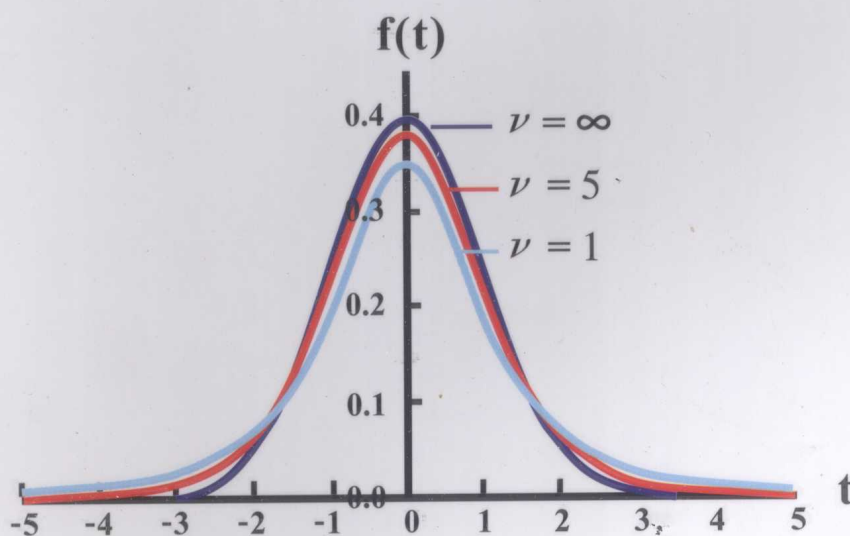


SPSS17.0 Yu Weishengtongjixue Yingyongzhinan

SPSS17.0

与卫生统计学 应用指南

◎ 主编 姚友平 ◎ 副主编 陶新 姚汝铨

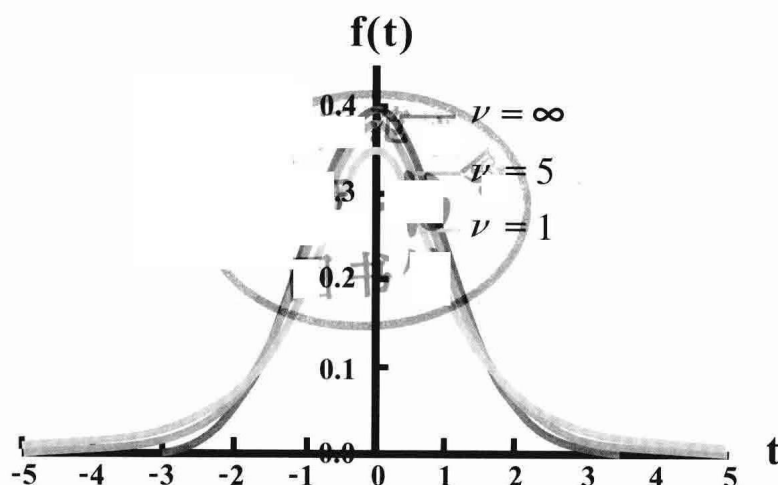


SPSS17.0 Yu Weishengtongjixue Yingyongzhinan

SPSS 17.0

与卫生统计学 应用指南

◎ 主编 姚友平 ◎ 副主编 陶新 姚汝铨



华中科技大学出版社

<http://www.hustp.com>

中国·武汉

图书在版编目(CIP)数据

SPSS 17.0 与卫生统计学应用指南/姚友平 主编. —武汉:华中科技大学出版社,2010 年 3 月

ISBN 978-7-5609-6094-4

I. S… II. 姚… III. 卫生统计-统计分析-软件包, SPSS 17.0-指南 IV. R195.1-39

中国版本图书馆 CIP 数据核字(2010)第 051670 号

SPSS 17.0 与卫生统计学应用指南

姚友平 主编

策划编辑:居 颖

责任编辑:余 涛

责任校对:周 娟

封面设计:范翠璇

责任监印:周治超

出版发行:华中科技大学出版社(中国·武汉)

武昌喻家山 邮编:430074 电话:(027)87557437

录 排:武汉佳年华科技有限公司

印 刷:湖北新华印务有限公司

开本:787mm×1092mm 1/16

印张:15.25

字数:357 000

版次:2010 年 3 月第 1 版

印次:2010 年 3 月第 1 次印刷

定价:36.00 元

ISBN 978-7-5609-6094-4/R·161

(本书若有印装质量问题,请向出版社发行部调换)

