

高等学校统计分析软件系列教材

S TATISTICA for Windows

统计与图表分析教程

102345678910235465284312046532687
425678-932301246385874697623221133
32564789564321300 546528431204653
456879311 93230124638587469762322
20456879311

洪 楠 主编

洪 楠 侯 军 李志辉 编著



清华大学出版社

<http://www.tup.tsinghua.edu.cn>



北方交通大学出版社

<http://press.njtu.edu.cn>



本书配光盘

高等学校统计分析软件系列教材

STATISTICA for Windows 统计与图表分析教程

洪 楠 主编

洪 楠 侯 军 李志辉 编著

清华大学出版社
北方交通大学出版社
北京·BEIJING

内 容 简 介

STATISTICA/Win(大型专业统计与图表分析软件包)是当今世界上最大型的统计与图表分析软件之一,得到了广大用户和评论家大量的赞誉和好评。STATISTICA/Win在每一项功能的评比与评价中都获得了最高的分数。

本书以统计分析与图表分析兼备的独特风格全面介绍了 STATISTICA /Win 的基本知识。统计分析内容包括:基本统计分析/制表,非参数统计检验/分布拟合,广义单因素方差分析与多因素方差分析,多重回归分析,非线性估计法,时间序列/预测预报分析,聚类分析,因子分析,典型分析,信度/项目分析,判别分析,对数线性模型分析以及生存分析等。此外,书中还用较大的篇幅介绍了 STATISTICA /Win 的各种作图功能与技巧。

本书的内容与方法广泛适于自然科学、社会科学,特别是经济学、生物学、金融学与医疗卫生保健等多学科、多专业、多层次的需要,可供高等院校、科研单位及相关专业的研究生、本科生、科技人员与企事业单位计算机工作者使用。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究。

图书在版编目(CIP)数据

STATISTICA for Windows 统计与图表分析教程/洪楠,侯军,李志辉编著. —北京:北方交通大学出版社, 2002. 10

高等学校统计分析软件系列教材

ISBN 7-81082-066-4

I. S... II. ①洪... ②侯... ③李... III. 统计分析—软件包, STATISTICA—高等学校—教材 IV. C819

中国版本图书馆 CIP 数据核定(2002)第 039566 号

丛 书 名: 高等学校统计分析软件系列教材

书 名: STATISTICA for Windows 统计与图表分析教程

主 编: 洪 楠

编 著: 洪 楠 侯 军 李志辉

责任编辑: 李 莉

排版制作: 海天计算机技术开发有限公司

印 刷 者: 北京东光印刷厂

装 订 者: 三河市桃园装订厂

出版发行: 北方交通大学出版社 邮编: 100044 电话: 010-62237564 51686045

清华大学出版社 邮编: 100084

经 销: 各地新华书店

开 本: 787×1092 1/16 印张: 28.75 字数: 732.8 千字 附光盘: 1 张

版 次: 2002 年 10 月第 1 版 2002 年 10 月第 1 次印刷

书 号: ISBN 7-81082-066-4
TP·25

印 数: 5000 册 定价: 49.00 元(含光盘)

高等学校统计分析软件系列教材

总 序

20 世纪以来,由于电子计算机应用技术的高速发展,使人类从各种繁重的体力和脑力劳动中得到解放,并带领人类昂首阔步地挺进当今 21 世纪的信息时代。如今人们可以借助电子计算机及其统计分析系统或统计软件包来实现繁冗的数据处理工作,在大量的信息中获得有科学价值的结果,帮助人们分析、决策和解决问题。用电子计算机实现统计数据处理,必须对采集的信息进行存储、整理、计算、统计分析和图表绘制,从而利用计算的结果揭示和阐明客观事物变化规律。

目前,经典的大型全功能的统计分析系统或统计软件包,例如 SAS,SPSS,SYSTAT,STATISTICA,BMDP, SPLM, MiniTab 等,对于广大科技工作者和教师、学生来说,需要一本既全面又实用还能及时追踪最新计算机软件包的教材来满足他们的科研、教学和工作需求。本系列教材的编写目的正在于此。

