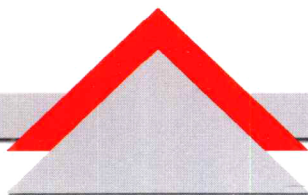




全国统计教材编审委员会“十一五”规划教材

统计实验教材系列 许涤龙 主编

统计学基础实验



★ 许涤龙 谭朵朵 沈春华 编著



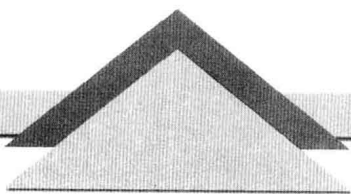
中国统计出版社
China Statistics Press



全国统计教材编审委员会“十一五”规划教材

统计实验教材系列 许涤龙 主编

统计学基础实验



★ 许涤龙 谭朵朵 沈春华 编著



中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

统计学基础实验 / 许涤龙, 谭朵朵, 沈春华编. --
北京: 中国统计出版社, 2010.8
ISBN 978-7-5037-6026-6

I. ①统… II. ①许… ②谭… ③沈… III. ①统计学
—实验 IV. ①C8-33

中国版本图书馆 CIP 数据核字(2010)第 149327 号

统计学基础实验

作 者/许涤龙 谭朵朵 沈春华
责任编辑/钟 钰
装帧设计/黄 晨
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 57 号 邮政编码/100826
办公地址/北京市丰台区西三环南路甲 6 号
网 址/www.stats.gov.cn/tjshujia
电 话/邮购(010)63376907 书店(010)68783172
印 刷/河北天普润印刷厂
经 销/新华书店
开 本/710×1000mm 1/18
字 数/280 千字
印 张/16.5
印 数/1—3000 册
版 别/2010 年 9 月第 1 版
版 次/2010 年 9 月第 1 次印刷
书 号/ISBN 978-7-5037-6026-6/C·2386
定 价/30.00 元

中国统计版图书, 版权所有。侵权必究。

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

“十一五”时期是继续深化教育改革、加强素质教育、努力建设有利于创新型科技人才生长的教育培训体系的关键时期。为了更好地培育统计创新型科技人才,适应统计教育培训的新形势,全国统计教材编审委员会制定了《“十一五”全国统计教材建设规划》(以下简称《规划》)。规划坚持“以人为本”的科学发展观,坚持统计教育与实践相结合,坚持统计教育同国际接轨,坚持培养创新型的统计人才的指导思想,编写符合国民经济发展需要和统计事业发展需要的统计教材。

这批教材是在深入分析统计教育形势和统计教材建设发展状况,总结多年来统计教材建设经验的基础上,本着以建设本科统计教材为主的方针,积极探索研究生层次的统计教材,力争使规划统计教材的编写做到层次分明,有针对性和实用性。建设精品教材,是编委会自成立以来就孜孜以求的目标。考虑到统计教材建设的实际情况,“十一五”期间,本科教材主要以修订为主,对以往规划统计教材中使用面广、得到广大教师和学生普遍认可的教材组织了修订。修订后的教材,淘汰了过时的内容和例子,增加了计算机操作和大量的案例,编写手法也做了一定的调整,在实用性、可操作性等方面有了较大的改进。

近年来,我国现代化建设快速发展,高等教育规模持续扩大,尤其是研究生教育规模的扩大,使得高等学校研究生统计教学工作面临着许多新情况、新问题,任务艰巨。因此,必须坚持科学发展观,在规模持续发展的同时,把提高研究生统计教学质量放在突出的位置,培养全面发展的创新型的统计人才。教材是统计教学的载体,建设高质量

的研究生层次的统计教材是统计教育发展的需要。因此,编委会在“十一五”期间对研究生的统计基础课教材编写做了些有益的探索。根据《规划》的要求,这批教材主要采取招标和邀请的方式组织有关院校的专家、学者编写。

值得特别提出的是,在这批教材中,有《非参数统计》、《概率论与数理统计》、《经济计量学教程》、《医学统计》、《应用时间序列分析》、《多元统计分析》、《统计学》、《现代指数理论》、《现代金融投资统计分析》9部教材入选国家教育部组织编写的“普通高等教育‘十一五’国家级规划教材”,更加充实和完善了“十一五”期间统计教材的建设。

