Stochastic Systems: an introduction

May 10, 2018

Stochastic processes

Suppose that a system has properties which can be described in terms of a single stochastic variable Y, for example, the number of molecules in a given volume of air, the number of people in a queue,..

Then we introduce the following quantities:

$$p(y,t) \equiv \quad \text{(the probability density that the stochastic variable } Y \text{ has value } y \text{ at time } t\text{)}$$

The expectation value of Y at time t is

$$\langle Y(t) \rangle = \int_{\text{all } y} dy\, y\, p(y,t)$$

Similarly

$$\langle Y^n(t) \rangle = \int_{\text{all } y} dy\, y^n\, p(y,t)\,, \qquad \langle f(Y(t)) \rangle = \int_{\text{all } y} dy\, f(y)\, p(y,t)\,,$$

$p(y_2, t_2; y_1, t_1) \equiv$ (the joint probability density that the stochastic variable Y has value y_1 at time t_1 and y_2 at time t_2)

So, for example, the expectation value of $Y(t_2)Y(t_1)$ is

$$\langle Y(t_2)Y(t_1)\rangle = \int dy_2\, dy_1\, y_2 y_1\, p(y_2, t_2; y_1, t_1)$$

If the value of Y at time t_2 is completely independent of the value of Y at time t_1, then

$$p(y_2, t_2; y_1, t_1) = p(y_2, t_2)p(y_1, t_1)\,,$$

and we find that

$$\langle Y(t_2)Y(t_1)\rangle = \int dy_2\, y_2\, p(y_2, t_2) \int dy_1\, y_1\, p(y_1, t_1) = \langle Y(t_2)\rangle\langle Y(t_1)\rangle$$

$$\Rightarrow \quad \langle Y(t_2)Y(t_1)\rangle - \langle Y(t_2)\rangle\langle Y(t_1)\rangle$$

is a measure of the *correlation* between Y at time t_2 and time t_1

$p(y_n, t_n; \ldots; y_2, t_2; y_1, t_1) \equiv$ (the joint probability density that the stochastic variable Y has value y_1 at time t_1, y_2 at time $t_2, \ldots, y_n$ at time t_n)

Marginal and conditional probabilities may be defined as follows:

$$\int dy_n \ldots dy_{m+1}\, p(y_n, t_n; \ldots; y_2, t_2; y_1, t_1)$$
$$= p(y_m, t_m; \ldots; y_2, t_2; y_1, t_1) \quad \text{[marginal pdf]}$$

$$
\begin{aligned}
&p(y_n, t_n; \ldots; y_{m+1}, t_{m+1} | y_m, t_m; \ldots; y_1, t_1) \\
&= \frac{p(y_n, t_n; \ldots; y_1, t_1)}{p(y_m, t_m; \ldots; y_1, t_1)} \quad \text{[conditional pdf]}
\end{aligned}
$$

Conditional means may also be defined as:

$$
\langle Y(t_2) \rangle_{Y(t_1)=y_1} = \int dy_2 \, y_2 \, p(y_2, t_2 | y_1, t_1)
$$

$$
\Rightarrow \quad p(y_2, t_2 | y_1, t_1) = \langle \delta \left(Y(t_2) - y_2 \right) \rangle_{Y(t_1)=y_1}
$$

We also introduce the notation of double angle brackets for cumulants. So for stochastic variables

$$
\langle\langle X^2 \rangle\rangle = \langle X^2 \rangle - \langle X \rangle^2 \, ,
$$

which is just the variance. Similarly,

$$
\langle\langle X^3 \rangle\rangle = \langle X^3 \rangle - 3\langle X \rangle^2 \langle X \rangle + 2\langle X \rangle^3
$$

For stochastic processes an example of an analogous result is

$$\langle\langle Y(t_2)Y(t_1)\rangle\rangle = \langle Y(t_2)Y(t_1)\rangle - \langle Y(t_2)\rangle\langle Y(t_1)\rangle\,,$$

which is just the two-point correlation function

Let us end this rather formal section by defining two special types of stochastic process we'll need later on