前 言

如果要求使用电视机的人必须懂得音频和视频信号的转换与传输原理,必须知道每一个电子元件的功能及其参数,那么就不会有如此多的人收看电视节目了。

在计算机没有普及的年代,统计学必须通过计算公式一步一步地计算出 t 值、 F 值、 χ^2 等统计量。复杂的计算公式、烦琐的计算步骤、深奥的数学理论使统计学成为一门难于掌握的工具学科。

在计算机已经普及的今天,SPSS、SAS 等高度专业化的统计软件已得到广泛应用。尤以 SAS 的声望最高,在世界上最著名的杂志上发表论文、在美国 FDA 审批药物均要求使用它作为统计分析工具,但由于它在运用中要求使用者编写处理程序,在掌握上有相当大的难度。而 SPSS 则不同,它有类似于 Windows 操作系统的界面,对于绝大多数统计处理均无需编写程序语句,通过对话框的操作就可实现。目前市面上有关 SPSS 软件的书籍一方面是由于受英语语句的影响,另一方面又受计算机专业术语表述的影响,其语句大多深奥难懂。传统的卫生统计学重在计算公式及其推导过程,必须十分关注手工计算统计量,公式的记忆和统计量计算的准确性要求,使学习者无暇顾及数据属性、分布类型、适用的统计分析方法等重要问题。

《现代统计学与 SAS 应用》的作者胡良平教授指出:“医学期刊中,误用统计学的现象十分严重。”有人认为我国 80% 医学杂志的论著中有统计学错误。多数错误的出现是由于作者不了解所研究数据的属性、分布类型,盲目套用统计分析方法,只要检测出“有显著性差异”就感到欣慰。

SPSS 就像一台电视机,它把统计量的计算公式、计算步骤“封装”成为程序模块。用好这台“电视机”,在医学科研中就会如虎添翼。

本书以最新版本的 SPSS 17.0 统计软件为工具,介绍卫生统计学的原理与方法。在 SPSS 的易用性基础上使用中文界面,更方便于理解和使用。在传统卫生统计学严谨理论和科学合理编排顺序的基础上,将复杂的数据计算过程交由 SPSS 软件完成,在理解统计学原理和高深的数理之间找到平衡点,避免了医学专业学生和医学工作者在高深数理方面所面临的困境,使其易用而不滥用。本书充分吸取和利用传统卫生统计学中经典、通俗的语言,避免晦涩难懂的语句,删除医学科研中派不上用场的软件功能介绍,极大地减少了阅读量和阅读难度。这些努力,但愿能解决学了卫生统计学而不会处理医学科研数据、学了 SPSS 软件而滥用统计分析方法的局面,以此提高医学科研论文的质量与可信度,推动我国医学科研工作的深入与普及。同时,也希望能借此推动卫生统计学教学模式的改革。

本书借助 SPSS 不仅解决了传统卫生统计学中的众多问题,而且还编入了传统模式无法讲解的、当今非常流行的、实用的多元统计分析方法。这对于全面提升医学专业学生和医学工作者对科研数据的处理能力非常重要。

学习卫生统计学的目的是使医学工作者在实际科研工作中运用统计学方法处理科研数据,而不是培养医学统计学方法的研究者或发明家。这一指导思想对于卫生统计学的编排方案起着重要作用,是提升医学科研数据处理水平和能力的必由之路。

本书得到华中科技大学同济医学院公共卫生学院余松林教授的阅评与作序,得到浠水县疾病预防控制中心领导和职工多方面的支持和帮助。在此,谨向他们致以衷心感谢!