为了便于教学和学习,这批教材里面包含了与之相配套的《学习指导与习题》,使得这批教材在编辑出版上形成了比较完整的体系。我们相信,这批教材的出版和发行,对于推动我国统计教育改革,加快我国统计教材体系和教材内容更新、改造的步伐,打造精品教材,都将起到积极的促进作用。

限于水平和经验,这批教材的编审、出版工作还会有缺点和不足,诚恳欢迎教材的使用单位、广大教师 and 同学们提出批评和建议。

全国统计教材编审委员会
2006年6月

总 序

随着我国经济、社会和科技的发展,统计应用的面越来越广,如何用好统计工具,成为人才培养、科学研究和实际工作中面临的共同问题。而科学地应用统计工具,至少要解决好三个问题:一是用好统计数据,即了解所要采用的统计数据的性质和特点,做到“用数据说话”;二是用好统计方法,即能够根据统计数据的性质和特点,采用恰当的统计方法对之进行分析研究,使得“用数据说话”说得恰当而有力;三是用好统计软件,即在现代信息技术条件下,能够运用统计软件对大量数据进行处理和分析,提高分析研究的精度与效率。在现代科技条件下,上述三个方面是统计工具应用的有机组成部分,三者必须相互结合、不可分割。而统计实验教学,就是上述三者相互结合的一种新的教学模式,它针对所要研究解决的实际问题,将相应的统计数据、统计方法和统计软件相结合,在专门实验室内完成统计过程的模拟、统计知识的学习和统计技能的培养,使学习者能够在模拟实践的环境下,掌握统计工具应用技能,提高统计实践能力。

传统的统计教学与通常的社会科学类专业教学一样,基本上属于理论教学模式。我们所倡导的实验教学模式与理论教学模式有很大的不同:第一,教学场所不同,理论教学是在教室进行,实验教学是在实验室进行(在具备条件的单台计算机上亦可进行个别学习);第二,教学工具不同,与一般教室的教学工具配给不同,实施统计实验教学的实验室至少需要配备计算机、统计软件及相关配套设施,并具备良好的网络运行环境;第三,教学方式不同,理论教学通常是教师讲、学生听、偶尔有交流,实验教学要求

学生与教师一起在计算机上进行操作,学习的自主性更强;第四,教学内容不同,与理论教学比较,实验教学使学生在统计理论与方法的同时也学习了统计软件应用技术,并且能够进行更复杂的统计方法的学习和更大规模的统计数据的处理;第五,教学效率不同,由于节省了板书时间,实验教学在单位时间内比理论教学能够提供更大的信息量,讲授更多的内容;第六,教学效果不同,实验教学是在模拟实践环境和遇到实际问题时通过实验项目来完成,一般情况下比理论教学更有利于培养学生的研究能力、创新能力和实践能力。当然,我们倡导实验教学,并不是要用实验教学完全取代理论教学,而是将统计课程中能够实验化的部分(通常就是统计方法及其应用问题)予以实验化,即:将之设计成为实验项目来实施教学,它既是统计教学模式的创新,也是对统计理论教学的延伸和发展。对于大多数统计课程来说,并不是所有的教学内容都可以实验化,因而实验教学应当与理论教学相结合。

然而,目前我国高校统计教学工作中,理论讲授过多、实际操作过少、缺少实验教学环节的问题仍然比较突出,统计实验教材少,而与相关课程理论教学相配套的实验教材更少。为此,我们组织编写了本套“统计实验教材系列”,其设计和编写的基本思路是:

一、按照统计专业主干课程设计实验教材。为了更好地体现统计理论教学与实验教学相结合并便于组织教学实施,我们基本上根据现阶段统计专业的主干课程来设计实验教材的种类,使得各门专业主干课程都有实验教材与之相配套。初步计划,系列实验教材包括八种,即:《统计学基础实验》、《数理统计学实验》、《计量经济学实验》、《多元统计分析实验》、《时间序列分析实验》、《数据挖掘实验》、《市场调查与分析实验》和《经济预测与决策实验》。

