

Preface

This is the first of a planned two-volume sequence discussing mathematical aspects of population genetics theory, with an emphasis on the evolutionary theory. This first volume is intended to discuss the more introductory aspects of the theory, with the second volume taking up more advanced and more recent aspects. Because of this, this first volume draws heavily on the first (1979) edition of this book, since the material in that edition may now be taken, to a large extent, as introductory to the contemporary theory. A second reason for drawing heavily on the 1979 edition is that many present-day students have asked for access to earlier material not now easily available. It is indeed remarkable how many results well-known in the 1970's, and appearing in the literature of the time, are rediscovered in the modern literature.

On the other hand, the subject has greatly expanded in scope and depth over the last twenty-five years. Many topics have been introduced during that time, or developed well beyond the level reached in the 1970's. No doubt the most important of these is the development of the theory of molecular population genetics. Introductory aspects of this theory molecular population genetics are taken up in the later chapters of this volume, but a far more extensive description of the molecular theory will be given in Volume II. As one example of this, the theory behind currently active haplotype mapping projects will be discussed. To this extent, Volume II will be largely data-based. It will thus also form connections between evolutionary genetics and currently active areas of problems of human genetics and bioinformatics. On the other hand, developments of the evolutionary theory itself will be considered also, taking up evolutionary questions relat-

ing to many species rather than evolutionary behavior within one species. Other evolutionary topics such as the game theoretical approach to evolution, the analysis of gene-environment interactions, gene conversion and the extended development of the concept of inclusive fitness, will also be discussed in Volume II.

Despite the emphasis on evolutionary population genetics in this volume, some material concerning human genetics, in particular those parts of the theory that are best discussed in evolutionary terms, has been included. One of the more pleasing developments over the last two decades has been a convergence of work in mathematical human genetics and mathematical evolutionary genetics, areas which in 1979 had very little overlap. A manifestation of this convergence is the recent volume on mathematical population genetics and human evolution by Donnelly and Tavaré (1997).

The aim of the 1979 edition, namely to focus on the purely mathematical aspects of population genetics theory, is retained in this book, even though it is recognized that this provides a narrow and distorted view on the subject of population genetics, and indeed of theoretical population genetics, as a whole. Thus, as in 1979, the book is intended as a complement to broader and more balanced accounts of population genetics generally. There are now many excellent books available devoted to this broader field, but these often do not attempt any depth of mathematical treatment, so that there is still a place for a narrowly focussed mathematical treatment.

Apart from this, there are now several excellent books on specific aspects of population genetics theory. Of these it is appropriate to mention that by Lynch and Walsh (1998) on quantitative traits, a topic not covered in this volume, Epperson (2003) on geographical genetics and books by Christiansen (2000) and Bürger (2000) on multilocus theory. All these books carry the theory beyond the introductory level aimed at in this volume.

One aim of the 1979 volume, not explicitly stated, was to induce mathematically-trained workers to enter the population genetics field. This aim is continued in this volume, and the mathematical beauty of many of the formulas in the molecular genetics chapters of this book should help in this endeavor.

The molecular nature of current data implies that statistical methods are used far more frequently than was the case in 1979, with the molecular data being used to test various hypotheses about the evolutionary process. For the statistical analyses discussed I have adopted the standard convention of employing upper-case letters to denote random variables and the corresponding lower-case letters to denote their observed values, except in cases where this seemed pedantic. This has also sometimes implied replacing Greek letters sometimes used in the literature for random variables by Roman letters. Probability distributions and density functions are written in lower case.

Despite the fact that the earlier chapters of this book are based heavily on the 1979 edition, the discussion does sometimes differ substantially from

that in the 1979 edition, especially where the 1979 viewpoint now seems to be misguided or out of date. As one example, the discussion of the Fundamental Theorem of Natural Selection is now quite different from that of the 1979 edition. The 1979 interpretation of the theorem, standard at the time, is now seen as incorrect and has been discarded. However I have no illusions about its ability to continue to exist as the textbook interpretation, offered to students, especially since the correct interpretation requires greater mathematical depth than does the textbook version.