由于编者的水平有限,本书难免存在错误之处,在统计学理论、统计学公式等的取舍上可能还存在一定的争议,诚挚地欢迎各位专家、学者的批评指教。

姚友平

2009 年 11 月

目 录

第一章 绪论	(1)
第一节 卫生统计学与 SPSS	(1)
第二节 统计学中的几个基本概念	(2)
第二章 SPSS 统计软件的应用基础	(4)
第一节 SPSS 数据编辑器	(4)
第二节 SPSS 查看器及其他窗口	(5)
第三节 常用的统计分析对话框及其使用	(5)
第四节 变量的分类、属性及其设置	(6)
第三章 刻度测量资料的描述性统计分析	(10)
第一节 描述集中趋势和离散程度的统计指标	(10)
第二节 刻度测量资料的频数表	(12)
第四章 正态分布及其应用	(17)
第一节 正态分布的概念和特征	(17)
第二节 正态分布的应用	(18)
第五章 总体均数的估计和假设检验	(21)
第一节 抽样研究与抽样误差	(21)
第二节 t 分布	(23)
第三节 总体均数的估计	(23)
第四节 假设检验的基本步骤	(26)
第五节 单样本均数的假设检验	(28)
第六节 两样本均数的假设检验	(29)
第七节 配对 T 检验	(31)
第八节 第一类错误和第二类错误	(33)
第九节 假设检验应注意的问题	(33)
第六章 方差分析	(35)
第一节 方差分析的基本思想	(35)
第二节 单因素方差分析	(37)
第三节 两因素方差分析	(41)
第四节 多因素方差分析及其实验设计介绍	(48)
第五节 协方差分析	(59)
第六节 多元方差分析	(62)
第七节 重复测量数据的方差分析	(66)
第八节 变量变换	(71)
第七章 分类资料的统计描述	(73)
第一节 常用相对数	(73)

第二节 应用相对数应注意的问题	(74)
第三节 率的标准化法	(75)
第四节 动态数列及其分析指标	(78)
第八章 二项分布、Poisson 分布及其应用	(80)
第一节 二项分布的概念与特征	(80)
第二节 率的可信区间	(83)
第三节 Poisson 分布的概念与特征	(83)
第四节 Poisson 分布的应用	(85)
第九章 卡方(χ^2)分布与交叉表卡方检验	(88)
第一节 χ^2 分布的特征	(88)
第二节 交叉表卡方检验	(88)
第三节 配对资料的 χ^2 检验	(94)
第四节 分层卡方检验	(96)
第十章 非参数检验	(98)
第一节 非参数卡方检验	(98)
第二节 二项式检验	(100)
第三节 游程检验	(101)
第四节 单样本 K-S 检验	(102)
第五节 两个独立样本与多个独立样本检验	(103)
第六节 两个相关样本与多个相关样本检验	(105)
第七节 非参数检验中需要注意的问题	(107)
第八节 秩变换分析方法	(107)
第十一章 直线回归与相关	(110)
第一节 直线回归	(110)
第二节 直线相关分析	(121)
第三节 曲线拟合	(125)
第四节 多元直线回归分析	(128)
第十二章 Logistic 回归分析	(131)
第一节 二项 Logistic 回归模型	(131)
第二节 “有序”Logistic 回归模型	(143)
第三节 Probit 分析	(147)
第十三章 生存分析	(151)
第一节 生存分析的基本概念	(151)
第二节 寿命表	(152)
第三节 Kaplan-Meier 生存分析	(155)
第四节 Cox 回归分析	(160)
第十四章 判别分析	(166)
第一节 判别分析的模型及用途	(166)
第二节 常用判别分析方法	(166)

第三节 实例分析·····	(168)
第四节 判别分析对话框介绍·····	(173)
第十五章 聚类分析·····	(177)
第一节 K 均值聚类分析 ·····	(177)
第二节 系统聚类分析·····	(180)
第十六章 统计表和统计图·····	(187)
第一节 常用统计表·····	(187)
第二节 用“OLAP 立方”生成计量资料汇总表 ·····	(188)
第三节 用“个案汇总”生成计量资料的汇总表·····	(191)
第四节 常用统计图·····	(193)
第十七章 实验设计·····	(199)
第一节 实验设计的特点及分类·····	(199)
第二节 实验设计的基本要素·····	(199)
第三节 实验设计的基本原则·····	(202)
第四节 常用的实验设计方法·····	(204)
第五节 实验结果的误差与偏倚·····	(210)
第六节 样本含量的计算·····	(214)
第十八章 调查设计·····	(224)
第一节 调查研究的特点·····	(224)
第二节 调查设计的基本原则与内容·····	(224)
第三节 常用的抽样方法·····	(227)
第四节 调查的质量控制·····	(228)
参考文献·····	(230)

第一章 绪 论

第一节 卫生统计学与 SPSS

1. 卫生统计学的定义

统计学是一门处理数据中变异性的科学,内容包括收集、分析、解释和表达数据,目的是获得最大可能性的结果。应该强调的是,收集、归类、分析和解释大量数据的各个环节都非常重要,如果在实际工作中,忽略了设计、收集和归类(整理)中的任何一个环节,所得到的数据是不能用于统计学分析的。卫生统计学是应用统计学的原理与方法,研究人体健康状况及卫生服务领域中数据的收集、整理和分析的一门科学。

