

IBM SPSS 在教育与心理学中的应用

刘学敏 孙 巍 许红梅 耿希峰 编 著



北京大学出版社
PEKING UNIVERSITY PRESS



黑龙江大学出版社
HEILONGJIANG UNIVERSITY PRESS

IBM SPSS 在教育学与心理学中的应用

刘学敏 孙 巍 许红梅 耿希峰 编 著



北京大学出版社
PEKING UNIVERSITY PRESS



黑龙江大学出版社
HEILONGJIANG UNIVERSITY PRESS

图书在版编目 (CIP) 数据

IBM SPSS 在教育学与心理学中的应用 / 刘学敏等编著. — 哈尔滨 : 黑龙江大学出版社 ; 北京 : 北京大学出版社, 2016.8

ISBN 978-7-5686-0004-0

I. ① I… II. ① 刘… III. ① 统计分析—统计程序—应用—教育统计②统计分析—统计程序—应用—心理统计
IV. ① G40-051 ② B841.2

中国版本图书馆 CIP 数据核字 (2016) 第 100553 号

IBM SPSS 在教育学与心理学中的应用

IBM SPSS ZAI JIAOYUXUE YU XINLIXUE ZHONG DE YINGYONG

刘学敏 孙 崑 许红梅 耿希峰 编著

责任编辑 李 卉

出版发行 北京大学出版社 黑龙江大学出版社

地 址 北京市海淀区成府路 205 号 哈尔滨市南岗区学府三道街 36 号

印 刷 哈尔滨市石桥印务有限公司

开 本 787×1092 1/16

印 张 27.75

字 数 542 千

版 次 2016 年 8 月第 1 版

印 次 2016 年 8 月第 1 次印刷

书 号 ISBN 978-7-5686-0004-0

定 价 48.00 元

本书如有印装错误请与本社联系更换。

版权所有 侵权必究

序 言

在心理学与教育学的教学和研究领域中,IBM SPSS 统计分析软件已经成为本科生、研究生必备的统计分析工具。SPSS 自 2010 年更名为 IBM SPSS,其功能更加强大,界面更加友好。本书的窗口界面是以 IBM SPSS Statistics 19.0 为参考撰写的,其操作程序同时也适用于先前和之后的中英文版本。

本书的内容架构包括 IBM SPSS 统计分析软件概述、数据文件的建立与编辑、描述性统计、t 检验、方差分析、相关分析、回归分析、非参数检验、聚类分析、问卷中多选题和排序题的处理、问卷的效度分析与信度分析、常用统计图表的绘制等。通过原理与实例相结合的形式,帮助初学者建立统计思想和思维方式,掌握 SPSS 数据获取和管理的方法;使初学者能够熟练运用 SPSS 实现频数表、均值、标准差、中位数、众数、极差等描述统计分析;使初学者掌握运用 SPSS 实现基本推断性统计分析的方法,如均值比较、方差分析、卡方检验、相关分析和回归分析等;教会初学者如何利用 SPSS 生成漂亮的表格和图形,以及表格和图形含义的解读方法;同时针对高级推断性统计分析方法的应用,系统地介绍了聚类分析、因子分析、主成分分析、信度分析等常用的统计方法。

本书主编多年从事心理学和教育学本科生、研究生的心理与教育统计学的理论和软件的教学工作,下面对学好 IBM SPSS 统计分析软件提出几点个人建议:要想学好统计分析软件,练习和实践是最重要的,只有经常用才能更牢固地掌握 SPSS;要将“如何做”(具体方法和步骤)与“为什么这样做”(方法论)相结合,统计模型的严格数学表达很复杂、烦琐,但其背后的思想往往很简单;要将原理与案例结合,作为统计学的使用者,重要的是掌握统计学的思想、解决问题的步骤和结果的解读方法;要学会举一反三、触类旁通、灵活运用,统计学只是一个工具,SPSS 更是工具的工具,不要指望它能够“自动”解决你面临的问题;复杂的方法未必是可行的方法,越是简单的方法,越容易得到广泛采用,也往往给使用者带来更多的价值。

