

高等院校统计学实验课系列教程

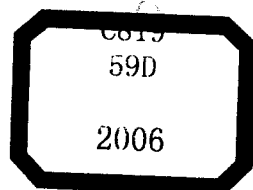
SAS应用统计实验

邵建利 主 编
任 琳 副主编

SAS YINGYONG TONGJI SHIYAN

上海财经大学出版社

高等院校统计学实验课系列教程



SAS 应用统计实验

邵建利 主 编
任 琳 副主编

 上海财经大学出版社

图书在版编目(CIP)数据

SAS应用统计实验/邵建利主编,任琳副主编. —上海:上海财经大学出版社,2006.10

高等院校统计学实验课系列教程

ISBN 7-81098-729-1/F·664

I. S… II. ①邵…②任… III. 统计分析-应用软件, SAS-高等学校-教材 IV. C812

中国版本图书馆 CIP 数据核字(2006)第 120106 号

☐ 责任编辑 刘光本

☐ 封面设计 周卫民

SAS YINGYONG TONGJI SHIYAN

SAS 应用统计实验

邵建利 主 编

任 琳 副主编

上海财经大学出版社出版发行
(上海市武东路 321 号乙 邮编 200434)

网 址: <http://www.sufep.com>

电子邮箱: webmaster@sufep.com

全国新华书店经销

上海市印刷十厂印刷

上海叶大装订厂装订

2006 年 10 月第 1 版 2006 年 10 月第 1 次印刷

700mm×960mm 1/16 11.75 印张 204 千字
印数: 0 001—3 000 定价: 19.00 元(本书含光盘 1 张)

前 言

SAS 是一个庞大的软件系统,它具有“信息系统”和“计算”两种应用属性。本书的实验内容主要是针对 SAS 的统计计算属性,其适用对象是高等院校的本科生、研究生,以及与此相关的大学教师和科研人员。

据统计,目前全球发表的经济与管理学术文献中,按研究方法出现频次排名,统计学方法位列第一,其次是运筹学。本书编写的目的在于通过实验,培养学生应用统计学知识的能力,同时提高他们的实践和创新能力。

SAS 涉及统计方法应用的模块主要有 SAS/BASE、SAS/STAT、SAS/GRAPH、SAS/ETS、SAS/ASSIST、SAS/IML、SAS/INSIGHT、SAS 分析员应用、SAS/Enterprise Miner 和 SAS/AF 程序语言等。本书主要针对以上模块,介绍与现代统计学方法相关的一些应用实验,具体包括认知性、设计性和研究创新性三类实验。其中,认知性实验是以传授 SAS 统计软件和统计学知识为主,在教学中以教师辅导为主,通过实践,使学生对 SAS 统计软件操作和统计学应用有感性的认识;设计性实验是以掌握解决问题的方法为主线,倡导自学,要求学生通过实验预习掌握实验的原理,在实验过程中解决所碰到的具体问题,通过实验报告总结提高;创新性实验以提高素质为主线,要求学生能结合所学专业或自行提出实验课题,通过以学生为主、以教师为辅的实验教学方法,逐步培养学生的独立研究能力。

本书分 3 篇,共 25 个实验。第一篇主要是关于统计软件实验的介绍。第二篇是 SAS 软件的认知性实验,主要包括 SAS 软件的操作性实验、SAS 软件的单变量统计分析实验、SAS 软件的多元统计分析实验。第三篇是关于 SAS 软件的设计性与创新性实验,主要涉及经济学、管理学、社会学等学科领域中的统计学应用,其目的是期望学生在通过实验掌握 SAS 软件操作的基础上,利用 SAS 软件进行科学研究。本书涉及的所有程序与数据集都收录在附赠光盘。

本书的出版得到上海财经大学的专项基金资助。上海财经大学黄林芳副校长,以及陈启杰教授、何玉长教授、陈信元教授等项目专家对本书出版给予充分肯定和大力支持,在此表示衷心感谢。同时,还要感谢艾春荣教授对应用统计学的倡导,以及统计实验室李晓玉副教授和出版社刘光本博士对本书出版的帮助。另外,厥开伟、符正真、王巍、韩娟娟等同学也参加了部分实验的整理和编写工作,在此表示感谢。

