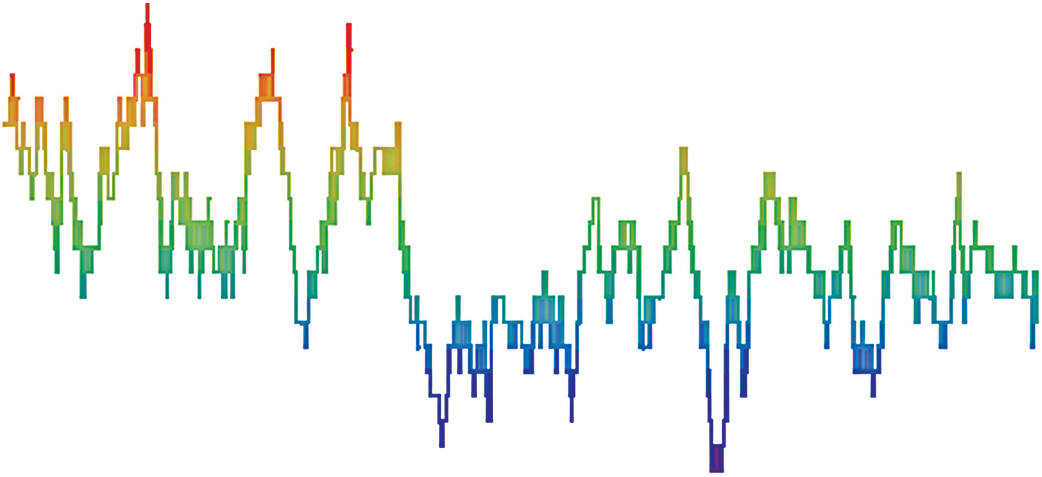
Graduate Student Series in Physics



The Dynamics of Discrete Populations and Series of Events

Keith Iain Hopcraft
Eric Jakeman
Kevin D. Ridley

CRC Press
Taylor & Francis Group

The Dynamics of Discrete Populations and Series of Events

Graduate Student Series in Physics

Series Editor:
Professor Derek Raine
Senior Lecturer, Department of Physics and Astronomy, University of Leicester

The Dynamics of Discrete Populations and Series of Events
Keith Iain Hopcraft, Eric Jakeman, and Kevin D. Ridley

Introduction to Surface and Superlattice Excitations, Second Edition
Michael G. Cottam and D.R. Tilley

Cosmology in Gauge Field Theory and String Theory
D. Bailin and Alexander Love

Introduction to Nuclear Reactions
C.A. Bertulani and P. Danielewicz

Gauge Theories in Particle Physics, Volume II: QCD and the Electroweak Theory, Third Edition
I.J.R. Aitchison and A.J.G. Hey

Gauge Theories in Particle Physics: Volume I: From Relativistic Quantum Mechanics to QED, Third Edition
I.J.R. Aitchison and A.J.G. Hey

Problems on Statistical Mechanics
D.A.R Dalvit, J. Frastai, and Ian Lawrie

Symmetries in Quantum Mechanics: From Angular Momentum to Supersymmetry (PBK)
M. Chaichian and R. Hagedorn

Supersymmetric Gauge Field Theory and String Theory
D. Bailin and Alexander Love

Introduction to Gauge Field Theory Revised Edition
D. Bailin and Alexander Love

Superfluidity and Superconductivity
D.R. Tilley and J. Tilley

The Dynamics of Discrete Populations and Series of Events

Keith Iain Hopcraft
University of Nottingham, UK
Eric Jakeman
University of Nottingham, UK
Kevin D. Ridley
Consultant Optical Physicist, UK

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Taylor & Francis
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20131203

International Standard Book Number-13: 978-1-4822-2765-9 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

This book is largely concerned with mathematical models that can be used to describe time-varying populations and series of events. These models are simple and generic, and the emphasis is on their general properties rather than their applicability to any particular real-world systems, although examples of the latter can be found in some of the further reading material listed at the end of each chapter. Although the models are relatively simple and analytically tractable, they provide a basic understanding of the effect of the competing processes governing the changing number of individuals present in a population and therefore a guide to the development of more complex models that can be accessible only through numerical modelling.

