

海外优秀数学类教材系列丛书

Mc
Graw
Hill

Educational

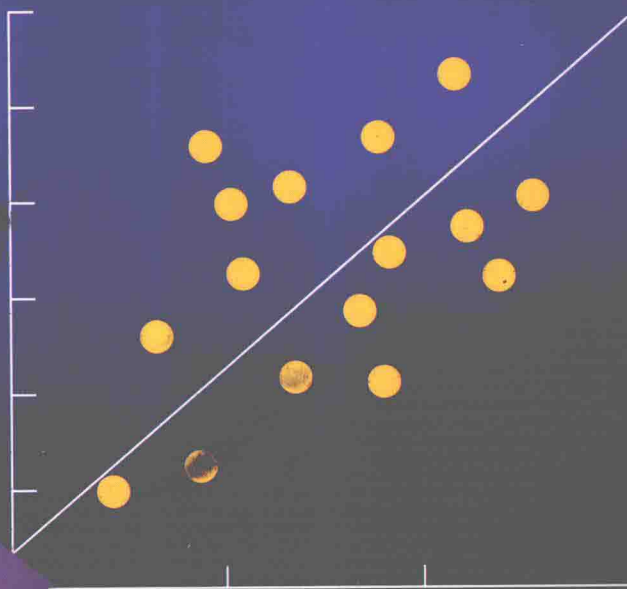
影印版

Applied Linear Regression Models (Fourth Edition)

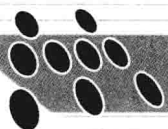
应用线性回归模型

(第4版)

- ☐ KUTNER
- ☐ NACHTSHEIM
- ☐ NETER



高等教育出版社
Higher Education Press



海外优秀数学类教材系列丛书

影印版

Applied Linear Regression Models (Fourth Edition)

应用线性回归模型

(第4版)

Michael H. Kutner
Emory University

Christopher J. Nachtsheim
University of Minnesota

John Neter
University of Georgia



高等教育出版社
Higher Education Press

图字：01-2004-4908 号

Kutner Nachtsheim Neter

Applied Linear Regression Models, Fourth Edition

ISBN: 007-123252-4

Copyright © 2004 by the McGraw-Hill Companies, Inc.

Original language published by The McGraw-Hill Companies, Inc. All Rights reserved. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Authorized English language reprint edition jointly published by McGraw-Hill Education (Asia) Co. and Higher Education Press. This edition is authorized for sale in the People's Republic of China only, excluding Hong Kong, Macao SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书英文影印版由高等教育出版社和美国麦格劳-希尔教育出版(亚洲)公司合作出版。此版本仅限在中华人民共和国境内(不包括香港、澳门特别行政区及台湾)销售。未经许可之出口,视为违反著作权法,将受法律之制裁。

未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有 McGraw-Hill 公司防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

应用线性回归模型 = Applied Linear Regression Models: 第4版/(美)库特纳等著. —影印本. —北京: 高等教育出版社, 2005.2

(海外优秀数学类教材系列丛书)

ISBN 7-04-016380-2

I. 应... II. 库... III. 线性回归—高等学校—教材—英文 IV. 0212.1

中国版本图书馆 CIP 数据核字(2004)第 143560 号

策划编辑 徐可 责任编辑 徐可 封面设计 王凌波 责任印制 孔源

出版发行 高等教育出版社
社 址 北京市西城区德外大街4号
邮政编码 100011
总 机 010-58581000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>

经 销 北京蓝色畅想图书发行有限公司
印 刷 北京铭成印刷有限公司

网上订购 <http://www.landaco.com>
<http://www.landaco.com.cn>

开 本 787×1092 1/16
印 张 45.5
字 数 1 000 000

版 次 2005年2月第1版
印 次 2005年2月第1次印刷
定 价 60.00元(含光盘)

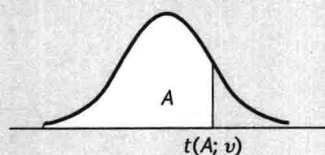
本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 16380-00

TABLE B.2
Percentiles
of the t
Distribution.

Entry is $t(A; \nu)$ where $P\{t(\nu) \leq t(A; \nu)\} = A$



ν	A						
	.60	.70	.80	.85	.90	.95	.975
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

TABLE B.2
(concluded)
Percentiles
of the *t*
Distribution.