2. 关于 SPSS

数据的统计学处理是一项十分复杂的数据演算过程,尤其是近年来兴起的多元统计方法,不是手工所能计算的。在计算机科学技术快速发展的今天,我们没有理由不去使用计算机和高度专业化的统计软件来代替手工计算。卫生统计学作为统计学的一门应用型学科,也没有理由去纠缠于统计学公式及其计算过程。本书将统计软件 SPSS 作为一种工具,使医学工作者从复杂的数理推导、计算中解脱出来,重点研究医学数据的收集(医学科研设计)和整理、统计方法的运用,以及统计推断的方法、结论和意义。

SPSS 软件是公认的最优秀的统计分析软件之一,它作为统计分析工具,理论严谨、内容丰富,具有数据管理、统计分析、趋势研究、生成统计表与统计图、文字处理等功能。

SPSS 具有类似于 Windows 的视窗操作界面,把难以学习和掌握的程序语句隐藏于一见便知的视窗之后,极大地降低了使用的难度,故以易于掌握和使用为其鲜明的特色。SPSS 自开发以来推出了多个版本,本书使用的是最新版本 SPSS 17.0 for Windows(以下简称 SPSS),也是 SPSS 第一个带中文的版本。

把这样一个为人们所能很快上手的统计软件与卫生统计学有机地结合,人们再也不需要去学习和掌握统计公式的推导和记忆及复杂的数学计算过程,也不需要计算出来的统计量(如 F 值、卡方值、 T 值等)去查找相应的概率。这对于需要记忆大量医学知识的医学专业学生和医学工作者来说,是一种极大的减负。

3. SPSS 17.0 与卫生统计学应用指南的内容

(1) SPSS 界面介绍、数据属性及其设置。

(2) 卫生统计学的基本原理和方法:统计描述,如定量资料和分类资料的描述性指标及常用统计图表;常见的理论分布及其应用,如正态分布、二项分布与 Poisson 分布;总体参数的估计,如总体均数、总体率和总体平均数;假设检验,如 u 检验、 t 检验、方差分析、 χ^2 检验、非参数检验等;回归与相关;多元统计,如多元线性回归、Logistic 回归、聚类分析、判别分析。

第二节 统计学中的几个基本概念

1. 同质与变异

同质是指被研究指标的影响因素完全相同。但在医学研究中,有些影响因素(如遗传、营养等)往往是难以控制的,甚至是未知的。所以,统计学中常把同质理解为对研究指标影响较大的、可以控制的主要因素尽可能要求相同,而对不易控制的影响因素忽略。

同质基础上的个体差异称为变异。客观世界充满了变异,生物医学领域更是如此。有了变异,就需要有统计学处理。

2. 总体与样本

任何统计研究都必须首先确定观察对象,也称个体或观察单位。观察对象是统计研究中最基本的单位,可以是一个人、一个家庭、一个地区、一个样品、一个采样点等。

总体是根据研究目的确定的同质观察对象的全体,或者说,是所有同质观察对象某种观察值(变量值)的集合。总体分为有限总体和无限总体。有限总体是指在某特定的时间与空间范围内,同质研究对象的所有观察单位的某变量值的个数为有限个。无限总体是抽象的,无时间和空间的限制,观察单位的数量是无限的。

样本是按照随机化原则,从总体中抽取的有代表性的部分观察对象的变量值的集合。抽样研究的目的是用样本信息推断总体特征。样本含量也称观察对象数,或样本大小,是指样本中所包含的观察对象或试验对象的数量。

统计学是用样本的信息去认识总体信息的工具,能帮助人们设计与实施如何从总体中科学地抽取样本,使样本中的观察对象数恰当,信息丰富,代表性好;能帮助人们挖掘样本中的信息,推断总体的规律性。

3. 资料与变量

总体确定之后,研究者要对每个观察对象的某项特征进行测量或观察,这些要测量或观察的特征称为变量,也称为变数、随机变量等。如身高、体重、性别、血型、疗效等均是变量。变量的测定值或观察值称为变量值或观察值;变量测定值的集合称为资料。

4. 随机事件与概率

医学研究的现象大多数是随机现象,对随机现象进行实验或观察称为随机试验。随机试验的各种可能结果的集合称为随机事件,也称偶然事件,简称事件。对于随机事件来说,在一次随机试验中,某个随机事件可能发生也可能不发生,但在一定数量的重复试验后,该随机事件的发生情况是有规律的。概率是描述随机事件发生的可能性大小的数值,常用 P 表示。