The book was originally stimulated by the recent development of models that can be used to characterise the evolution of populations with fluctuations governed by long-tailed probability distributions such as are commonly observed in so-called complex systems. It was clear to the authors that a transparent description of such models would be greatly facilitated by tracing the way that they had emerged from the more familiar population processes described in the existing literature. The main part of this book is therefore devoted to the logical development of the theory of first-order discrete Markov population processes starting from a few basic assumptions. Note the term *discrete*, here; it is also possible to characterise populations by approximating them with continuous variables, and this is a method that has been widely used. However, this book is concerned solely with the discrete approach. More information about these two contrasting methods can be found in Chapter 1.

The book also covers work inspired by the authors' background in quantum optics. This led them to extend the more familiar models to situations where the populations have novel properties, including sub-Poisson statistics and odd–even effects that have no continuum counterpart. Perhaps more importantly, it prompted them to investigate processes by which the populations may be monitored and to calculate how the measured and intrinsic statistics differ. This leads naturally to the generation and characterisation of time series that record, for example, when individuals leave a population.

A substantial review of the history of the subject and of the contents of each chapter is given in the introductory Chapter 1, but the book does not aim to give a comprehensive coverage of what is a very large area of science with a long and distinguished history and with a wide range of applications. Rather, it is intended to provide some basic mathematical tools, physical insight, and some novel ideas that will assist students to develop their own stochastic modelling capabilities. To aid this, Chapter 11 describes methods for numerical simulation of the models developed in the rest of the book, with examples given in the mathematical programming language *Mathematica*®. These examples are, however, kept sufficiently simple that even those with no in-depth knowledge of *Mathematica* should be able to re-write them in a programming language of their choice.

The authors are particularly indebted to Terry Shepherd, and to Eric Renshaw who has authored two comprehensive volumes of related work. They are also indebted, of course, to a number of research students for their work in this area, including S. Phayre, J.O. Matthews, O.E. French, J.M. Smith, and W.H. Lee.

1

Introduction

This book has a dual purpose. The first is to provide an introductory text that collects various discrete stochastic population models. This will furnish the reader with a systematic methodology for the formulation of the models, and it will explore their dynamical properties and the ways of characterising their behaviours in terms of customary measurements that can be made upon them. The second purpose is to then use these models as a tool for generating a series of events, or *point processes,* that have distinct properties according to which population model is used as a motor. The two qualifiers, *discrete* and *stochastic*, simultaneously provide the subject with a richness of phenomenology and technical challenges for formulating, describing, and extracting those behaviours. This is because the population can only change by an integer amount, and those changes are triggered to occur at times not governed by the invariable ticks of a clock. These strictures are absent when adopting a more straightforward *continuous* and *deterministic* approach, whereby the population is described as a continuous density and time marches uniformly onward at a regular pace. This distinction between the two approaches also delineates between the source, character and strength of fluctuations. In the stochastic formulation the fluctuations are an *intrinsic* property of the population itself. The mechanisms causing the changes in population size are essentially non-perturbative in nature because they change the state of the population by values of finite size, and this is true even if the mathematical formulation of a particular mechanism is linear. In the deterministic approach, intrinsic fluctuations can only arise through non-linearity. Both approaches can be affected by the presence of extrinsic noise, but again this is essentially non-perturbative in the stochastic formulation.