• A process is called *stationary* when all the probability densities depend on the time differences alone:

$$\begin{aligned} &p(y_n, t_n+\tau; \ldots; y_2, t_2+\tau; y_1, t_1+\tau) \\ &= p(y_n, t_n; \ldots; y_2, t_2; y_1, t_1) \ \text{ for all } n \text{ and } \tau \end{aligned}$$

So, for example, taking $\tau = -t_1$, then

$$p(y_1, t_1 - t_1) = p(y_1, t_1) \quad \Rightarrow \quad p(y_1, t_1) = p(y_1, 0) \equiv p(y_1)$$

is time-independent

Similarly by taking different values for τ,

$$p(y_2, t_2; y_1, t_1) = p(y_2, t_2 - t_1; y_1, 0) = p(y_2, 0; y_1, t_1 - t_2)\,,$$

and so depends only on the time difference

But

$$\langle Y(t_2)Y(t_1)\rangle = \int dy_2\, dy_1\, y_2 y_1\, p(y_2, t_2; y_1, t_1)$$

is symmetric under $t_1 \leftrightarrow t_2$ and so $\langle Y(t_2)Y(t_1)\rangle$ depends only on $|t_2 - t_1|$ when the process is stationary

- A process is called *Gaussian* if all the cumulants beyond the second vanish

$\Rightarrow$ A Gaussian process is fully specified by $\langle\langle Y(t_2)Y(t_1)\rangle\rangle$ and $\langle Y(t_1)\rangle$, or equivalently by $\langle Y(t_2)Y(t_1)\rangle$ and $\langle Y(t_1)\rangle$

Markov processes

A process is *Markov* if $p(y_{k+1}, t_{k+1}|y_k, t_k; \ldots; y_1, t_1)$ depends on the *state* $Y(t_k) = y_k$, but not on $Y(t_{k-1}) = y_{k-1}, \ldots, Y(t_1) = y_1$

i.e., $p(y_{k+1}, t_{k+1}|y_k, t_k; \ldots; y_1, t_1) = p(y_{k+1}, t_{k+1}|y_k, t_k)\,\forall k$

So the conditional pdfs are affected only by the state of the system at a given time, and not by the state of the system at times prior to this.

$$
\begin{aligned}
\text{(a)} \qquad & p(y_n, t_n; \ldots; y_1, t_1) \\
& = p(y_n, t_n|y_{n-1}, t_{n-1}; \ldots; y_1, t_1)p(y_{n-1}, t_{n-1}; \ldots; y_1, t_1) \\
& p(y_n, t_n|y_{n-1}, t_{n-1})p(y_{n-1}, t_{n-1}|y_{n-2}, t_{n-2}; \ldots; y_1, t_1) \\
& \times p(y_{n-2}, t_{n-2}; \ldots; y_1, t_1) \\
& = \ldots = \prod_{i=1}^{n-1} p(y_{i+1}, t_{i+1}|y_i, t_i)p(y_1, t_1)
\end{aligned}
$$

$$
\text{(b)} \qquad \begin{aligned} & p(y_{k+\ell}, t_{k+\ell}; \ldots; y_{k+1}, t_{k+1} | y_k, t_k; \ldots; y_1, t_1) \\ &= \frac{\prod_{i=1}^{k+\ell-1} p(y_{i+1}, t_{i+1} | y_i, t_i) p(y_1, t_1)}{\prod_{i=1}^{k-1} p(y_{i+1}, t_{i+1} | y_i, t_i) p(y_1, t_1)} \\ &= \prod_{i=k}^{k+\ell-1} p(y_{i+1}, t_{i+1} | y_i, t_i) \end{aligned}
$$

(a) tells us that for Markov processes all joint pdfs can be written down in terms of the functions $p(y', t' | y, t)$ and $p(y, t)$ and (b) tells us that for Markov processes all conditional pdfs can be written down in terms of $p(y', t' | y, t)$