<i>ν</i>	A						
	.98	.985	.99	.9925	.995	.9975	.9995
1	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	2.197	2.336	2.528	2.661	2.845	3.153	3.849
21	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	2.054	2.170	2.326	2.432	2.576	2.807	3.291

The McGraw-Hill/Irwin Series: Operations and Decision Sciences

BUSINESS STATISTICS

Alwan

Statistical Process Analysis
First Edition

Acel and Souderpandian
Complete Business Statistics
Fifth Edition

Bowerman and O'Connell
Business Statistics in Practice
Third Edition

Bryant and Smith
**Practical Data Analysis:
Case Studies in Business Statistics,
Volumes I, II, III**
Second Edition

Cooper and Schindler
Business Research Methods
Eighth Edition

Delurgio
**Forecasting Principles
and Applications**
First Edition

Doane
LearningStats
First Edition

Doane, Mathieson, and Tracy
Visual Statistics
Second Edition, 2.0

Gitlow, Oppenheim, and Oppenheim
**Quality Management: Tools and
Methods for Improvement**
Second Edition

Lind, Marchal, and Wathen
**Basic Statistics for Business
and Economics**
Fourth Edition

Lind, Marchal, and Mason
**Statistical Techniques in Business
and Economics**
Eleventh Edition

Merchant, Goffinet, and Koehler
**Basic Statistics Using Excel
for Office 2000**
Third Edition

Sahai and Khurshid
Pocket Dictionary of Statistics
First Edition

Siegel
Practical Business Statistics
Fifth Edition

Wilson, Keating, and John Galt
Solutions, Inc.

Business Forecasting
Fourth Edition

Zagorsky
Business Information
First Edition

QUANTITATIVE METHODS AND MANAGEMENT SCIENCE

Bodily, Carraway, Frey, Pfeifer
**Quantitative Business Analysis:
Text and Cases**
First Edition

Bonini, Hausman, and Bierman
**Quantitative Analysis for Business
Decisions**
Ninth Edition

Hillier and Hillier
**Introduction to Management Science:
A Modeling and Case Studies Approach
with Spreadsheets**
Second Edition

出版者的话

在我国已经加入 WTO、经济全球化的今天,为适应当前我国高校各类创新人才培养的需要,大力推进教育部倡导的双语教学,配合教育部实施的“高等学校教学质量与教学改革工程”和“精品课程”建设的需要,高等教育出版社有计划、大规模地开展了海外优秀数学类系列教材的引进工作。

高等教育出版社和 Pearson Education, John Wiley & Sons, McGraw-Hill, Thomson Learning 等国外出版公司进行了广泛接触,经国外出版公司的推荐并在国内专家的协助下,提交引进版权总数 100 余种。收到样书后,我们聘请了国内高校一线教师、专家、学者参与这些原版教材的评介工作,并参考国内相关专业的课程设置和教学实际情况,从中遴选出了这套优秀教材组织出版。

这批教材普遍具有以下特点:(1)基本上是近 3 年出版的,在国际上被广泛使用,在同类教材中具有相当的权威性;(2)高版次,历经多年教学实践检验,内容翔实准确、反映时代要求;(3)各种教学资源配套整齐,为师生提供了极大的便利;(4)插图精美、丰富,图文并茂,与正文相辅相成;(5)语言简练、流畅、可读性强,比较适合非英语国家的学生阅读。

本系列丛书中,有 Finney、Weir 等编的《托马斯微积分》(第 10 版, Pearson),其特色可用“呈传统特色、富革新精神”概括,本书自 20 世纪 50 年代第 1 版以来,平均每四五年就有一个新版面世,长达 50 余年始终盛行于西方教坛,作者既有相当高的学术水平,又热爱教学,长期工作在教学第一线,其中,年近 90 的 G.B. Thomas 教授长年在 MIT 工作,具有丰富的教学经验;Finney 教授也在 MIT 工作达 10 年;Weir 是美国数学建模竞赛委员会主任。Stewart 编的立体化教材《微积分》(第 5 版, Thomson Learning)配备了丰富的教学资源,是国际上最畅销的微积分原版教材,2003 年全球销量约 40 余万册,在美国,占据了约 50%~60% 的微积分教材市场,其用户包括耶鲁等名牌院校及众多一般院校。本系列丛书还包括 Anton 编的经典教材《线性代数及其应用》(第 8 版, Wiley); Jay L. Devore 编的优秀教材《概率论与数理统计》(第 5 版, Thomson Learning)等。在努力降低引进教材售价方面,高等教育出版社做了大量和细致的工作,这套引进的教材体现了一定的权威性、系统性、先进性和经济性等特点。

