Information Geometry and Population Genetics: The Mathematical Structure of the Wright-Fisher Model

by Julian Hofrichter, Jurgen Jost and Tat Dat Tran
Dordrecht: Springer, 2017
ISBN 978-33-1952-044-5
Softcover $109, 331 pp.

Reviewed by Ranjita Pandey
University of Delhi

Population genetics is the study of evolutionary behaviour of alleles caused by genetic drift, mutations, natural selection, or recombination. *Information Geometry and Population Genetics* masterfully explores the stochastic dynamics of the progressive distribution of alleles over generations through a geometric perspective on the traditional Wright-Fisher model. *Information geometry* is the study of the geometry of probability distributions under parametric set up. The authors of this book provide a unified source of information on the mathematics of the Wright-Fisher model. The book exhaustively covers mathematical concepts related to analysis based on the Wright-Fisher model, while exploring its potential to connect to other mathematically amenable disciplines. The novel contribution of the present book is in providing a new analytic approach for computing coexistence probabilities of different alleles, a geometric perspective to assess and analyze resultant impact of recombination, free energy constructions for asymptotic inferences, and in hierarchical modelling of allele loss.

The book comprises of three broad sections, arranged in ten chapters, with the objective of translating biological processes into mathematical form. The first two chapters are devoted to an introduction of the elementary Wright-Fisher model and its adaptations. The next five chapters develop the geometry behind the model through Kolmogorov equations and computational tools of moments and free-energy functionals. The last three chapters focus on hierarchical approaches to solutions.

Chapter 2 begins with an explanation of background assumptions for the basic Wright-Fisher model. Extensions of the model to accommodate mutation and selection effects are given. Probability distributions, utilized in the development of a modelling approach through Kolmogorov equations, are described, addressing the evolution of allele distribution for the future as well as in the ancestral states. Probabilistic preliminaries and visualization of their extension to notion of cubes and their boundary instances are covered in a simplistic manner.

Chapter 3 deals with Riemannian metrics for the probability simplex which is utilised in the future chapters to analyse relative allele frequencies. Related mathematical concepts of basic Calculus operators, connection map, curvature tensor, torsion tensor, Fisher metric, geometry of probability simplex, affine and Beltrami Laplacian, and Brownian motion on the sphere are presented in a simple and understandable manner.

Chapter 4 begins with the justification of Markovian diffusion limits, which is derived for the Wright-Fisher model. The corresponding moment evolution equations of the rudimentary two-allele model without selection or mutation, to the more general combinatorial forms are described. The idea of *moment duality* is presented with Kingman coalescent model which chalks out the ancestral trajectory of a population generation resulting from Wright-Fisher models.

Chapter 5 explores dynamics of allele frequency probabilities through the tenets of gamete recombination, linkage and stochastic combinations of gametes and zygotes. Mathematics for the hierarchy in the single locus case, compositionality in the two locus case and geometry of the state space of recombination and linkage equilibria are elaborately covered for various loci-allelic permutations.

In chapter 6 construction of moment generating function and free energy functionals for Kolmogorov forward equations of Wright Fisher model are explained in context of two alleles case with mutation and selection, which are then generalised to N individuals with $n+1$ alleles. Authors deftly handle mathematics of the evolution of free energy functional along the flow of specified densities and curvature dimension conditions based on the assumption of positive uniform mutation rates.

Under chapter 7, Large Deviation Principle (LDP) for a sequence of probability measures are discussed with illustrative examples. This is followed by systematic reconstruction of minimizers for action functionals for various adaptations of Wright Fisher model. Two alleles case, without and with mutations (one way and two way) and selection, is covered with graphical explanation.

Chapter 8 is an in-depth exposition on the diffusion approximation of the ($n+1$) allelic one locus Wright-Fisher model, without mutation and selection based on Kolmogorov forward equation. Construction of local solution, determination of moments and weak formulation of the Kolmogorov forward equation is followed by hierarchical solution and their extensions leading to deeper comprehension of boundary transitions.

Asymmetry between forward and backward equations implies that some special treatment is required for configuration on boundary of the probability simplex for the latter. Chapter 9 undertakes assessment of influence of boundary composition on the interior and its subsequent extension to the entire stratified boundary of the domain. Probabilistic interpretation and iterated extensions are followed by exhaustive stochastic treatment of a suitably iterated application of the basic blow-up transformation with intention to translate the extended Kolmogorov backward equation into a corresponding differential equation. Stationary solutions to the Kolmogorov backward equation are developed with the objective of determining expected times for exit of one or several alleles.

Chapter 10 sums up the contributions of the text by focussing on computation of biological metrics for two alleles and k alleles population. Hierarchical solutions to determine rate of loss of one allele in a three allele population are provided.

The authors have given exhaustive mathematical derivation for each of the stochastic tools proposed, with reference to the application, analysis, and interpretation of Wright-Fisher model. Related literature reviews are presented adequately in the context of specific topics as the book unfolds. Building stochastic solutions for the complex loci structure and linkage equilibria are handled remarkably well in the book, with a foresight to identify the expected distribution of alleles in future generations. Overall, the book provides exhaustive coverage of the Wright-Fisher model and its application to expected genetic transformations in future generations. The present book is a useful piece of literature for applied biologists with a fair understanding of calculus, who are looking toward the exploration of new dimensions in research on genetic evolution.