

# 医学统计学 与 SPSS应用

供医学类本科各专业用

主 编 李娟生



兰州大学出版社

医学统计学  
SPSS应用



供医学类本科各专业用

# 医学统计学 与 SPSS应用

主 编 李娟生

副主编 申希平



兰州大学出版社

图书在版编目 (CIP) 数据

医学统计学与 SPSS 应用/李娟生主编. —兰州:

兰州大学出版社, 2011. 7

ISBN 978-7-311-03711-6

I. ①医… II. ①李… III. ①医学统计—统计分析—  
软件包, SPSS—教材 IV. ①R195. 1-39

中国版本图书馆 CIP 数据核字 (2011) 第 136382 号


策划编辑 陈红升  
责任编辑 郝可伟 陈红升  
封面设计 刘杰

---

书 名 医学统计与 SPSS 应用  
作 者 李娟生 主编  
出版发行 兰州大学出版社 (地址:兰州市天水南路 222 号 730000)  
电 话 0931-8912613(总编办公室) 0931-8617156(营销中心)  
0931-8914298(读者服务部)  
网 址 <http://www.onbook.com.cn>  
电子信箱 [press@lzu.edu.cn](mailto:press@lzu.edu.cn)  
印 刷 甘肃北辰印务有限公司  
开 本 787×1092 1/16  
印 张 9.5  
字 数 221 千  
版 次 2011 年 8 月第 1 版  
印 次 2011 年 8 月第 1 次印刷  
书 号 ISBN 978-7-311-03711-6  
定 价 20.00 元

---

(图书若有破损、缺页、掉页可随时与本社联系)

 本书由兰州大学教材建设基金资助

主 审 白亚娜

参 编 (按姓氏音序排序)

胡晓斌 兰州大学

李娟生 兰州大学

刘小宁 兰州大学

秦林原 兰州大学

任晓卫 兰州大学

申希平 兰州大学

杨 燕 甘肃省人民医院

岳 雄 甘肃省妇幼保健院

# 前 言

医学统计学作为医学生建立统计分析思维、进行医学科研及处理医学资料的工具课程是必不可少的。对一个医学生来说,掌握了医学统计方法,就意味着在从事医学专业的同时,能够将医学实践与医学研究科学地结合起来,为更好地探索医学难题奠定基础,而一本精简而不失其应用功能的医学统计学教材对医学生而言无疑是一得力的助手。计算机科学技术的迅猛发展,医学统计软件包的日益完善和广泛应用,大大促进了医学统计学的发展,使繁杂的统计运算得以轻松地实现;因此,掌握好统计软件技能又成为医学生的另一任务。将统计学方法与统计软件有机地结合起来,使其更具实用效果,是我们编写此教材的根本思想。

SPSS即统计产品和服务解决方案(Statistical Product and Service Solutions, SPSS),是世界最著名的统计分析软件之一。SPSS最鲜明的特点是易学易用,其应用范围非常广泛,对医学生也非常适宜。本书的配套软件SPSS兼容于各版本。

可以说有关SPSS的教材或参考书不少,也不乏优秀的作品,但大多数是以单纯地介绍SPSS的内容和操作为主,而没有相应的统计学方法与之配套,应用起来不太方便,特别是那些不熟悉统计概念、不能正确选择统计方法的用户更会对之感到困惑和不知所措。因此,本书的总体编写思路是:首先阐明医学统计学理论与方法,内容尽可能通俗易懂,然后介绍SPSS软件应用方法,从基本的操作方式、数据文件的建立、结果文件的输出、结果的统计分析到结论等都给予完整、清晰的解释,为学生上机应用、完成实验操作提供了方便。

本书共有11章内容,包括数据类型与各种类型的统计学描述;数据的统计推断;简单线性回归与相关统计图表等。书中应用了大量的图、表等多种描述形式,使内容更加生动、形象、通俗易懂。每一种统计学方法均配有研究实例,并在SPSS中得以实现。因此,通过学习本书的内容,医学生不仅可以学到统计学方法,而且通过对实例的学习,可以利用SPSS解决有关数据的分析问题。

