外国优秀统计学教材系列丛书·影印版

中国统计出版社
China Statistics Press

THOMSON

作者：

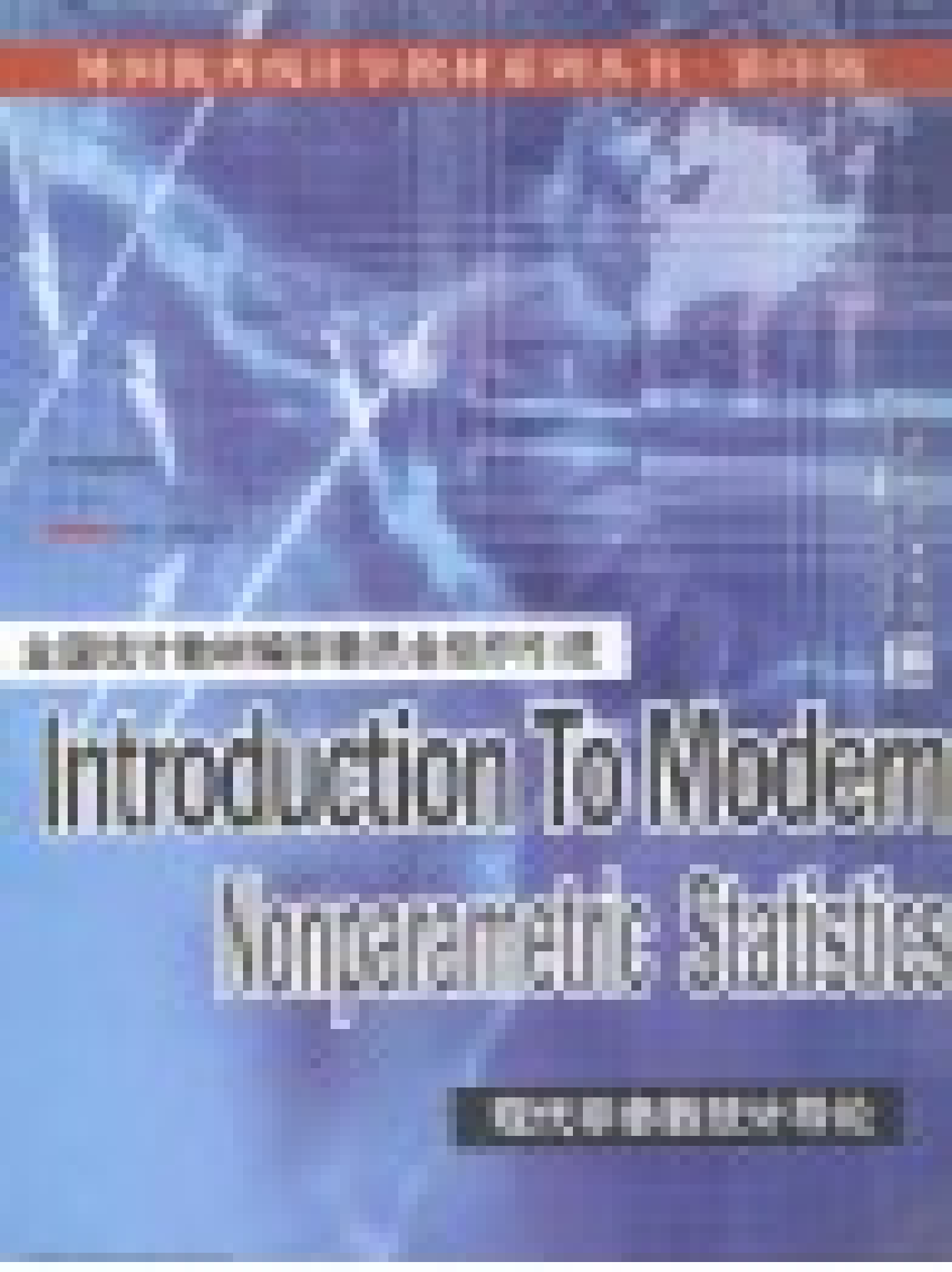
James J.Higgins

全国统计教材编审委员会组织引进

Introduction To Modern Nonparametric Statistics

现代非参数统计导论

http://www.stats.gov.cn


Introduction To Modern
Nonparametric Statistics

外国优秀统计学教材系列丛书.影印版

全国统计教材编审委员会组织引进

现代非参数统计导论

Introduction To Modern Nonparametric Statistics

[美] Higgings J.J 著

中国统计出版社
China Statistics Press

COPYRJGHT©2004 Brooks/Cole,a division of Thomson Learning,Inc.Thomson Learning™ is a trademark used herein under license.
ISBN 0-534-38775-6
ALL RIGHTS RESERVED.No part of this work covered by the copyright hereon may be reproduced or used in any form or by any means — graphic,electronic,or mechanical,including but not limited to photocopying, recording, taping,Web distribution,information networks,or information storage and retrieval systems — without the written permission of the publisher.
981-254-980-3

本书英文影印版由汤姆森学习出版集团 THOMSON 授权中国统计出版社独家出版。
未经出版者书面许可，不得以任何方式翻印、拷贝、仿制或转载。
此版本仅限在中华人民共和国境内（不包括中国香港、澳门特别行政区及台湾）销售
版权所有，侵权必究。
北京市版权局著作权合同登记号：01-2004-4361

（京）新登字041号