本书可作为教育学、心理学核心课程“教育统计学”和“心理统计学”课程实验部分的教材,或教育学类、心理学类专业研究生和本科生“教育统计软件”和“心理统计软件”课

程的教材和参考书。本书还适合作为社会科学统计软件培训教材,书中案例的操作过程和输出结果的解释都非常详尽,特别适合初学者学习使用。

本书第一章、第二章、第三章由刘学敏编写,第四章、第五章、第六章由孙崑编写,第七章、第八章、第九章由许红梅编写,第十章、第十一章、第十二章由耿希峰编写。佳木斯大学教育科学学院应用心理学专业和教育技术学专业的葛欣、伍雪、周昱希、郭赞、李王伟、曲芷萱、陈月、姜婷婷等同学参与了数据的前期筛选和准备工作,全书由刘学敏负责策划、修改统稿。

各位作者在撰写过程中参阅并引用了大量的文献,在此一并向所有文献作者致谢,同时要感谢佳木斯大学有关领导的支持,感谢佳木斯大学教材中心领导的鼓励与帮助,感谢为本书的出版、印刷和发行等工作付出艰辛劳动的工作人员。

由于编者学识和能力有限,本书不当之处还请读者批评指正。

编者

2015年12月

目 录

第一章 IBM SPSS 统计分析软件概述	1
第一节 IBM SPSS 统计分析软件的发展历史和特点	1
第二节 IBM SPSS 软件的安装与启动	5
第三节 IBM SPSS 软件窗口及菜单功能概述	11
第四节 IBM SPSS 软件系统参数设置	20
第二章 数据文件的建立、编辑、加工与处理	28
第一节 数据文件的建立	28
第二节 定义变量、数据录入和保存	39
第三节 数据的编辑与整理	46
第四节 数据的加工和处理	74
第三章 描述性统计	96
第一节 描述性统计的基本概念和原理	96
第二节 次数分析	106
第三节 描述统计	119
第四节 探索分析	123
第五节 Means 过程及应用	131
第四章 t 检验	137
第一节 假设检验的原理	137
第二节 单样本 t 检验	140
第三节 两独立样本 t 检验	146
第四节 两配对样本 t 检验	154
第五章 方差分析	163
第一节 方差分析概述	163
第二节 单因素方差分析	166
第三节 多因素方差分析	179

第四节	随机区组设计的方差分析	192
第五节	协方差分析	195
第六节	多元方差分析	199
第七节	重复测量的方差分析	206
第八节	方差成分分析	217
第六章	相关分析	223
第一节	相关分析概述	223
第二节	皮尔逊积差相关	226
第三节	斯皮尔曼等级相关	231
第四节	肯德尔和谐系数	234
第五节	偏相关分析	238
第七章	回归分析	244
第一节	回归分析的原理	244
第二节	一元线性回归分析	251
第三节	多元线性回归分析	260
第四节	曲线估计	268
第八章	非参数检验	279
第一节	非参数检验概述	279
第二节	卡方检验	285
第三节	二项分布检验	289
第四节	两个独立样本检验	292
第五节	多个独立样本检验	296
第六节	两个相关样本检验	301
第七节	多个相关样本检验	305
第九章	聚类分析	313
第一节	聚类分析概述	313
第二节	距离分析	315
第三节	系统聚类分析	320
第四节	快速聚类分析	329
第十章	多选题和排序题的 SPSS 处理	335
第一节	多选题的录入与分析	335
第二节	排序题的录入与分析	349
第十一章	问卷的效度分析与信度分析	361
第一节	问卷的效度分析	361
第二节	问卷效度 SPSS 应用及操作实例	374

第三节	问卷的信度分析	395
第四节	SPSS 信度分析实例	398
第十二章	绘制统计图	405
第一节	ROC 曲线	405
第二节	图表构建程序	413
第三节	图形画板模板选择程序	421
第四节	旧对话框	425
参考文献		433

第一章 IBM SPSS 统计分析软件概述

本章主要介绍 IBM SPSS 统计分析软件的特点,发展历史,系统的安装、启动和退出,以及窗口功能和系统参数的设置等操作方法。其中软件的安装、窗口的类型及其功能是学习的重点。建议初学者重点学习软件的安装步骤及常用参数的设置,简单了解窗口的类型及其功能。