本教材可供各类医学本科学生使用。本书既可作为医学统计学理论学习的参考书,也可作为SPSS数据分析的实习教材,还可以作为其他SPSS版本教学与科研的参考书。

编者

2011年5月

## 目 录

|                                       |    |
|---------------------------------------|----|
| <b>第一章 医学统计学概念与 SPSS 统计软件简介</b> ..... | 1  |
| 第一节 医学资料的类型 .....                     | 1  |
| 第二节 医学统计学中的基本概念 .....                 | 2  |
| 第三节 医学统计学工作的基本步骤 .....                | 3  |
| 第四节 SPSS 统计软件简介 .....                 | 4  |
| <b>第二章 定量资料的统计描述</b> .....            | 13 |
| 第一节 频数分布表与频数分布图 .....                 | 13 |
| 第二节 集中趋势指标 .....                      | 15 |
| 第三节 离散趋势指标 .....                      | 18 |
| 第四节 正态分布及其应用 .....                    | 20 |
| 第五节 定量资料统计描述在 SPSS 中的实现 .....         | 23 |
| <b>第三章 总体均数的估计与假设检验</b> .....         | 28 |
| 第一节 均数的抽样误差与标准误 .....                 | 28 |
| 第二节 $t$ 分布 .....                      | 29 |
| 第三节 总体均数的估计 .....                     | 30 |
| 第四节 假设检验 .....                        | 31 |
| 第五节 $t$ 检验与 $Z$ 检验 .....              | 33 |
| 第六节 $t$ 检验的注意事项 .....                 | 36 |
| 第七节 假设检验中的两类错误 .....                  | 37 |
| 第八节 $t$ 检验在 SPSS 中的实现 .....           | 37 |
| <b>第四章 方差分析</b> .....                 | 42 |
| 第一节 完全随机设计的方差分析 .....                 | 42 |
| 第二节 随机区组设计的方差分析 .....                 | 45 |
| 第三节 方差分析在 SPSS 中的实现 .....             | 47 |
| <b>第五章 定性资料的统计描述</b> .....            | 53 |
| 第一节 频数分布表和频数分布图 .....                 | 53 |
| 第二节 常用相对数 .....                       | 54 |

|             |   |            |
|-------------|---|------------|
| 第三节         | 应用相对数时的注意事项 .....                         | 55         |
| 第四节         | 率的标准化法 .....                              | 56         |
| <b>第六章</b>  | <b>定性资料的统计推断 .....</b>                    | <b>59</b>  |
| 第一节         | 总体率估计 .....                               | 59         |
| 第二节         | 率的假设检验 .....                              | 60         |
| <b>第七章</b>  | <b><math>\chi^2</math> 检验 .....</b>       | <b>62</b>  |
| 第一节         | $\chi^2$ 检验的基本思想 .....                    | 62         |
| 第二节         | 独立样本资料的 $\chi^2$ 检验 .....                 | 62         |
| 第三节         | 配对样本资料的 $\chi^2$ 检验 .....                 | 67         |
| 第四节         | 行列表资料 $\chi^2$ 检验注意事项 .....               | 69         |
| 第五节         | $\chi^2$ 检验在 SPSS 中的实现 .....              | 70         |
| <b>第八章</b>  | <b>秩和检验 .....</b>                         | <b>76</b>  |
| 第一节         | 配对设计的 Wilcoxon 符号秩检验 .....                | 76         |
| 第二节         | 完全随机设计两样本比较的 Wilcoxon 秩和检验 .....          | 77         |
| 第三节         | 完全随机设计多个样本比较的 Kruskal-Wallis $H$ 检验 ..... | 79         |
| 第四节         | 秩和检验在 SPSS 中的实现 .....                     | 81         |
| <b>第九章</b>  | <b>两变量间的线性相关与回归 .....</b>                 | <b>87</b>  |
| 第一节         | 两变量间的线性相关 .....                           | 87         |
| 第二节         | 两变量间的线性回归 .....                           | 90         |
| 第三节         | 两变量间的线性相关与回归在 SPSS 中的实现 .....             | 94         |
| <b>第十章</b>  | <b>统计图表 .....</b>                         | <b>98</b>  |
| 第一节         | 统计表 .....                                 | 98         |
| 第二节         | 统计图 .....                                 | 101        |
| 第三节         | 统计表在 Word 文档中的制作 .....                    | 101        |
| 第四节         | 常用统计图在 SPSS 中的实现 .....                    | 103        |
| <b>第十一章</b> | <b>调查研究设计 .....</b>                       | <b>112</b> |
| 第一节         | 调查研究概述 .....                              | 112        |
| 第二节         | 调查研究的设计与实施 .....                          | 114        |
| 第三节         | 调查研究的组织保障与质量控制 .....                      | 132        |
| <b>参考文献</b> | <b>.....</b>                              | <b>134</b> |
| <b>附表</b>   | <b>.....</b>                              | <b>135</b> |