图书在版编目（CIP）数据

现代非参数统计导论 /（美）希金斯（Higgins,J.J.）著.
- 北京：中国统计出版社，2005.3
ISBN 7-5037-4550-9
Ⅰ.现… Ⅱ.希… Ⅲ.非参数统计-英文 Ⅳ.0212.7

中国版本图书馆CIP数据核字(2005)第022543号

责任编辑/徐占波
封面设计/刘国宁　赵　辉
E-mail/cbsebs@stats.gov.cn
出版发行/中国统计出版社
通信地址/北京市西城区月坛南街75号
邮政编码/100826
办公地址/北京市丰台区西三环南路甲6号
电　　话/（010）63459084　63266600-22500(发行部)
印　　刷/北京兴华印刷厂
经　　销/新华书店
开　　本/850×1092毫米　1/16
字　　数/580千字
印　　张/24.75
印　　数/1-3000册
版　　别/2005年6月第1版
版　　次/2005年6月北京第1次印刷
书　　号/ISBN 7-5037-4550-9/O·55
定　　价/46.00元

中国统计版图书，如有印装错误，本社发行部负责调换。

出版说明

21世纪的竞争是人才的竞争，是全球性人才培养机制的较量。如何培养面向现代化、面向世界、面向未来的高素质的人才成为我国人才培养的当务之急。为此，教育部发出通知，倡导在全国普通高等学校中使用原版外国教材，进行双语教学，培养适应经济全球化的人才。

为了响应教育部的号召，促进统计教材的改革，培养既懂统计专业知识又具备较高英语语言能力的统计人才，全国统计教材编审委员会在国家统计局领导的大力支持下，组织引进了这套“外国优秀统计学教材”。

为了做好“外国优秀统计学教材”引进工作，全国统计教材编审委员会将其列入了“十五”规划，并成立了由海内外统计学家组成的专家委员会。在对国外统计学教材的使用情况进行了充分了解，对国内高等院校使用外国统计学教材的需求情况进行了仔细分析，并对从各种渠道推荐来的统计教材进行了认真审定的基础上，制定了引进教材书目。在确定引进教材书目的过程中，我们得到了国内外有关专家、有关院校和外国出版公司及其北京办事处的支持和帮助，在此致谢。中国人民大学统计学系的吴喜之教授不仅推荐了大量的优秀候选书目，而且校译了影印教材的翻译目录，为这套教材的及早出版作了大量的工作，我们表示衷心的谢意。

这套引进教材多数是国外再版多次、反响良好，又比较适合国内情况、易于教学的统计教材。我们希望这套引进教材的出版对促进我国统计教材的改革和高校统计学专业双语教学的发展能够起到重要的推动作用。