The dynamics of populations has a long and varied history, both in terms of the subjects' development per se and for the stimulus it has provided to other seemingly unconnected areas. Indeed, the subject has a pervasiveness that can cross between scientific boundaries and even beyond the ambit of the physical sciences by virtue of its utility. It was the cleric and scholar Robert Malthus who, in his *Essay on the Principle of Population* in 1798 [1], considered the brakes to unfettered linear exponential population growth. His conclusion that 'the power of population is indefinitely greater than the power in the earth to produce subsistence in man' resonates down the ages in spirit if not through the predictive accuracy of his analysis. This essay stimulated, in 1801, the establishment of the first decennial national census in the United Kingdom in order to ascertain how much corn was required to feed the nation (and, serendipitously, to quantify the number of able-bodied men available to fight in the Napoleonic wars). The essay provided one of many stimuli to Charles Darwin's development of the theory of natural selection. Darwin recognised that the effects of competition for resources by species as described by Malthus provided a mechanism for their diversification, whereby 'favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The results of this would be the formation of a new species. Here, then I had at last got a theory by which to work' [2].

Pierre-Francois Vehulst sought to model and thereby quantify the limitations to exponential growth to which Malthus alluded [3]. This introduced to the equation describing the continuous and deterministic evolution of population size the notion of a *carrying capacity* that embodies an ecosystem's ability to sustain such a population. The key ingredient of this 'logistic' equation is its non-linearity, which predicted that the population's size saturated eventually and established the timescale on which this occurred. It is not so much this result but a technical reinterpretation of the logistic equation itself, over a century after Vehulst, which is of interest and profound consequence. This reinterpretation of the governing equation recognised that there is a season in which an animal or plant species breeds or grows. To encapsulate this observation, time was no longer treated continuously but as a discrete variable, although still deterministically. In this way the logistic differential equation is transformed to a difference equation or iterative mapping. The mapping contains a parameter, related to the carrying capacity, whose value leads to very different classes of behaviour. These range from solutions similar to those obtained from the continuous equation, to periodic bi-stability where the population oscillates between two values, period-doubling and aperiodic behaviour sensitive to the initial conditions. The logistic mapping therefore provides a paradigm for deterministic chaos, as described in the review article by Robert May [4].

Descriptions of the events that are associated with various population models are also ubiquitous because of their usefulness. The model with which all others are compared is the Poisson process. The distribution for the number of these birth events occurring in an interval of time was introduced by Siméon-Denis Poisson but with totally different context and application, namely to the deliberations of juries in criminal and civil legal trials [5]. The Poisson process' reach is very wide because of its elemental and simple nature. The Poisson distribution is an example belonging to the discrete stable class of distributions. This means that the sum of Poisson distributed random variables is itself a Poisson distributed random variable. This makes the Poisson the discrete analogue of the continuous Gaussian random variable, which also possesses the stability property. The widening of the discrete stable class through the agency of population models, and the events associated with them, forms the subject matter of Chapter 7.

The intended readership of this book is the advanced undergraduate or postgraduate student, but we also have in mind the experienced researcher who may be changing fields, or who needs to know about a particular class of population model and to access key information for its characterisation. The mathematical machinery required to understand the development will have been encountered by most chemistry, mathematics, physics and theoretical biology students. Moreover, we have adhered to pursuing the development using a limited bag of tools, even in those instances when application of another method can produce a result with greater economy of effort. If a particular model admits an analytical solution, then that solution can usually be obtained by more than one means. The text is concerned with exploring the stochastic formulation of population models and the qualitative differences between them rather than the methods available for obtaining their solution. The following chapters include a brief summary, some problems that explore both the basics and some more interesting aspects of the development, and a 'Further Reading' list. A manual with solutions to the problems is also available. The bibliography is not meant to be exhaustive but rather a first port of call for the reader to explore applications, subtleties, and further techniques.

Chapter 2 provides a primer for the probabilistic and statistical tools that will be used throughout. It introduces the discrete distributions that will feature in subsequent chapters and also their alternative representation in terms of a generating function. The technique of using generating functions to represent the population dynamics may, at first encounter, appear as an unwelcome abstraction and diversion. But their use enables the equations to be solved in a systematic fashion, and once obtained, the probabilities and moments that correspond to observable quantities can be derived from the generating function by elementary means.