通过影印、翻译、编译这批优秀教材,我们一方面要不断地分析、学习、消化吸收国外优秀教材的长处,吸取国外出版公司的制作经验,提升我们自编教材的立体化配套标准,使

我国高校教材建设水平上一个新的台阶；与此同时，我们还将尝试组织海外作者和国内作者合编外文版基础课数学教材，并约请国内专家改编部分国外优秀教材，以适应我国实际教学环境。

这套教材出版后，我们将结合各高校的双语教学计划，开展大规模的宣传、培训工作，及时地将本套丛书推荐给高校使用。在使用过程中，我们衷心希望广大高校教师和同学提出宝贵的意见和建议。

高等教育出版社高等理科分社联系电话：010-58581384，E-mail: xuke@hep.com.cn。

高等教育出版社
2004年4月20日

To
Nancy, Michelle, Allison,
Maureen, Abigael, Andrew, Henry G.,
Dorothy, Ron, David

Preface

Linear regression models are widely used today in business administration, economics, engineering, and the social, health, and biological sciences. Successful applications of these models require a sound understanding of both the underlying theory and the practical problems that are encountered in using the models in real-life situations. While *Applied Linear Regression Models*, Fourth Edition, is basically an applied book, it seeks to blend theory and applications effectively, avoiding the extremes of presenting theory in isolation and of giving elements of applications without the needed understanding of the theoretical foundations.

The fourth edition differs from the third in a number of important respects.

1. We have reorganized the chapters for better clarity and flow of topics. Material from the old Chapter 15 on normal correlation models has been integrated throughout the text where appropriate. Much of the material is now found in an expanded Chapter 2, which focuses on inference in regression analysis. Material from the old Chapter 7 pertaining to polynomial and interaction regression models and from old Chapter 11 on qualitative predictors has been integrated into a new Chapter 8 called, "Regression Models for Quantitative and Qualitative Predictors." Material on model validation from old Chapter 10 is now fully integrated with updated material on model selection in a new Chapter 9 entitled, "Building the Regression Model I: Model Selection and Validation."
2. We have added material on important techniques for data mining, including regression trees and neural network models, in Chapters 11 and 13, respectively.
3. We have made extensive revisions to the problem material. Problem data sets are generally larger and more challenging, and we have included a large number of new case data sets in Appendix C. In addition, we have added a new category of chapter exercises, called Case Studies. These are open-ended problems that require students, given an overall objective, to carry out complete analyses of the various case data sets in Appendix C. They are distinct from the material in the Problems and Projects sections, which frequently ask students to simply carry out specific analytical procedures.
4. The chapter on logistic regression (Chapter 14) has been extensively revised and expanded to include a more thorough treatment of logistic, probit, and complementary log-log models, logistic regression residuals, model selection, model assessment, logistic regression diagnostics, and goodness of fit tests. We have also developed new material on polytomous (multicategory) nominal logistic regression models and polytomous ordinal logistic regression models.
5. We have expanded the discussion of model selection methods and criteria. The Akaike information criterion and Schwarz Bayesian criterion have been added, and a greater emphasis is placed on the use of cross-validation for model selection and validation.
6. We have substantially expanded the amount of graphic presentation, including much greater use of scatter plot matrices, three-dimensional rotating plots, three-dimensional response surface and contour plots, and conditional effects plots.
7. Throughout the text, we have made extensive revisions in the exposition on the basis of classroom experience to improve the clarity of the presentation.

We have included in this book not only the more conventional topics in regression, but also topics that are frequently slighted, though important in practice. We devote three chapters (Chapter 9–11) to the model-building process for regression, including computer-assisted selection procedures for identifying good subsets of predictor variables and validation of the chosen regression model. Two chapters (Chapters 8 and 14) are devoted to indicator variables, covering both response and predictor indicator variables. The use of residual analysis and other diagnostics for examining the appropriateness of a regression model is a recurring theme throughout this book. So is the use of remedial measures that may be helpful when the model is not appropriate. In the analysis of the results of a study, we give greater emphasis to the use of estimation procedures than to significance tests, because estimation is often more meaningful in practice. Also, since practical problems seldom are concerned with a single inference, we stress the use of simultaneous inference procedures.