Using (a) and (b) we can show that the hierarchy of pdfs related through the definition of marginal and conditional pdfs collapse down to just two relations between the functions $p(y', t' | y, t)$ and $p(y, t)$

These are:

$$\text{(i)} \qquad p(y_2, t_2) = \int dy_1\, p(y_2, t_2 | y_1, t_1) p(y_1, t_1)$$

<u>*Proof*</u>

$$\begin{aligned} p(y_2, t_2) &= \int dy_1\, p(y_2, t_2; y_1, t_1) \\ &= \int dy_1\, p(y_2, t_2 | y_1, t_1) p(y_1, t_1) \end{aligned}$$

In fact we have not used the Markov assumption to derive the, rather obvious, result above

$$\text{(ii)} \qquad p(y_3, t_3 | y_1, t_1) = \int dy_2\, p(y_3, t_3 | y_2, t_2) p(y_2, t_2 | y_1, t_1)$$

<u>*Proof*</u>

$$\begin{aligned} & p(y_3, t_3; y_2, t_2; y_1, t_1) \\ &= p(y_3, t_3 | y_2, t_2) p(y_2, t_2 | y_1, t_1) p(y_1, t_1) \end{aligned}$$

Integrating over y_2 leads to

$$p(y_3, t_3; y_1, t_1) = \left\{ \int dy_2 \, p(y_3, t_3 | y_2, t_2) p(y_2, t_2 | y_1, t_1) \right\} p(y_1, t_1)$$

But $p(y_3, t_3; y_1, t_1) = p(y_3, t_3 | y_1, t_1) p(y_1, t_1)$ and the result is proved

The result (ii) is called the *Chapman-Kolmogorov equation*. It is the starting point for the study of Markov processes

For a Markov process the functions $p(y', t' | y, t)$ and $p(y, t)$ are not arbitrary; they must satisfy conditions (i) and (ii)

Conversely, any two non-negative functions $p(y', t' | y, t)$ and $p(y, t)$ that obey (i) and (ii) uniquely define a Markov process

The CK equation tells us that we can break up the probability of transition from state y_1 at time t_1 to state y_3 at time t_3 into a process involving two successive steps which are statistically independent; the probability of the transition from y_2 to y_3 is not affected by the fact that it was preceded by a transition from y_1 to y_2

Markov chains

In this case the stochastic variables are discrete and will be labelled by an integer n

Then the two equations (i) and (ii) governing Markov processes take the form

$$\begin{aligned} p(n_2, t_2) &= \sum_{n_1} p(n_2, t_2 | n_1, t_1) p(n_1, t_1) \\ p(n_3, t_3 | n_1, t_1) &= \sum_{n_2} p(n_3, t_3 | n_2, t_2) p(n_2, t_2 | n_1, t_1) \quad t_1 < t_2 < t_3 \end{aligned}$$

In addition, we take time to be discrete, so that t also takes on integer values $t = 0, 1, \ldots$

If time is discrete, the Chapman-Kolmogorov (CK) equation tells us that the conditional probability at any time, $p(n, t' | m, t) \equiv p(n, t + \ell | m, t)$ $(\ell = 2, 3, \ldots)$, can be found if the function $p(n, t + 1 | m, t)$ is known for all t

This follows because

$$
\begin{aligned}
p(n, t+2|m, t) &= \sum_{n'} p(n, t+2|n', t+1) p(n', t+1|m, t)\,, \\
p(n, t+3|m, t) &= \sum_{n'} p(n, t+3|n', t+1) p(n', t+1|m, t)\,, \quad \text{etc}
\end{aligned}
$$

This fundamental conditional pdf can be thought of as a matrix:

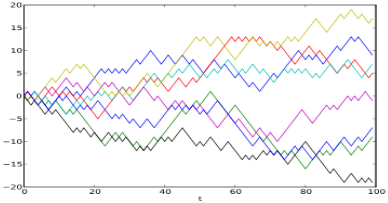
$$
Q_{nm}(t) \equiv p(n, t+1|m, t)
$$

Such pdfs are called *transition probabilities* since they give the probability of the system making a transition from state m to state n