Chapter 3 commences by establishing the Markov property, which is a further preliminary probabilistic foundation that will inform the development of most of the population models we go on to describe. This property pertains to how the future evolution of a system is affected by its immediate or more distant past. The chapter continues by exploring the three important elements of births, deaths and immigration as a cause for population change, and shows, with the assistance of the Markov property, how these separate processes are represented mathematically in the equation describing the stochastic evolution. Rather than solve for the complete dynamics, the evolution of measurables, such as the mean size and the correlation in size between one instant and the next, is obtained. The reason for this is that these microscopic causes for change in a populations' size are easy to intuit in terms of their averaged manifestation; births and deaths occurring in a population cause it to either grow or diminish in size unless the birth and death rates are identical, and so this contains no particular surprises. The stochastic treatment of the birth process was originally treated by Yule [6], in the context of mathematical biology, and independently by Furry [7] with regard to transmutation of elements by radioactive decay. The combination of deaths with immigration leads to the important concept of a stationary or equilibrium state for the population. Indeed, this equilibrium forms something of a benchmark against which all others can be compared, for it is the ubiquitous Poisson distribution. The simple mechanisms of birth, death and immigration provide a useful ground in which to explore the effect of measurement. The idea of making a measurement on a fluctuating quantity is not as straightforward as one might expect, or indeed hope. Any measurement takes a finite time to perform, during which the population continues to change. If that change is small compared with the population's size at the commencement of the measurement, then it will approximate the instantaneous state of the population itself. If the converse is true, the measurement *process* will sample and aggregate the various intermediate states through which the population evolves during the measurement time. The type of measurement made can also affect the population itself. Counting the number of individuals leaving (say) an airport is fundamentally different from counting the number of photons leaving a cavity. The detection of a photon is synonymous with its destruction and it can no longer participate in the dynamics. The measurement process also furnishes a more abstract but utilitarian function for turning the fluctuations in population *size* into a series of events in *time* that mark when those changes occur. It is this connection between the primary process by which the population evolves in time, and the secondary process by which it is monitored, that forms an important and recurrent theme throughout the text.

Chapter 4 considers the simultaneous combination of births, deaths, and immigration into a process, and this leads to a broader class of equilibrium solution provided that the death rate exceeds the rate at which births occur. This problem has surfaced in numerous branches of science, with the three mechanisms for change being co-opted to represent genetic mutations, through the production of photons in a cavity to the fluctuating state of the sea's surface. The full machinery of the generating function method of solution is deployed here for the first time. This serves to illustrate, in a natural way, how the relative sizes of the rates affect the structure of the equilibrium solution, how these combine to form the timescale that governs the approach to this steady state and the intrinsic fluctuations in it. The monitoring of this process is explored in more detail with two scenarios being treated. The first of these is where the counted 'individuals' are removed and no longer participate in the evolution; the second corresponds to when they are replaced and so continue to contribute to the evolution. Although one might consider that such distinctions would lead to minor differences, the discrepancies are nevertheless significant and the underlying physical interpretation for their occurrence is explored. The equilibrium solution of the birth–death–immigration process is characterised by fluctuations whose relative sizes are greater than the benchmark Poisson process. The series of events that are generated by this process are characterised by their occurring in *clusters* or *bunches* and therefore have an intermittent quality.

There are processes whose fluctuations are of lesser size than those associated with the Poisson, and events whose occurrences show a greater regularity than the purely random, manifesting a propensity for *anti-bunching*. Such populations can be generated using the mechanisms already introduced but with the important modification that the population size is capped but dynamically coupled with the extant population size. A limiting mechanism of this kind is redolent of the carrying capacity of an environment, but crucially it has no continuum analogue and is therefore quite distinct from the logistic model. A *lower limit* can also be applied so that the population cannot fall below some base value, and even when the cap is removed, this process can exhibit sub-Poisson effects. Chapter 5 provides the details of these models that exhibit sub-Poisson traits whilst retaining similar mathematical properties to those processes already encountered. An altogether different process that also has no continuum equivalent arises from relaxing the condition that the population can only increase though singletons. If immigrants enter the population in pairs but die singly, then the fluctuations display different properties according to whether there are an even or odd number present. Moreover, the monitoring without replacement of such a population serves to amplify the odd–even parity effects. Although this process is of interest to a rather arcane amplifier effect in quantum optics, its formulation prompts an important generalisation of the death–immigration process to one possessing great flexibility for modelling arbitrary populations with a prescribed steady state.

