

# Research on the Application and Design of SPSS Statistics Software in the Teaching Course of Mathematical Statistics

Hui Jiang

Department of Mathematics, Qinghai Normal University, Xining  
Email: [jh655@163.com](mailto:jh655@163.com)

Received: Sep. 14<sup>th</sup>, 2014; revised: Oct. 17<sup>th</sup>, 2014; accepted: Oct. 25<sup>th</sup>, 2014

Copyright © 2014 by author and Hans Publishers Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this paper, it discusses the corresponding relationship between teaching content of mathematical statistics and applied module of SPSS statistics software. There are some cases about a combination of thoughts, methods of mathematical statistics and application and outputs in SPSS. Now, there are a lot of books about SPSS statistics software. However, they are still limited on studies of applications and designs of SPSS statistics software in the teaching course of mathematical statistics.

## Keywords

Mathematical Statistics, SPSS, Statistics Software

---

# SPSS统计软件在数理统计课程教学中应用及设计的研究

蒋 卉

青海师范大学, 数学系, 西宁  
Email: [jh655@163.com](mailto:jh655@163.com)

收稿日期: 2014年9月14日; 修回日期: 2014年10月17日; 录用日期: 2014年10月25日

## 摘要

本文探讨了数理统计的教学内容与SPSS统计软件应用模块的对应关系,及数理统计思想和方法与SPSS软件的应用及输出结果相结合的案例。目前,已有大量关于SPSS统计软件方面的书籍,但是数理统计与SPSS统计软件内容一一对应的设计及教学的研究还是非常有限的。

## 关键词

数理统计, SPSS, 统计软件

## 1. 引言

现今整个世界处于一个信息高度发达、经济活动频繁的时期。有报道称,当前全球的数据总量正呈指数增长。移动互联网、物联网等的迅速发展,使新数据源不断出现,GPS、传感器等数据持续、大量产生。而数据获取成本、存储成本和处理成本的下降,也推动了数据量的膨胀。数据正成为与物质资产和人力资本相提并论的重要生产要素。

各领域产生的大量信息、数据需要人们综合分析和处理。如何从这些大量数据中发现经济活动规律并提取有价值的信息,以进一步指导人们的社会实践活动显得越发迫切和需要,也为数理统计理论方法的应用提供了非常广泛的需求空间。给我们学习数理统计方法也提出了新要求,必须加强数理统计学习,充实数理统计实践内容及实际应用方法。

现代技术教育环境的发展,特别是计算机的普遍应用,使得数理统计的应用范围愈来愈广泛,成为科学研究不可缺少的工具。数理统计学又是一门实践性很强的应用科学,它同大量的数据打交道。处理数据是一件十分重要而又繁琐的工作,SPSS软件是世界上应用最广泛的统计软件之一,它具有强大的处理数据功能。在《数理统计》[1]学习过程中应用统计软件,加强理论联系实际的学习方法,能够提高学生处理实际问题的能力和学习理论知识的效率,可以为社会实践提供更高效的服务。在《数理统计》的教学中,加入SPSS的实验课已是刻不容缓的需求。目前,已有大量关于SPSS统计软件方面的书籍,但是数理统计与SPSS统计软件内容一一对应的课件设计及教学的研究还是非常有限的。

## 2. 数理统计教与学中存在问题

在本科《数理统计》课程教与学过程中,教师常常遇到的一个实际而又棘手的问题是,如何使枯燥的统计理论、计算公式与社会经济实践活动结合起来而变得更加具体实际;同时,在课程学习过程中,教师的教学只能使学生掌握理论知识与原理,学生在学习时,用了大量时间精力来学习统计问题的数学基础理论知识和简单的手工计算,但在学习其他相关学科中用数理统计方法处理大量专业数据时,需要的是解决专业问题所需要的知识和能力,这恰恰是以往数理统计类课程教与学中所没有关注到的,仅凭教师课堂授课的知识,对于在生活工作实践中遇到的问题,很难使学生学以致用。更甚者所学的理论知识无法与现实生活工作的实际问题接轨。掌握统计软件的使用技能是将统计方法应用于社会实践的基础,因为大部分的实际问题都涉及到大量数据的计算和处理,没有统计软件的辅助,手工计算简直无法想象,甚至无法完成。数理统计与统计软件结合的教学,可使学生全面掌握数理统计理论原理与基本方法与生活工作实践中的应用有机结合,很快将所学知识融入工作生活实践之中。

### 3. SPSS 统计软件特点

SPSS(Statistics Package for Social Science)是一种运行在 Windows 系统下的社会科学统计软件软件包,是世界著名的统计分析软件之一。目前,SPSS 在全球约有几十万家产品用户,分布与通讯、医疗、银行、证券、保险、商业、市场研究、科研教育等多个行业和领域,已成为世界上最流行、应用最广泛的专业统计分析软件。