本软件系列教材的编写队伍由广州中山大学从事多年教学研究的统计学专家教授、从事医疗卫生保健工作多年的专家教授、电子计算机专家教授以及富有多年临床实际经验的老中青医务工作者组成。软件系列教材的内容包括基本统计方法,常用多元统计方法,高等统计学方法以及作图方法,并介绍每一个软件包的独特风格及其在实际应用中的技巧。作者力求理论联系实际,深入浅出,重点和难点以许多来自于社会实践中的实际例子进行操作和计算,对计算出的每一个结果进行了详尽分析和讨论。在编写过程中,作者使用普通语言进行描述,并配以丰富的插图,详细的中英文对照解说,力图使对于数学、外语或计算机科学了解不多的实际工作者得以受益。本系列教材还追踪各统计软件包的最新发展,在编写过程中不断更新教材使用的软件包版本,保持教材与世界上的最新成果同步。本系列教材凝结着许许多多科技工作者的辛勤劳动和智慧。

本系列教材的内容与方法广泛适于自然科学、社会科学,特别是经济学、生物学、金融学与医疗卫生保健等多学科、多专业、多层次的需要,可供高等院校、科研单位相关专业的研究生、本科生、科技人员与企事业单位计算机实际工作者使用。

洪 楠
广州·中山大学
2002 年 8 月

前 言

人类已进入信息时代,要在大量的信息中获得有科学价值的结果,必须对采集的信息进行存储、整理、计算、统计分析(Analysis)和图表(Graphs/Tables)绘制,这可以借助于计算机统计图表软件包来实现繁冗的数据处理工作,从而达到揭示与阐明客观事物变化规律的目的。

STATISTICA/Win(STATISTICA for Windows,大型专业统计图表分析软件包)是当今世界上流行的统计图表软件包之一,也是世界上用户较多的统计图表分析软件系统。

美国俄克拉荷马州(Oklahoma)StatSoft 软件公司继 1993 年推出 Windows 版本大型专业统计图表分析软件包 STATISTICA for Windows(v4.5)之后,近年来又推出 v5.0, v6.0 的 Windows 版本。新版本中除保持了原有版本的诸多优点外,还汇集了 SPSS, SAS, SYSTAT 与 BMDP for Windows 的长处,增加了更多的统计分析与图表分析功能,可以说是当今国际上最新的 Windows 版本统计分析与图表分析软件包之一。

STATISTICA/Win 不仅统计分析方法丰富,能绘制高清晰度图形图表,而且具有统计分析与图表分析兼备的独特风格,这些新特点得到了广大用户和评论家大量的赞誉和好评。STATISTICA/Win 在每一项功能的评比与评价中都获得最高分,因而深受广大用户的欢迎。

本书的内容与方法广泛适应于自然科学、社会科学,特别是经济学、生物学、金融学与医疗卫生保健等多学科、多专业、多层次的需要,也可供高等院校、科研单位及相关专业的研究生、本科生、科技人员与企事业单位计算机工作者使用。

本书的大多数数据资料是经过严格的科学设计,从多年众多科学实验中获得的,编著者将这些数据资料以文件的形式复制在本书附带光盘的 \STADATA 目录中,供广大读者享用。

本书由洪楠、侯军与李志辉编著,洪楠主编。

在本书的编写过程中,中山大学胡孟璇教授、方积乾教授、邹赛德教授与袁衡新博士均给予了热情鼓励与指导,卫生统计学教研室与计算机中心的同事也给予了大力支持,在此深表感谢。