全国统计教材编审委员会

2002年8月28日

此为试读，需要完整PDF请访问：www.ertongbook.com

引进外国优秀统计学教材专家委员会

主　任：贺　铿　国家统计局副局长

副主任：王吉利　国家统计局统计教育中心主任
谢鸿光　国家统计局办公室主任
严建辉　中国统计出版社副社长

委　员：（按姓氏笔画排序）

文兼武　国家统计局统计科学研究所　所长
孙山泽　北京大学数学科学学院概率统计系　教授
刘　科　中国统计出版社　总编
吴喜之　中国人民大学统计学系　教授
何书元　北京大学数学科学学院概率统计系　教授
张尧庭　上海财经大学经济学院　教授
张润楚　南开大学数学学院统计学系　教授
邱　东　东北财经大学　教授
陈　江　美国雪城大学管理学院　教授
陈家鼎　北京大学数学科学学院概率统计系　教授
杨振海　北京工业大学应用数理学院　教授
孟晓犁　哈佛大学统计系　教授
林正炎　浙江大学理学院数学系　教授
茆诗松　华东师范大学统计系　教授
郑　明　复旦大学管理学院统计系　副教授
柯惠新　北京广播学院　教授
徐国祥　上海财经大学统计学系　教授
袁　卫　中国人民大学　教授
蒋　萍　东北财经大学统计系　教授
谢邦昌　台湾辅仁大学统计资讯学系　教授
濮晓龙　华东师范大学统计系　副教授

办公室：
温　明　国家统计局统计教育中心教材处　处长
杨映霜　中国统计出版社教材编辑部　编审
孙洪娟　国家统计局统计教育中心教材处　统计师
吕　军　中国统计出版社教材编辑部　副编审

Preface

Modern nonparametric statistics comprises a broad range of methods for data analysis. Rank-based methods, permutation tests, bootstrap methods, and curve smoothing fall under this heading. Most normal-theory methods have nonparametric counterparts that may be used when the normal-theory assumptions are violated. Nonparametric methods may also solve problems for which conventional methods do not apply or are difficult to implement. These problems include the analysis of sparse contingency tables, tests for ordered alternatives, one-sided multivariate tests, and the analysis of censored data.

This book presents a wide array of nonparametric methods that researchers will find useful in analyzing their data. It is appropriate for advanced undergraduates and beginning graduate students in fields such as the life sciences, engineering, medicine, social sciences, and statistics. Students are expected to be familiar with topics typically covered in an introductory methods course such as *t* tests, inferences for proportions, simple linear regression, and one-way analysis of variance. A few topics in the latter part of the book require a more advanced background, including multiple regression and two-way analysis of variance.

At one time nonparametric methods were thought of as quick, hand-calculation methods, suitable only for simple designs and small data sets. However, the availability of high-speed computing to carry out time-consuming or difficult computations has enabled nonparametric statistics to come into its own as a field of statistics. Computationally intensive methods such as permutation tests, bootstrap estimation, curve smoothing, and robust methods, along with the classic methods based on ranks, have become part of the statistician's tool kit.

The selection of material for this book was guided by data analysis problems that commonly arise in practice. In addition to the familiar one-sample, two-sample, and *k*-sample procedures, nonparametric counterparts to analysis of variance, multiple regression, and multivariate analysis are considered. Categorical data methods include exact permutation tests for contingency tables and tests for tables with ordinal classifications. A chapter on censored data covers common nonparametric techniques

for the analysis of survival data. Bootstrap methods, nonparametric density estimation, and curve smoothing are also considered.

Where possible, the connection among methods is stressed. For instance, rank tests are introduced as special cases of permutation tests applied to ranks, and methods for censored data are treated as special cases of permutation tests applied to general scores. Some statistical theory is sketched where it is felt that it would help illuminate the methods. In most cases, the theory is set apart from the methods and may be omitted at the option of the instructor. Selected computer output and code show how methods are implemented. Although generally not needed in light of the availability of statistical packages, statistical tables are given for some of the more common tests.

Chapter 1 treats one-sample methods based on the binomial distribution. The issue of the power of statistical tests is discussed, and an example is given to show how a simple nonparametric test for a location parameter can have substantially greater power than the one-sample t test.