例如,投掷一枚均匀的硬币,随机事件 A 表示“正面向上”,用 n 表示投掷次数, k 表示随机事件 A 发生的次数, P 表示随机事件 A 发生的概率,显然 $P=k/n, 0 \leq k \leq n, 0 \leq P \leq 1$ 。当投掷次数 n 足够大时, $P=k/n \rightarrow 0.5$,即得到硬币正、反两面的结果各为 50%,称 $P(A)=0.5$,或简写为 $P=0.5$ 。当 n 足够大时,可以用 k 估计总体的 P 。

随机事件概率的大小在 $0 \sim 1$ 之间,即 $0 < P < 1$,常用小数或百分数表示。 P 越接近 1,表示某事件发生的可能性越大; P 越接近 0,表示某事件发生的可能性越小。 $P=1$ 表示事件必然发生, $P=0$ 表示事件不可能发生,它们是确定性的,不是随机事件,但可以把它看成随机事件的特例。

若随机事件 A 的概率 $P(A) \leq \alpha$,习惯上当 $\alpha=0.05$ 时,就称 A 为小概率事件。其统计学

意义是小概率事件在一次随机试验中不可能发生。“小概率”的标准 α 是人为规定的,对于可能引起严重后果的事件,如术中大出血等,可规定 $\alpha=0.01$,甚至更小。

5. 随机化与抽样误差

能使总体中每一观察对象以同等机会(概率)进入样本,或分配到实验组或对照组的过程,称为随机化。随机化的特点是机会均等地选取或分配研究对象的公平性法则。由于组成总体的各个个体间存在着差异性,因抽样过程的随机性导致样本的统计量与总体的参数不等,样本与样本的统计量之间存在差异的特性称为抽样误差。

6. 参数与统计量

描述总体特征的数量称为参数,如总体均数(μ)、总体标准差(σ)和总体率(π)等。描述样本特征的数量称为统计量,如样本均数($\bar{X}$)、样本标准差(Std, s)、样本率(P)。

第二章 SPSS 统计软件的应用基础

熟悉 SPSS 统计软件的界面及其操作方法是利用这一平台学习、掌握统计学处理方法的基础。SPSS 统计软件有数据编辑器、结果输出的查看器、语法编辑器、图表编辑器和帮助等五个窗口。前三个窗口有基本相同的菜单,利用这些菜单可以打开执行统计分析任务的多个对话框。

第一节 SPSS 数据编辑器

1. SPSS 的启动及数据编辑界面

在 Windows 操作系统(以 Windows XP 为例,下同)中,鼠标点击开始/程序/SPSS Inc/Statistics 17.0/SPSS Statistics 17.0,即可启动 SPSS 17.0。SPSS 启动成功后出现 SPSS 的封面,片刻后封面消失,呈现 SPSS Statistics 数据编辑器,如图 2-1 所示。

数据编辑器由“数据视图”和“变量视图”两个视图组成,两个视图可通过点击窗口左下角的“数据视图”和“变量视图”选项卡、点击菜单视图/变量或视图/数据、**Ctrl** + **T** 键三种方式之一进行切换。

“数据视图”中自上而下有:标题栏、菜单栏、工具栏、编辑数据显示栏、数据二维表、视图切换选项卡和状态栏。位于中部的数据二维表是一个可扩展的表格,由若干行和列组成,每一行对应一条记录,每一列对应一个变量,其中可进行数据的录入、修改、删除等数据编辑。



图 2-1 SPSS“数据视图”

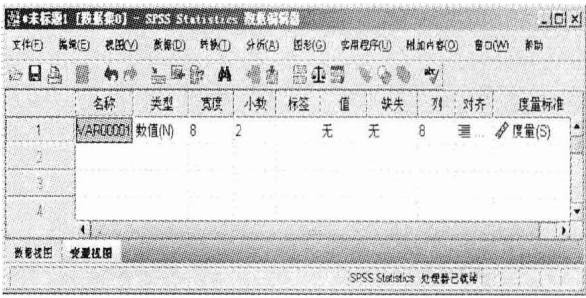


图 2-2 SPSS“变量视图”

注 SPSS 17.0 是第一个具有中文界面的版本,在安装时应注意选择安装的语言选项为 simplified Chinese,安装了简体中文版后,初次显示的仍然是英文界面。显示中文的设置方法为:在“Data Editor”窗口选择 Edit/Options 菜单,打开“Options”对话框,单击“General”选项卡;在“Output”组合框中,单击“Language”右侧的下拉箭头,选择“simplified Chinese”选项;在“User interface”(用户界面)组合框中,单击“Language”右侧的下拉箭头,选择“simplified Chinese”,单击 **Apply** 按钮。

