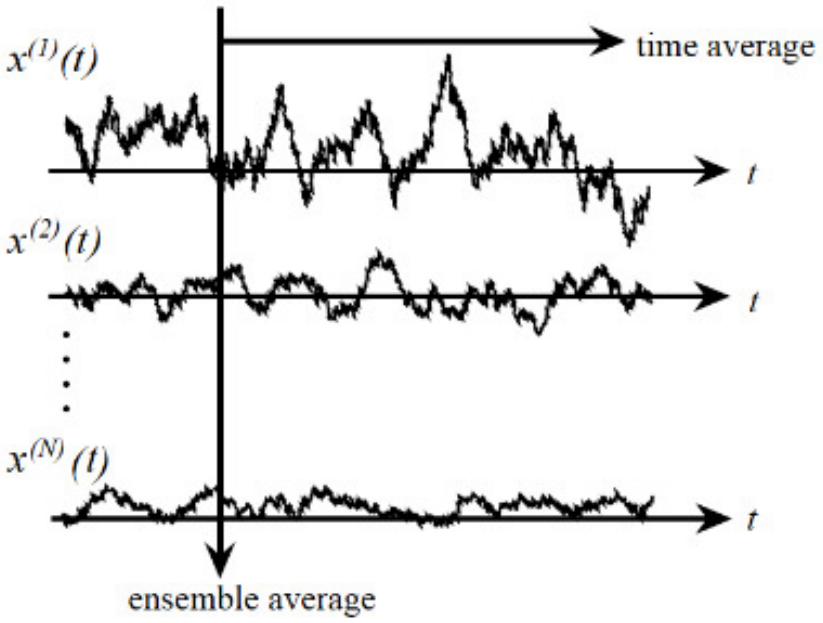

Figure 2: $x^{(i)}$, $i = 1, ..., N$ are different sample realizations of a stochastic process.

So in a numerical calculation the ensemble average would be

$$\langle x(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} x^{(i)}(t), \tag{1}$$

where N is the number of sample realizations. The time average would be

$$\overline{x} = \frac{1}{M} \sum_{i=1}^{M} x^{(1)}(t_i), \tag{2}$$

where M is the number of discrete timesteps.

2 Markov processes

Given the present state of a process, the future state is independent of the past. This property is referred to as the Markov property. The master equation (4.15) in the compendium is derived under the assumption of Markovianity.

Markov processes are classified according to the nature of the time parameter and the nature of the state space. With respect to state space, a Markov process can be either a discrete-state Markov process or continuous-state Markov process. A discrete-state Markov process is called a Markov chain. Similarly, with respect to time, a Markov process can be either a discrete-time Markov processor a continuous-time Markov process. Thus, there are four basic types of Markov processes:

- Discrete-time Markov chain
- Continuous-time Markov chain
- Discrete-time Markov process
- Continuous-time Markov process

We will look closer at Markov chains.

3 Markov chains

3.1 Transition matrix

The evolution of a Markov process is determined by the transition matrix W, where the matrix elements W_{nm} is the transition probability from state m at timestep s, to state n:

$$W_{nm} = W_{n \leftarrow m} = \text{Prob}(n, s+1 | m, s). \tag{3}$$

This so-called *left-stochastic* notation is different from the one used for the continuous master equation, in which the right variable $\vec{X}_1$ in $w(\vec{X}, \vec{X}_1)$ instead labels the final state. Our motivation for this choice is that in modern literature, left-stochastic notation for discrete-state master equations has become somewhat more popular.

We require

$$\sum_{n=1}^{N} W_{nm} = 1, \qquad W_{nm} \geq 0. \tag{4}$$

Matrices that have these two properties are called *stochastic matrices.*

3.2 Probability vector

If there are N states in a Markov chain, then we describe the probability that the process is in a given state by a probability *column vector*

$$\boldsymbol{P}(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_N(t) \end{pmatrix}, \tag{5}$$

where p_i is the probability that the system is in state i. The two requirements for a probability vector are $p_i \geq 0$ and $\sum_{i=1}^{N} p_i = 1$.