If we also write $p(n, t)$ as $P_n(t)$, then the first equation for Markov processes can be written as

$$
P_n(t+1) = \sum_m Q_{nm}(t) P_m(t)
$$

Example (1) – *A one-dimensional random walk*



Basic idea:- A walker starts out at a given point, n_0, at time $t = 0$, and moves to the right with a given probability, p, and to the left with a probability $q = 1 - p$. Then at $t = 1$ the walker again moves to the right or left with the same probabilities. What is the probability that the walker is at position n at time t?

This is simplest formulation, but there are many variants: the walker can also stay put instead of having to move to the left or right, the walker can take steps of any length, or the rules by which he moves can be time dependent, etc

- Physicists usually state that random walks were first introduced by Einstein in his discussion of Brownian motion in 1905, but Bachelier used the idea to model stock market fluctuations in 1900. The problem of a random walk was discussed more generally by Pearson in 1905
- There are many versions of the random walk problem. The simplest is the "nearest neighbour random walk", alluded to above, which is formally defined as follows:
  If the walker is at position n', he moves one step to the right (to $n = n' + 1$) with probability p and one step to the left (to $n = n' - 1$) with probability q, where $p + q = 1$. This means that

$$Q_{n\,n'} = \begin{cases} p, & \text{if } n = n' + 1 \\ q, & \text{if } n = n' - 1 \\ 0, & \text{otherwise}, \end{cases}$$

  So in this case Q is time-independent and non-zero only if $n = n' \pm 1$

Notice also that the probability of making a transition only depends on the current state of the system, not on the previous states: the probability of moving to the right does not depend on whether a move to the right was made at a previous time step, or where the walker came from before arriving at his current position

If any of the above were true then the process would not be Markov, and the techniques we are using would not be applicable

If we write $Q_{nn'}$ as a matrix, then

$$Q = \begin{pmatrix} 0 & q & 0 & 0 & \ldots \\ p & 0 & q & 0 & \ldots \\ 0 & p & 0 & q & \ldots \\ .. & .. & .. & .. & \ldots \end{pmatrix} ,$$

where we have not specified what happens at the boundaries (see later)

- Another simple version of the nearest neighbour model has the possibility of staying put with a probability r

$$
Q_{n\,n'} = \begin{cases} p, & \text{if } n = n' + 1 \\ r, & \text{if } n = n' \\ q, & \text{if } n = n' - 1 \\ 0, & \text{otherwise}\,, \end{cases}
$$

where $p + q + r = 1$

In this case,

$$
Q = \begin{pmatrix} r & q & 0 & 0 & \dots \\ p & r & q & 0 & \dots \\ 0 & p & r & q & \dots \\ .. & .. & .. & .. & \dots \end{pmatrix} ,
$$

— a tridiagonal matrix

- We are interested in predicting $P_n(t)$ given $P_n(0)$. Usually this initial condition will be that the walker starts (with certainty) at a given position, say $n = n_0$, at an initial time, say $t = 0$

That is,

$$P_n(0) = \delta_{n,n_0} \,,$$

i.e. with probability 1 the walker is at position n_0 at time $t = 0$. Note that $\sum_n P_n(0) = \sum_n \delta_{n,n_0} = 1$, as required
We can think of $P_n(t)$ as the elements of a column vector:

$$P(t) = \left(\begin{array}{c} P_1(t) \\ P_2(t) \\ \dots \end{array} \right)$$

- A general random walk would have $Q_{nn'}$ as a general matrix — any position could be reached from any other position in one jump, if the relevant entry in the matrix Q was non-zero. The entry could also depend on n or n' or both. For example, the probability of making a transition could depend on $|n' - n|$: this could model the situation that the further the points are from each other, the less likely is the walker to jump to n from n'
- Frequently in applications $Q_{nn'}(t)$ does not actually depend on time — the "rules of the game" are fixed and time-independent