Allowing the immigrants to enter not just singly or in pairs, but in addition as triplets, …, *r*-tuplets, with rates of immigration particular to *r*, results in a mathematical structure with a simple and appealing interpretation. The population may be thought as being coupled to a separate equilibrium population of potential immigrants; it provides a coupling to an environment. Crucially, the formulation, which is described in Chapter 6, admits an inverse problem to be performed, whereby a population with a desired equilibrium state can be constructed through tailoring the rates at which the multiple immigrants are introduced.

The utility of the death–multiple immigration model is demonstrated in Chapter 7 to generate the discrete analogue of the stable process. Continuous stable random variables were treated by Paul Lévy in 1937 [8] as a generalisation of the Gaussian random variable, which is ubiquitous in pure and applied science. Nevertheless, the generalisation resulted in distributions whose variance is infinite and for which there is no characteristic associated scale size. Consequently, the generalisation was regarded as mathematical pathology and of little practical importance. In 1963, Beniot Mandelbrot showed that the fluctuations in cotton prices were described by such distributions [9]. So too were the fluctuations in diverse systems whose correlation scale size become commensurate with that of the system, irrespective of the details of the dynamics governing the system itself. Such behaviours are observed in systems close to some change in its phase or state. Throughout the late 1980s onward, various manifestations of the ‘sandpile’ paradigm were developed with the aim of demonstrating manifestations of these phenomena without recourse to fine-tuning a systems' parameters in order to adopt particular outcomes, in other words to achieve *emergent* behaviours from a parsimonious and generic dynamics [10]. This was achieved using computational models, and these possessed the wild fluctuations and scale-free properties of the non-Gaussian stable distributions. Since these were computational results, they are necessarily discrete in nature. A discrete stable *process* can be generated using the death–multiple immigration population model. The particular problems associated with monitoring such a process are acute, for the detecting 'instrument' will have a finite range or bandwidth, whereas the population has fluctuations of arbitrary size. Saturation effects in sampling, which are therefore inevitable, must be accounted for. When monitored, the series of events so formed present fractal properties irrespective of the properties of the detector—the scale-free attributes being transferred from a cascade in increasing population size to an inverse cascade in the time between events.

The coupling of populations in the stochastic framework presents technical challenges that are absent in the continuous and deterministic picture. It is an important area with much scope for future work in diverse application areas. A celebrated continuous formulation was pioneered independently by Alfred Lotka and Vito Volterra and has become known as the *predator–prey equations* [11], through which the dynamics of an idealised ecosystem comprising the hunters and hunted can be explored. The developments of this paradigm to treat a more diverse, realistic, and necessarily more complex set of ecosystems is described by Robert May [12]. From the stochastic perspective, the incorporation of immigrations into a population, be they single or multiple, is essentially modelling the situation of the population being affected and adjusting to accommodate the influence of an external environment. This occurs without the population reacting back onto those stimuli. Chapter 8 examines how the multiple immigration picture can be adapted so that a population and its environment become fully coupled with one another, evolving self-consistently according to their internal and mutual dynamics. This approach admits treating increasingly complex situations, where the environment can be regarded as a heterogeneous ensemble of different populations that connected in a chain. But the interactions between populations can occur in a greater variety of ways. The situation where members of one population are exchanged and rebadged as members of another provides another mechanism that displays distinct characteristics.

