Preface

Linear Model theory plays a fundamental role in the foundation of mathematical and applied statistics. It has a base in distribution theory and statistical inference, and finds application in many advanced areas in statistics including univariate and multivariate regression, analysis of designed experiments, longitudinal and time series analysis, spatial analysis, multivariate analysis, wavelet methods, etc. Most statistics departments offer at least one course on linear model theory at the graduate level. There are several excellent books on the subject, such as "Linear Statistical Inference and its Applications" by C.R. Rao, "Linear Models" by S.R. Searle, "Theory and Applications of the Linear Model" by F.A. Graybill, "Plane Answers to Complex Questions: The Theory of Linear Models" by R. Christiansen and "The Theory of Linear Models" by B. Jorgensen.

Our motivation has been to incorporate general principles of inference in linear models to the fundamental statistical education of students at the graduate level, while our treatment of contemporary topics in a systematic way will serve the needs of professionals in various industries. The three salient features of this book are: (1) developing standard theory of linear models with numerous applications in simple and multiple regression, as well as fixed, random and mixed-effects models, (2) introducing generalized linear models with examples, and (3) presenting some current topics including Bayesian linear models, general additive models, dynamic linear models and longitudinal models. The first two chapters introduce to the reader requisite linear and matrix algebra. This book is therefore a self-contained exposition of the theory of linear models, including motivational and practical aspects. We have tried to achieve a healthy compromise between theory and practice, by providing a sound theoretical basis, and indicating how the theory works in important special cases in practice. There are several examples throughout the text. In addition, we provide summaries of many numerical examples in different chapters, while a more comprehensive description of these is available in the first author's web site (http://www.stat.uconn.edu/∼nalini). There are several exercises at the end of each chapter that should serve to reinforce the methods.

Our entire book is intended for a two semester graduate course in linear models. For a one semester course, we recommend essentially the first eight chapters, omitting a few subsections, if necessary, and supplementing a few selected topics from chapters 9-11, if time permits. For instance, section 5.5, section 6.4, sections 7.5.2-7.5.4, and sections 8.5, 8.7 and 8.8 may be omitted in a one semester course. The first two chapters, which present a review on vectors and matrices specifically as they pertain to linear model theory, may also be assigned as background reading if the students had previous exposure to these topics. Our book requires some knowledge of statistics; in particular, a knowledge of elementary sampling distributions, basic estimation theory and hypothesis testing at an undergraduate level is definitely required. Occasionally, more advanced concepts of statistical inference are invoked in this book, for which suitable references are provided.

The plan of this book follows. The first two chapters develop basic concepts of linear and matrix algebra with a view towards application in linear models. Chapter 3 describes generalized inverses and solutions to systems of linear equations. We develop the notion of a general linear model in Chapter 4. An attractive feature of our book is that we unify full-rank and non full-rank models in the development of least squares inference and optimality via the Gauss-Markov theorem. Results for the full-rank (regression) case are provided as special cases. We also introduce via examples, balanced ANOVA models that are widely used in practice. Chapter 5 deals with multivariate normal and related distributions, as well as distributions of quadratic forms that are at the heart of inference. We also introduce the class of elliptical distributions that can serve as error distributions for linear models. Sampling from multivariate normal distributions is the topic of Chapter 6, together with assessment of and transformations to multivariate normality. This is followed by inference for the general linear model in Chapter 7. Inference under normal and elliptical errors is developed and illustrated on examples from regression and balanced ANOVA models. In Chapter 8, topics in multiple regression models such as model checking, variable selection, regression diagnostics, robust regression and nonparametric regression are presented. Chapter 9 is devoted to the study of unbalanced designs in fixed-effects ANOVA models, the analysis of covariance (ANACOVA) and some nonparametric test procedures. Random-effects models and mixed-effects models are discussed in detail in Chapter 10. Finally in Chapter 11, we introduce several special topics including Bayesian linear models, dynamic linear models, linear longitudinal models and generalized linear models (GLIM). The purpose of this chapter is to introduce to the reader some new frontiers of linear models theory; several references are provided so that the reader may explore further in these directions. Given the exploding nature of our subject area, it is impossible to be exhaustive in a text, and cover everything that should ideally be covered. We hope that our judgment in choice of material is appropriate and useful.

Most of our book was developed in the form of lecture notes for a sequence of two courses on linear models which both of us have taught for several years in the Department of Statistics at the University of Connecticut. The numerical examples in the text and in the web site were developed by NR over many years. In the text, we have acknowledged published work, wherever appropriate, for the use of data in our numerical examples, as well as for some of the exercise problems. We are indeed grateful for their use, and apologize for any inadvertent omission in this regard.

In writing this text, discussions with many colleagues were invaluable. In particular, we thank Malay Ghosh, for several suggestions that vastly improved the structure and content of this book. We deeply appreciate his time and goodwill. We thank Chris Chatfield and Jim Lindsey for their review and for the suggestion about including numerical examples in the text. We are also very grateful for the support and encouragement of our statistical colleagues, in particular Joe Glaz, Bani Mallick, Alan Gelfand and Yazhen Wang. We thank Ming-Hui Chen for all his technical help with Latex.

Many graduate students helped in proof reading the typed manuscript; we are especially grateful to Junfeng Liu, Madhuja Mallick and Prashni Paliwal. We also thank Karen Houle, a graduate student in Statistics, who helped with "polishing-up" the numerical examples in NR's web site. We appreciate all the help we received from people at Chapman & Hall/CRC – Bob Stern, Helena Redshaw, Gail Renard and Sean Davey.

Ravi, what can I say, except thanks for the warm smiles and hot dinners! N.R. Rita and Debosri, without your sacrifice, the project wouldn't be completed on time. D.K.D.

Nalini Ravishanker and Dipak K. Dey
Department of Statistics
University of Connecticut
Storrs, CT