鉴于作者水平有限,书中错误与缺点难免,恳请读者批评指正。

邵建利

2006年9月于上海财经大学

目 录

前言/1

第一篇 应用统计软件实验概述/1

第一节 统计分析方法与统计软件/1

第二节 应用统计软件实验的优点/9

第三节 基本框架及与光盘的程序对应/10

实验一 透视线性回归模型/12

实验二 观察参数对分布的影响/14

第二篇 SAS 软件的认知性实验/21

第一节 SAS 软件的操作/21

实验一 SAS 语言与函数/21

实验二 SAS 的数组和矩阵/26

实验三 SAS 的作图/31

实验四 SAS/INSIGHT/36

实验五 SAS/ANALYST/41

实验六 统计数据的描述性分析/47

实验七 统计数据的探索性分析/49

第二节 单变量统计分析/54

实验一 置信估计和假设检验/54

实验二 方差分析/57

实验三 一元回归分析/61

实验四 属性数据分析/66

第三节 多元统计分析/71

实验一 回归分析与回归诊断/71

实验二 主成分分析/82

实验三 因子分析/91

实验四 聚类分析/97

实验五 典型相关分析/106

第三篇 SAS 软件的设计与创新性实验/114

实验一 股票市场价格不遵循随机游走/114

实验二 风险、收益与均衡:数据模型检验/120

实验三 人口分布变动的影响因素模型/126

实验四 高校体育活动资金分配标准研究/136

实验五 体质测定标准的实证研究/146

实验六 长途客运客流量预测模型/157

实验七 数据特征对预测分类数据挖掘模型选择的影响/165

参考文献/179

第一篇 应用统计软件实验概述

第一节 统计分析方法与统计软件

一、统计分析方法

现代统计学发展至今积累了许多行之有效的数据分析方法,这些统计分析方法可以帮助我们发现大量数据中蕴藏的规律性,帮助我们在数量上把握客观事物的本质,进一步帮助我们掌握客观事物的发展趋势。所以,统计分析方法是当今数据分析的主要应用方法之一。这些统计分析方法一般有以下特点:

1. 定量性

统计分析方法处理的对象是各种各样的数据,通过处理这些数据达到对客观事物定量的认识。这些统计分析方法建立在严谨的数学推导基础之上,并将不断有新的有效的方法出现。

2. 复杂性

统计分析方法的复杂性首先表现在处理的数据量可能很大,数据之间的关系复杂,只有基于对大量数据的分析,才能全面、准确地认识事物。其次表现在许多统计分析方法需要进行繁琐的计算。再次表现在多种统计分析方法的综合运用。在实际统计分析中要发现大量数据中的规律性,可能需要使用多种方法、多个参数反复实验比较,不断发现新的问题,探索性地逐步逼近事物的本质规律。

3. 广泛性

统计分析方法是一种通用的分析处理数据的方法,具有广泛的应用性。它可以成为统计、金融、企业经营、经济管理、科学研究以至医药、卫生、农业、气象等各行各业相关人员的服务工具。

现实中,统计分析方法的复杂性成为制约统计分析方法广泛应用的关键

问题。统计分析软件可以避免繁琐演算,也可以避免用户在计算机上自行编制复杂的程序,是普及应用统计分析方法的重要工具。

二、统计分析软件

统计分析软件根据其功能可以大致分为专用统计分析软件和综合统计分析软件两大类。

(一)综合统计分析软件

综合统计分析软件功能全面,它较系统地集成了多种成熟的统计分析方法,具有较完善的数据定义、操作和管理功能,可以方便地生成各种统计图形和统计表格,同时提供各种简便的软件使用方式,并带有完备的错误提示及联机帮助功能。综合统计分析软件是开放的软件系统,能方便地和其他各种常用的软件进行数据交换。常在微机上使用的综合统计分析软件有以下几种:

1. SAS

它是美国 SAS 软件研究所研制的一套大型集成应用软件系统,具有完备的数据存取、数据管理、数据分析和数据展现功能。尤其是其统计分析系统部分,由于其具有强大的数据分析能力,一直奉为业界著名软件,在数据处理和统计分析领域,被誉为国际上的标准软件和最权威的优秀统计软件包,广泛应用于政府行政管理、科研、教育、生产和金融等不同领域,发挥着重要的作用。SAS 系统中提供的主要分析功能包括统计分析、经济计量分析、时间序列分析、决策分析、财务分析和全面质量管理工具等。