Current theory in mathematical population genetics emphasizes retrospective analyses rather than the prospective analyses making up much of the classical theory. In particular, theory surrounding the Kingman coalescent process forms, quite appropriately, a significant part of current research. An introduction to this theory is given in Chapter 10, and a more extensive discussion will be given in Volume II. One of the aims of this book is to make connections between the prospective theory that much of the book considers with this retrospective theory. Apart from this, the classical prospective theory, considering properties of forward-going evolutionary processes, is still relevant to retrospective analyses. As one example of this, the theory surrounding the coalescent is often best developed by considering a process moving forward in time from a common ancestor to a sample of genes in the present generation, rather than by starting with the contemporary sample and moving backward in time to the common ancestor.

Despite the natural current emphasis on the retrospective theory, there are several reasons for discussing the prospective theory in some detail in this book. The Darwinian theory of evolution continues to be attacked by various interest groups, and these attacks are sometimes helped by incorrect statements about the prospective evolutionary theory made sometimes even by biologists. The many extraordinary statements made by Løvtrup (1987), for example, illustrate this.

Even professionals in population genetics contribute to this problem. Arguments against evolution as a Darwinian process have been based the concept of the substitutional genetic load, which I believe has been deleterious concept that should be dropped from into the theory. Substitutional load “theory”, as well as segregational load “theory”, is discussed, and I hope debunked, in Section 2.11. The “blind watchmaker” paradigm, periodically raised by outsiders to population genetics theory as refuting the Darwinian process and indeed evolution generally, is discussed, and I hope also debunked, in Section 1.6. On two more narrow points where those active in areas close to population genetics theory frequently abuse the theory, the correct as opposed to the textbook version of the Fundamental Theorem of Natural Selection, mentioned above, is described in Sections 2.9 and 7.4.5. A discussion of the much-misunderstood expression “effective population size”, often incorrectly used in with reference to the history of the human population, is given in Section 3.7.

The recent and welcome infusion of population genetics theory into a variety of disciplines associated with the evolutionary process has not been without some problems. Perhaps the most important of these is that it has led to an uncritical use of some formulas from the theory without due assessment of whether the formulas are appropriate to the situation at hand. All formulas in population genetics theory derive some model of the evolutionary process, and in some cases this model can be no more than a very rough approximation to reality. For this reason a new section has been added, in this volume, discussing the modeling process and what may reasonably be concluded from the models discussed in the population genetics literature.

On more technical matters, it has not always been possible to use the notation of various published papers whose results are described here, since in some cases the notation used in different papers for the same quantity differ, and in other cases different authors use the same symbol for different quantities. As in the 1979 version of this book, the notation is not consistent, so that the symbol " x_i " might variously mean the frequency of an allele in generation i , in subpopulation i , the frequency of the allele A_i , and so on. On a similar point, I have adopted American spelling but English punctuation conventions: The latter are more suited to a mathematical text.

It is a pleasure to acknowledge the inspiration I have received from my long-time colleagues Bob Griffiths and Geoff Watterson. It is also a pleasure to thank Peter Donnelly, John Kingman and Simon Tavaré for an equally close, albeit long-range, collegial association. I thank various colleagues for pointing out typographical errors in the 1979 edition of this book, and Alan Rogers for pointing out an error concerning exchangeable model calculations. Any errors observed in this volume will be gratefully received at wewens@sas.upenn.edu and an archive of these will be maintained at www.textbook-errata.org

Philadelphia, Pennsylvania, USA
October 2003

Warren J. Ewens



<http://www.springer.com/978-0-387-20191-7>

Mathematical Population Genetics 1

Theoretical Introduction

Ewens, W.J.

2004, XX, 418 p., Hardcover

ISBN: 978-0-387-20191-7