Stochastic Processes for Physicists:

Understanding Noisy Systems

Kurt Jacobs

物理学家用的随机过程

CAMBRIDGE

世界图书出版公司
www.wpcbj.com.cn

STOCHASTIC PROCESSES FOR PHYSICISTS

Understanding Noisy Systems

KURT JACOBS
University of Massachusetts

CAMBRIDGE
UNIVERSITY PRESS

图书在版编目（CIP）数据

物理学家用的随机过程 = Stochastic Processes for Physicists : Understanding Noisy Systems : 英文 /（美）K. 雅各布斯（Kurt Jacobs）著 . —影印本 . —北京 : 世界图书出版有限公司北京分公司，2018.5
ISBN 978-7-5192-4466-8

Ⅰ . ①物… Ⅱ . ① K… Ⅲ . ①物理学—随机过程—英文 Ⅳ . ① O4

中国版本图书馆 CIP 数据核字（2018）第 043647 号

中文书名： 物理学家用的随机过程
英文书名： Stochastic Processes for Physicists: Understanding Noisy Systems
著　　者： Kurt Jacobs

责任编辑： 刘　慧　高　蓉
装帧设计： 蔡　彬
出版发行： 世界图书出版有限公司北京分公司
地　　址： 北京市东城区朝内大街 137 号
邮　　编： 100010
电　　话： 010-64038355（发行）　64015580（客服）　64033507（总编室）
网　　址： http://www.wpcbj.com.cn
邮　　箱： wpcbjst@vip.163.com
销　　售： 新华书店
印　　刷： 三河市国英印务有限公司
开　　本： 787mm × 1092mm　1/16
印　　张： 13
字　　数： 250 千字
版　　次： 2018 年 5 月第 1 版
印　　次： 2018 年 5 月第 1 次印刷
版权登记： 01-2016-9936
国际书号： ISBN 978-7-5192-4466-8
定　　价： 48.00 元

STOCHASTIC PROCESSES FOR PHYSICISTS

Understanding Noisy Systems

Stochastic processes are an essential part of numerous branches of physics, as well as biology, chemistry, and finance. This textbook provides a solid understanding of stochastic processes and stochastic calculus in physics, without the need for measure theory.

In avoiding measure theory, this textbook gives readers the tools necessary to use stochastic methods in research with a minimum of mathematical background. Coverage of the more exotic Levy processes is included, as is a concise account of numerical methods for simulating stochastic systems driven by Gaussian noise. The book concludes with a non-technical introduction to the concepts and jargon of measure-theoretic probability theory.

With over 70 exercises, this textbook is an easily accessible introduction to stochastic processes and their applications, as well as methods for numerical simulation, for graduate students and researchers in physics.

KURT JACOBS is an Assistant Professor in the Physics Department at the University of Massachusetts, Boston. He is a leading expert in the theory of quantum feedback control and the measurement and control of quantum nano-electro-mechanical systems.

Preface

This book is intended for a one-semester graduate course on stochastic methods. It is specifically targeted at students and researchers who wish to understand and apply stochastic methods to problems in the natural sciences, and to do so without learning the technical details of measure theory. For those who want to familiarize themselves with the concepts and jargon of the "modern" measure-theoretic formulation of probability theory, these are described in the final chapter. The purpose of this final chapter is to provide the interested reader with the jargon necessary to read research articles that use the modern formalism. This can be useful even if one does not require this formalism in one's own research.

This book contains more material than I cover in my current graduate class on the subject at UMass Boston. One can select from the text various optional paths depending on the purpose of the class. For a graduate class for physics students who will be using stochastic methods in their research work, whether in physics or interdisciplinary applications, I would suggest the following: Chapters 1, 2, 3 (with Section 3.8.5 optional), 4 (with Section 4.2 optional, as alternative methods are given in 7.7), 5 (with Section 5.2 optional), 7 (with Sections 7.8 and 7.9 optional), and 8 (with Section 8.9 optional). In the above outline I have left out Chapters 6, 9 and 10. Chapter 6 covers numerical methods for solving equations with Gaussian noise, and is the sort of thing that can be picked-up at a later point by a student if needed for research. Chapter 9 covers Levy stochastic processes, including exotic noise processes that generate probability densities with infinite variance. While this chapter is no more difficult than the preceding chapters, it is a more specialized subject in the sense that relatively few students are likely to need it in their research work. Chapter 10, as mentioned above, covers the concepts and jargon of the rigorous measure-theoretic formulation of probability theory.