This normalized probability vector takes the same role as the phase space distribution f in the continuous space case.

3.3 Discrete time

The master equation can be compactly written as

$$\boldsymbol{P}(s+1) = \mathrm{W}\boldsymbol{P}(s). \tag{6}$$

We can see that after s steps, the probability vector is

$$\boldsymbol{P}(s) = \mathrm{W}^s\boldsymbol{P}(0). \tag{7}$$

3.3.1 Stationary state

The stationary distribution $\boldsymbol{P}^{\mathrm{st}}$ is reached as $\boldsymbol{P}(s \to \infty) = \boldsymbol{P}^{\mathrm{st}}$. This state satisfies

$$1\boldsymbol{P}^{\mathrm{st}} = \mathrm{W}\boldsymbol{P}^{\mathrm{st}}. \tag{8}$$

This is an eigenvalue equation and it implies that the stationary state corresponds to the eigenvector of W with eigenvalue 1. It can be proven that a stochastic matrix has an eigenvalue of 1 and no eigenvalue greater than 1. i.e. go for the max eig

Also, for a positive transition matrix, the limit $\lim_{s\to\infty} W^s$ exists. If we denote the limit by matrix A, then the columns of A are identical, and represent the stationary probability vector.

3.3.2 Detailed balance

The condition of equilibrium in closed isolated systems is much stronger that the condition of stationarity. Here we require as an additional constraint a balance between pairs of states n and n' separately. This *detailed balance* relation is written for the equilibrium distribution $\boldsymbol{P}^{\mathrm{eq}}{}_n$ as

$$W_{nn'}\boldsymbol{P}^{\mathrm{eq}}{}_{n'} - W_{n'n}\boldsymbol{P}^{\mathrm{eq}}{}_{n} = 0. \tag{9}$$

By definition, an equilibrium state is stationary. But the opposite implication does not hold: a stationary distribution does not have to be an equilibrium distribution.

3.3.3 Time evolution

W can be symmetrized if it obeys detailed balance. But in general it is not symmetric. This means that its left and right eigenvectors will be different. Here it is convenient to use Dirac notation:

$$W\left|\chi_i\right\rangle = \lambda_i\left|\chi_i\right\rangle, \quad \left\langle\chi_i\right| W = \left\langle\chi_i\right|\lambda_i, \tag{10}$$

where $\langle\chi_i|$ and $|\chi_i\rangle$ are the left and right eigenvectors, respectively. The eigenvectors are orthogonal: $\langle\chi_i|\chi_j\rangle = \delta_{ij}$, and form a complete set: $\sum_{i=1}^{M} = |\chi_i\rangle\langle\chi_i| = \mathbb{1}$, where M is the size of W. We can write

$$W = \sum_{i=1}^{M} \lambda_i\left|\chi_i\right\rangle\left\langle\chi_i\right| = \sum_i \lambda_i B^{(i)}, \tag{11}$$

where we define the matrix

$$B^{(i)} = |\chi_i\rangle \langle\chi_i| . \tag{12}$$

For s number of steps we obtain

$$W^s = \sum_{i=1}^{M} \lambda_i^s B^{(i)}, \tag{13}$$

and

$$\boldsymbol{P}(s) = \sum_{i=1}^{M} \lambda_i^s B^{(i)} \boldsymbol{P}(0) = \boldsymbol{P}^{\text{st}} + \sum_{i=2}^{M} \lambda_i^s B^{(i)}. \tag{14}$$

We see that independent on the initial distribution $\boldsymbol{P}(0)$, the probability distribution approaches the stationary distribution $\boldsymbol{P}^{\text{st}}$ at long times (sufficiently large number of steps s). The rate of the approach is determined by the eigenvalues λ_i (with $|\lambda_i| < 1$, for $i \geq 2$).