Chapter 2 sets a pattern that is used for the next six chapters. A permutation test is introduced to solve a problem, and then rank tests or other tests are developed as special cases or modifications of the permutation test. Chapter 2 introduces two-sample methods. K-sample methods, including multiple comparisons, follow in Chapter 3. Then come paired comparisons and blocked designs in Chapter 4, tests for trends and association including contingency table analysis in Chapter 5, multivariate tests in Chapter 6, and tests for censored data in Chapter 7. Asymptotic approximations are introduced in Chapter 2 and appear throughout. However, the need for such approximations is not as compelling as it once was as a result of the availability of computer software to carry out exact tests.

Chapter 8 introduces bootstrap methods for one-sample, two-sample, k-sample, and regression problems. In Chapter 9, both bootstrap and aligned-rank methods are given for two-factor analysis of variance with the extension to the multifactor setting being straightforward. A section on lattice-ordered alternatives extends tests for ordered alternatives to the multifactor setting. In Chapter 10, nonparametric density estimation, curve smoothing, and robust model fitting are introduced. The availability of computer software to carry out the computations makes this material accessible to the practitioner.

The material in this text is sufficient for a one-semester course. Problems at the end of each chapter are included to provide practice with the methods and, in some cases, to extend theory or methods not covered in the text.

Statistical software is indispensable for the implementation of modern statistical methods, and the area of nonparametrics is no exception. We feature four statistical packages in this book: Resampling Stats, StatXact, S-Plus, and MINITAB.

Resampling Stats

Resampling Stats comes in three versions: a stand-alone package, an add-in to Microsoft Excel, and an add-in to Matlab. The stand-alone version is a simple-to-use programming language that is specifically designed to do the type of re-

sampling of data needed for permutation tests and bootstrap sampling. The programming language has functions that enable programmers to construct a variety of statistical procedures. With intuitive commands such as "shuffle," "sample," "repeat," and "score," Resampling Stats code may also serve as pseudocode for describing the steps needed to program nonparametric procedures in other languages. The add-ins are designed to work in conjunction with the features of the host software. Most of the methods in Chapters 2–7 can be done with this software. If students have the option of only one statistical software package to use with this book, Resampling Stats is the one to choose. The web address is www.resample.com.

StatXact

StatXact is a powerful, simple-to-use, menu-driven program that has more than 80 nonparametric procedures for continuous and categorical data. The software also has data manipulation capabilities. Power and sample size computations are available for tests for one and two binomial populations. Most of the procedures in Chapters 2–5 and Chapter 7 are included in this software package, as well as a number of procedures that we do not cover. Those who make use of exact nonparametric procedures in their statistical consulting and research will find this program extremely useful. The web address is www.cytel.com.

S-Plus

S-Plus is a powerful programming language that can be used to carry out the procedures in this book. We limit our use to those procedures that can be accessed through the S-Plus menu options. This program is especially useful for the smoothing techniques and robust methods discussed in Chapter 10. The web address is www.insightful.com.

MINITAB

MINITAB has long been a favorite of statisticians for both instruction and consulting. Some of the standard statistical procedures in Chapters 1–4 and Chapter 7 are included in the MINITAB menu. Data manipulation capabilities are included. MINITAB also has commands for carrying out rank-based regression as discussed in Section 10.3, although these are not presently documented in the help menu. The web address is www.minitab.com.

Other

The SAS® programming language, a standard for applied statisticians, has several capabilities for nonparametric statistics. PROC NPAR1WAY has a variety of two-sample and k-sample tests, and there is an option for doing exact tests. PROC FREQ has an option to do exact tests for contingency tables. PROC MULTTEST allows

bootstrap and permutation sampling and may be applied in the multivariate setting. The web address is www.sas.com.

One other source of software that we have used is at a web site created by J. W. McKean. Here one can carry out rank-based regression and analysis of variance online. The web address is www.stat.wmich.edu/slab/RGLM.