SAS 系统是一个组合软件系统,它由多个功能模块组合而成,其基本部分是 BASE SAS 模块。BASE SAS 模块是 SAS 系统的核心,承担着主要的数据管理任务,并管理用户使用环境,进行用户语言的处理,调用其他 SAS 模块和产品。也就是说,SAS 系统的运行,首先必须启动 BASE SAS 模块,它除了本身所具有的数据管理、程序设计及描述统计计算功能以外,还是 SAS 系统的中央调度室。它除可单独存在外,也可与其他产品或模块共同构成一个完整的系统。各模块的安装及更新都可通过其安装程序非常方便地进行。SAS 系统具有灵活的功能扩展接口和强大的功能模块,在 BASE SAS 的基础上,还可以增加如下不同的模块进而增加不同的功能:SAS/STAT(统计分析模块)、SAS/GRAPH(绘图模块)、SAS/QC(质量控制模块)、SAS/ETS(经济计量学和时间序列分析模块)、SAS/OR(运筹学模块)、SAS/IML(交互式矩阵程序设计语言模块)、SAS/FSP(快速数据处理的交互式菜单系统模块)、SAS/AF(交互式全屏幕软件应用系统模块)等。SAS 有一个智能型绘图系

统,可以绘制各种统计图和地图。SAS 提供多个统计过程,每个过程均含有极丰富的任选项。用户还可以通过对数据集的一连串加工,实现更为复杂的统计分析。此外,SAS 还提供了各类概率分析函数、分位数函数、样本统计函数和随机数生成函数,使用户能方便地实现特殊统计要求。

在国际学术界有条不成文的规定,即在国际学术交流中凡是用 SAS 软件完成的计算和统计分析,可以不必说明算法。虽然在我国 SAS 的广泛应用还是近几年的事,但是随着计算机应用的普及和信息事业的不断发展,越来越多的单位采用了 SAS 软件。尤其在教育、科研领域,SAS 软件已成为专业研究人员进行统计分析的标准软件。

然而,由于 SAS 系统是从大型机上的系统发展而来,其操作至今仍以编程为主,人机对话界面不太友好,系统地学习和掌握 SAS,需要花费一定的时间和精力。而对大多数科技工作者而言,需要掌握的仅是如何利用 SAS 来解决自己的实际问题,因此往往会与 SAS 软件失之交臂。但不管怎样,SAS 作为专业统计软件,现在还很难有其他统计软件能与之抗衡。

2. SPSS

它是社会科学统计软件包,也是世界著名的统计分析软件之一。20 世纪 60 年代末,美国斯坦福大学的三位研究生研制开发了最早的统计分析软件 SPSS,同时成立了 SPSS 公司,并于 1975 年在芝加哥组建了 SPSS 总部。20 世纪 80 年代以前,SPSS 统计软件主要应用于企事业单位。目前的 SPSS 是一个组合式软件包,它集数据整理、分析功能于一身。用户可以根据实际需要和计算机的功能选择模块,以降低对系统硬盘容量的要求,有利于该软件的推广应用。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。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。

1994~1998 年间,SPSS 公司陆续购并了 SYSTAT 公司、BMDP 软件公司、Quantime 公司、ISL 公司等,并将各公司的主打产品纳入 SPSS 旗下,从而使 SPSS 公司由原来的单一统计产品开发与销售转为向企业、教育科研及政府机构提供全面信息统计决策支持服务,成为走在了最新流行的“数据仓库”和“数据挖掘”领域前沿的一家综合统计软件公司。

由于在 SPSS 公司的产品线中, SPSS 软件属于中、低档(SPSS 公司共有二十余个产品), 因此从战略的观点来看, SPSS 显然是把相当的精力放在了用户界面的开发上。该软件只吸收较为成熟的统计方法, 而对于最新的统计方法, SPSS 公司的做法是为之发展一些专门软件, 如针对树结构模型的 Answer Tree、针对神经网络技术的 Neural Connection、专门用于数据挖掘的 Clementine 等, 而不是直接纳入 SPSS, 因此它们在 SPSS 中均难觅芳踪。