Theoretical ideas are presented to the degree needed for good understanding in making sound applications. Proofs are given in those instances where we feel they serve to demonstrate an important method of approach. Emphasis is placed on a thorough understanding of the regression models, particularly the meaning of the model parameters, since such understanding is basic to proper applications. A wide variety of examples and cases is presented to illustrate the use of the theoretical principles, to show the great diversity of applications of regression models, and to demonstrate how analyses are carried out for different problems.

We use “Comments” sections in each chapter to present additional discussion and matters related to the mainstream of development. In this way, the basic ideas in a chapter are presented concisely and without distraction.

Applications of regression models frequently require extensive computations. We take the position that a computer is available in most applied work and that almost every computer user has access to program packages for regression analysis. Hence, we explain the basic mathematical steps in fitting a regression model but do not dwell on computational details. This approach permits us to avoid many complex formulas and enables us to focus on basic principles. We make extensive use in this text of computer capabilities for performing computations and preparing graphic plots, and we illustrate a variety of computer printouts and plots and explain how they are used for analysis.

A selection of problems is provided at the end of each chapter. Here readers can reinforce their understanding of the methodology and use the concepts learned to analyze data. We have been careful to supply data-analysis problems that typify genuine applications. In most problems the calculations are best handled on a computer. To facilitate data entry, a compact disk is provided with the text that includes the data sets for all examples, problems, exercises, projects, and case studies, as well as for the data sets in Appendix C. The README.TXT file on the compact disk provides information about the identification of the data sets.

We assume that the reader of *Applied Linear Regression Models*, Fourth Edition, has had an introductory course in statistical inference, covering the material outlined in Appendix A. Should some gaps in the reader’s background exist, the relevant portions of an introductory text can be studied, or the instructor of the class may use supplemental materials for covering the missing segments. Appendix A is primarily intended as a reference of basic statistical results for continuing use as the reader progresses through the book.

Calculus is not required for reading *Applied Linear Regression Models*, Fourth Edition. In a number of instances, we use calculus to demonstrate how some important results are obtained, but these demonstrations are confined to supplementary comments and can

be omitted without any loss of continuity. Readers who do know calculus will find these comments in natural sequence so that the benefits of the mathematical developments are obtained in their immediate context. Some basic elements of matrix algebra are needed for multiple regression. Chapter 5 introduces these elements of matrix algebra in the context of simple regression for easy learning.

Applied Linear Regression Models, Fourth Edition, is intended for use in undergraduate and graduate courses in regression analysis and in second courses in applied statistics. The extent to which material presented in this text is used in a particular course depends upon the amount of time available and the objectives of the course. The basic elements of regression are covered in Chapters 1, 2, 3, 4 (Sections 4.1–4.3), 5, 6, 7, 8, 9, and 10. Chapters 11, 12, 13, and 14 can be covered as time permits and interests dictate.

This book can also be used for self-study by persons engaged in the fields of business administration, economics, engineering, and the social, health, and biological sciences who desire to obtain competence in the application of regression models.

An *Instructor Solutions Manual*, containing detailed solutions to all numerical problems and analytical exercises, is available from the publisher, McGraw-Hill/Irwin, for use by instructors.

A *Student Solutions Manual* is included on the compact disk for use by students. The Student Solutions Manual provides intermediate and final numerical results for easy self-checking of solutions for selected problems. We use an asterisk (*) in front of the problem number to designate the problems for which the solutions appear in the Student Solutions Manual.

Available new for this Fourth Edition is a *SAS and SPSS Program Solution Manual* for examples contained in the text. The manual was prepared by Dr. William Replogle and Dr. William Johnson, both of the University of Mississippi Medical School. The manual uses data files to present solutions, output, and discussion of necessary steps for students to solve examples selected from the book using SAS and SPSS.