统计分析软件是数据分析的主要工具,完整的数据分析过程包括数据的收集、整理和分析,统计学为数据分析过程提供了一套完整的方法论,统计软件为数据分析提供了实现手段。SPSS 全称为 Statistical Product and Service Solutions,即统计产品与服务解决方案,是美国 SPSS 公司开发的大型社会科学软件,2010 年被 IBM 公司收购,改名为 IBM SPSS。它与 SAS(Statistical Analysis System,统计分析系统)、BMDP(Bio Medical Data Processing,生物医学程序)并称为国际上最有影响的“三大统计软件”。

第一节 IBM SPSS 统计分析软件的发展历史和特点

一、SPSS 软件的诞生和发展

20 世纪 60 年代末,美国斯坦福大学三位研究生研制了“社会科学统计软件”,简称 SPSS(Statistical Package for the Social Science),该软件采用 FORTRAN 语言编写,这代表着 SPSS 统计分析软件的诞生。1975 年,SPSS 公司总部在芝加哥成立,推出 SPSS 中小型机版,即 SPSSX。

20 世纪 80 年代,SPSS 推出微机版,即 SPSS/PC + 诞生,最初的微机版是基于 DOS 系统的,在用户界面和输入、输出环境等方面不是十分方便。

20 世纪 90 年代,SPSS 推出 Windows 版,这是基于 MS Windows 操作系统的 5.0、6.0、6.1 版本,在操作界面上改善许多。1995 年以后,SPSS 推出 7.0 版本;2000 年,SPSS 公司由于产品升级和业务拓展的需要,将其产品正式更名为 SPSS(Statistical Product and

Service Solutions),即统计产品与服务解决方案;2000年以后,随着 Windows 操作系统的升级,SPSS 先后推出 10.0、11.0、12.0……18.0 等版本。2010 年以后,SPSS 公司被 IBM 公司收购,更名为 IBM SPSS,先后推出了 19.0、20.0、21.0 等版本。

二、IBM SPSS 统计分析软件的特点

IBM SPSS 功能全面,系统地集成了多种成熟的统计分析方法,如频数分析、描述统计、交叉分组下的频数分析、多选项分析、方差分析、聚类分析、判别分析、因子分析、回归分析、生存分析、时间序列分析等。其特点如下。

有完善的数据定义、操作和管理功能;可以对数据变量进行定义,如设置变量类型、值标签、显示格式、测度类型等。

可以方便地生成各种统计图形和统计表格,如直方图、直条图、折线图、饼状图、散点图、概率图等。

使用方式简单,有完备的联机帮助功能;大多数软件都采用可视化的图文操作界面,人机对话方便快捷,不需要编程知识也可以方便使用。它为初学者提供了多种类型的帮助功能。

软件开放性好,能方便地和其他软件进行数据交换。例如,打开和保存成文本格式、MS Excel 等多种格式的数据文件。

三、IBM SPSS 软件的统计分析功能

(一) 基本统计分析功能

包括频数分析、描述统计、交叉分组下的频数分析、多选项分析等具体功能。利用频数分析功能可以粗略把握变量的总体分布情况。利用描述统计功能可以了解数据的集中趋势、离散趋势、对称程度、陡峭程度,并精确把握变量的总体分布情况。利用交叉分组下的频数分析功能可以了解不同变量在不同水平下的数据分布情况,产生交叉列联表,分析列联表中变量间的关系。多选项分析可以用来分析多项选择题,通过定义多选项变量集,进行多选项频数分析和多选项交叉分组下的频数分析及卡方检验。

(二) 均值检验

包括单样本 t 检验、两独立样本 t 检验、两配对样本 t 检验等参数检验。单样本 t 检验,用来检验某变量的总体均值与指定的检验值之间是否存在显著差异。两独立样本 t

检验,可以根据两独立样本的数据,对两总体均值是否有显著差异进行推断。两配对样本 t 检验,可以根据配对样本对两总体均值是否有显著差异进行推断。

(三) 方差分析

包括单因素方差分析、多因素方差分析、协方差分析、多元方差分析、重复测量的方差分析和方差成分分析等。单因素方差分析用来检验某一个控制因素(自变量)的改变是否会给观察变量带来显著影响;相应地,多因素方差分析用来检验多个自变量对因变量的影响是否显著。协方差分析是将无法或很难控制的因素作为协变量,在排除协变量影响的条件下更精确地分析控制变量对观察变量的影响。多元方差分析用于解决控制变量的不同水平同时对多个因变量产生影响的情况。重复测量的方差分析可以用来处理被试的观测指标重复 3 次或 3 次以上的情况。方差成分分析是对混合效应模型的分析,用来分析混合效应模型中各随机效应对因变量变异贡献的大小。