2. 数据编辑器的变量视图

点击数据编辑器左下角的“变量视图”选项卡,切换到变量视图,如图 2-2 所示。该界面中

也是一个二维表格,每一行对应一个变量,每一列对应变量的一个属性。在数据录入前,应当对这些属性进行设置。

第二节 SPSS 查看器及其他窗口

1. SPSS 查看器

SPSS 查看器(见图 2-3)是在打开数据文件时产生的,用于记录一些重要的操作过程、出错信息和显示统计分析结果。也可以通过菜单中的文件/打开/输出和文件/新建/输出菜单打开和新建一个查看器。SPSS 查看器的左侧是一个大纲视图,主要用于方便浏览右侧视图中的信息。当完成一项处理后,在该窗口显示处理过程和计算结果等信息。这些信息可以通过单击选中,如果双击选中的部分,则可以对其中的内容进行编辑。

2. SPSS 语法编辑器

SPSS 语法编辑器是 SPSS 的辅助性功能窗口,可利用该窗口进行 SPSS 命令的输入、编辑和运行。由于其功能超出本书深度,故不作详细介绍。

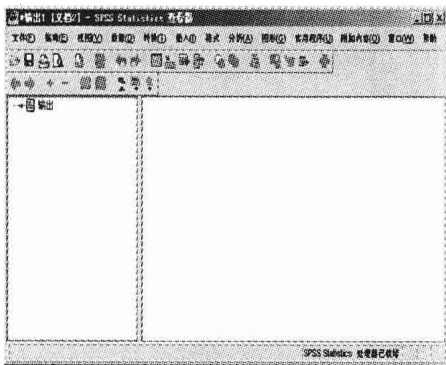


图 2-3 SPSS“查看器”

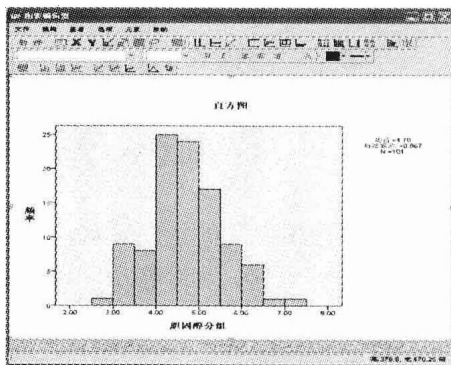


图 2-4 “图表编辑器”

3. 图表编辑器

在 SPSS 查看器中,双击所要编辑的图形,则可以打开图表编辑器,如图 2-4 所示。在此窗口中可以对统计图形进行修改、编辑。

4. 帮助窗口

SPSS 提供了丰富、详尽的多种语言的帮助信息。帮助窗口可由下列两种途径打开。

(1) 主窗口的“帮助”菜单:在软件运行的任何时候,点击“帮助”菜单,选择相关的子菜单,可得到所需的各种帮助。

(2) 各种对话框中的 **帮助** 按钮:在具体操作过程中,当打开某一对话框时,一般总有 **帮助** 按钮,点击该按钮,可得到这一对话框相关内容的详细帮助。

第三节 常用的统计分析对话框及其使用

“打开数据”对话框和“将数据保存为”对话框属文件操作对话框,与 Windows 文件操作对话框相同,故不作介绍。

凡 SPSS 数据编辑器的菜单栏中所有带省略号“...”的菜单项,均打开一个执行相应功能的对话框,通过点击对话框中的某些按钮,还可以打开二级对话框。在数据编辑器的“分析”菜

单下的每一个菜单项对应一个统计分析方法,点击一个菜单项就打开一个对话框,通过选择对话框中的选项完成相应的统计分析任务。如点击菜单分析/描述统计/描述...,打开“描述性”对话框,如图 2-5 所示。

点击图 2-5 中的“选项...”按钮,打开“描述:选项”对话框,如图 2-6 所示。在 SPSS 中,不同的统计分析方法有不同的对话框,但这两种类型的对话框在 SPSS 中较为典型。