# 第一章 医学统计学概念与 SPSS 统计软件简介

当前,医学统计方法在医学研究中的应用越来越广泛,越来越深入,一个好的医学研究设计或一份医学研究报告都是离不开统计学方法的。正确地把握医学统计理论和方法这个坚实的基础,就能以科学、严谨的手段探索医学领域中的未知现象,揭示和阐明现象变化的规律,并指导医学的理论和实践。

医学统计学是应用概率论和数理统计的基本原理和方法,结合医学实际,对研究资料和信息进行搜集、整理与分析的一门学科。要学好医学统计学,首先应该对基本的概念有很好的领会和理解。本章对一些常用的、重要的、贯穿于整个统计学的基本概念及定义进行系统的介绍。同时,本章也将对 SPSS 统计软件的基本功能和有关概念作简要的介绍。

## 第一节 医学资料的类型

医学资料是通过测量或观察所研究的观察单位的某项特征指标得到的数据。这种特征指标能表现观察单位的变异性,又称为变量(Variable)。变量的观察值形式是决定变量类型的关键。根据观察值的形式可将医学数据分为定量资料和定性资料两大类。

### 一、定量资料(Quantitative Data)

又称数值变量资料,即变量的观察值(变量值)的大小是数值化的,通过用定量的方法,测量得到的数值大小,具有度量衡单位。这种变量,常作为统计分析变量。如:身高(cm)、体重(kg)、血压(kPa)等。

### 二、定性资料(Qualitative Data)

又称分类变量资料,即变量的观察值(变量值)的大小是定性化的,可按照性质或类别清点个数,进行分组,没有度量衡单位。这种变量,常作为统计分组变量。

定性资料还可以分为:

1. 计数资料(Enumeration Data)或无序分类变量(Unordered Categorical Variable)资料。即不同性质或类别的变量值间无程度高低之分,相互独立。如某人群的血型变量,变量值可按照 A、B、AB 与 O 四种分布类型分组。

2. 等级资料(Ranked Data)或有序分类变量(Ordinal Categorical Variable)资料。即不同性质或类别的变量值间具有程度高低之分,其变量值表现出等级大小或属性程度,具有半定量性质。如某人群某血清反应变量,其结果可分为-、±、+、++、+++、++++六个级别;又如某疗效变量,其结果可分为治愈、显效、好转、无效四个级别。

不同数据类型,所用统计方法是不一样的。因此,要能够正确地判断数据的类型以便正确地确定统计分析方法。

## 第二节 医学统计学中的基本概念

### 一、总体与样本

总体 (Population) 是根据研究的目的确定的同质观察单位数的全体。观察单位 (Observed Unit), 亦称个体 (Individual), 是统计研究中最基本的单位。如研究对象是个人, 其观察单位即为一个人; 如研究对象是动物, 其观察单位即为一只 (头) 动物; 同样地, 如研究对象是某一类人 (一个家庭、一个班级、一个自然村等), 则其观察单位即为这一类人。

总体包括无限总体和有限总体。无限总体即总体当中的观察单位数虽理论上存在, 但无法确定这一数值到底为多少; 有限总体即总体当中的观察单位数目是有限的。对一个总体来讲, 无论是有限的, 还是无限的, 都由于各种各样的原因, 如人力、物力、财力等各种条件的制约, 在现实中无法直接进行研究, 即不可能把总体当中的所有个体都用来做研究, 而只能按照科学的方法从中随机抽取一部分个体进行研究, 所抽取的这部分个体就是样本。