Most sections of the book have a subsection entitled "Computer Analysis" that outlines the use of selected statistical packages. We do not intend this to be a comprehensive discussion, but rather it is to give an indication of the capabilities of various software packages.

Acknowledgments

I would like to thank the reviewers of this book for their valuable suggestions and comments. They are: R. Clifford Blair, University of South Florida; Joshua D. Naranjo, Western Michigan University; and Thomas H. Short, Villanova University. I would also like to thank Scott J. Richter, University of North Carolina at Greensboro, who taught from the original manuscript and provided valuable input.

James J. Higgins

THOMSON
BROOKS/COLE

Editor: *Carolyn Crockett*
Assistant Editor: *Ann Day*
Editorial Assistant: *Julie Bliss*
Technology Project Manager: *Burke Taft*
Marketing Manager: *Joseph Rogrove*
Project Manager, Editorial Production: *Ellen Brownstein*
Print/Media Buyer: *Barbara Britton*

Permissions Editor: *Sue Ewing*
Production Service: *Scratchgravel Publishing Services*
Copy Editor: *Carol Reitz*
Cover Designer: *Vernon Boes*
Compositor: *Scratchgravel Publishing Services*
Printer: *Phoenix Color Corp.*

COPYRIGHT © 2004 Brooks/Cole, a division of Thomson Learning, Inc. Thomson Learning™ is a trademark used herein under license.

ALL RIGHTS RESERVED. No part of this work covered by the copyright hereon may be reproduced or used in any form or by any means—graphic, electronic, or mechanical, including but not limited to photocopying, recording, taping, Web distribution, information networks, or information storage and retrieval systems—without the written permission of the publisher.

Printed in the United States of America

1 2 3 4 5 6 7 08 07 06 05 04 03

For more information about our products, contact us at:

Thomson Learning Academic Resource Center
1-800-423-0563

For permission to use material from this text, contact us by:

Phone: 1-800-730-2214
Fax: 1-800-730-2215
Web: http://www.thomsonrights.com

Library of Congress Control Number: 2002117217

ISBN 0-534-38775-6

Brooks/Cole–Thomson Learning
511 Forest Lodge Road
Pacific Grove, CA 93950
USA

Asia
Thomson Learning
5 Shenton Way #01-01
IC Building
Singapore 068808

Australia/New Zealand
Thomson Learning
102 Dodds Street
Southbank, Victoria 3006
Australia

Canada
Nelson
1120 Birchmount Road
Toronto, Ontario M1K 5G4
Canada

Europe/Middle East/Africa
Thomson Learning
High Holborn House
50/51 Bedford Row
London WC1R 4LR
United Kingdom

Latin America
Thomson Learning
Seneca, 53
Colonia Polanco
11560 Mexico D.F.
Mexico

Spain/Portugal
Paraninfo
Calle Magallanes, 25
28015 Madrid
Spain

此为试读,需要完整PDF请访问: www.ertongbook.com

Duxbury Titles of Related Interest

Daniel, *Applied Nonparametric Statistics,* 2nd ed.

Derr, *Statistical Consulting: A Guide to Effective Communication*

Durrett, *Probability: Theory and Examples,* 2nd ed.

Graybill, *Theory and Application of the Linear Model*

Johnson, *Applied Multivariate Methods for Data Analysts*

Kuehl, *Design of Experiments: Statistical Principles of Research Design and Analysis,* 2nd ed.

Larsen, Marx, & Cooil, *Statistics for Applied Problem Solving and Decision Making*

Lohr, *Sampling: Design and Analysis*

Lunneborg, *Data Analysis by Resampling: Concepts and Applications*

Minh, *Applied Probability Models*

Minitab, Inc., *MINITAB™ Student Version 12 for Windows*

Myers, *Classical and Modern Regression with Applications,* 2nd ed.

Newton & Harvill, *StatConcepts: A Visual Tour of Statistical Ideas*

Ramsey & Schafer, *The Statistical Sleuth,* 2nd ed.

SAS Institute Inc., *JMP-IN: Statistical Discovery Software*

Savage, *INSIGHT: Business Analysis Software for Microsoft® Excel*

Scheaffer, Mendenhall, & Ott, *Elementary Survey Sampling,* 5th ed.