3.3.4 Example

Consider the daily weather in Göteborg. It can be in three different states: (1) sunny (2) cloudy, or (3) rain. Construct the transition matrix based on the following observations. A sunny day is never followed by another sunny day. Rainy or cloudy weather is equally probable after a sunny day. A rainy or cloudy day is followed by 50% probability by another day with the same weather. If, on the other hand, the weather is changing from cloudy or rainy weather, the following day will be sunny only in half of the cases. From this the following transition matrix can be constructed:

$$\mathrm{W} = \begin{pmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{pmatrix}. \tag{15}$$

The eigenvector of W corresponding to the eigenvalue 1 is

$$|\chi_1\rangle = \boldsymbol{P}^{\text{st}} = \begin{pmatrix} 0.2 \\ 0.4 \\ 0.4 \end{pmatrix}. \tag{16}$$

The other two eigenvalues are $\lambda_2 = 0.25$ and $\lambda_2 = -0.25$ The right and left eigenvectors are

$$|\chi_2\rangle = \begin{pmatrix} 0 \\ 0.5 \\ -0.5 \end{pmatrix}, \quad |\chi_3\rangle = \begin{pmatrix} 0.2 \\ -0.1 \\ -0.1 \end{pmatrix}, \tag{17}$$

$$\langle\chi_1| = (1\,1\,1), \quad \langle\chi_2| = (0\,1\,{-1}), \quad \langle\chi_3| = (4\,{-1}\,{-1}). \tag{18}$$

Using Eq. (14), we get for the time evolution of the probability vector

$$\boldsymbol{P}(s) = \begin{pmatrix} 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} + \left(\frac{1}{4}\right)^s \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.5 & -0.5 \\ 0 & -0.5 & 0.5 \end{pmatrix} \boldsymbol{P}(0) + + \left(-\frac{1}{4}\right)^s \begin{pmatrix} 0.8 & -0.2 & -0.2 \\ -0.4 & 0.1 & 0.1 \\ -0.4 & 0.1 & 0.1 \end{pmatrix} \boldsymbol{P}(0). \tag{19}$$

3.4 Continuous time

In discrete-time Markov chains, a change of state could only occur at specific points in time. For continuous time Markov chains, a change of state can occur at any time. The construction of a continuous-time Markov chain involves combining a discrete Markov chain and a random process that chooses when the change of state occurs. The transition mechanism must be linked to an exponential distribution because of the memoryless property. The simplest example of a continuous-time Markov chain is a Poisson process.

Definition

The matrix elements of the infinitesimal generator of the Markov chain are the combination of a gain/influx and a loss/outflux term,

$$\mathrm{G}_{nm} = g_{nm} - \delta_{nm} \sum_{k=1}^{N} g_{km}, \quad g_{n\neq m} \geq 0, \quad g_{nn} = 0. \tag{20}$$

This is a matrix whose entries are the transition probabilities per unit time. It obeys

$$G_{nm} \geq 0 \text{ for } n \neq m, \text{ and } \sum_{n} G_{nm} = 0 \text{ for each } m. \tag{21}$$

The master equation for continuous time is

$$\frac{d}{dt}\boldsymbol{P}(t) = \mathrm{G}\boldsymbol{P}(t). \tag{22}$$

The solution of this equation with given initial distribution $\boldsymbol{P}(0)$ is written formally (in the left-stochastic formulation) as

$$\boldsymbol{P}(t) = \mathrm{e}^{t\mathrm{G}}\boldsymbol{P}(0). \tag{23}$$

3.4.1 Stationary state

For any probability vector $\boldsymbol{P}$ we have

$$\lim_{t\to\infty} \mathrm{e}^{t\mathrm{G}}\boldsymbol{P} = \boldsymbol{P}^{\mathrm{st}}, \tag{24}$$

where $\boldsymbol{P}^{\mathrm{st}}$ is the unique probability vector for which $G\boldsymbol{P}^{\mathrm{st}} = 0$. We see that the stationary state is given by the eigenvector of G corresponding to the eigenvalue 0.