(四) 相关分析

包括皮尔逊积差相关、斯皮尔曼等级相关、肯德尔相关和偏相关分析等。通过皮尔逊积差相关可以对两成对、正态、连续并具有直线关系的数据进行相关分析,得出其相关系数的值和显著性水平。斯皮尔曼等级相关可以用于两列具有等级顺序关系的数据间关系的分析。肯德尔等级相关可以用于两列或多列数据的相关性分析。偏相关又称静相关,可以在控制了其他变量的影响下计算两变量的相关系数。

(五) 回归分析

包括一元线性回归分析、多元线性回归分析、曲线估计等统计分析功能。一元线性回归分析可以对一个自变量与一个因变量之间的线性关系进行分析,得出回归方程并进行显著性检验。多元线性回归分析可以用于分析多个自变量与一个因变量间的关系,得出回归方程并进行显著性检验。曲线估计可以在根据观测量本身无法确定最佳回归模型时,采用多种曲线进行拟合,通过比较,找出最好的拟合模型。

(六) 非参数检验

SPSS 支持的非参数检验功能包括卡方检验、二项分布检验、两个或多个独立样本检验、两个或多个相关样本检验等。卡方检验可以用于检验因素的两项或多项分类的实际观测频数与期望频数是否有显著差异。二项分布检验用于检验样本分布与指定的二项分布是否存在显著差异。两个或多个独立样本检验用于检验两个或多个数据分布不明的独立样本的显著差异。两个或多个相关样本检验用于检验两个或多个数据分布不

明确的相关样本的显著差异。

(七) 主成分分析

SPSS 可以分析给出特征根、贡献率、因子载荷、因子正交旋转矩阵、碎石图等。当需要对问卷调查所得数据或测验所得数据进行探索性因素分析和结构效度分析时,可以采用主成分分析的方法。

(八) 聚类分析

SPSS 支持变量聚类和样本聚类,可以通过快速聚类、系统聚类等功能实现聚类分析,并可以给出树形图和冰柱图等图形。快速聚类可以很快将观测量分到各类中,适用于对正态分布的等间隔测度的连续的量纲相同的观测值进行聚类。系统聚类可以对样本聚类,也可以对变量聚类,它对数据要求较低,不要求数据具有等间隔,也不要求数据量纲相同。

(九) 判别分析

略。

(十) 对应分析

略。

(十一) 生存分析

略。

四、利用 IBM SPSS 做数据分析的一般步骤

1. 建立 IBM SPSS 数据文件

建立数据文件是利用 SPSS 进行数据分析的第一步,新建的数据文件格式为 .sav,也可以保存成其他格式。但是,在刚开始使用时保存成 .sav 文件格式是比较合适的。有关建立数据文件的具体操作和文件的类型在第二章中会有详细介绍。

2. 定义数据文件的结构

在录入数据信息以前,先要对研究的自变量和因变量进行分析,确定有几个变量以及每个变量的具体属性,包括变量名、变量类型(数值型、字符型等)、值标签等。注意:一般情况下,SPSS 中一列表示的是一个变量,一行表示的是一个个案(记录)。

3. 录入、修改和编辑待分析的数据

在录入数据时,SPSS 支持对数据的复制、粘贴、剪切、清除等具体的编辑操作。因此,在录入数据量较大时,可以由多人同时进行数据录入后,再进行数据的汇总和整理。

4. 统计分析之前的数据预处理

统计分析之前的数据预处理工作非常重要,因为录入数据的工作量往往比较大,有时需要聘请非专业人员录入数据。因此,数据很可能存在录入错误,例如录入的年龄数据本应该为 30,结果录成了 300,如果不加以修改将直接影响分析结果。数据预处理工作包括很多内容:利用探索分析、描述统计等功能,通过频次分布表和箱线图查找疑似异常值;利用数据计算或编码等功能,计算总分和平均分;将“男”“女”转换成“1”“2”。总之,为了后续统计分析的方便,数据预处理工作要尽可能做得细致些。