二、各种实验教材的内容由系列实验项目组成。实验项目是根据课程教学内容进行实验设计而成,一般包括三类:一是演示性实验,通常用来介绍所选用的统计软件,包括统计软件运行的硬件条件与环境;二是设计性实验,用

来展示课程中各种统计方法的应用技术及其实现过程,一般按课程的主要统计方法类别分别设计实验项目,是一门课程实验项目的主要形式;三是综合性实验,属于一门课程各种统计方法综合运用的大型实验项目,主要培养学生综合运用各种统计方法分析和解决实际问题的能力。

三、实验项目的设计体现教学规律的要求。实验教材包括的各个实验项目,大多都是编著教师在长期的教学实践中经过不断探索、反复试验而总结形成的,不仅在内容和形式上具有创新性,而且力求体现教学规律的要求。一个实验项目按照教学程序大致包括以下几个部分内容:(1)实验目的,概述一个实验项目设计的意图与思路,应达到的目的与要求;(2)实验原理,简要介绍实验项目所应用的理论、方法与模型;(3)实验数据,给出实验项目要用到的数据;(4)实验过程,具体说明实验项目操作的环节、步骤及其要求,是实验项目的主体;(5)实验小结,归纳实验的要点,给出类推实验和相关提示;(6)练习实验,给出同类或相关方法实验的练习题。在课时设计上,一般一个实验项目对应2个课时,大型综合性实验项目可安排4个课时。

由此可见,本套教材是统计教学模式改革和统计实验教学建设的结果。

本套教材的编写和出版,各位编著教师付出了艰辛的努力,全国统计教材编审委员会、中国统计出版社和国家统计局统计教育中心给予了大力的支持,并得到湖南省普通高等学校教学改革研究项目《经济管理类专业实验教学模式的创新与应用》的支持,在此一并致以诚挚的感谢!

统计实验教学及其教材编写在我国还是一项具有探索性的工作,加之编者的学识水平和研究能力有限,教材中的问题和错漏在所难免,敬请各位专家和读者批评指正,也祈盼有关教学研究和教材编写工作能够不断推进。

许涤龙

2009年3月

前言

统计学是一门应用性、操作性极强的方法论科学。随着我国经济的快速发展,统计的应用范围越来越广泛,社会不但需要大量的统计专业人才,也要求非统计专业人才尤其是经济管理类专业人才能够正确的使用统计数据、基本的统计方法和常用的统计软件,并具备一定的分析问题和解决问题的综合能力。统计实验教学是将统计数据、统计方法、统计软件和实际问题相结合,是统计理论学习的继续、补充和发展。

在统计学基础课程的教学中,我们发现,如果缺乏实验教学的支撑,学生对于统计学的学习不但仅停留在对基础理论知识的掌握,而且对实际统计数据不会处理,缺乏运用统计软件对实际问题具体分析的能力。《统计学基础实验》作为基础课程或核心课程“统计学”(或“统计学基础”、“统计学导论”、“统计学原理”)的实验教材,其编写目的在于,在讲授统计学基础理论和基本方法的同时,组织学生做同步的实验,通过实验操作,使学生加深对基础理论和方法的理解,掌握统计方法的实际应用,培养学生的创新能力。

本书可作为高等院校统计、经济、管理、人文及其他社会科学、自然科学类专业的本科生学习统计学入门课程的配套教材,也可以作为相关专业的研究生和广大实际工作者学习相关课程的参考用书。

本书由许涤龙教授负责提出写作构思,设计全书的主体框架;由许涤龙、谭朵朵、沈春华编著。全书共包括八个实验项目,许涤龙负责项目1和项目7的编写;沈春华负责

项目 2、项目 3 和项目 8 的编写；谭朵朵负责项目 4、项目 5 和项目 6 的编写。在编写中，候鹏、周光洪、谢敏、吴朝平和王仁花对数据资料的搜集、整理做了大量的工作，并参与撰写了部分实验项目的初稿。本书中的实验数据可在 <http://www.stats.gov.cn/tjshujia/tjjc> 下载或查询。

限于编者的知识和水平，书中的错误及不当之处在所难免，恳请广大读者不吝批评指正。

本书在编著过程中，参考了许多专家、学者的有关教材、专著、论文，并引用了部分资料，从中获益匪浅。同时，本书的出版还得到了中国统计出版社的大力支持，在此一并表示衷心的感谢。