A brief overview of this book is as follows: Chapters 1 (probability theory without measure theory) and 2 (ordinary differential equations) give background material that is essential for understanding the rest of course. Chapter 2 will be almost

all revision for students with an undergraduate physics degree. Chapter 3 covers all the basics of Ito calculus and solving stochastic differential equations. Chapter 4 introduces some further concepts such as auto-correlation functions, power spectra and white noise. Chapter 5 contains two applications (Brownian motion and option pricing), as well as a discussion of the Stratonovich formulation of stochastic equations and its role in modeling multiplicative noise. Chapter 6 covers numerical methods for solving stochastic equations. Chapter 7 covers Fokker–Planck equations. This chapter also includes applications to reaction–diffusion systems, and pattern formation in these systems. Chapter 8 explains jump processes and how they are described using master equations. It also contains applications to population dynamics and neuron behavior. Chapter 9 covers Levy processes. These include noise processes that generate probability densities with infinite variance, such as the Cauchy distribution. Finally Chapter 10 introduces the concepts and jargon of the "modern" measure-theoretic description of probability theory.

While I have corrected many errors that found their way into the manuscript, it is unlikely that I eliminated them all. For the purposes of future editions I would certainly be grateful if you can let me know of any errors you find.

Acknowledgments

Thanks to ...

Aric Hagberg for providing me with a beautiful plot of labyrinthine pattern formation in reaction–diffusion systems (see Chapter 7) and Joshua Combes and Jason Ralph for bringing to my attention Edwin Jaynes' discussion of mathematical style (see Chapter 10). I am also grateful to my students for being patient and helping me iron-out errors and omissions in the text.

I am very grateful to Aaron Dörr, Andy Chia, Anita Dąbrowska, Chuck Yeung, Chris Chudzicki, David Pearson, David Sussillo, Jon Armond, Tom Chou, and Pu Ke, as well as the students in my graduate class, for sending me the errors they found in the first and second printings.

Contents

1

A review of probability theory

In this book we will study dynamical systems driven by noise. Noise is something that changes randomly with time, and quantities that do this are called *stochastic processes*. When a dynamical system is driven by a stochastic process, its motion too has a random component, and the variables that describe it are therefore also stochastic processes. To describe noisy systems requires combining differential equations with probability theory. We begin, therefore, by reviewing what we will need to know about probability.

1.1 Random variables and mutually exclusive events

Probability theory is used to describe a situation in which we do not know the precise value of a variable, but may have an idea of the relative likelihood that it will have one of a number of possible values. Let us call the unknown quantity X. This quantity is referred to as a *random variable*. If X is the value that we will get when we roll a six-sided die, then the possible values of X are $1, 2, \ldots, 6$. We describe the likelihood that X will have one of these values, say 3, by a number between 0 and 1, called the *probability*. If the probability that $X = 3$ is unity, then this means we will *always* get 3 when we roll the die. If this probability is zero, then we will never get the value 3. If the probability is $2/3$ that the die comes up 3, then it means that we expect to get the number 3 about two thirds of the time, if we roll the die many times.

The various values of X, and of any random variable, are an example of *mutually exclusive* events. That is, whenever we throw the die, X can have only one of the values between 1 and 6, no more and no less. Rather obviously, if the probability for X to be 3 is $1/8$, and for X to be 6 is $2/8$, then the probability for X to be *either* 3 *or* 6 is $1/8 + 2/8 = 3/8$. That is, the total probability that one of two or more mutually exclusive events occurs is the *sum* of the probabilities for each event. One usually states this by saying that "mutually exclusive probabilities sum". Thus, if

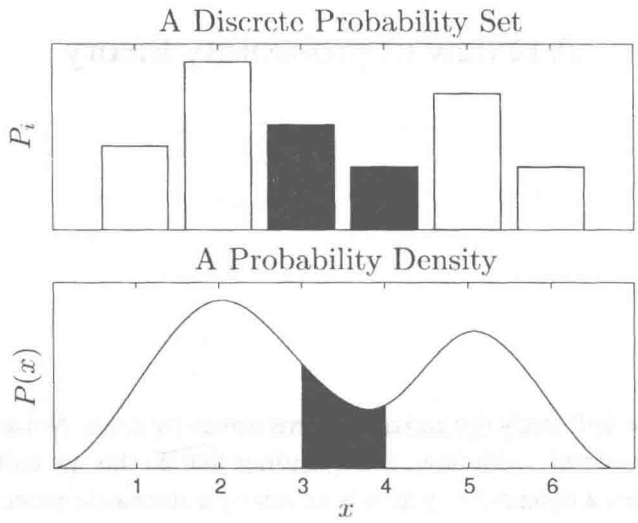

Figure 1.1. An illustration of summing the probabilities of mutually exclusive events, both for discrete and continuous random variables.