3.5 Generating function

The probability generating function is used in the case where the random variable takes values in the nonnegative integers. If X is such a random variable, the probability generating function for X, denoted $G(z)$, is defined by

$$G(z) = \sum_{n=0}^{\infty} p(n) z^n, \tag{25}$$

where $p(n) = P(X = n)$. Note that

$$p(n) = \frac{G^{(n)}(0)}{n!}, \tag{26}$$

where $G^{(n)}$ denotes the nth derivative of G. So the probability density function p can be recovered from the probability generating function G.

You can also calculate various moments by taking derivatives of G. For example,

$$\frac{\mathrm{d}}{\mathrm{d}z} G(z)\Big|_{z=1} = \langle n \rangle , \tag{27}$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2} G(z)\Big|_{z=1} = \langle n^2 \rangle - \langle n \rangle . \tag{28}$$

Higher moments can be obtained in an analogous manner.

3.5.1 Example: radioactive decay

Consider a sample of radioactive atoms. Let $p_n(t)$ be the probability to find n nuclei surviving at time $t > 0$. For a constant decay rate γ (probability of decaying per nucleus per unit time), the master equation can be written as

$$\frac{\partial}{\partial t} p_n(t) = \gamma(n+1)p_{n+1}(t) - \gamma n p_n(t). \tag{29}$$

Assume that there are n_0 excited nuclei in the sample at time t_0, that is, $p_n(t_0) = \delta_{n,n_0}$. The boundary conditions are

prob. rate of p_0 depends only on decay rate of 1-nuc state

$$\frac{\partial}{\partial t} p_{n_0}(t) = -n_0\gamma p_{n_0}(t), \qquad \frac{\partial}{\partial t} p_0(t) = \gamma p_1(t) \tag{30}$$

We solve this by using the generating function (25): We multiply (29) by z^n and sum over all n:

$$\sum_{n=0}^{\infty} z^n \frac{\partial p_n(t)}{\partial t} = \sum_{n=0}^{\infty} z^n \gamma(n+1)p_{n+1}(t) - \sum_{n=0}^{\infty} z^n \gamma n p_n(t), \tag{31}$$

We assume the sum converges uniformly, so we are allowed to exchange the order of summation and differentiation.

$$\frac{\partial}{\partial t} G(z,t) = \gamma \frac{\partial}{\partial z} G(z,t) - \gamma z \frac{\partial}{\partial z} G(z,t). \tag{32}$$

The boundary conditions are

Insert p(n_0)=1 into def. of G

Insert z=1 into def. of G -> just tot. prob = 1

$$G(z,t_0) = z^{n_0}, \qquad G(1,t) = 1. \tag{33}$$

The solution of Eq. (32) with boundary conditions (33) is

$$G(z,t) = \left(1 - (1-z)e^{-\gamma(t-t_0)}\right)^{n_0} \tag{34}$$

Now we use (26) to get the probability distribtution:

$$p_n(t) = \frac{1}{n!}\frac{\partial^n}{\partial z^n} G(z,t)\bigg|_{z=0} = \frac{1}{n!}\frac{n_0!}{(n_0-n)!} e^{-\gamma n(t-t_0)} \left(1 - e^{-\gamma(t-t_0)}\right)^{(n-n_0)}, \tag{35}$$

which has the form of a binomial distribution.

4 Additional literature

These relevant books can be found as ebooks via the library. Even more can be found there: `www.lib.chalmers.se`.

- Nonequilibrium Statistical Physics, Gerd Röpke
- Elements of Random Walk and Diffusion Processes, Oliver Ibe
- Physics of Stochastic Processes, R. Mahnke, J. Kaupuzs, and I. Lubashevsky
- Markov Processes for Stochastic Modeling, Oliver Ibe
- Stochastic Processes in Physics and Chemistry, N.G. Van Kampen
- Understanding Markov Chains, Nicolas Privault