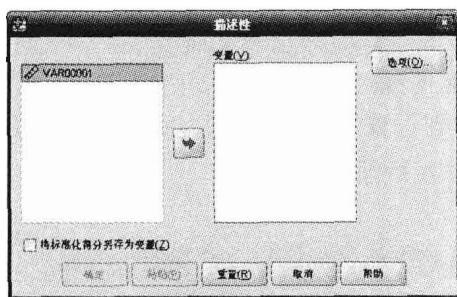


图 2-5 “描述性”对话框

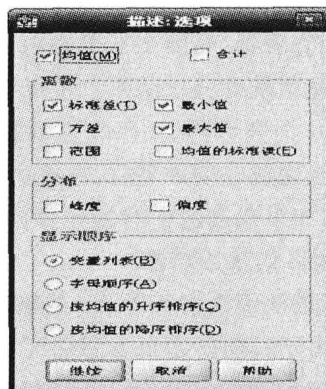


图 2-6 “描述:选项”对话框

第四节 变量的分类、属性及其设置

变量是统计学研究的主要对象,它也像物质世界的任何其他事物一样,具有不同的类别和属性,掌握变量的类别、属性等基本知识是正确应用 SPSS,正确选择统计分析方法的前提。

1. 变量的度量标准

变量的度量标准不同,其分布规律不同,对它们采用的统计分析方法也不同。在统计分析开始之前,首先要分清变量的资料类型。按变量的值是定量的还是定性的不同,或者是按变量的度量标准不同,将变量分为下列三种。

1) 刻度测量变量

刻度测量变量有两个特点。一个特点是变量值是以数值表示的。有的是经测量取得的数值,其值为任意可能的连续型数值,也可表现为带小数的数值,如身高、体重、血压等;而有的刻度测量变量测定值只能是正整数,如脉搏、白细胞计数等。所以,在卫生统计学中刻度测量变量也称为连续型变量。另一个特点是有度量衡单位。如身高为米、厘米;体重为千克;血压为 mmHg、kPa;脉搏为次/分;白细胞计数为 $\times 10^9$ 个/L 等。这种由数值变量的测量值构成的资料称为刻度测量资料,也称为定量资料。

2) 有序测量变量

有序测量变量在过去的统计学书籍中又称为有序分类变量,其变量值是定性的,在各类别之间有一定程度上的差别。如尿糖化验结果按一、±、+、++、+++ 分类;疗效按治愈、显效、好转、无效分类。有序测量的资料称为等级资料。

3) 名义测量变量

名义测量变量又称为标定测量变量、无序分类变量,其变量值在各类别或属性之间无程度和顺序上的差别,表现为互不相容的类别。名义测量变量的资料称为无序分类资料,也称为计数资料。它又可分为二项分类资料,如性别(男、女)、药物反应(阴性、阳性)等,以及多项分类

资料,如血型(O、A、B、AB)、职业(工、农、商、学、兵)等。

变量的类型是可以改变的。根据研究目的的需要,刻度测量变量可以转化为有序测量变量和名义测量变量。例如,血红蛋白量(g/L)原是数值变量(刻度测量级别),若按血红蛋白正常与偏低分为两类时,即成为二项分类资料;若按重度贫血、中度贫血、轻度贫血、正常、血红蛋白增高分为五个等级时,即成为有序分类资料。但是,这种数据类型的转化会降低统计分析的效能。有时也可将分类资料数量化,如将病人的恶心反应的有、无(二项分类)资料,用 0、1 来表示,则成为数值变量资料(定量资料)。

2. 变量的数据属性及其设置

在 SPSS 中,变量具有名称、类型、宽度、小数位数、标签、数值、缺失值、列宽、对齐方式和度量标准等 10 个方面的属性(或称为格式),前 9 个属性对数据在电脑中的存储起着重要作用,最后一个属性对数据资料采用何种统计分析方法起着重要作用。

1) 变量名称的设置

变量名长度一般为 8 位,SPSS 17.0 以英文字母或拼音字母命名,也可以使用中文变量名。如果在数据录入前没有设置变量名,系统则以 var 加一个 5 位数字作为默认的变量名。

2) 变量类型的设置

选中“类型”列的单元格时,该单元格中右侧会出现如图 2-7 所示的图形,单击它会弹出“变量类型”对话框,如图 2-7 所示。

“变量类型”对话框的左侧列出了 8 种数据类型,可以通过点击选中其中一个选项。

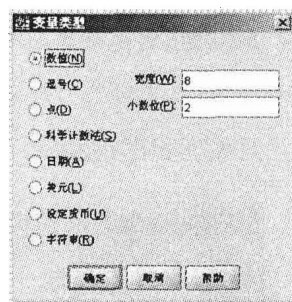


图 2-7 “变量类型”对话框

(1) 数值 选中后系统默认为刻度测量变量。在对话框的右侧可以定义数值的宽度和小数位。宽度默认为 8 个字符,可以根据实际需要进行增减,它是整数部分、小数点、小数部分的位数之和,小数位数默认为 2 位。