we want to know the probability for X to be in the range from 3 to 4, we sum all the probabilities for the values from 3 to 4. This is illustrated in Figure 1.1. Since X *always* takes a value between 1 and 6, the probability for it to take a value in this range must be unity. Thus, the sum of the probabilities for all the mutually exclusive possible values must always be unity. If the die is *fair*, then all the possible values are equally likely, and each is therefore equal to $1/6$.

Note: in mathematics texts it is customary to denote the unknown quantity using a capital letter, say X, and a variable that specifies one of the possible values that X may have as the equivalent lower-case letter, x. We will use this convention in this chapter, but in the following chapters we will use a lower-case letter for both the unknown quantity and the values it can take, since it causes no confusion.

In the above example, X is a *discrete random variable*, since it takes the discrete set of values $1, \ldots, 6$. If instead the value of X can be any real number, then we say that X is a *continuous* random variable. Once again we assign a number to each of these values to describe their relative likelihoods. This number is now a function of x (where x ranges over the values that X can take), called the *probability density*, and is usually denoted by $P_X(x)$ (or just $P(x)$). The probability for X to be in the range from $x = a$ to $x = b$ is now the area under $P(x)$ from $x = a$ to $x = b$. That is

$$\text{Prob}(a < X < b) = \int_a^b P(x)dx. \tag{1.1}$$

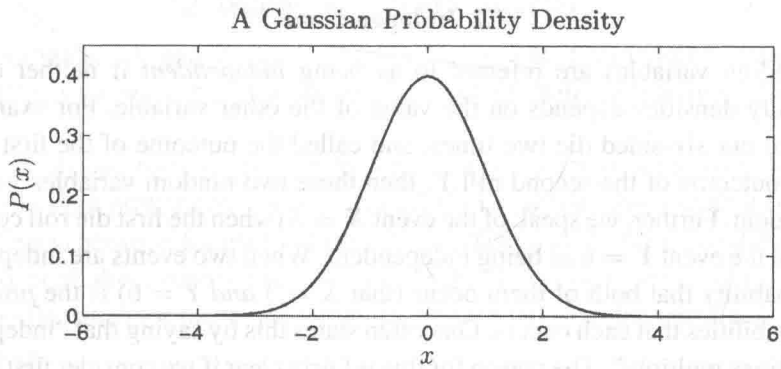

Figure 1.2. A Gaussian probability density with variance $V = 1$, and mean $\langle X \rangle = 0$.

This is illustrated in Figure 1.1. Thus the integral of $P(x)$ over the whole real line (from $-\infty$ to ∞) must be 1, since X must take one of these values:

$$\int_{-\infty}^{\infty} P(x)dx = 1. \tag{1.2}$$

The average of X, also known as the *mean*, or *expectation value*, of X is defined by

$$\langle X \rangle \equiv \int_{-\infty}^{\infty} P(x)x\,dx. \tag{1.3}$$

If $P(x)$ is symmetric about $x = 0$, then it is not difficult to see that the mean of X is zero, which is also the center of the density. If the density is symmetric about any other point, say $x = a$, then the mean is also a. This is clear if one considers a density that is symmetric about $x = 0$, and then shifts it along the x-axis so that it is symmetric about $x = a$: shifting the density shifts the mean by the same amount.

The *variance* of X is defined as

$$V_X \equiv \int_{-\infty}^{\infty} P(x)(x - \langle X \rangle)^2\,dx = \int_{-\infty}^{\infty} P(x)x^2\,dx - \langle X \rangle^2 = \langle X^2 \rangle - \langle X \rangle^2. \tag{1.4}$$

The *standard deviation* of X, denoted by σ_X and defined as the square root of the variance, is a measure of how broad the probability density for X is – that is, how much we can expect X to deviate from its mean value.

An important example of a probability density is the Gaussian, given by

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{1.5}$$

The mean of this Gaussian probability density is $\langle X \rangle = \mu$ and the variance is $V(x) = \sigma^2$. A plot of this probability density in given in Figure 1.2.