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,具体内容包括描述统计、列联分析,总体的均值比较、相关分析、回归模型分析、聚类分析、主成份分析、因子分析、非参数检验等多个大类,每个类中还有多个专项统计方法。“易学,易用,易普及”已成为 SPSS 软件最大的竞争优势之一,也是广大数据分析人员对其偏爱有加的主要原因。而大量成熟的统计分析方法、完善简便的数据定义操作管理、开放的数据接口以及灵活的统计表格和统计图形,使得 SPSS 获得更多用户的喜爱。

然而,由于对数理统计基础理论和方法缺乏正确理解和认识,这类软件的应用可能带来的问题是,一方面它们有时候使统计方法处理数据时变得过于容易,从而可能导致对统计方法的误用和滥用;另一方面,由于过分信赖软件,许多用户会简单地满足于软件输出的结果,而对其中的错误不加分析和关注。鉴于此,我们需要将两者结合起来,将数理统计思想和方法,与 SPSS 软件的应用及输出结果相结合,以使数据结果更加符合实际的意义。

### 4. 《数理统计》辅以相应的 SPSS 统计软件内容的教学

在《数理统计》教与学中应用 SPSS 统计软件配合教学。软件教学起到如下作用:使学生能通过学习中应用统计软件处理相关数据,加深对数理统计基本理论的认识和理解,提高学习《数理统计》课程的兴趣。同时,通过对数理统计原理和知识的学习,在统计软件的应用时,鉴别数据结果输出时的正误,纠正应用统计软件时产生的错误。

《数理统计》课程教学辅以 SPSS 统计软件教学的主要内容,按照《数理统计》课程教学大纲的要求及社会实际的需求,SPSS 统计软件课件设计共分为六部分:

第一部分:数据的预处理。即是数据的排序、变量计算、数据的选取、计数、分类汇总、数据分组、拆分、转置、加权处理等。

依照《数理统计》课程教学大纲,没有数据预处理相关内容的介绍。SPSS 作为辅助教学,对这部分内容的简单介绍可弥补《数理统计》教学的不足。通过计算机与软件结合的实用性和易操作的特点,使得学生掌握基础的数据处理,对数据有初步的感性认识。

第二部分:数据的基本统计分析。即是统计分布基础理论部分(包括样本均值、样本方差等常用统计量及其分布理论)、交叉分组下的频数分析、多选项分析和比率分析。对应于《数理统计》相应的内容(统计分布基础理论部分),应用统计软件 SPSS 主菜单“分析”下的描述统计模块下的频数分析子模块、描述统计子模块、探索分析子模块等子模块辅助学习本章内容。

作为 SPSS 课程的教学,在该部分中的交叉分组下的频数分析、多选项分析和比率分析应是数据处理与分析中,比较有实际应用的内容。但在《数理统计》教材中没有这部分的内容介绍。可根据学生的专业取向,介绍学生应自学的相应内容。

第三部分:参数估计,包括参数的矩估计、极大似然估计和估计量的优良性质,参数的区间等估计方法。应用 SPSS 主菜单“分析”下的描述统计模块下的频数分析子模块、描述统计子模块、探索分析子模块进行参数估计和置信区间的计算。

第四部分:统计假设检验,包括假设检验的基本思想、基本步骤,正态总体参数均值、方差、标准差的检验的原理与方法,检验的 P-值和  $\chi^2$ -验;应用 SPSS 主菜单“分析”下均值比较模块进行参数假设

检验。应用 SPSS 主菜单“分析”下非参数检验模块进行非参数假设检验。

第五部分：方差分析，包括单因素的方差分析，双因素的无交互作用的方差分析和考虑交互作用的方差分析；应用 SPSS 主菜单“分析”下的均值比较模块的方差分析子模块(单因素 ANOVA)进行单因素方差分析，“分析”下的一般线性模型模块中的单变量子模块进行多因素方差分析。

第六部分：回归分析，包括一元线性模型、多元线性模型参数的最小二乘估计方法，回归系数的统计分布，模型的检验、检验，以及预测问题与控制问题。应用 SPSS 主菜单“分析”下相关模块的双变量子模块进行相关分析或应用偏相关子模块进行偏相关分析，应用回归模块中的线性子模块进行回归分析。

## 5. 数理统计方法、SPSS 操作及案例分析的有机结合

### 1) 针对《数理统计》课程的 SPSS 统计软件课件内容的补充

为了学生将来工作的实际需求及对数据的感性认识，我们补充了数据预处理部分的相关内容。其中，数据分组是有些《数理统计》教材中求做直方图的传统教学中内容。要求排序数据，确定组距、组数、组内频数，最后制作直方图，过程比较繁琐。利用 SPSS 处理就简单多了。并且附带正态曲线的直方图，对数据分布与正态分布的拟合程度，具有直观的认识。