(2) 逗号 加显逗号的数值型变量,即整数部分每 3 位数加一逗号,其余定义方式同数值型。

(3) 点 3 位加点数值型变量,无论数值大小,整数位均以每 3 位加一小点(但不是小数点)标志;可定义小数位数,小数点用逗号表示。例如,1234.50 显示为 1.234,50(实际是 1.2345E3)。

(4) 科学计数法 科学记数型在对话框的右侧可以定义数值宽度和小数位数,在数据编辑器、查看器等窗口中以指数形式显示。例如,定义数值宽度为 9,小数位数为 2,则 345.678 显示为 3.46E+02。

(5) 日期 对于日期型变量,当选中该项时,在对话框的右侧列出全部可供选择的日期格式,可根据需要点击选取其中一种。例如,选择 mm/dd/yyyy 格式,则 2009 年 6 月 25 日显示为 06/25/2009。

(6) 美元 对于美元货币型变量,可定义数值宽度和小数位数,显示形式为数值前有 \$。

(7) 设定货币 显示为整数部分每 3 位加一逗号,用户可定义数值宽度和小数位数。例如,12345.67 显示为 12,345.67。

(8) 字符串 对于字符串型变量,选中后,无论输入的是字符还是数字,系统均将其作为

字符处理,默认为名义测量变量。在对话框的右侧可以定义字符长度。

选择完成后,点击[确定]按钮返回。

3) 变量宽度的设置

变量宽度的设置一般取默认值,如数值型、字符串型均为 8,必要时可做更改。

4) 小数位数的设置

小数位数一般取默认值,也可以根据显示小数位数的要求进行更改,显示的小数根据设定进行了四舍五入。但无论小数位数显示如何,并不影响数据的实际精度。

5) 变量标签的变量

变量标签是变量名的别名设置。别名是为了更好地表达变量名的含义而设计的,最好使用较为习惯的中文作别名。

6) 变量值标签的变量

点击某变量与“值”对应的单元格,该单元格中右侧会出现如[...]的图形,单击它会弹出“值标签”对话框,如图 2-8 所示。

如果变量值是重复的、有限的几个数值,可以将这些值用中文来表示,这样在数据录入时显得特别方便,出现录入错误时也很容易被发觉。其操作方法是,分别在名称为“值”的框中输入变量的值,在名称为“标签”的框中输入想要输入的中文文字,点击[添加]按钮;依次按上述方法输入其他值和标签名,最后点击[确定]按钮。

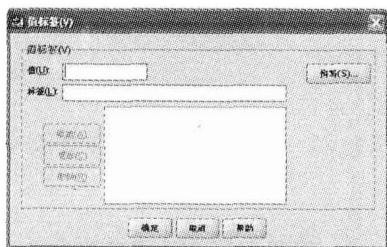


图 2-8 “值标签”对话框

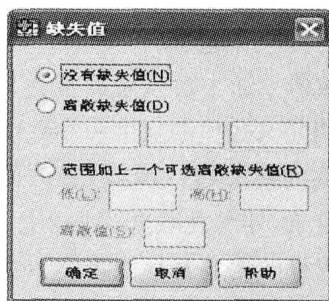


图 2-9 “缺失值”对话框

7) 缺失值的设置

缺失值是在调查或实验中所产生的空白值或漏测值。在统计学中,缺失值不能作为数字“0”来处理,因此必须以一种特定的方式来表示。在图 2-2 中点击某变量与“缺失”对应的单元格,该单元格中右侧会出现如[...]的图形,单击它会弹出“缺失值”对话框,如图 2-9 所示。

第一个单选钮为“没有缺失值”设置;选中第二个单选钮时,可以在其下面的三个框中分别输入三个不同的缺失值;选中第三个单选钮时,可以在其下面的两个框中分别输入一个下限值和一个上限值,这两个值间的任何值均是缺失值。

8) 列宽的设置

数据列宽默认为 8,当数据值超过这一宽度时,变量在数据录入界面中不能全部显示出来,这时可以调整其宽度。

9) 对齐方式的设置

对齐方式可以选取居中、靠左、靠右三种,以便于在数据编辑器中查看。

10) 度量标准

系统默认是将字符型变量设置为名义测量精度,将数值型变量作为刻度测量精度。为了便于统计分析处理,应将数值型变量明确指定为是刻度测量精度,还是有序测量精度。设置方法是:在变量视图(见图 2-2)中,点击某“变量”行对应的“度量标准”列的单元格,该单元格中左侧出现一个向下箭头,点击箭头,在弹出的下拉列表中选中需要的度量标准,该单元格中出现相应的文字和徽标。