The *Student Solutions Manual* and all of the data files on the compact disk can also be downloaded from the book's website at: www.mhhe.com/KutnerALRM4e. A list of errata for the book as well as some useful, related links will also be maintained at this address.

A book such as this cannot be written without substantial assistance from numerous persons. We are indebted to the many contributors who have developed the theory and practice discussed in this book. We also would like to acknowledge appreciation to our students, who helped us in a variety of ways to fashion the method of presentation contained herein. We are grateful to the many users of *Applied Linear Statistical Models* and *Applied Linear Regression Models*, who have provided us with comments and suggestions based on their teaching with these texts. We are also indebted to Professors James E. Holstein, University of Missouri, and David L. Sherry, University of West Florida, for their review of *Applied Linear Statistical Models*, First Edition; to Professors Samuel Kotz, University of Maryland at College Park, Ralph P. Russo, University of Iowa, and Peter F. Thall, The George Washington University, for their review of *Applied Linear Regression Models*, First Edition; to Professors John S. Y Chiu, University of Washington, James A. Calvin, University of Iowa, and Michael F. Driscoll, Arizona State University, for their review of *Applied Linear Statistical Models*, Second Edition; to Professor Richard Anderson-Sprecher, University of Wyoming, for his review of *Applied Linear Regression Models*, Second Edition; and to Professors Alexander von Eye, The Pennsylvania State University, Samuel Kotz, University of

Maryland at College Park, and John B. Willett, Harvard University, for their review of *Applied Linear Statistical Models*, Third Edition; to Professors Jason Abrevaya, University of Chicago, Frank Alt, University of Maryland, Vitoria Chen, Georgia Tech, Rebecca Doerge, Purdue University, Mark Henry, Clemson University, Jim Hobert, University of Florida, Ken Koehler, Iowa State University, Chii-Dean Lin, University of Massachusetts Amherst, Mark Reiser, Arizona State University, Lawrence Ries, University of Missouri Columbia, and Ehsan Soofi, University of Wisconsin Milwaukee, for their reviews of *Applied Linear Regression Models*, Third Edition, or *Applied Linear Statistical Models*, Fourth Edition. These reviews provided many important suggestions, for which we are most grateful.

In addition, valuable assistance was provided by Professors Richard K. Burdick, Arizona State University, R. Dennis Cook, University of Minnesota, W. J. Conover, Texas Tech University, Mark E. Johnson, University of Central Florida, Dick DeVeaux, Williams College, and by Drs. Richard I. Beckman, Los Alamos National Laboratory, Ronald L. Iman, Sandia National Laboratories, and Brad Jones, SAS Institute. We are most appreciative of their willing help. We are also indebted to the 88 participants in a survey concerning *Applied Linear Regression Models*, Second Edition, the 76 participants in a survey concerning *Applied Linear Statistical Models*, Third Edition, and the 73 participants in a survey concerning *Applied Linear Regression Models*, Third Edition, or *Applied Linear Statistical Models*, Fourth Edition. Helpful suggestions were received in these surveys, for which we are thankful.

Weiyong Zhang and Vincent Agboto assisted us in the development of new problem material, and Lexin Li and Yingwen Dong prepared the revised *Instructor Solutions Manual* and *Student Solutions Manual* under considerable time pressure. Amy Hendrickson provided much-needed LaTeX expertise. We are most grateful to these individuals for their invaluable help and assistance. We also wish to thank the various members of the Carlson Executive MBA Program classes of 2003 and 2004, notably Mike Ohmes, Trevor Bynum, Baxter Stephenson, Zakir Salyani, Sanders Marvin, Trent Spurgeon, Nate Ogzawalla, David Mott, Preston McKenzie, Bruce DeJong, and Tim Kensok, for their contributions of interesting and relevant case study data and materials.

Finally, our families bore patiently the pressures caused by our commitment to complete this revision. We are appreciative of their understanding.