本书的出版得到北方交通大学出版社总编辑卢先河博士的大力支持,他付出了很多智慧与劳动,作者在此一并感谢。

由于编者的水平有限,错误之处,在所难免,敬请读者批评指正。

反馈意见,请发电子邮件(E-mail)至: hopejun@21cn.com 或 hongnan@163.net。

洪 楠
广州·中山大学
2002 年 8 月

目 录

第 1 章	STATISTICA/Win 基础	(1)
1.1	STATISTICA/Win 概述	(1)
1.1.1	统计分析	(1)
1.1.2	图表分析	(4)
1.1.3	STATISTICA/Win 的运行环境	(5)
1.1.4	STATISTICA/Win 的启动	(6)
1.1.5	STATISTICA/Win 的主画面	(7)
1.1.6	STATISTICA/Win 的退出	(13)
1.2	STATISTICA/Win 入门	(14)
1.2.1	鼠标的操作	(14)
1.2.2	数据/大数据文件管理	(14)
1.2.3	建立新的数据文件	(15)
1.3	STATISTICA/Win 的基本操作与文件类型	(21)
1.3.1	STATISTICA/Win 的基本操作	(21)
1.3.2	STATISTICA/Win 的文件类型	(23)
1.4	中文注释编辑	(23)
第 2 章	STATISTICA/Win 的数据文件与变量	(29)
2.1	合并两个数据文件	(29)
2.1.1	合并变量	(29)
2.1.2	合并个案	(31)
2.2	建立子数据文件与个案排序	(31)
2.2.1	建立子数据文件	(31)
2.2.2	个案排序	(33)
2.3	修改变量/个案	(34)
2.3.1	修改变量	(34)
2.3.2	修改个案	(40)
2.4	缺失数据处理方法	(41)
2.4.1	用均数替代缺失数据	(41)
2.4.2	缺失数据的常用删除方法	(43)
2.4.3	缺失数据的其他替代方法	(43)
2.5	标准化变量	(44)
2.5.1	计量资料	(44)
2.5.2	频数表资料	(47)
2.6	统计数据类型	(49)
第 3 章	基本统计分析与制表	(51)
3.1	描述性统计分析	(51)
3.2	相关矩阵分析	(60)

3.3	独立样本(成组)t-检验	(66)
3.3.1	独立样本 t-检验(1)(记录/个案格式)	(66)
3.3.2	独立样本 t-检验(2)(组/变量记录格式)	(69)
3.3.3	多变量霍特林(Hotelling) T^2 检验	(71)
3.4	相依(相关)样本 t 检验	(74)
3.5	细分与单因素方差分析	(79)
3.6	频数表分析	(89)
3.7	制表与标题	(96)
3.7.1	交叉分组列联表	(96)
3.7.2	多重反应项目表	(102)
3.8	概率分布计算	(108)
3.8.1	贝塔(Beta)分布	(109)
3.8.2	χ^2 (Chi-Square)分布	(109)
3.8.3	F 分布	(110)
3.8.4	Z(Normal)分布	(111)
3.8.5	t(Student)分布	(112)
3.9	其他显著性检验	(113)
3.9.1	两个相关系数的差异性检验	(113)
3.9.2	两个正态分布均数的差异性检验	(115)
3.9.3	单一均数与总体均数的差异性检验	(116)
3.9.4	两个百分数的差异性检验	(117)
第 4 章	非参数统计检验/分布拟合	(119)
4.1	2×2 表与费歇尔精确检验	(119)
4.2	观察与期望频数的对照检验	(122)
4.3	非参数相关分析	(123)
4.3.1	斯皮尔曼(Spearman)等级相关分析	(123)
4.3.2	肯德尔(Kendall)相关分析	(125)
4.3.3	肯德尔(Kendall, 计数资料)相关分析	(127)
4.4	沃尔德-沃尔佛维茨游程检验	(129)
4.5	曼-惠特尼 U 检验	(132)
4.5.1	计量资料的曼-惠特尼 U 检验	(132)
4.5.2	频数表(或等级)资料的曼-惠特尼 U 检验	(134)
4.6	柯尔莫哥罗夫-斯米尔诺夫两样本检验	(136)
4.7	克鲁斯凯-沃利斯方差分析(中位数检验)	(138)
4.8	符号检验	(143)
4.9	威尔科克森配对检验	(145)
4.10	弗里德曼方差分析与肯德尔一致性	(147)
4.11	科克伦 Q 检验	(150)
4.12	常用描述性统计分析	(151)
4.13	拟合连续型分布	(153)
4.13.1	正态分布(Normal)(计量资料)	(154)
4.13.2	正态分布(Normal)(频数表资料)	(158)

