

Warren J. Ewens

Mathematical Population Genetics

I. Theoretical Introduction

Second Edition



Springer

Warren J. Ewens
Department of Biology
University of Pennsylvania
Philadelphia, PA 19104
USA
wewens@sas.upenn.edu

Editors

S.S. Antman
Department of Mathematics
and
Institute for Physical Science and Technology
University of Maryland
College Park, MD 20742
USA
ssa@math.umd.edu

J.E. Marsden
Control and Dynamical Systems
Mail Code 107-81
California Institute of Technology
Pasadena, CA 91125
USA
marsden@cds.caltech.edu

L. Sirovich
Division of Applied Mathematics
Brown University
Providence, RI 02912
USA
chico@camelot.mssm.edu

S. Wiggins
School of Mathematics
University of Bristol
Bristol BS8 1TW
UK
s.wiggins@bris.ac.uk

Mathematics Subject Classification (2000): 92-02, 92D10, 60J70

ISBN 978-1-4419-1898-7 ISBN 978-0-387-21822-9 (eBook)
DOI 10.1007/978-0-387-21822-9

© 2004 Springer Science+Business Media New York

Originally published by Springer-Verlag New York, Inc in 2004

Softcover reprint of the hardcover 2nd edition 2004

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Science+Business Media, LLC), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Contents

Preface	vii
Introduction	xvii
1 Historical Background	1
1.1 Biometricians, Saltationists and Mendelians	1
1.2 The Hardy–Weinberg Law	3
1.3 The Correlation Between Relatives	6
1.4 Evolution	11
1.4.1 The Deterministic Theory	11
1.4.2 Non-Random-Mating Populations	18
1.4.3 The Stochastic Theory	20
1.5 Evolved Genetic Phenomena	31
1.6 Modelling	35
1.7 Overall Evolutionary Theories	38
2 Technicalities and Generalizations	43
2.1 Introduction	43
2.2 Random Union of Gametes	44
2.3 Dioecious Populations	44
2.4 Multiple Alleles	49
2.5 Frequency-Dependent Selection	54
2.6 Fertility Selection	54
2.7 Continuous-Time Models	57

2.8	Non-Random-Mating Populations	62
2.9	The Fundamental Theorem of Natural Selection	64
2.10	Two Loci	67
2.11	Genetic Loads	78
2.12	Finite Markov Chains	86
3	Discrete Stochastic Models	92
3.1	Introduction	92
3.2	Wright–Fisher Model: Two Alleles	92
3.3	The Cannings (Exchangeable) Model: Two Alleles	99
3.4	Moran Models: Two Alleles	104
3.5	K-Allele Wright–Fisher Models	109
3.6	Infinitely Many Alleles Models	111
3.6.1	Introduction	111
3.6.2	The Wright–Fisher Infinitely Many Alleles Model	111
3.6.3	The Cannings Infinitely Many Alleles Model	117
3.6.4	The Moran Infinitely Many Alleles Model	117
3.7	The Effective Population Size	119
3.8	Frequency-Dependent Selection	129
3.9	Two Loci	129
4	Diffusion Theory	136
4.1	Introduction	136
4.2	The Forward and Backward Kolmogorov Equations	137
4.3	Fixation Probabilities	139
4.4	Absorption Time Properties	140
4.5	The Stationary Distribution	145
4.6	Conditional Processes	146
4.7	Diffusion Theory	148
4.8	Multi-dimensional Processes	151
4.9	Time Reversibility	153
4.10	Expectations of Functions of Diffusion Variables	153
5	Applications of Diffusion Theory	156
5.1	Introduction	156
5.2	No Selection or Mutation	158
5.3	Selection	165
5.4	Selection: Absorption Time Properties	167
5.5	One-Way Mutation	171
5.6	Two-Way Mutation	174
5.7	Diffusion Approximations and Boundary Conditions	176
5.8	Random Environments	181
5.9	Time-Reversal and Age Properties	188
5.10	Multi-Allele Diffusion Processes	192