编者

2010 年 3 月

目 录

项目 1	数据文件的建立与管理	1
1.1	实验目的	1
1.2	实验原理	1
1.2.1	SPSS 软件概述	1
1.2.2	SPSS 的运行方式	3
1.2.3	SPSS 中的两个常用基本窗口	4
1.3	实验数据	7
1.4	实验过程	8
1.4.1	SPSS 软件的安装	8
1.4.2	SPSS 的启动和退出	10
1.4.3	SPSS 数据文件的创建	12
1.4.4	数据的管理	22
1.5	实验小结	49
1.6	练习实验	50
项目 2	描述性统计分析	53
2.1	实验目的	53
2.2	实验原理	53
2.2.1	统计表	53
2.2.2	统计图	54
2.2.3	基本描述统计指标	55
2.4	实验过程	61
2.5	实验小结	78
2.6	练习实验	79
项目 3	列联分析	81
3.1	实验目的	81
3.2	实验原理	81
3.2.1	列联表的构造	81
3.2.2	χ^2 分布与 χ^2 检验	83
3.2.3	品质相关程度的测度	85
3.3	实验数据	87
3.4	实验过程	87
3.4.1	各检验方法在 SPSS 中的实现	87

目 录

3.4.2 实例分析	94
3.5 实验小结	98
3.6 练习实验	98
项目 4 相关与回归分析	101
4.1 实验目的	101
4.2 实验原理	101
4.2.1 相关分析	101
4.2.2 回归分析	103
4.3 实验数据	104
4.4 实验过程	107
4.4.1 相关分析	107
4.4.2 回归分析	113
4.5 实验小结	141
4.6 练习实验	141
项目 5 时间序列分析	143
5.1 实验目的	143
5.2 实验原理	143
5.2.1 时间数列的分析指标	143
5.2.2 长期趋势分析	147
5.2.3 季节变动分析	150
5.3 实验数据	152
5.4 实验过程	153
5.4.1 时间数列的分析指标	153
5.4.2 长期趋势分析	156
5.4.3 季节变动分析	167
5.5 实验小结	176
5.6 练习实验	176
项目 6 指数分析	178
6.1 实验目的	178
6.2 实验原理	178
6.2.1 指数的编制	179
6.2.2 指数因素分析	181
6.2.3 指数数列的短期预测	185

目 录

6.3	实验数据	187
6.4	实验过程	188
6.4.1	综合指数的编制	188
6.4.2	平均数指数的编制	190
6.4.3	总量指标变动的指数因素分析	191
6.4.4	平均指标变动的指数因素分析	192
6.4.5	指数的短期预测	193
6.5	实验小结	202
6.6	练习实验	203
项目 7	抽样推断	204
7.1	实验目的	204
7.2	实验原理	204
7.2.1	随机数的产生	204
7.2.2	总体平均数的推断	206
7.2.3	总体成数的推断	208
7.3	实验数据	210
7.4	实验过程	211
7.4.1	随机数的产生	211
7.4.2	总体平均数的推断	215
7.5	实验小结	223
7.6	练习实验	223
项目 8	综合评价	225
8.1	实验目的	225
8.2	实验原理	225
8.2.1	综合评价	225
8.2.2	模糊综合评价	231
8.3	实验数据	237
8.4	实验过程	239
8.4.1	实验的基本过程	239
8.5	实验小结	244
8.6	练习实验	245
参考文献		247

项目 1

数据文件的建立与管理

1.1 实验目的

统计学是一门旨在探索数据内在数量规律的方法论科学,创建数据文件并进行管理是统计分析的基础和前提。SPSS 软件在数据准备、整理及统计分析方面具有强大的功能,本实验项目的主要目的是使学生熟悉并掌握运用 SPSS 进行数据文件的建立和管理。其内容主要包括:

1. 了解 SPSS 软件的基本特点、模块式结构、SPSS 软件的安装及运行和退出等基础知识,学会 SPSS 的入门操作,方便以后对该软件的运用。
2. 运用 SPSS 软件创建数据文件,包括数据的直接录入以及对 EXCEL 形式的已有数据文件进行导入。
3. 学会根据统计分析目的和要求的不同对数据进一步的加工和管理,包括数据的合并、拆分、转置和汇总。