因此, 样本 (Sample) 是从总体当中随机抽取的有代表性的一部分。样本中包含的观察单位数目称为样本含量 (Sample Size)。从总体当中随机抽取样本的过程叫做随机抽样 (Randomization Sampling), 统计研究的过程就是以有限的样本信息的研究结果去推断 (估计) 总体信息。

### 二、同质与变异

同质是指一批性质相同、条件相似的观察单位组成的范围, 即同质范围; 而在同质的条件下, 就同一观察条件来说, 各观察单位表现出来的数量间存在着差异, 这种客观存在的差异性为变异, 变异通常来源于一些未加控制或无法控制的、甚至不明原因的因素。

### 三、参数与统计量

参数 (Parameter) 是决定总体变量值分布规律的特征指标, 是未确定的常数, 若给定了这个常数, 总体的分布就明确了。因此, 参数通常用来描述总体的指标, 常用希腊字母来表示, 如总体均数  $\mu$ 、总体标准差  $\sigma$ 、总体率  $\pi$  等。

一般情况下, 总体参数是未知的, 需要通过观察样本的数据, 计算相应的指标大小, 用以估计未知总体参数, 这个样本指标就是统计量 (Statistics)。因此, 统计量通常用来描述样本的指标, 如样本均数  $\bar{x}$ 、样本标准差  $s$ 、样本率  $P$  等。

参数与统计量之间的关系是: 参数是固定不变的, 统计量随着研究数据的不同而有所不同, 根据样本统计量的分布规律, 可推断总体参数。

### 四、误差

统计学所说的误差, 泛指观察值与真实值之差, 以及样本指标与总体指标之差, 主要包括以下三类:

#### 1. 系统误差

在收集资料的过程中, 由于系统原因所产生的误差叫系统误差。如仪器初始状态未调零、标准试剂未经校正、医生掌握疗效标准偏高或偏低等原因, 可造成观察结果倾向性地偏大或偏小, 这类误差叫系统误差 (或定向误差)。系统误差影响原始数据的准确性, 可以克服, 如果已经发生, 要尽力查明原因并予以校正。

## 2. 随机测量误差

在收集数据的过程中,即使系统误差已经完全消除,各种偶然因素的影响也会造成同一对象多次测定结果不一致。如实验操作员操作技术不稳定、不同实验操作员之间操作差异、电压不稳、温度不恒定等因素可使测量结果存在误差。这种误差往往没有固定的倾向,有时高有时低,称为随机测量误差,这种误差是不可避免的,应采取措施,尽可能减少其发生,至少应将其控制在一定的允许范围内。一般可以通过采用技术培训、指定固定的实验操作员、加强责任感教育及配置一定精度的稳压器、恒温器等措施达到控制目的。

## 3. 随机抽样误差

即使消除了系统误差,并把随机测量误差控制在允许范围之内,样本均数(或其他统计量)与总体均数(或其他参数)之间仍会有误差。这种误差产生的原因主要是个体之间存在变异,由此样本数据构成的统计指标(如均数)就会与总体的该统计指标有误差,这种误差是由抽样引起的,故叫随机抽样误差。这种随机抽样误差是不可避免的,但其分布具有一定规律,通过这种规律可估计它的大小。

一般来说,样本含量越大,则抽样误差越小,样本的观察指标与总体指标接近,即能说明总体的规律。反之,样本含量越小,则抽样误差相对较大。

## 五、频率与概率

频率(Relative Frequency)是指一个随机试验,在 $n$ 次试验当中,某事件出现阳性次数的比例。重复这个随机试验往往会出现几种可能的结果,这些不同的结果都可以看做是偶然发生的,但当重复试验次数相当大时,某种规律就会出现,这种规律就被认为是这个事件发生的概率。