5. 统计分析和建模

统计分析过程并不是一蹴而就的,它需要我们进行反复的尝试,最后找到最合适的统计方法和数据表现形式。例如:在多元化回归分析中,我们首先要建模,再对模型进行拟合,根据各个模型的拟合优度和显著性水平选择回归模型。

6. 结果的说明和解释

统计分析的结果只是一些图表,需要利用专业的统计学知识进行解释和说明,否则,它们只是一堆符号,没有任何意义,非专业人士也根本看不明白。因此,对统计分析结果的解释说明非常重要,这部分工作需要统计学原理的支持。

7. 数据和分析结果的保存

在数据统计分析完成以后,一般需要将录入的数据文件和输出结果文件都保存起来。数据文件一般存储为 .sav 格式,也可以存储为文本、Excel、网页等格式;输出结果文件一般保存为 .spo 格式。

第二节 IBM SPSS 软件的安装与启动

一、IBM SPSS 软件的安装

1. 打开电源,启动计算机,软件安装前最好不要启动其他应用程序。
2. 将 SPSS 安装光盘放入光驱中,或将安装文件拷贝到机器某一个空闲盘符下。
3. 在安装文件夹中找到安装文件图标,名称可能为“SETUP. exe”或“SPSS_win32. exe”,不同版本安装文件名称可能不同,但是扩展名都是 .exe。双击安装图标,启动安装程序,进入如图 1-2-1 所示的安装程序向导(InstallShield Wizard)。

4. 按照提示选择“下一步”或者“Next”。

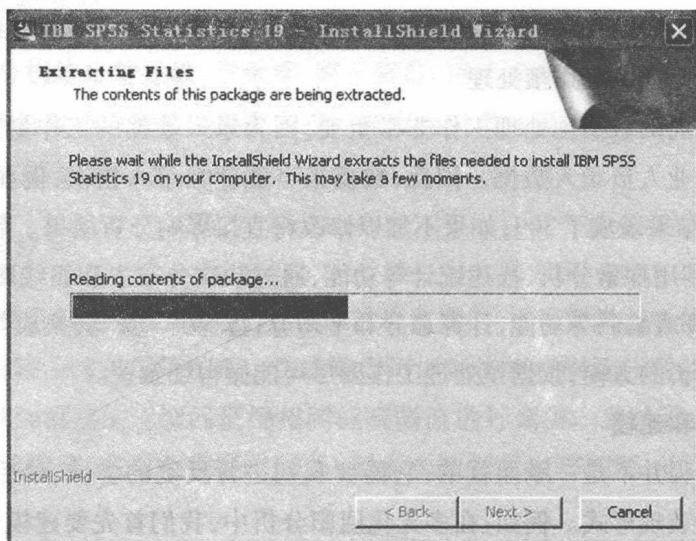


图 1-2-1 IBM SPSS 安装界面

5. 如果是单个用户,建议在用户许可证设置上选择“单个用户许可证”;如果公司或组织购买的产品,每个用户按照分配的代码注册时,建议选择“站点许可证”;如果是组织购买,通过网络 IP 获得使用权限,建议选择“网络许可证”。一般情况下,建议用户选择第一项“单个用户许可证”,如图 1-2-2 所示。

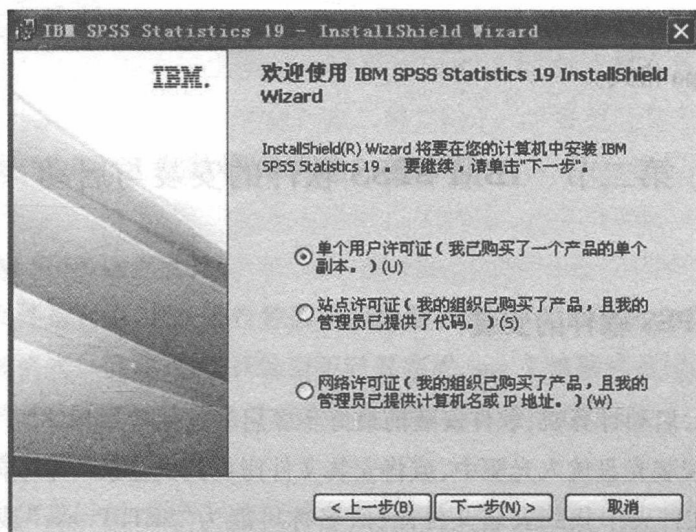


图 1-2-2 选择“单个用户许可证”