6	Two Loci	201
6.1	Introduction	201
6.2	Evolutionary Properties of Mean Fitness	202
6.3	Equilibrium Points	208
6.4	Special Models	209
6.5	Modifier Theory	221
6.6	Two-Locus Diffusion Processes	227
6.7	Associative Overdominance and Hitchhiking	230
6.8	The Evolutionary Advantage of Recombination	235
6.9	Summary	239
7	Many Loci	241
7.1	Introduction	241
7.2	Notation	242
7.3	The Random Mating Case	243
7.3.1	Linkage Disequilibrium, Means and Variances	243
7.3.2	Recurrence Relations for Gametic Frequencies	245
7.3.3	Components of Variance	246
7.3.4	Particular Models	249
7.4	Non-Random Mating	254
7.4.1	Introduction	254
7.4.2	Notation and Theory	255
7.4.3	Marginal Fitnesses and Average Effects	256
7.4.4	Implications	258
7.4.5	The Fundamental Theorem of Natural Selection	259
7.4.6	Optimality Principles	261
7.5	The Correlation Between Relatives	266
7.6	Summary	274
8	Further Considerations	276
8.1	Introduction	276
8.2	What is Fitness?	276
8.3	Sex Ratio	277
8.4	Geographical Structure	278
8.5	Age Structure	282
8.6	Ecological Considerations	283
8.7	Sociobiology	285
9	Molecular Population Genetics: Introduction	288
9.1	Introduction	288
9.2	Technical Comments	290
9.3	Infinitely Many Alleles Models: Population Properties	292
9.3.1	The Wright–Fisher Model	292
9.3.2	The Moran Model	294
9.4	Infinitely Many Sites Models: Population Properties	297

9.4.1	Introduction	297
9.4.2	The Wright–Fisher Model	298
9.4.3	The Moran Model	300
9.5	Sample Properties of Infinitely Many Alleles Models	301
9.5.1	Introduction	301
9.5.2	The Wright–Fisher Model	301
9.5.3	The Moran Model	306
9.6	Sample Properties of Infinitely Many Sites Models	308
9.6.1	Introduction	308
9.6.2	The Wright–Fisher Model	308
9.6.3	The Moran Model	314
9.7	Relation Between Infinitely Many Alleles and Infinitely Many Sites Models	316
9.8	Genetic Variation Within and Between Populations	319
9.9	Age-Ordered Alleles: Frequencies and Ages	320
10	Looking Backward in Time: The Coalescent	328
10.1	Introduction	328
10.2	Competing Poisson and Geometric Processes	329
10.3	The Coalescent Process	330
10.4	The Coalescent and Its Relation to Evolutionary Genetic Models	331
10.5	Coalescent Calculations: Wright–Fisher Models	333
10.6	Coalescent Calculations: Exact Moran Model Results	338
10.7	General Comments	341
10.8	The Coalescent and Human Genetics	342
11	Looking Backward: Testing the Neutral Theory	346
11.1	Introduction	346
11.2	Testing in the Infinitely Many Alleles Models	349
11.2.1	Introduction	349
11.2.2	The Ewens and the Watterson Tests	349
11.2.3	Procedures Based on the Conditional Sample Frequency Spectrum	353
11.2.4	Age-Dependent Tests	354
11.3	Testing in the Infinitely Many Sites Models	355
11.3.1	Introduction	355
11.3.2	Estimators of θ	356
11.3.3	The Tajima Test	358
11.3.4	Other “Tajima-like” Testing Procedures	361
11.3.5	Testing for the Signature of a Selective Sweep	362
11.3.6	Combining Infinitely Many Alleles and Infinitely Many Sites Approaches	364
11.3.7	Data from Several Unlinked Loci	365
11.3.8	Data from Unlinked Sites	368

11.3.9 Tests Based on Historical Features	369
12 Looking Backward in Time:	
Population and Species Comparisons	370
12.1 Introduction	370
12.1.1 The Reversibility Criterion	372
12.2 Various Evolutionary Models	373
12.2.1 The Jukes–Cantor Model	373
12.2.2 The Kimura Model and Its Generalizations	374
12.2.3 The Felsenstein Models	375
12.3 Some Implications	377
12.3.1 Introduction	377
12.3.2 The Jukes–Cantor Model	377
12.3.3 The Kimura Model	380
12.4 Statistical Procedures	381
Appendix A: Eigenvalue Calculations	384
Appendix B: Significance Levels for \hat{F}	385
Appendix C: Means and Variances of \hat{F}	386
References	387
Author Index	409
Subject Index	413