概率(Probability,  $P$ )是表示某随机事件发生可能性大小的一个指标。随机事件概率的大小介于0与1之间,即 $0 \leq P \leq 1$ ,常用小数或百分数表示。 $P$ 越接近1,表示事件发生的可能性越大; $P$ 越接近0,表示事件发生的可能性越小。 $P=1$ 表示事件必然发生,称为必然事件; $P=0$ 表示事件不可能发生,称为不可能事件。这两类事件具有确定性,不是随机事件,但可视为随机事件的特例。统计分析中的很多结论都基于一定可信程度下的概率推断,习惯上将 $P \leq 0.05$ 的事件称为小概率事件,表示在一次试验或观察中该事件发生的可能性很小,可视为很可能不发生。

## 第三节 医学统计学工作的基本步骤

医学统计学工作的实践经验总结出医学统计学工作可以分为统计设计、收集资料、整理资料和分析资料四个部分。一个具体的医学研究就是以这四个环节的医学统计知识为指导,以专业知识为内容方法开展研究工作,如此才可达到预期的效果。因此,医学统计工作贯穿于整个医学科研的全过程,具有重要的地位和作用。

### 一、统计设计

统计设计就是根据研究目的,制定总的研究方案。它是影响研究能否成功的最关键环节,是提高观察或试验质量的重要保证。在统计设计过程中,要正确运用统计学原则、方法和技术,对研究对象(总体)的标准、样本量和样本的获取方法、研究对象分组的原则和对照设置的原则、观察指标的确定、试验过程中的质量控制、研究进度及人力、物力、财力进行合

理安排;同时要根据研究变量的类型考虑数据处理的方法及统计分析方法。因此,只有明确了设计要求和分析方法,才能有目的地、有针对性地做后面的具体工作,而不是盲目地去做。

## 二、收集资料

研究数据是统计分析工作的基础。按照设计要求,取得准确、完整的原始数据,是具体的统计工作的开始阶段,在研究方案中要明确资料来源。

资料根据来源可分为:

### 1. 试验数据

试验数据指在试验过程中获得的数据,试验包括临床试验和动物试验。

### 2. 现场调查数据

现场调查数据指有针对性地设计调查表,通过调查获得的数据。

### 3. 医疗卫生工作记录

这些记录包括门诊病历、住院病历、化验报告、传染病报告、出生报告、死亡报告等。

前两项必须进行全面的统计设计以得到可用的统计资料;后面一项是比较简单的获取统计资料的方法,无须进行全面设计,其特点是容易得到,但缺点是常常不适合特定的研究目的。

## 三、整理资料

整理资料就是对原始资料用正确的统计方法,进行科学加工,即通过分组、归纳、核对等,使杂乱的数据系统化、条理化。根据研究目的和设计方案,具体可列出相应的整理表格,以便分析处理。还要注意整理的技巧,如进行逻辑检查,检查报表(或报告)的纵向、横向合计与总计是否吻合;进行专业检查,如退休年龄不应小于 20 岁;男性患者的调查表中不应出现妇科疾病等。

## 四、分析资料

分析资料就是对经过整理的资料,计算有关指标,进行相应的统计描述或统计推断,得出结果,做出结论等。不同的资料,不同的分析目的,所用到的统计方法是不一样的。

因此,统计工作的步骤是相互联系的,任何一个步骤产生错误都会影响最后结论的正确性。

# 第四节 SPSS 统计软件简介

SPSS 原为 Statistical Package for the Social Sciences(社会科学统计软件包)的简称,是由 SPSS 公司出品的大型通用专业分析软件。2000 年 SPSS 公司重新定义了 SPSS 的含义,即 Statistical Product and Service Solutions(统计产品与服务解决方案)。SPSS 是目前国际上最流行并具有权威性的统计分析软件之一,其最显著的特点是通过菜单和对话框操作方式完成统计过程,绝大多数操作过程仅通过击鼠标即可完成,简单易行,可轻松获得所需结果。因此,特别适用于非计算机专业的医学类学生使用。

## 一、SPSS 界面及菜单简介

在已经安装了 SPSS 11.0 软件的计算机中,从“开始”菜单中选择“程序”,从中选择“SPSS for Windows”,在弹出的子菜单中点击“SPSS 11.0 for Windows”,则系统开始启动 SPSS 11.0(如果桌面上有 SPSS 11.0 的快捷方式,点击运行也可启动),即可进入 SPSS Data