Michael H. Kutner

Christopher J. Nachtsheim

John Neter

Contents

PART ONE

SIMPLE LINEAR REGRESSION 1

Chapter 1

Linear Regression with One Predictor Variable 2

- 1.1 Relations between Variables 2
 - Functional Relation between Two Variables* 2
 - Statistical Relation between Two Variables* 3
- 1.2 Regression Models and Their Uses 5
 - Historical Origins* 5
 - Basic Concepts* 5
 - Construction of Regression Models* 7
 - Uses of Regression Analysis* 8
 - Regression and Causality* 8
 - Use of Computers* 9
- 1.3 Simple Linear Regression Model with Distribution of Error Terms Unspecified 9
 - Formal Statement of Model* 9
 - Important Features of Model* 9
 - Meaning of Regression Parameters* 11
 - Alternative Versions of Regression Model* 12
- 1.4 Data for Regression Analysis 12
 - Observational Data* 12
 - Experimental Data* 13
 - Completely Randomized Design* 13
- 1.5 Overview of Steps in Regression Analysis 13
- 1.6 Estimation of Regression Function 15
 - Method of Least Squares* 15
 - Point Estimation of Mean Response* 21
 - Residuals* 22
 - Properties of Fitted Regression Line* 23
- 1.7 Estimation of Error Terms Variance σ^2 24
 - Point Estimator of σ^2* 24
- 1.8 Normal Error Regression Model 26
 - Model* 26
 - Estimation of Parameters by Method of Maximum Likelihood* 27

Cited References 33

Problems 33

Exercises 37

Projects 38

Chapter 2

Inferences in Regression and Correlation Analysis 40

- 2.1 Inferences Concerning β_1 40
 - Sampling Distribution of b_1* 41
 - Sampling Distribution of $(b_1 - \beta_1)/s\{b_1\}$* 44
 - Confidence Interval for β_1* 45
 - Tests Concerning β_1* 47
- 2.2 Inferences Concerning β_0 48
 - Sampling Distribution of b_0* 48
 - Sampling Distribution of $(b_0 - \beta_0)/s\{b_0\}$* 49
 - Confidence Interval for β_0* 49
- 2.3 Some Considerations on Making Inferences Concerning β_0 and β_1 50
 - Effects of Departures from Normality* 50
 - Interpretation of Confidence Coefficient and Risks of Errors* 50
 - Spacing of the X Levels* 50
 - Power of Tests* 50
- 2.4 Interval Estimation of $E\{Y_h\}$ 52
 - Sampling Distribution of \hat{Y}_h* 52
 - Sampling Distribution of $(\hat{Y}_h - E\{Y_h\})/s\{\hat{Y}_h\}$* 54
 - Confidence Interval for $E\{Y_h\}$* 54
- 2.5 Prediction of New Observation 55
 - Prediction Interval for $Y_{h(\text{new})}$ when Parameters Known* 56
 - Prediction Interval for $Y_{h(\text{new})}$ when Parameters Unknown* 57
 - Prediction of Mean of m New Observations for Given X_h* 60
- 2.6 Confidence Band for Regression Line 61
- 2.7 Analysis of Variance Approach to Regression Analysis 63
 - Partitioning of Total Sum of Squares* 63
 - Breakdown of Degrees of Freedom* 66

	<i>Mean Squares</i>	66
	<i>Analysis of Variance Table</i>	67
	<i>Expected Mean Squares</i>	68
	<i>F Test of $\beta_1 = 0$ versus $\beta_1 \neq 0$</i>	69
2.8	General Linear Test Approach	72
	<i>Full Model</i>	72
	<i>Reduced Model</i>	72
	<i>Test Statistic</i>	73
	<i>Summary</i>	73
2.9	Descriptive Measures of Linear Association between X and Y	74
	<i>Coefficient of Determination</i>	74
	<i>Limitations of R^2</i>	75
	<i>Coefficient of Correlation</i>	76
2.10	Considerations in Applying Regression Analysis	77
2.11	Normal Correlation Models	78
	<i>Distinction between Regression and Correlation Model</i>	78
	<i>Bivariate Normal Distribution</i>	78
	<i>Conditional Inferences</i>	80
	<i>Inferences on Correlation Coefficients</i>	83
	<i>Spearman Rank Correlation Coefficient</i>	87
	<i>Cited References</i>	89
	<i>Problems</i>	89
	<i>Exercises</i>	97
	<i>Projects</i>	98