Rather than a series of events being driven through the agency of a discrete population process, it is possible for the events to be caused by some underlying fluctuation in a continuous quantity. Useful ideas frequently emerge in unconnected application areas and at broadly similar times. David Cox formulated this idea and illustrated it with application to the periods when a weaving loom was operational, dormant or had broken down [13]. By contrast, Leonard Mandel independently arrived at the same formulation with regard to the intervals between detection of photons using photoelectric multiplier tubes [14]. Such a representation has become known as *doubly stochastic* or *compound* because the fluctuations in the continuous variable stimulate an associated discrete process, the two forming a transform pair. This mechanism is explored in Chapter 9 and furnishes a method of great power and flexibility for generating populations. Indeed, the method can be demonstrated using the discrete stable process detailed in Chapter 7 to deduce the continuous process that produces it. Not surprisingly (perhaps) this is a continuous stable *process*, and the underlying Markov assumption enables the entire joint statistical properties of stable distributed variables to be deduced within the doubly stochastic context.

Chapter 10 discusses another mechanism for generating a series of events defined by an underlying continuous random process exceeding a predefined threshold: the *level crossing problem*. This problem has a huge literature associated with it [15], which has continued to expand. The reason for the large volume and diversity of work is principally because of the problem's close association with diverse applications in communications theory and technology, instrumentation and signal processing, and its relevance to the size, frequency and duration of extremal events. The principal difference between this problem and all the others treated in the text stems from its non-Markovian nature; the future evolution of the continuous process is affected not just by its immediate past but also further back into its history. This complexity through the mechanism of 'memory' is inherited by the level crossings and properties associated with them. The mathematical manifestation of this occurs through the spectrum of the continuous process having a non-Lorenztian form, which leads to profound technical challenges but also considerable richness in phenomenology. The statistical properties of the events display all the attributes that particular models describe: bunching, anti-bunching, fractal and sub-fractal.

Despite the emphasis placed on analytical methods, computational simulation of populations and the properties derived from them provides an important complementary approach to understanding, model development and the treatment of real-world situations. Often it is the only way to make progress. Chapter 11 describes the techniques for the population models previously described, with examples of *Mathematica*® code for their implementation. The essentials of any stochastic population model are to determine when the next event occurs and its nature, both of these being conditioned on what happened at the preceding event. This procedure can be elaborated without difficulty to incorporate an arbitrary number of different mechanisms that influence the dynamics of the population. The algorithmic method usually carries Daniel Gillespie's name [16], following his treatment of chemical reactions, but it can be traced back further to a computational implementation made by David Kendall [17] of a birth–death population model that flowed from rigorous work in probability theory carried out by Andrey Kolmogorov, William Feller and Joseph Doob. The other numerical technique describes the simulation of a Gaussian random process with zero mean, unit variance and allowable but otherwise arbitrary autocorrelation function, hence a non-Markov situation. The higher-order statistics of a Gaussian process is completely defined through this single function. New continuous processes can be generated as functions of a Gaussian process, and so the ability to simulate them provides utility and flexibility for producing a range of continuous random processes.

The construction of models is the most basic activity in science. A 'good' model is one which captures the essentials of a particular phenomenon and which possesses a predictive capacity, both of these being achieved with the fewest of underlying assumptions. If those assumptions are transgressed or applied inappropriately, the model will fail to correspond to the reality. Hence, a model should not be confused with the actual phenomenon it purports to explain, rather it is a simulacrum of some aspect of its manifestation. This book contains a collection of models that have found a direct applicability in a range of application areas. But they also supply a more general way of thinking about the generation of events having different characters. We have aimed to show how these models can be fabricated in a hierarchical and systematic way from a few basic elements, and have arranged them according to the different phenomena they present. The reader can therefore select whether a given model has the appropriate character for describing their application of interest. If not, then a further aim is to provide for the reader the foundations and techniques to apply, adapt, and augment the models described here together with the ability and confidence to create new models for their own purposes.