Editor——SPSS 数据编辑窗口(如图 1-1 所示)。

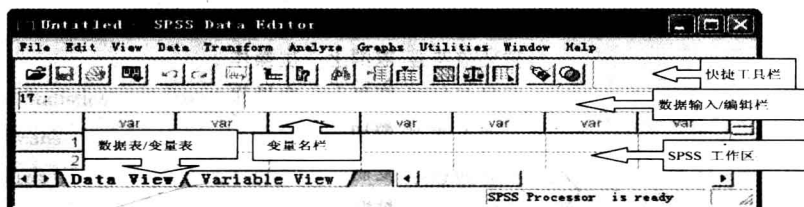


图 1-1 SPSS 11.0 数据编辑窗口

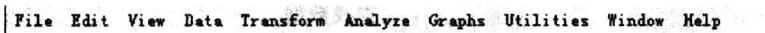


图 1-2 菜单栏

图 1-2 是图 1-1 数据编辑窗口中的 SPSS 应用菜单栏,利用菜单,可实现 SPSS 的数据文件建立、数据编辑、统计分析等各种应用功能。具体如下:

1.File(文件菜单)包括以下功能:

| 一级菜单                | 二级菜单         | 功能说明              |
|---------------------|--------------|-------------------|
| New                 | Data         | 清除屏幕所有数据          |
|                     | Syntax       | 新建 SPSS 程序编辑窗口    |
|                     | Output       | 新建 SPSS 输出指南窗口    |
|                     | Draft Output | 新建 SPSS 草图输出指南窗口  |
|                     | Script       | 新建 SPSS 手稿编辑窗口    |
| Open                | Data         | 打开新的数据文件          |
|                     | Syntax       | 打开 SPSS 程序文件      |
|                     | Output       | 打开 SPSS 结果输出文件    |
|                     | Draft Output | 打开 SPSS 草图输出文件    |
|                     | Script       | 打开 SPSS 手稿编辑文件    |
| Open Database       | New Query    | 新建数据库查询           |
|                     | Edit Query   | 编辑数据库查询           |
|                     | Run Query    | 运行数据库查询           |
| Read Text Data      |              | 打开以文本格式保存的数据      |
| Save                |              | 存储当前数据文件          |
| Save as             |              | 将当前数据文件另存为新的名称或路径 |
| Display Data info   |              | 显示数据信息            |
| Run Data info       |              | 应用数据字典            |
| Cache Data          |              | 数据缓存              |
| Print               |              | 打印                |
| Print Preview       |              | 打印预览              |
| Switch Serve        |              | 转换服务器             |
| Stop Processor      |              | 停止 SPSS 处理器       |
| Exit SPSS           |              | 退出 SPSS           |
| Recently Used Data  |              | 最近使用过的数据库列表       |
| Recently Used Files |              | 最近使用过的文件列表        |
| Exit                |              | 退出 SPSS           |

2.Edit(编辑菜单)包括以下功能:

| 命令      | 功能说明      |
|---------|-----------|
| Undo    | 取消上一次操作   |
| Redo    | 恢复上一次操作   |
| Cut     | 剪切        |
| Copy    | 复制        |
| Paste   | 粘贴        |
| Clear   | 清除选定数据或变量 |
| Find    | 查找数据      |
| Options | SPSS 参数设定 |

3.View(视图菜单)包括以下功能:

| 命令         | 功能说明   |
|------------|--------|
| Status Bar | 显示状态栏  |
| Toolbars   | 显示工具栏  |
| Fonts      | 定义字体   |
| Grid Lines | 显示单元格线 |
| Value      | 显示数值标签 |

4.Data(数据建立和整理菜单)包括以下功能:

| 命令              | 功能说明     |
|-----------------|----------|
| Define Dates    | 定义日期及格式  |
| Insert Variable | 插入变量     |
| Insert Case     | 插入病例(个体) |
| Go to Case      | 转到某病例    |
| Sort Case       | 病例排序     |
| Transpose       | 行列转置     |
| Merge Files     | 合并数据文件   |
| Aggregate       | 整合数据     |
| Split File      | 分割文件     |
| Select Cases    | 选择个体     |
| Weigh Cases     | 加权个体     |