Shapiro, *Modeling the Supply Chain*

Winston, *Simulation Modeling Using @RISK*

To order copies, contact your local bookstore or call 1-800-354-9706. For more information, contact Duxbury Press at 10 Davis Drive, Belmont, CA 94002, or go to www.duxbury.com.

目录

0 预备知识 1
0.1 累积分布和概率密度函数 1
0.2 常用连续概率分布 1
0.3 Bernoulli分布 4
0.4 置信区间及假设检验 5
0.5 参数方法及非参数方法的比较 7
0.6 非参数方法的分类 7

1 单个样本方法 11
1.1 关于中位数的非参数假设检验和置信区间 11
1.1.1 Bernoulli检验 11
1.1.2 置信区间 13
1.1.3 计算机分析 14
1.2 对总体累积分布函数及百分位数的估计 14
1.2.1 总体累积分布函数的置信区间 14
1.2.2 关于百分位数的推断 16
1.2.3 计算机分析 17
1.3 统计检验的比较 17
1.3.1 第一类错误 18
1.3.2 势 18
1.3.3 公式推导 20
练习 21

2 两样本方法 23
2.1 两样本置换检验 23
2.1.1 置换检验 24
2.1.2 两样本置换检验的步骤 27

2.1.3 两样本置换检验的假设 28
2.1.4 计算机分析 29
2.2 基于中位数及修整均值的置换检验 30
2.2.1 基于中位数的置换检验 30
2.2.2 修整均值 31
2.3 置换的随机抽样 31
2.3.1 置换随机抽样的近似 p-值 32
2.3.2 应用Resampling Stats的计算机分析 33
2.3.3 应用StatXact 的计算机分析 35
2.4 Wilcoxon秩和检验 35
2.4.1 Wilcoxon秩和检验的步骤 37
2.4.2 关于Wilcoxon秩和检验的评论 38
2.4.3 Wilcoxon秩和检验的统计表 38
2.4.4 计算机分析 39
2.5 在存在结点时调整的Wilcoxon秩和检验 41
2.5.1 在存在结点时调整的Wilcoxon秩和检验步骤 41
2.5.2 计算机分析 42
2.6 Mann-Whitney 检验及其置信区间 43
2.6.1 Mann-Whitney统计量 43
2.6.2 Mann-Whitney统计量及Wilcoxon 秩和检验统计量的等价性 44
2.6.3 移位参数的置信区间及Hodges-Lehmann统计量 45
2.6.4 计算机分析 48
2.7 记分系统 49
2.7.1 三个常用记分系统 49
2.7.2 计算机分析 51
2.8 尺度等同检验及Omnibus检验 51
2.8.1 Siegel-Tukey 及 Ansari-Bradley 检验 52
2.8.2 偏差的检验 53
2.8.3 Kolmogorov-Smirnov 检验 57
2.8.4 计算机分析 58
2.9 两样本检验方法的选择 58
2.9.1 t-检验 60
2.9.2 Wilcoxon 秩和检验与t-检验的比较 61
2.9.3 相对效率 62
2.9.4 置换检验的势 63
2.10 大样本近似 65
2.10.1 抽样公式 65
2.10.2 在Wilcoxon秩和检验中的应用 66

2.10.3 存在结点时的Wilcoxon 秩和检验 68
2.10.4 $E(W_{ties})$ 及 $\text{var}(W_{ties})$ 的计算公式 70
2.10.5 基于Mann-Whitney 检验的大样本置信区间 71
2.10.6 置换分布的正态近似 72
练习 73