6. 软件许可协议,选择“我接受许可协议中的全部条款”,如图 1-2-3 所示。

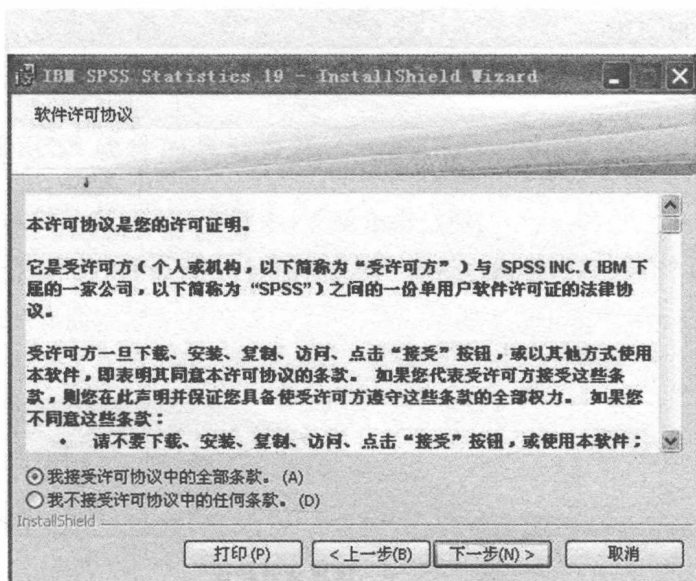


图 1-2-3 软件许可协议

7. 填写“用户姓名”和“单位”,如图 1-2-4 所示。

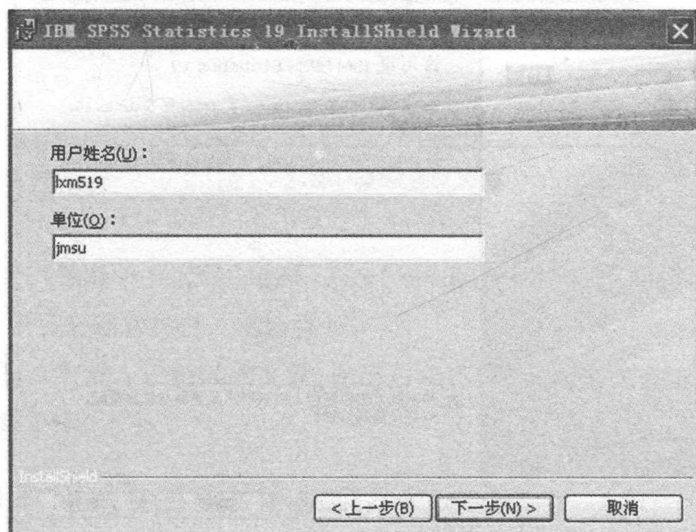


图 1-2-4 用户姓名和单位

8. 设置安装路径,可以接受系统默认的安装路径,也可以选择“更改”按钮,重新选择安装路径。如果重新选择安装路径,建议记住安装位置,以便后期注册时使用,如图 1-2-5 所示。