Chapter 3

Diagnostics and Remedial Measures 100

3.1	Diagnostics for Predictor Variable	100
3.2	Residuals	102
	<i>Properties of Residuals</i>	102
	<i>Semistudentized Residuals</i>	103
	<i>Departures from Model to Be Studied by Residuals</i>	103
3.3	Diagnostics for Residuals	103
	<i>Nonlinearity of Regression Function</i>	104
	<i>Nonconstancy of Error Variance</i>	107
	<i>Presence of Outliers</i>	108
	<i>Nonindependence of Error Terms</i>	108
	<i>Nonnormality of Error Terms</i>	110
	<i>Omission of Important Predictor Variables</i>	112
	<i>Some Final Comments</i>	114

3.4	Overview of Tests Involving Residuals	114
	<i>Tests for Randomness</i>	114
	<i>Tests for Constancy of Variance</i>	115
	<i>Tests for Outliers</i>	115
	<i>Tests for Normality</i>	115
3.5	Correlation Test for Normality	115
3.6	Tests for Constancy of Error Variance	116
	<i>Brown-Forsythe Test</i>	116
	<i>Breusch-Pagan Test</i>	118
3.7	F Test for Lack of Fit	119
	<i>Assumptions</i>	119
	<i>Notation</i>	121
	<i>Full Model</i>	121
	<i>Reduced Model</i>	123
	<i>Test Statistic</i>	123
	<i>ANOVA Table</i>	124
3.8	Overview of Remedial Measures	127
	<i>Nonlinearity of Regression Function</i>	128
	<i>Nonconstancy of Error Variance</i>	128
	<i>Nonindependence of Error Terms</i>	128
	<i>Nonnormality of Error Terms</i>	128
	<i>Omission of Important Predictor Variables</i>	129
	<i>Outlying Observations</i>	129
3.9	Transformations	129
	<i>Transformations for Nonlinear Relation Only</i>	129
	<i>Transformations for Nonnormality and Unequal Error Variances</i>	132
	<i>Box-Cox Transformations</i>	134
3.10	Exploration of Shape of Regression Function	137
	<i>Lowess Method</i>	138
	<i>Use of Smoothed Curves to Confirm Fitted Regression Function</i>	139
3.11	Case Example—Plutonium Measurement	141
	<i>Cited References</i>	146
	<i>Problems</i>	146
	<i>Exercises</i>	151
	<i>Projects</i>	152
	<i>Case Studies</i>	153

Chapter 4

Simultaneous Inferences and Other Topics in Regression Analysis 154

- 4.1 Joint Estimation of β_0 and β_1 154
 - Need for Joint Estimation* 154
 - Bonferroni Joint Confidence Intervals* 155
- 4.2 Simultaneous Estimation of Mean Responses 157
 - Working-Hotelling Procedure* 158
 - Bonferroni Procedure* 159
- 4.3 Simultaneous Prediction Intervals for New Observations 160
- 4.4 Regression through Origin 161
 - Model* 161
 - Inferences* 161
 - Important Cautions for Using Regression through Origin* 164
- 4.5 Effects of Measurement Errors 165
 - Measurement Errors in Y* 165
 - Measurement Errors in X* 165
 - Berkson Model* 167
- 4.6 Inverse Predictions 168
- 4.7 Choice of X Levels 170
 - Cited References* 172
 - Problems* 172
 - Exercises* 175
 - Projects* 175

Chapter 5

Matrix Approach to Simple Linear Regression Analysis 176

- 5.1 Matrices 176
 - Definition of Matrix* 176
 - Square Matrix* 178
 - Vector* 178
 - Transpose* 178
 - Equality of Matrices* 179
- 5.2 Matrix Addition and Subtraction 180
- 5.3 Matrix Multiplication 182
 - Multiplication of a Matrix by a Scalar* 182
 - Multiplication of a Matrix by a Matrix* 182
- 5.4 Special Types of Matrices 185
 - Symmetric Matrix* 185
 - Diagonal Matrix* 185