5.Transform(变量变换菜单)包括以下功能:

| 命令                     | 功能说明     |
|------------------------|----------|
| Compute                | 新变量赋值    |
| Random Number Seed     | 产生随机数    |
| Count                  | 记数       |
| Recode                 | 原变量重新赋值  |
| Categorize Variables   | 生成新分类变量  |
| Rank Cases             | 个体排序     |
| Automatic Recode       | 自动重新赋值   |
| Create Time Series     | 生成时间序列   |
| Replace Missing Values | 替代缺失值    |
| Run Pending Transform  | 运行未执行的变量 |

## 6. Analyze(统计分析菜单)包括以下功能:

| 命令                     | 功能说明   |
|------------------------|--------|
| Reports                | 数据概况统计 |
| Descriptive Statistics | 描述性统计  |
| Compare Means          | 均数比较   |
| General Linear Model   | 一般线性模型 |
| Correlate              | 相关分析   |
| Regression             | 回归分析   |
| Loglinear              | 对数线性分析 |
| Classify               | 分类分析   |
| Data Reduction         | 数据简化分析 |
| Scale                  | 尺度分析   |
| Nonparametric Tests    | 非参数检验  |
| Survival               | 生存分析   |
| Multiple Response      | 多重响应分析 |

## 7. Graphs(作图菜单)包括以下功能:

| 命令          | 功能说明     |
|-------------|----------|
| Gallery     | 图形特性描述   |
| Interactive | 交互图形     |
| Bar         | 条图       |
| Line        | 线图       |
| Area        | 区域图      |
| Pie         | 圆图       |
| High-Low    | 高一低图     |
| Pareto      | Pareto 图 |
| Control     | 控制图      |
| Boxplot     | 箱式图      |
| Error Bar   | 误差条图     |
| Scatter     | 散点图      |
| Histogram   | 直方图      |
| P-P         | 正态 P—P 图 |
| Q-Q         | 正态 Q—Q 图 |
| Sequence    | 序列图      |
| ROC Curve   | ROC 曲线   |
| Time Series | 时间序列图    |

此外,还有 Window(窗口菜单)、Utilities(工具菜单)、Help(帮助菜单),考虑到这些菜单的功能不常用,相对也不太重要,故略过。

## 二、SPSS 11.0 数据文件的建立

通常情况下,要建立一个完整的数据文件需要进行下面的操作步骤:

1. 启动 SPSS 11.0

进入如图 1-1 所示的数据编辑窗口(SPSS Data Editor)。

2. 定义变量

- (1)定义变量名(Variable Name);
- (2)定义变量类型(Variable Type);
- (3)定义变量的数值位数(Variable Width);
- (4)定义变量的小数位数(Variable Decimals)(仅对数值变量有效);
- (5)定义变量标签(Labels)和数值标签(Values)(此步通常省略)。

这五步操作都在变量定义窗口(Variable View)中完成。

3. 输入数据

在数据编辑窗口(Data View)中完成。

4. 保存数据

可以将数据存为 SPSS(\*.sav)、Excel(\*.xls)、dBASE(\*.dbf)、ASCII(\*.dat,\*.txt)等多种数据文件形式。

【例 1-1】表 1-1 为某一单位员工的部分体检数据,请将下列表格输入 SPSS 中,并保存到文件名为“例 1-1.sav”的数据文件中。

表 1-1 某单位部分体检数据

| 姓名  | 性别 | 年龄 | 出生日期       | 体重/kg |
|-----|----|----|------------|-------|
| 王 鹏 | 男  | 35 | 1965.10.10 | 69    |
| 张 三 | 男  | 40 | 1950.12.11 | 74    |
| 李 四 | 女  | 40 | 1950.08.09 | 64    |
| 王 五 | 女  | 42 | 1948.06.06 | 74    |
| 孙 六 | 男  | 37 | 1963.05.05 | 66    |
| 李 丽 | 女  | 38 | 1962.10.14 | 72    |
| 田 燕 | 女  | 45 | 1945.02.06 | 68    |


操作步骤

1. 启动 SPSS 11.0 进入 SPSS Data Editor(数据编辑窗口)