1.2 Independence

Two random variables are referred to as being *independent* if neither of their probability densities depends on the value of the other variable. For example, if we rolled our six-sided die two times, and called the outcome of the first roll X, and the outcome of the second roll Y, then these two random variables would be independent. Further, we speak of the event $X = 3$ (when the first die roll comes up as 3) and the event $Y = 6$ as being independent. When two events are independent, the probability that both of them occur (that $X = 3$ *and* $Y = 6$) is the *product* of the probabilities that each occurs. One often states this by saying that "independent probabilities multiply". The reason for this is fairly clear if we consider first making the die roll to obtain X. Only if $X = 3$ do we then make the second roll, and only if that comes up 6 do we get the result $X = 3$ and $Y = 6$. If the first roll only comes up 3 one eighth of the time, and the second comes up 6 one sixth of the time, then we will only get both of them $1/8 \times 1/6 = 1/48$ of the time.

Once again this is just as true for independent random variables that take a continuum of values. In this case we speak of the "joint probability density", $P(x, y)$, that X is equal to x and Y is equal to y. This joint probability density is the product of the probability densities for each of the two independent random variables, and we write this as $P(x, y) = P_X(x)P_Y(y)$. The probability that X falls within the interval $[a, b]$ *and* Y falls in the interval $[c, d]$ is then

$$\begin{aligned}\text{Prob}(X \in [a,b] \text{ and } Y \in [c,d]) &= \int_a^b \int_c^d P(x,y)dydx \\ = \int_a^b \int_c^d P_X(x)P_Y(y)dydx &= \left(\int_a^b P_X(x)dx\right)\left(\int_c^d P_Y(y)dy\right) \\ &= \text{Prob}(X \in [a,b]) \times \text{Prob}(Y \in [c,d]).\end{aligned}$$

In general, if we have a joint probability density, $P(x_1, \ldots, x_N)$, for the N variables $X_1, \ldots, X_N$, then the expectation value of a function of the variables, $f(X_1, \ldots, X_N)$, is given by integrating the joint probability density over all the variables:

$$\langle f(X_1, \ldots, X_N)\rangle = \int_{-\infty}^{\infty} f(x_1, \ldots, x_N)P(x_1, \ldots, x_N)\,dx_1 \ldots dx_N. \quad (1.6)$$

It is also worth noting that when two variables are independent, then the expectation value of their product is simply the product of their individual expectation values. That is

$$\langle XY\rangle = \langle X\rangle\langle Y\rangle. \quad (1.7)$$

1.3 Dependent random variables

Random variables, X and Y, are said to be *dependent* if their joint probability density, $P(x, y)$, does not factor into the product of their respective probability densities.

To obtain the probability density for one of the variables alone (say X), we integrate the joint probability density over all values of the other variable (in this case Y). This is because, for each value of X, we want to know the total probability summed over all the mutually exclusive values that Y can take. In this context, the probability densities for the single variables are referred to as the *marginals* of the joint density.

If we know nothing about the value of Y, then our probability density for X is just the marginal

$$P_X(x) = \int_{-\infty}^{\infty} P(x, y)dy. \tag{1.8}$$

If X and Y are dependent, and we learn the value of Y, then in general this will change our probability density for X (and vice versa). The probability density for X *given* that we know that $Y = y$, is written $P(x|y)$, and is referred to as the *conditional* probability density for X *given* Y.

To see how to calculate this conditional probability, we note first that $P(x, y)$ with $y = a$ gives the *relative* probability for different values of x given that $Y = a$. To obtain the conditional probability density for X given that $Y = a$, all we have to do is divide $P(x, a)$ by its integral over all values of x. This ensures that the integral of the conditional probability is 1. Since this is true for any value of y, we have

$$P(x|y) = \frac{P(x, y)}{\int_{-\infty}^{\infty} P(x, y)dx}. \tag{1.9}$$

Note also that since

$$P_Y(y) = \int_{-\infty}^{\infty} P(x, y)dx, \tag{1.10}$$

if we substitute this into the equation for the conditional probability above (Eq. (1.9)) we have

$$P(x|y) = \frac{P(x, y)}{P_Y(y)}, \tag{1.11}$$

and further that $P(x, y) = P(x|y)P_Y(y)$.

As an example of a conditional probability density consider a joint probability density for X and Y, where the probability density for Y is a Gaussian with zero