
Preface

This book is intended for use as the textbook in a second course in applied statistics that covers topics in multiple regression and analysis of variance at an intermediate level. Generally, students enrolled in such courses are primarily graduate majors or advanced undergraduate students from a variety of disciplines. These students typically have taken an introductory-level statistical methods course that requires the use a software system such as SAS for performing statistical analysis. Thus students are expected to have an understanding of basic concepts of statistical inference such as estimation and hypothesis testing.

Understandably, adequate time is not available in a first course in statistical methods to cover the use of a software system adequately in the amount of time available for instruction. The aim of this book is to teach how to use the SAS system for data analysis. The SAS language is introduced at a level of sophistication not found in most introductory SAS books. Important features such as SAS data step programming, pointers, and line-hold specifiers are described in detail. The powerful graphics support available in SAS is emphasized throughout, and many worked SAS program examples contain graphic components.

The basic theory of those statistical methods covered in the text is discussed briefly and then is extended beyond the elementary level. Particular attention has been given to topics that are usually not included in introductory courses. These include models involving random effects, covariance analysis, variable subset selection in regression methods, categorical data analysis, and graphical tools for residual diagnostics. However, a thorough knowledge of advanced theoretical material such as linear model theory will not be assumed or required to assimilate the material presented.

SAS programs and SAS program outputs are used extensively to supplement the description of the analysis methods. Example data sets are from the biological and physical sciences and engineering. Exercises are included in each chapter. Most exercises involve constructing SAS programs for the analysis of given observational or experimental data. Complete pdf files of all

SAS examples used in the book can be downloaded from the Springer website www.springer.com/978-0-387-77371-1. The text versions of all data sets used in examples and exercises are available from the website. Statistical tables are not reprinted in the book. These are also on the website.

The first author has taught a one-semester course based on material from this book for several years. The coverage depends on the preparation and maturity level of students enrolled in a particular semester. In a class mainly composed of graduate students from disciplines other than statistics, with adequate knowledge of statistical methods and the use of SAS, most of the book is covered. Otherwise, in a mixed class of undergraduate and graduate students with little experience using SAS, the coverage is usually 5 weeks of introduction to SAS, 5 weeks on regression and graphics, and 5 weeks of anova applications. This amounts to approximately 60% of the material in the textbook. The structure of sections in the chapters facilitates this kind of selective coverage.

SAS for Data Analysis: Intermediate Statistical Methods, although intended to be used as a textbook, may also be useful as a reference to researchers and data analysts both in the academic setting and in industry.

The first author wishes to thank his former instructors at Kansas State University, Professor George Milliken and Professor Dallas Johnson, for introducing him to the SAS language as well as to the fundamentals of statistical data analysis. He is also grateful to Professor Kenneth Koehler, Chair of the Department of Statistics at Iowa State University for steady encouragement and allowing him the time and resources to complete this work. The students in his data analysis course were responsible for locating numerous errors and typos in earlier versions of the manuscript. The authors also wish to thank Professor Dan Nettleton, Laurence H. Baker Endowed Chair in Biological Statistics and Professor of Statistics at Iowa State University for valuable discussions concerning inference from the mixed model and providing a nice exercise problem and Dr. Grace H. Liu of Pioneer Hi-Bred for reading an earlier version of the manuscript and offering constructive suggestions.

MERVYN G. MARASINGHE

Associate Professor

Department of Statistics

Iowa State University, Ames, IA 50011,
USA

WILLIAM J. KENNEDY

Professor Emeritus

Department of Statistics

Iowa State University, Ames, IA 50011,
USA

SAS for Data Analysis

Intermediate Statistical Methods

Marasinghe, M.G.; Kennedy, W.J.

2008, XII, 558 p. With 100 SAS Programs., Hardcover

ISBN: 978-0-387-77371-1