1.2 实验原理

1.2.1 SPSS 软件概述

SPSS 是软件英文名称(Statistical Package for the Social Sciences,即“社会科学统计软件包”)的首字母缩写,但是随着 SPSS 产品服务领域的扩大和服务深度的增加,SPSS 公司已于 2000 年正式将英文全称更改为 Statistical Product and Service Solutions,意为“统计产品与服务解决方案”,标志着 SPSS 的战略方向正在做出重大调整。

SPSS 是由美国斯坦福大学的三位研究生于 20 世纪 60 年代末研制的,并于 1975 年创立了 SPSS 公司。SPSS 是世界上最早的统计分析软件,迄今已有 30 余年的成年历史,全球约有 25 万家产品用户,它们分布于通讯、医疗、银行、证券、保险、制造、商业、市场研究、科研教育等多个领域和行业,是世界上应用最广泛的专业统计软件。SPSS 的基本功能包括数据管理、统计分析、图表分析、输出管理等等。SPSS 统计分析过程包括描述性统计、均值比较、一般线性模型、相关分析、回归分析、对数线性模型、聚类分析、数据简化、生存分析、时间序列分析、多重响应等几大类,每类中又分好几个统计过程,比如回归分析中又分线性回归分析、曲线估计、Logistic 回归、Probit 回归、加权估计、两阶段最小二乘法、非线性回归等多个统计过程,而且每个过程中又允许用户选择不同的方法及参数。SPSS 也有专门的绘图系统,可以根据数据绘制各种图形。正是由于 SPSS 的强大功能和广泛应用,使得 SPSS 成为国际上最有影响力的三大统计软件之一。

SPSS 软件具有如下一些特点:

1. 操作简单,无需编程

在使用 SPSS 的过程中,除了数据录入及部分命令程序等少数输入工作需要键盘键入外,大多数操作可通过“菜单”、“按钮”和“对话框”来完成。SPSS 具有第四代语言的特点,告诉系统要做什么,无需告诉怎样做。只要了解统计分析的原理,无需通晓统计方法的各种算法,即可得到需要的统计分析结果。对于常见的统计方法,SPSS 的命令语句、子命令及选择项的选择绝大部分由“对话框”的操作完成。因此,用户无需花大量时间记忆大量的命令、过程、选择项。另外,SPSS 也向一些高级用户提供了编程功能。

2. 功能强大

具有完整的数据输入、编辑、统计分析、报表、图形制作等功能。自带 11 种类型 136 个函数,为统计分析用户提供了全方位的统计学算法。SPSS 提供了从简单的统计描述到复杂的多因素统计分析方法,囊括了各种成熟的统计方法与模型,为各种研究提供了相应的统计学方法。比如数据的探索性分析、统计描述、列联表分析、二维相关、秩相关、偏相关、方差分析、非参数检验、多元回归、生存分析、协方差分析、判别分析、因子分析、聚类分析、非线性回归、Logistic 回归等,方法体系覆盖全面。SPSS 还提供了各种数据准备与数据整理技术。如利用值标签来快捷地录入数据,从而为数据审核与分析提供了便利条件;生产新的变量,对连续性变量进行离散性转换;将几个小类别合并为一个类别等。利用 SPSS 强大的数据整理技术,可使数据结构、内容

更易于分析。此外,SPSS 还提供了自由灵活的表格功能以及各种统计学图形,并且可将表格和图形直接拷贝到 Word 文档、幻灯片中,直接进行结构的展现。

3. 方便的数据接口

SPSS 能够读取及输出多种格式的文件。比如由 dBASE、FoxBASE、FoxPRO 产生的 *.dbf 文件,文本编辑器软件生成的 ASCII 数据文件,Excel 的 *.xls 文件等均可转换成可供分析的 SPSS 数据文件。这不仅节省了相当大的工作量,并且避免了因拷贝粘贴可能引起的错误。SPSS 的表格、图形结构可直接导出为 Word、文本、网页、Excel 格式等,也可以将表格、图形作为对象选择粘贴到 Word、PowerPoint 等中,并在其中利用 SPSS 对它们进行编辑。

