

Preface

The importance of genetics can be felt almost daily. Genetics news not only appears in scientific magazines such as *Science*, *Nature*, *Discovery* and *Scientific American* where genetics occupies a very large proportion, but also in daily newspapers where news of genes, cloning and genetically engineered new species attracts much public attention. Though I learned some statistics with genetic applications in my graduate study, they seemed to be very far from the excitement of present day genetics. I began to read papers in modern genetics to see what new statistical tools are used. I also attended several short courses and did a sequence of experiments in DNA recombination and protein manipulation. Out of the knowledge I gathered and the data analyses I performed for the medical school, I taught a genetic data analysis course several times. This book grew out of my lecture notes for the course.

There are several recent books on statistical methods in genetics, such as *Genetic Data Analysis II* by Bruce Weir, *Analysis of Human Genetic Linkage* by Jurg Ott and *Mathematical and Statistical Methods for Genetic Analysis* by Kenneth Lange. All are excellent books and I have used them in my class. Though there is some overlap between this book and theirs, the overlap is not very much. My emphasis is on understanding how and why statistical concepts and methods are needed and used.

There are numerous topics that one can select in modern genetics. The topics I chose are undoubtedly biased, but these were the topics I wanted to know more about when I got into this field. I feel that many beginners will share my interest in these topics. They include:

- How is a gene found?
- How have scientists separated the genetic and environmental aspects of a person's intelligence?
- How have genetics been used in agriculture so that the domestic animals and crops are constantly improving?
- What is a DNA fingerprint and why are there controversies about it?
- How did they use genes to rebuild the evolutionary history?

I endeavored to understand them and I believe I do now. I hope that in this book you will find no gaps in how these questions are answered. Though the statistical theory behind basic modern genetics is not difficult, most statistical genetics papers are not easy to read for a beginner. Formulae quickly become very tedious to fit a particular application of genetics. This book tries to distinguish between the necessary and unnecessary complexity in presenting the material to a statistician. The details for specific applications are referred to as references.

Since this book is written mainly for statistics students, a certain statistical background is assumed. The necessary background statistical knowledge beyond elementary statistics is given in Appendix A. One can easily see whether one has enough background for this book. One year of graduate study in most statistics departments should be sufficient.