但是, 作为三大综合性统计软件之一, 其统计分析功能与另外两个软件即 SAS 和 BMDP 相比仍有一定欠缺。虽然如此, 目前的 SPSS 由于其操作简单, 已经在我国的社会科学、自然科学的各个领域发挥了巨大作用。该软件还可以应用于经济学、生物学、心理学、医疗卫生、体育、农业、林业、商业、金融等各个领域。与 SAS 一样, 在国际学术交流中用 SPSS 软件完成的计算和统计分析, 也不必说明其算法。

3. SYSTAT

它由美国 SYSTAT 公司于 20 世纪 70 年代推出, 因方法齐全、速度快、精度高、软件小、处理数据量大而大受欢迎, 成为目前较为流行的通用数据分析软件包之一。1994 年 9 月为 SPSS 公司兼并。但比 BMDP 幸运的是, SPSS 公司并没有放弃 SYSTAT 的开发研究, 它通过调整产品布局, 利用 SYSTAT 较为突出的图形优势, 发展 MAC 平台上的产品系列使得 SYSTAT 焕然一新, 并且在 SPSS 吞并 BMDP 后, 拟将其 UNIX 平台版本划入 SYSTAT 发展。目前 SYSTAT 主要针对科学研究者, SYSTAT 几乎可以完成统计研究者所需要的任何统计方法, 软件包含有包括世界地图、三维图、经纬图等普通及奇特的图像模型, 它虽然还没有 SPSS 这样的软件包先进, 但比 SPSS 便宜得多, 而且硬盘容量要求更小。

4. Statistica

它是一套完整的统计资料分析、图表、资料管理、应用程序开发系统, 以及对其他技术、工程、工商企业资料挖掘应用等进阶分析的应用程序。此系统不仅包含统计上的一般功能及制图程序, 还包含特殊的统计应用, 可供社会统计人员或工程师使用。全新的 Statistica 在功能上提供了四种线性模型的分析工具, 包括 VGLM、VGSR、VGLZ 与 VPLS; 对使用者而言, 提供完整且具可选择性的使用者界面; 或整合 Statistica 与其他应用程序进行计算, 这些都是非常方便好用的模组。Statistica 能提供使用者所有需要的统计及制图程序。另外, 能够在图表视窗中显示各种分析及有别于传统统计范畴外的最新统计制图技术, 皆获得使用者的好评。Statistica 为基本系列产品, 也可独立使用

此模块,或搭配 Statistica 其他产品系列使用。

5. S-plus

S 语言是由 AT&T 贝尔实验室开发的一种用来进行数据探索、统计分析、制图的解释型语言。它丰富的数据类型(向量、数组、列表、对象等)特别有利于实现新的统计算法,其交互式运行方式及强大的图形和交互图形功能使得我们可以方便地探索数据。目前 S 语言的实现版本主要就是 S-plus。

S-plus 是基于 S 语言并由 MathSoft 公司的统计科学部进一步完善的统计软件。作为统计学家及一般研究人员的通用方法工具箱, S-plus 强调演示图形、探索性数据分析、统计方法、开发新统计工具的计算方法以及可扩展性。S-plus 是一个商业软件,可以直接用来进行标准的统计分析以得到所需结果。但是它的主要特点还是在于它可以交互地从各个方面去发现数据中的信息,并可以很容易地实现一个新的统计方法。另外, Auckland 大学的 Robert Gentleman 和 Ross Ihaka 及其他志愿人员开发了一个 R 系统,其语法形式与 S 语言基本相同,但实现方法不同,两种语言的程序有一定的兼容性。R 是一个 GPL 自由软件,它比 S-plus 还少许多功能,但已经具有了很强的实用性。

6. Stata

它作为一个小型的统计软件,其统计分析能力非常强。Stata 也是采用命令行方式来操作,但使用上远比 SAS 简单。其生存数据分析、纵向数据(重复测量数据)分析等模块的功能甚至超过了 SAS。用 Stata 绘制的统计图形相当精美,很有特色。Stata 的另一个特点是它的许多高级统计模块均是编程人员用其宏语言写成的程序文件,这些文件可以自行修改、添加和下载。用户可随时到 Stata 网站寻找并下载最新的升级文件。事实上, Stata 的这一特点使其成为几种统计软件中升级最多、最频繁的一个。Stata 的缺点是数据接口太简单,只能读入文本格式的数据文件,其数据管理界面也过于单调。

(二)专用统计分析软件

专用统计分析软件着重实现综合统计分析软件的部分功能,突出某种特色处理。常见的专用统计分析软件有:

1. EViews

它是美国 GMS 公司于 1981 年发行第一版的 Micro TSP 的 Windows 版本,通常称为计量经济学软件包。EViews 是 Econometrics Views 的缩写,它的本意是对社会经济关系与经济活动的数量规律,采用计量经济学方法与技术进行“观察”。计量经济学研究的核心是设计模型、收集资料、估计模型、检验模型、运用模型进行预测、求解和运用模型。EViews 是完成上述任务的必

不可少的工具。正是由于 EViews 等计量经济学软件包的出现,使计量经济学取得了长足的进步,发展成为实用与严谨的经济学科。使用 EViews 软件包可以对时间序列和非时间序列的数据进行分析,建立序列(变量)间的统计关系式,并用该关系式进行预测、模拟等。虽然 EViews 是由经济学家开发的,并且大多数被用于经济学领域,但并非意味着必须限制该软件包只用于处理经济方面的时间序列。EViews 处理非时间序列数据照样得心应手。实际上,相当大型的非时间序列(截面数据)的项目也能在 EViews 中进行处理。

2. Minitab

它同样是国际上流行的一个统计软件包,其特点是简单易懂。在国外大学统计学系开设的统计软件课程中,Minitab 与 SAS、BMDP 并列。Minitab 统计软件比 SAS、SPSS 等小得多,但其功能并不弱,特别是它的试验设计及质量控制等功能。Minitab 目前的最高版本为 V14.1,它提供了对存储在二维表中的数据进行分析的多种功能,包括:基本统计分析、回归分析、方差分析、多元分析、非参数分析、时间序列分析、试验设计、质量控制、模拟、绘制高质量三维图形等。从功能来看,Minitab 除各种统计模型外,还具有许多统计软件不具备的功能——矩阵运算。

3. BMDP

BMDP 第一版诞生于 1961 年,1968 年由 BMDP 公司发行,是最早的综合专业统计分析软件。以前,BMDP 与 SAS、SPSS 被并称为三大统计软件包,在国际上影响很大,客户达 1 万户以上,它方法全面、灵活,早期曾有很多独具特色的分析方法。但是今非昔比,由于 BMDP 发展路途不畅,从 1991 年的 7.0 版以后就没有新版本,最后被 SPSS 公司收购。由于 SPSS 公司在开发与推广 BMDP 统计软件方面的积极性也不大,因此现在 BMDP 的影响已不如从前。但不可否认,作为优秀的统计软件,BMDP 在国外仍然影响巨大,著名大学统计学系开设的多变量分析课程当中就有 BMDP 软件的教学内容,而且大型学术研究机构的服务器上也通常安装着 BMDP 软件供终端用户使用,国外许多临床医生进行神经网络分析用的就是 BMDP。

4. LISREL(Linear Structural Relationship)

这个名称有两个意思:一是由 Joreskog 提出的线性结构方程模型,二是指由 Joreskog 与其合作研究者 Sorbom 共同开发的软件系统。20 世纪 60 年代起,协方差结构分析(Covariance Structure Analysis)的方法开始萌芽,这是一种从变量间或变量群间的协方差结构出发,积极地探讨和确认因果关系的分析方法。心理学、行为科学、经济学等领域中常用的路径分析(Path A-

analysis)和结构方程模型(Structural Equation Model)分析等都属于协方差结构分析的范畴。20世纪70年代,Joreskog提出了LISREL模型和方法。几十年来,对这一模型研究的逐步完善和改进已使LISREL成为协方差结构模型分析的代表性的手法之一。特别是LISREL软件的研制及实现商品化以来,该方法的应用受到越来越多国家的研究者的重视。

5. AMOS

该软件开发者是James L. Arbuckle。AMOS是SmallWaters公司开发的路径分析软件,用户可以访问<http://www.smallwaters.com/>或<http://www.spss.com/>来了解更多的情况。AMOS与LISREL软件一样,它可以用来绘制路径图形,分析协方差结构。AMOS3.6及以后的版本可以以多种格式读取二进制文件,支持ASCII及文本文件。新的AMOS版本还可以从SPSS统计软件的“Statistics”菜单启动以读取SPSS格式文件。

