
Statistical Genetics

Theory, Practice and Application

Statistical Genetics

Theory, Practice and Application

Olúwáfémi Samuel Oyámákin

Stirling-Horden Publishers Ltd.
No 6, Adonai Estate, Opp Crown
Heights College, Arulogun Road
Ojoo, Ibadan.
Phone: 08023519154/07031967972
stirlinghorden@ymail.com, info@stirlinghorden.com
www.stirlinghorden.com

University Post Office	7 Ikioda Street
Box 20984	Jattu-Uzairue
Ibadan	Edo State, Nigeria.
Oyo State, Nigeria.	

© Olúwáfémi Samuel Oyámákin

Published 2026

ISBN 978-978-032-

All Rights Reserved

Apart from any fair dealing for the purpose of research or private study, criticism or review, as permitted under the Copyright, Designs and Patents Act of Nigeria, this publication may not be reproduced, stored or transmitted in any form, or by any means, except with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

Stirling-Horden Publishers Ltd.
Lagos, Ibadan, Benin City, Jattu-Uzairue.

Dedication

Acknowledgements

Preface

Introduction to Statistical Genetics has been written to bridge the gap between classical genetics and modern statistical practice, for students of Statistics, Biostatistics, Genetics, Public Health and the Biological Sciences in Nigeria and across Africa. Every chapter draws its illustrative material from conditions, populations and health problems that are familiar to the Nigerian reader — the sickle-cell gene among the Yoruba, Igbo and Hausa/Fulani peoples, G6PD deficiency and its interaction with malaria, the ABO and Rhesus blood group systems in Nigerian blood banks, oculocutaneous albinism, consanguinity in parts of Northern Nigeria, the *APOLI* kidney-disease risk variants common in West African ancestry, and quantitative traits such as birth weight, packed cell volume (PCV) and height measured in Nigerian cohorts.

This edition, titled *Statistical Genetics: Theory, Practice and Application*, substantially expands the first: six new chapters — on Pedigree Analysis, Segregation Analysis and Genetic Model Fitting, Candidate Gene Association Studies, Family-Based Association Tests (the TDT and FBAT), Population Substructure, Stratification, and Admixture, Genotype Imputation, and Genome Sequencing, Rare-Variant Methods, and Simulation-Based Approaches — have been added to the original fourteen, bringing the book's coverage from classical transmission and population genetics through to the family-based, genome-wide, and sequencing-based methods that dominate contemporary human statistical genetics. Every equation in this edition has been reset

in proper mathematical notation, and illustrative figures — pedigree charts, meiosis and recombination diagrams, linkage-disequilibrium heatmaps, coalescent gene trees, population-structure plots, and more — accompany the theory throughout, so that the visual and symbolic language of statistical genetics reinforces the written explanation at every step. Every chapter, old and new, now closes with a Computational Corner giving parallel, runnable R and Python code for the chapter's key calculations, worked wherever possible on simulated data reflecting Nigerian populations, diseases, and cohort studies, so that the reader not only learns the theory but sees, immediately, how to compute it.

To keep each chapter's narrative flow intact, worked solutions to every end-of-chapter Practice Question have been collected together into a single Master Answer Key (Appendix D) at the end of the book, with short R and Python verification scripts accompanying every computational question. A new Appendix B provides a practical guide to setting up the R and Python computing environments used throughout the book.

The book follows the classical sequence used in graduate courses in Statistical Genetics (such as STA 762 at the University of Ibadan): from the biological foundations of the gene, through Mendelian and population genetics, pedigree and family-based analysis, to the modern statistical machinery of likelihood inference, association testing (candidate-gene, family-based, and genome-wide), linkage disequilibrium, coalescent theory, linkage analysis, genotype imputation, and high-throughput sequencing. This book assumes a first course in probability and statistical inference, and at least an elementary acquaintance with R or Python (both are introduced gently, side by side, from the first Computational Corner onward). It is intended for use both as a primary text in a one- or two-semester course in Statistical

Genetics, and as a reference for researchers in biostatistics, epidemiology, and public health genetics working with Nigerian and other African data.

Samuel Oluwafemi Oyamakin, PhD
*Department of Statistics,
University of Ibadan, Ibadan, Nigeria*

Contents

Dedication	v
Acknowledgements	vi
Preface	vii
1. Basic Biological Concepts in Genetics	1
2. Mendel's Laws of Inheritance	19
3. Hardy–Weinberg Equilibrium	33
4. Estimation of Allele Frequency: Approach to Equilibrium for an X-Linked Gene	49
5. Non-Random Mating: Phenotypic Assortative Mating	61
6. Pedigree Analysis: Structure, Notation, and Relatedness	73
7. Segregation Analysis and Genetic Model Fitting	89
8. Estimation of the Recombination Fraction	101
9. Inheritance of Quantitative Traits	115
10. Likelihood Methods in Statistical Genetics	129

11. Population-Based Association Tests	143
12. Candidate Gene Association Studies	155
13. Family-Based Association Tests: The TDT and FBAT	169
14. Population Substructure, Stratification, and Admixture	181
15. Linkage Disequilibrium and Haplotype Inference	197
16. Coalescent Models	211
17. Parametric and Non-Parametric Linkage Analysis	225
18. Genome-Wide Association Studies (GWAS)	241
19. Genotype Imputation	257
20. Genome Sequencing, Rare-Variant Methods, and Simulation-Based Approaches in Statistical Genetics	271
Appendix A: Selected Statistical Tables	290
Appendix B: R and Python Computing Setup Guide	293
Appendix C: Glossary of Key Terms	296
Appendix D: Master Answer Key to End-of-Chapter Exercises	301