2) 针对《数理统计》课程学科特点，在应用 SPSS 统计软件学习数理统计的过程中，介绍数理统计基本原理时，配合相关领域的分析案例，紧紧围绕 SPSS 的输出结果教学，一方面使学生可以理解数理统计中统计思想和基本原理，掌握统计方法的应用范围。使学生真正理解 SPSS 为什么要输出这些统计量及它们对分析结论的重要性或必要性，进而会正确合理和完整地应用分析结果。另一方面也使学生能够快速熟练掌握 SPSS 统计软件，在理解分析方法的基础上，对分析结果以正确合理的解释。

例如 1 [2]，对某保险公司具有高等教育水平的员工比列的平均值不低于 0.8 的假设检验。

本题的原假设  $H_0: \mu \geq \mu_0 = 0.8$

利用单样本的 t-检验，SPSS 的输出结果如表 1，单个样本 t 检验输出中 t 统计量的双尾概率 P-值为 0.161，比例差分的 95%置信区间为(-1.373, 0.247)，比例总体均值 95%置信区间为(0.8 - 0.1373, 0.80 + 0.0247) = (0.6627, 0.8247)。给定显著性水平 0.05，由于应进行单尾(或单侧)检验，取概率 P-值的一半， $P/2 = 0.161/2 = 0.08$ ，大于显著性水平 0.05，于是接受原假设，不能认为保险公司具有高等教育水平与员工的比例平均值不显著高于 0.8 的假设。同时，0.8 也在比例总体均值 95%的置信区间内，进而从另一角度证实了这个结论。

在该例的教学中，一是，对概率 P-值的统计意义作介绍，与《数理统计》传统教学查表的概念的区

Table 1. T-test  
表 1. T-检验

单个样本统计量						
	N	均值	标准差	均值的标准误		
受高等教育比例	19	0.7437	0.16810	0.03856		

  

单个样本检验						
检验值 = 0.8						
	t	df	Sig.(双侧)	均值差值	差分的 95%置信区间	
					下限	上限
受高等教育比例	-1.460	18	0.161	-0.05632	-0.1373	0.0247

别在于,查表是统计量的值与临界值(查表的值)比较大小,确定统计量的值是否在临界域里,概率 P-值是统计量大于或小于临界值对应的概率,用概率 P-值大于或小于显著性水平,确定统计量的值与临界值(查表的值)的关系,从而判断接受或拒绝原假设。二是,利用 95%的置信区间与概率 P-值,互相印证结论的正确性。另外,需注意的是,进行单样本 t 检验前提是样本来自的总体应服从或近似服从正态分布。如果数据不服从正态分布,可取大样本,将数据标准化,根据中心极限定理知数据近似服从正态分布。如何是大样本,一般认为样本容量大于 30 就算大样本,随着计算机技术的发展,有人认为 50,或 80。关于大样本数的确定仍在探讨中。

例 2 [2],某企业在制定某商品的广告策略时,收集了该商品在不同地区采用不同广告形式促销后的销售额数据,希望 1) 对广告形式和地区是否对商品销售额产生影响进行分析; 2) 广告形式和地区的交互作用的交互作用是否对商品销售额产生影响进行分析。

本题是要进行方差分析,包括单因素的方差分析,双因素的无交互作用的方差分析和考虑交互作用的方差分析;单因素的方差分析应用 SPSS 主菜单“分析”下的均值比较模块的方差分析子模块(单因素 ANOVA)进行单因素方差分析,多因素方差分析在“分析”菜单下的一般线性模型模块中的单变量子模块进行多因素方差分析。

SPSS 进行方差分析或假设检验中的分析基础是基于对原假设的肯定,因此一定要清楚提出的原假设。本题的三个原假设分别为:

$H_{01}$ : 不同广告形式没有对销售额产生显著影响;

$H_{02}$ : 不同地区没有对销售额产生显著影响;

$H_{03}$ : 广告形式和地区对销售额没有产生显著的交互影响。

本题若仅进行广告形式  $H_{01}$  和地区  $H_{02}$  是否对商品销售额产生影响进行分析,可利用单因素的方差分析分别进行单因素的方差分析。在“分析”菜单下的一般线性模型模块中的单变量子模块进行多因素方差分析时,不仅能够分析多个因素对观测变量的独立影响,更能够分析多个控制因素的交互作用,即可同时进行  $H_{01}$ 、 $H_{02}$  和  $H_{03}$  的方差分析。

总之,应用统计软件学习数理统计,学以致用。使学生掌握研究群体现象的基本方法,获得处理实际问题的本领,通过应用来正确理解抽象繁杂的数理统计基本原理。在学习过程中使 SPSS 统计软件的学习和数理统计基本理论学习相结合,相互促进,提高学习效率。

## 参考文献 (References)

- [1] 魏宗舒,等 (1983) 概率论与数理统计教程. 第 2 版,高等教育出版社,北京.
- [2] 薛薇 (2006) 基于 SPSS 的数据分析. 中国人民大学出版社,北京.