在SEM分析方面,AMOS以图形界面见长,模型中的观测变量、潜变量、误差项都有不同的图形来表示,这些图形都可以从左边的工具栏中使用相应的工具在右边绘图区直接手工绘制,模型中的数据也可以通过菜单“File→Data Files→File Name”导入,在设定完参数的属性以及模型属性(包括标题、模型估计方法、模型输出结果选项、Bootstrap选项等)之后,通过菜单“Model-Fit→Calculate Estimates”即可对模型求解;而LISREL一般是通过执行程序语句(Syntax)来求解计算的,观测变量、潜变量、协方差阵等在LISREL程序语句中都有相应的代码,如LY、LX、BE、GA等。在数据要求上,LISREL与AMOS有所不同:AMOS中可以直接使用原始数据(Raw Data),而LISREL软件要求输入的是原始数据的协方差阵(Covariance Matrix)或者相关系数阵(Correlation Matrix),不能直接使用原始数据,通常使用PRELIS将原始数据读入然后转化为协方差阵或者相关系数阵;在参数估计方法上,LISREL可以使用工具变量法和两阶段最小二乘法。

总之,综合统计分析软件和专用统计分析软件今后将会有更多的新版本,因此应在实践中针对统计分析软件的功能和特点不断“打磨”,这样才能最终实现提高统计分析水平的目的。

三、统计软件的使用方式

1. 尽管统计软件非常之多,但是它的使用方式与其他各种计算机软件使用方式一样,是有基本规律可循的。一般来说,统计分析软件有以下几种使用方式:

(1) 人机对话方式(交互式方式)

统计分析软件提供给用户多种命令,由用户选择使用,去完成相应的处理。人机对话方式的基本特征是首先由用户键入一条命令,由统计分析软件判断正误,如错误则显示提示信息,然后等待用户键入下一条命令。如正确则执行此命令并显示结果,显示完毕后,等待用户键入下一条命令。

人机对话方式的优点是方便初学者一步步深入学习,熟悉各命令;人机对话方式的缺点是当统计分析软件提供的命令多且复杂时,用户上机很难记住,从而使操作困难。另一方面,由于计算机必须等待人工逐条录入命令,然后逐条执行,使工作效率低下。

(2) 批命令方式

先输入一批统计分析软件的命令在一个程序文件中,然后调用这个程序文件,由计算机自动逐条顺序地执行文件中的各条命令。此种工作方式可解决人机对话方式处理效率低的问题。

(3) 菜单方式

将统计分析软件提供的命令按菜单的方式列出来,由用户使用简单的键选择相应的命令,由计算机自动执行。菜单方式可以避免用户记忆命令,使用户对统计分析软件的命令一目了然,所以菜单方式对用户来说是一种较好的工作方式。根据选择命令所使用的键,菜单又分为字母菜单、数字菜单和光条菜单,根据多层菜单的显示方式,菜单又分为弹出式菜单和下拉式菜单。

SAS 软件主要有人机对话方式、批命令方式和菜单方式三种。总之,好的统计分析软件往往采用多种软件工作方式,这样可扬长避短,方便各种程度的用户使用和操作。

2. 统计分析软件是统计专家和计算机专家共同合作的成果,所以它必然会涉及统计和计算机两个领域的知识,所以学习统计分析软件应从几个方面入手:

(1) 弄清相应的统计概念和统计含义。明确相应的统计方法是解决什么问题的,此种方法适用的范围及其计算结果的含义,这样才能保证采用适当的统计分析方法处理具体问题,而避免无的放矢。由于是使用计算机处理,所以用户不必牢记其运算公式和步骤。

(2) 正确收集待处理的数据。首先,输入在一个计算机数据文件中,最终形成统计分析软件可处理的数据格式。其次,选择适当的一种或几种统计分析的方法,探索性地分析处理数据。

(3) 搞清并弄懂输出结果。使用者应在具备相应的统计学方面知识的基

基础上,掌握 SAS 统计软件进行各种数据处理时命令的基本格式、命令中各种参数的含义,并能读懂计算机输出的各种处理结果和完成相应的统计分析。

综上所述,对统计分析方法的理解和计算机处理能力的培养是学习统计分析软件的关键,因此这也是统计实验教学的目的所在。

第二节 应用统计软件实验的优点

自计算机问世以来,使用计算机完成统计分析一直是统计学家和计算机程序员面临的重大课题。利用计算机进行抽样设计和统计模拟实验,建立数据库,进行各种计算,绘制统计图表,是当今衡量统计工作和科学研究水平的重要标志。以往摸球、抽卡片、随机分组等统计模拟实验相当费时,且由于重复抽样次数不够,难以得到理想结果。在经济与管理学教育中,统计软件实验是培养学生统计科学思维、正确认识所研究问题的“钥匙”,有利于培养学生掌握新知识的能力。归纳起来,统计软件应用实验具有以下一些好处:

1. 有助于对统计学的认知

学生的知识面相对较窄的现象并不罕见。统计应用软件实验课是在教师指导下,学生进行的一种有助于消化和巩固所学知识,培养其应用理论解决问题的教学手段。实验课能启发思维,开阔视野,激发学生的学习兴趣和求知欲望,锻炼学生对统计学的应用能力。实验课上,各报告中都有一段较精练的专业背景介绍,使得学生在较短的时间内汲取了浓缩的本专业以外的知识,这对学生触类旁通、产生一些创造性的思维等很有帮助。

实验课上,学生了解或经历了统计分析的全过程,即从实验设计到资料的收集、整理与分析。报告引用的都是实际资料,这些资料形式多样,统计处理时既有正面的例子,也有反面的例子,而且常会遇到教科书以外的一些处理方法(特别是从外文文献中引用的)。教师根据具体资料,适当展开,重点讲解认识资料的一般规律和如何选用适当的统计方法。特别在范文讲解中选用典型的佳作,贯穿实验设计、资料收集和整理、资料分析、结果表达等每一个环节,给学生以深刻的印象。通过以上直接或间接的实践过程,有助于提高学生运用所学知识处理实际问题的能力。

2. 有助于科研论文的规范

通过统计软件应用实验,可使学生举一反三,提高独立解决常见及疑难问题的能力,并通过综合应用的操作演示扩大学生的知识面。在统计学教学时,培养学生用信息技术掌握学科新进展、解决实际问题的能力。在完成实验报告后,结合科

研与工作实际,可以培养学生了解统计学科论文的规范。另外,由于报告内容来源广泛,且资料性质各异,表达方式也不尽相同,丰富的信息亦可扩大学生的视野。报告所涉及的某些内容对其他同学来讲也属陌生,可促进学生互相学习。学生上交的分析报告积累了丰富的资料,可作为教学和科研实例。

3. 有助于提高学生的科研能力

计算机技术的飞速发展,给统计学分析带来很多新的方法和理论。复杂的多元统计分析技术在计算机上可轻而易举地完成。当教师将科研成果介绍给学生,就可以初步培养他们的创造能力。教学实践表明,这种模式十分有利于提高学生的科研和工作能力,缩短“实习期”。

第三节 基本框架及与光盘的程序对应

(一) 本书的基本框架

本书将从三个层面(即认知性、设计性和研究创新性实验角度)来介绍 SAS 软件应用统计实验。其中,认知性实验是以传授 SAS 软件操作知识为主,在教学中以教师辅导为主,学生通过实践对 SAS 统计软件操作和统计学有感性的认识;设计性实验是以解决问题为主线,倡导自学。要求学生通过实验预习掌握实验的原理,在实验过程中解决所碰到的具体问题,通过实验报告总结提高;创新性实验是以提高素质为主线,要求学生能结合所学的知识自行提出实验课题。通过以学生为主、以教师为辅的实验教学方法,逐步培养学生的独立工作能力。本书的基本框架如图 1-1 所示。

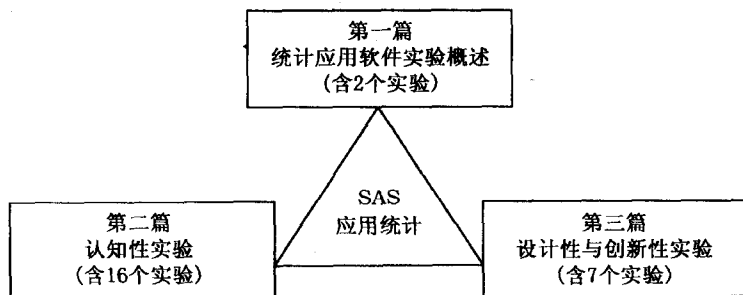


图 1-1 本书的基本框架

(二) 书与光盘的程序对应

书与光盘的程序对应见表 1-1。