Vector and Matrix with All Elements

Unity 187

Zero Vector 187

- 5.5 Linear Dependence and Rank of Matrix 188
 - Linear Dependence* 188
 - Rank of Matrix* 188
- 5.6 Inverse of a Matrix 189
 - Finding the Inverse* 190
 - Uses of Inverse Matrix* 192
- 5.7 Some Basic Results for Matrices 193
- 5.8 Random Vectors and Matrices 193
 - Expectation of Random Vector or Matrix* 193
 - Variance-Covariance Matrix of Random Vector* 194
 - Some Basic Results* 196
 - Multivariate Normal Distribution* 196
- 5.9 Simple Linear Regression Model in Matrix Terms 197
- 5.10 Least Squares Estimation of Regression Parameters 199
 - Normal Equations* 199
 - Estimated Regression Coefficients* 200
- 5.11 Fitted Values and Residuals 202
 - Fitted Values* 202
 - Residuals* 203
- 5.12 Analysis of Variance Results 204
 - Sums of Squares* 204
 - Sums of Squares as Quadratic Forms* 205
- 5.13 Inferences in Regression Analysis 206
 - Regression Coefficients* 207
 - Mean Response* 208
 - Prediction of New Observation* 209
 - Cited Reference* 209
 - Problems* 209
 - Exercises* 212

PART TWO

MULTIPLE LINEAR REGRESSION 213

Chapter 6

Multiple Regression I 214

- 6.1 Multiple Regression Models 214

- Need for Several Predictor Variables 214
 - First-Order Model with Two Predictor Variables 215
 - First-Order Model with More than Two Predictor Variables 217
 - General Linear Regression Model 217
 - 6.2 General Linear Regression Model in Matrix Terms 222**
 - 6.3 Estimation of Regression Coefficients 223**
 - 6.4 Fitted Values and Residuals 224**
 - 6.5 Analysis of Variance Results 225**
 - Sums of Squares and Mean Squares 225
 - F Test for Regression Relation 226
 - Coefficient of Multiple Determination 226
 - Coefficient of Multiple Correlation 227
 - 6.6 Inferences about Regression Parameters 227**
 - Interval Estimation of β_k 228
 - Tests for β_k 228
 - Joint Inferences 228
 - 6.7 Estimation of Mean Response and Prediction of New Observation 229**
 - Interval Estimation of $E\{Y_h\}$ 229
 - Confidence Region for Regression Surface 229
 - Simultaneous Confidence Intervals for Several Mean Responses 230
 - Prediction of New Observation $Y_{h(\text{new})}$ 230
 - Prediction of Mean of m New Observations at X_h 230
 - Predictions of g New Observations 231
 - Caution about Hidden Extrapolations 231
 - 6.8 Diagnostics and Remedial Measures 232**
 - Scatter Plot Matrix 232
 - Three-Dimensional Scatter Plots 233
 - Residual Plots 233
 - Correlation Test for Normality 234
 - Brown-Forsythe Test for Constancy of Error Variance 234
 - Breusch-Pagan Test for Constancy of Error Variance 234
 - F Test for Lack of Fit 235
 - Remedial Measures 236
 - 6.9 An Example—Multiple Regression with Two Predictor Variables 236**
 - Setting 236
 - Basic Calculations 237
 - Estimated Regression Function 240
 - Fitted Values and Residuals 241
 - Analysis of Appropriateness of Model 241
 - Analysis of Variance 243
 - Estimation of Regression Parameters 245
 - Estimation of Mean Response 245
 - Prediction Limits for New Observations 247
 - Cited Reference 248
 - Problems 248
 - Exercises 253
 - Projects 254
- ## Chapter 7
- ### Multiple Regression II 256
- 7.1 Extra Sums of Squares 256**
 - Basic Ideas 256
 - Definitions 259
 - Decomposition of SSR into Extra Sums of Squares 260
 - ANOVA Table Containing Decomposition of SSR 261
 - 7.2 Uses of Extra Sums of Squares in Tests for Regression Coefficients 263**
 - Test whether a Single $\beta_k = 0$ 263
 - Test whether Several $\beta_k = 0$ 264
 - 7.3 Summary of Tests Concerning Regression Coefficients 266**
 - Test whether All $\beta_k = 0$ 266
 - Test whether a Single $\beta_k = 0$ 267
 - Test whether Some $\beta_k = 0$ 267
 - Other Tests 268
 - 7.4 Coefficients of Partial Determination 268**
 - Two Predictor Variables 269
 - General Case 269
 - Coefficients of Partial Correlation 270
 - 7.5 Standardized Multiple Regression Model 271**
 - Roundoff Errors in Normal Equations Calculations 271
 - Lack of Comparability in Regression Coefficients 272
 - Correlation Transformation 272
 - Standardized Regression Model 273
 - $X'X$ Matrix for Transformed Variables 274