3 *K*-样本方法 79
3.1 *K*-样本置换检验 79
3.1.1 *F*统计量 80
3.1.2 置换*F*-检验的步骤 81
3.1.3 置换*F*-检验统计量的替代形式 83
3.1.4 计算机分析 84
3.2 Kruskal-Wallis检验 86
3.2.1 Kruskal-Wallis 统计量 86
3.2.2 打结时的调整 88
3.2.3 KW的卡方近似的直观导出 90
3.2.4 一般记分检验 91
3.2.5 计算机分析 91
3.3 多重比较 92
3.3.1 假定数据无结点时为控制试验误差率的三种秩方法 92
3.3.2 一般记分系统的多重比较（存在结点） 96
3.3.3 多重比较置换检验 96
3.3.4 从有限总体抽样时均值之差的方差 98
3.3.5 计算机分析 99
3.4 有序备择假设 99
3.4.1 Jonckheere-Terpstra 检验 101
3.4.2 大样本近似 101
3.4.3 计算机分析 103
练习 105

4 成对比较及区组设计 109
4.1 成对比较置换检验 109
4.1.1 成对比较置换检验的步骤 110
4.1.2 随机选择置换 113
4.1.3 大样本近似 113
4.1.4 对称总体中位数的检验 114
4.1.5 计算机分析 114
4.2 符号秩检验 117

4.2.1 数据无结点时的Wilcoxon符号秩检验 118
4.2.2 大样本近似 118
4.2.3 存在结点时的调整 120
4.2.4 计算机分析 124
4.3 其他成对比较检验 125
4.3.1 符号检验 125
4.3.2 一般记分系统 126
4.3.3 对成对比较检验方法的选择 127
4.4 随机化完全区组设计的置换检验 127
4.4.1 随机化完全区组设计的F检验统计量、128
4.4.2 随机化完全区组设计的置换F检验 129
4.4.3 多重比较 130
4.4.4 计算机分析 131
4.5 随机化完全区组设计的Friedman检验 132
4.5.1 无结点时的Friedman检验 133
4.5.2 有结点时的调整 134
4.5.3 Cochran's Q及Kendall's W 136
4.5.4 FM或FM_{ties}的卡方近似 136
4.5.5 计算机分析 137
4.6 随机化完全区组设计的有序备择假设 138
4.6.1 Page检验 139
4.6.2 计算机分析 140
练习 141

5 趋势及关联的检验 145
5.1 相关性及斜率的置换检验 145
5.1.1 相关系数 145
5.1.2 最小平方直线的斜率 146
5.1.3 置换检验 147
5.1.4 r的置换分布的大样本近似 151
5.1.5 计算机分析 152
5.2 Spearman秩相关 153
5.2.1 Spearman秩相关的统计检验 154
5.2.2 大样本近似 154
5.2.3 计算机分析 158
5.3 Kendall τ 158
5.3.1 数据无结点时对Kendall τ的估计 159
5.3.2 大样本近似 160

5.3.3 数据有结点时的调整 161
5.3.4 计算机分析 163
5.4 列联表的置换检验 163
5.4.1 待检验的假设及卡方统计量 163
5.4.2 置换卡方检验 165
5.4.3 列联表的多重比较 167
5.4.4 计算机分析 169
5.5 2×2列联表的Fisher精确检验 172
5.5.1 零假设下的概率分布 173
5.5.2 计算机分析 175
5.6 有序分类变量的列联表 176
5.6.1 单有序表 176
5.6.2 双有序表 178
5.6.3 计算机分析 179
5.7 Mantel-Haenszel 检验 179
5.7.1 待检验的假设 179
5.7.2 分层列联表的假设检验 181
5.7.3 优比的估计 183
5.7.4 计算机分析 184
5.8 McNemar检验 186
5.8.1 符号及假设 187
5.8.2 检验统计量及置换分布 187
5.8.3 计算机分析 188
练习 189

6 多变量检验 195
6.1 两样本的多变量置换检验 195
6.1.1 符号及假设 196
6.1.2 Hotelling T_2 的置换检验形式 197
6.1.3 其他多变量统计量 198
6.1.4 计算机分析 200
6.2 两样本多元秩检验 202
6.2.1 Wilcoxon 统计量的概括 203
6.2.2 计算机分析 205
6.3 多元成对比较 206
6.3.1 基于T_2的置换检验 207
6.3.2 置换的随机抽样 207
6.3.3 其他多元统计量 209

此为试读,需要完整PDF请访问: www.ertongbook.com