2. 定义变量

(1)在图 1-1 所示的数据编辑窗口中,点击左下方的 Variable View,进入变量编辑窗口。

(2)定义变量名(Name):单击 Name 选项下方的第一格定位后,双击,输入变量名“姓名”(系统默认为 VAR00001)。

(3)定义变量类型(Type):单击 Type 选项下方第一格定位,单击方格右侧按钮 ,进入变量类型定义对话框(如图 1-3 所示),选定 String,单击  确认。

(4)定义变量的数值位数(Width)、小数位数(Decimals)、变量标签(Label)、数值标签(Values)等,均取默认值。

此时第一个变量“姓名”已经定义完毕。第二个变量“性别”的定义过程和第一个变量“姓名”的定义过程类似,其类型可以定义为字符类型(如用汉字的男和女表示),也可定义为数值类型(如用 0 表示“女”,1 表示“男”,本例选择为数值类型,且推荐这样使用)。

变量“年龄”为数值型变量,在定义变量类型时默认数值型(Numeric)或者在图 1-3 所示



对话框选择  Numeric，在定义数值位数 (Width) 时默认为 8 即可，在定义小数位数时 (Decimals) 将 2 改为 0 (因为是整数, 不带小数点) 即可。

变量“出生日期”在定义类型时选择日期类型, 定义日期类型时, 在图 1-3 中选择  Data 并选择一种日期格式即可, 如图 1-4 所示。

变量“体重”的定义过程参考变量“年龄”的定义过程。所有变量定义完之后的变量编辑窗口如图 1-5 所示。

Width 栏中的数字表示字符串长度为 8 位 (一个中文字符=2 个英文字符, 也就是说如果定义字符串长度为 8, 则最多可输入 8 个英文字母或者 4 个中文字符)。

### 3. 输入数据

点击左下方的 Data View, 进入数据编辑窗口。将例 1-1 的数据记录, 逐个录入对应的变量名下。录入完成后的数据编辑窗口如图 1-6 所示。

### 4. 保存数据

选择“File”→“Save”, 出现如图 1-7 所示的对话框。在文件名栏中输入文件名“例 1-1”, 选择文件的保存路径, 点击“保存”按钮即可。

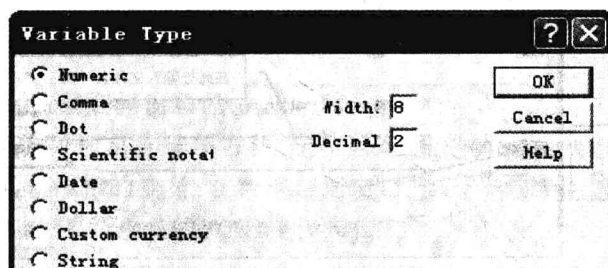


图 1-3 变量类型选项对话框

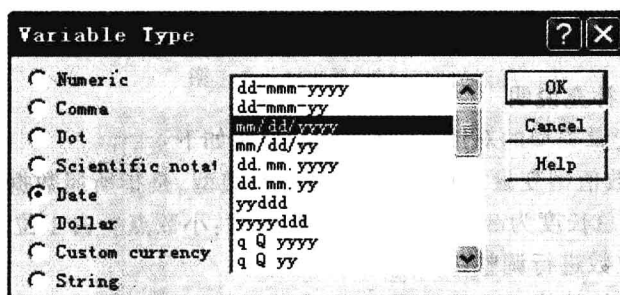


图 1-4 日期类型定义对话框

|   | Name | Type    | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|------|---------|-------|----------|-------|--------|---------|---------|-------|---------|
| 1 | 姓名   | String  | 8     | 0        |       | None   | None    | 8       | Left  | Nominal |
| 2 | 性别   | Numeric | 8     | 0        |       | None   | None    | 8       | Right | Nominal |
| 3 | 年龄   | Numeric | 8     | 0        |       | None   | None    | 8       | Right | Scale   |
| 4 | 出生日期 | Date    | 10    | 0        |       | None   | None    | 8       | Right | Scale   |
| 5 | 体重   | Numeric | 8     | 0        |       | None   | None    | 8       | Right | Scale   |

图 1-5 定义完成后的变量表