References

1. Thomas Robert Malthus, *An Essay on the Principle of Population*, Oxford World's Classics, Oxford, 1798, 2008.
2. Nora Barlow, ed., *The Autobiography of Charles Darwin*, W.W. Norton & Co., New York, 1969.
3. Pierre-Francois Vehulst, 'Notice sur la loi que la population pursuit dans son accroissement,' *Correspondance mathematique et physique*, **10**, 113–121 (1838).
4. R.M. May, 'Simple mathematical models with very complicated dynamics,' *Nature,* **261**, 459–467 (1976).
5. Simeon-Denis Poisson, '*Recherches sur la probabilité des jugements en maitère criminelle et en matière civile,*' *Procédés des Règles Générales du Calcul Des Probabilités,* Bachelier, Imprimeur-Libraire pour les Mathématiques, Paris (1837).
6. G.U. Yule, 'A mathematical theory of evolution, based on the conclusions of Dr J.C. Willis,' *F.R.S. Philosophical Transactions of the Royal Society (London) Series B*, **213**, 21–87 (1924).
7. W.H. Furry, 'On fluctuation phenomena in the passage of high energy electrons through lead,' *The Physical Review*, **52**, 569–581 (1937).

8. P. Lévy, *Théorie de l'Addition des Variables Aléatories*, Gaultier-Villars, Paris, 1937.
9. B. Mandelbrot, 'The variation of certain speculative prices,' *The Journal of Business*, **36**, 394–419 (1963).
10. P. Bak, C. Tang, K. Weisenfeld, 'Self-organised criticality: An explanation of $1/f$ noise,' *Physical Review Letters,* **59**, 381–384 (1987).
11. A.J. Lotka, 'Contribution to the theory of periodic reaction,' *The Journal of Physical Chemistry*, **14**, 271–274, (1910). V. Volterra, 'Variations and fluctuations of the number of individuals in animal species living together,' in *Animal Ecology: With Special Reference to Insects*, R.N. Chapman, ed., McGraw-Hill, New York, 1931.
12. R.M. May, *Stability and Complexity in Model Ecosystems*, Princeton Landmarks in Biology Edition, Princeton University Press, Princeton, 2001.
13. D.R. Cox, 'Statistical methods connected with series of events,' *Journal of the Royal Statistical Society, Series B*, **17**, 129–164 (1955).
14. L. Mandel, 'Fluctuations of photon beams: The distribution of the photo-electrons,' *Proceedings of the Physical Society (London)*, **74**, 233–243 (1959).
15. I. Blake, W. Lindsay, 'Level-crossing problems for random processes,' *IEEE Transactions on Information Theory*, **19**, 295–315 (1973).
16. D.T. Gillespie, 'Exact stochastic simulation of coupled chemical reactions,' *The Journal of Physical Chemistry*, **81**, 2340–2361 (1977).
17. D.G. Kendall, 'An artificial realisation of a simple "birth-and-death" process,' *Journal of the Royal Statistical Society, Series B*, **12**, 116–119 (1950).

2

Statistical Preliminaries

2.1 Introduction

The purpose of this chapter is to provide a brief introduction to the statistical quantities and notation that will be used in the remainder of the book. The treatment is oriented to the case of *discrete* variables and is not intended to be exhaustive. The reader is referred to books listed under 'Further Reading' at the end of the chapter for more comprehensive treatments of the basics and particularly for situations in which the quantities of interest are not quantised.

After defining the basic concept of probability in the context of discrete *variables*, we shall describe some commonly encountered discrete probability distributions and characteristic measures by which their properties can be recognised and assessed. The *generating functions* that can often be used to simplify calculation will be introduced, and it will be shown how these can be used to recover both the distributions themselves and other useful properties. Discrete *processes* will be introduced through the concept of a population of individuals that is evolving with time. The number in the population at any given instant may then be related to that at previous times, giving rise to the notion of *correlation*. Another kind of discrete process is the *series of events* and in later chapters we shall show how simple mathematical models for the evolution of the number of individuals in a population can also be used to generate time sequences of this type. In the present chapter, two related measures of such a process will be discussed: the probability distribution of the number of events that occur in a given time interval and the statistics of the separation of the points in time at which the events occur.

2.2 Probability Distributions

Consider a large collection or *ensemble* of similar populations. We can work out the fraction of the populations containing exactly N individuals. This represents the *probability* $P_N \equiv P(N)$ of finding a population of N individuals in such an ensemble and we can plot this as a function of N as illustrated in