4. 灵活的功能模块组合

SPSS for Windows 软件分为若干功能模块。用户可以根据自己的分析需要和计算机的实际配置情况灵活选择。

1.2.2 SPSS 的运行方式

SPSS 的运行方式主要有三种:完全窗口菜单运行方式、程序运行方式和混合运行方式。

1. 完全窗口菜单运行方式

完全窗口菜单运行方式是指用户根据要运行的任务,在窗口中选择菜单,在下拉菜单中选择要运行的命令,在打开的各级对话框中选择要设置的参数或者要分析的变量,直至所有的对话框中的选择完毕,点击“OK”按钮交付系统运行。此种方式可完成 SPSS 大部分的统计分析,无需编辑且较少需要输入,简单易用。本书以介绍这种方式为主。

2. 程序运行方式

程序运行方式是在 Syntax(语句)窗口中直接运行编写好的程序的一种方式或者在脚本窗口(Script)中运行脚本过程。在 Syntax 窗口中输入 SPSS 命令组成的程序,利用键盘或主菜单中的 Edit 菜单项对窗口中的程序进行修改、编辑。在 Syntax 窗中的程序可以分析数据窗口中的数据,也可以用有关的语句指定外部数据文件,对其进行分析,分析结果仍然是显示在结果管理窗口中。这种方式对一些高级用户来说仍有用武之地。

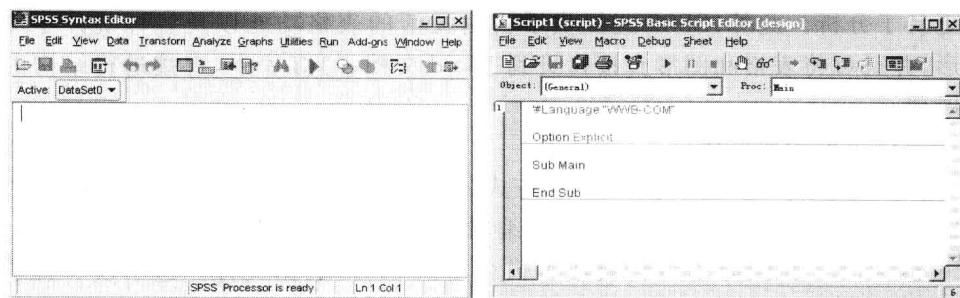


图 1.1 Syntax 和 Script 窗口

3. 混合运行方式

混合运行方式是上述两种方式的结合,首先在数据窗口中输入数据或利用 File 菜单项打开已经存在的数据文件,然后利用对话框选择分析过程和分析参数。选择完成后,不是按“OK”按钮提交系统立即运行,而是按“Paste”按钮将整个操作过程转换成相应的 SPSS 命令语句,粘贴到 Syntax 窗口中。在该窗口中增加对话框中没有包括的语句和参数,或修改子命令中的参数,然后按窗口中的“Run”功能按钮,将命令提交系统执行。结果显示在结果管理窗口中。混合运行方式,既能简化操作,又可以补充单纯窗口运行方式的不足。

1.2.3 SPSS 中的两个常用基本窗口

SPSS 运行时使用的窗口种类有:数据编辑窗口、结果管理窗口、语法编辑窗口、脚本窗口和草稿结果窗口。其中,数据编辑窗口和结果管理窗口是 SPSS 两个最常用的窗口,下面分别对这两个窗口进行简单介绍。

1. SPSS 数据编辑窗口(SPSS Data Editor)

SPSS 的数据编辑窗口同 Excel 窗口类似,并且还有些相同的功能,但 SPSS 的数据统计功能要比 Excel 强得多。SPSS 处理数据的主要工作全在该窗口中进行,该窗口由标题栏、菜单栏、工具栏、编辑栏、变量名称栏、内容区、窗口切换标签页和状态栏组成。

标题栏中显示编辑的数据文件的名称。

菜单栏上有 11 个菜单选项,下面介绍常用的几个菜单项:

File:可以新建、打开、导入/导出、保存数据文件,还可以进行对数据文件重命名等操作,通过该菜单也可以退出 SPSS 软件。

Edit:是 SPSS 文件的编辑菜单,可改变 SPSS 的默认设置,也可以插入变量或者观测个案。