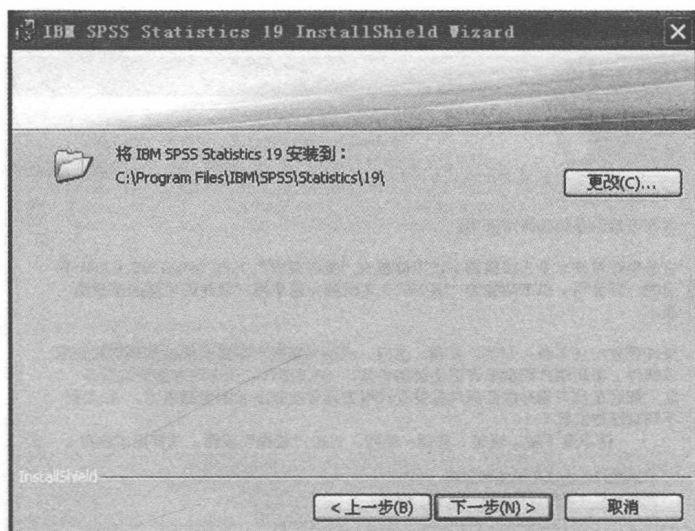


图 1-2-5 设置安装路径

9. 安装完成后,会提示对软件进行注册,可以通过软件配套的注册码进行注册,如果没有注册码可以启动免费试用,试用期一般为一个月,如图 1-2-6 所示。

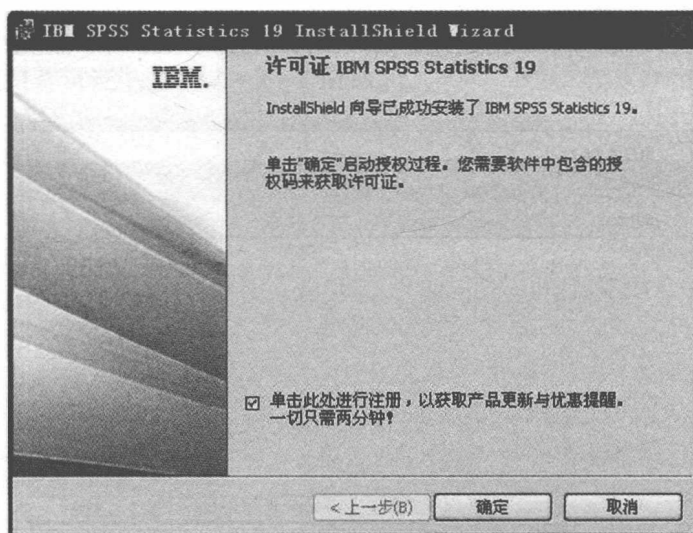


图 1-2-6 软件注册

10. 注册完成以后,软件就可以使用了。

二、IBM SPSS 软件的启动

(一) IBM SPSS 软件的启动方法

IBM SPSS 软件的启动方法有很多,下面介绍三种:

方法 1:双击桌面 IBM SPSS 快捷图标,会出现版本提示界面,如图 1-2-7 所示,之后,进入软件初始界面。

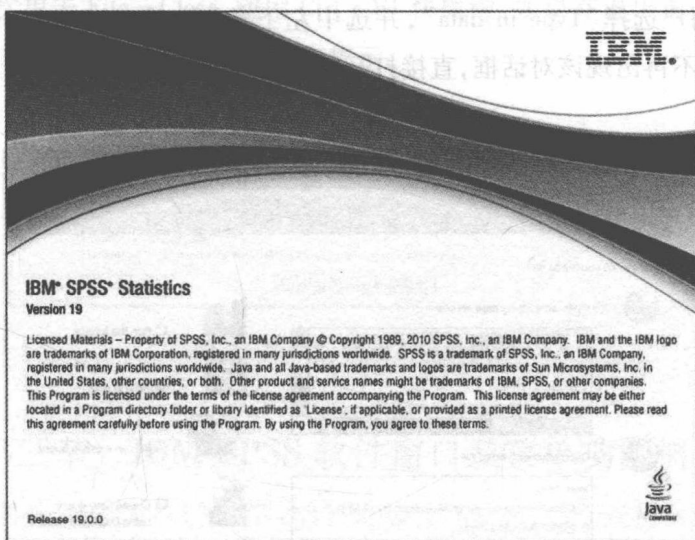


图 1-2-7 版本提示界面

方法 2:通过开始菜单—程序—IBM SPSS Statistics—IBM SPSS Statistics 19,也会出现版本提示界面,然后启动软件进入初始界面。

方法 3:通过现有的 .sav 格式的文件,即通过双击 SPSS 数据文件也可以直接启动软件并打开此数据文件。

(二) IBM SPSS 软件启动初始界面

初始界面由六个功能和一个复选框组成。界面提示:“What would you like to do?”它在提示:“你想要做什么?”如图 1-2-8 所示。

1. Run the tutorial,运行自学指导。选择此选项,可以找到与 SPSS 各类操作相关的信息。
2. Type in data,输入数据。选择此项,会出现数据编辑窗口,数据编辑区域是空着