4.14	拟合离散型分布	(160)
4.14.1	二项分布(Binomial)	(161)
4.14.2	泊松分布(Poisson)	(163)
第5章	广义单因素方差分析与多因素方差分析	(166)
5.1	多因素方差分析	(166)
5.2	协方差分析	(171)
第6章	多重回归分析	(179)
6.1	直线回归分析	(179)
6.2	过原点的直线回归分析	(184)
6.3	多重线性回归分析	(188)
6.4	前向逐步线性回归分析	(195)
6.5	后向逐步线性回归分析	(200)
6.6	岭回归分析	(205)
第7章	固定非线性回归分析	(210)
7.1	多项式回归分析	(210)
7.2	拟合对数曲线	(217)
7.3	生长曲线回归分析	(221)
第8章	非线性估计法	(227)
8.1	用户指定回归分析	(227)
8.2	逻辑斯谛(Logistic)回归分析	(234)
8.3	成组资料的逻辑斯谛(Logistic)回归分析	(241)
8.4	指数增长回归分析	(245)
8.5	分段线性回归分析	(248)
第9章	时间序列/预测预报分析	(254)
9.1	博克斯-詹金斯模型分析	(254)
9.2	直线趋势预测分析	(260)
9.3	指数趋势平滑法	(266)
9.4	时间序列季节性调整分解法	(270)
第10章	聚类分析	(279)
10.1	联接(树状聚类)分析	(279)
10.1.1	指标(R型)聚类分析	(279)
10.1.2	样品(Q型)聚类分析	(283)
10.2	逐步聚类分析	(286)
10.3	双向聚类分析	(294)
第11章	因子分析	(300)
第12章	典型分析	(308)
第13章	信度/项目分析	(317)
第14章	判别分析	(322)
第15章	对数线性模型分析	(330)
第16章	生存分析	(335)

16.1	寿命表与分布	(335)
16.2	卡卜兰-迈尔乘积极限法	(344)
16.3	两样本生存率比较	(348)
16.4	多样本生存率比较	(355)
16.5	比例危险(Cox)回归分析	(361)
第 17 章	图表分析	(369)
17.1	快速统计图	(369)
17.1.1	直方图(快速统计图)	(369)
17.1.2	二维直方图(快速统计图)	(370)
17.1.3	箱形-触须线图(快速统计图)	(371)
17.2	二维统计图	(372)
17.2.1	散点图/带直方图(二维统计图)	(373)
17.2.2	箱形图(二维统计图)	(376)
17.2.3	饼分图(二维统计图)	(379)
17.3	三维序列统计图	(383)
17.3.1	二元直方图(三维序列统计图)	(383)
17.3.2	箱形图(三维序列统计图)	(385)
17.4	三维 XYZ 统计图	(387)
17.4.1	散点图(三维 XYZ 统计图)	(387)
17.4.2	曲面图(三维 XYZ 统计图)	(389)
17.5	统计矩阵图	(392)
17.5.1	方形散点阵图(统计矩阵图)	(393)
17.5.2	长方形散点阵图(统计矩阵图)	(395)
17.6	统计分类图	(396)
17.6.1	直方图(统计分类图)	(396)
17.6.2	饼分图(统计分类图)	(400)
17.7	常用图形	(404)
17.7.1	三维序列图(常用图形)	(404)
17.7.2	三维 XYZ 图(常用图形)	(405)
17.8	块状统计图	(408)
17.8.1	纵直方图(块状统计图)	(408)
17.8.2	纵均数箱形图(块状统计图)	(408)
17.8.3	纵中位数箱形图(块状统计图)	(409)
17.8.4	纵正态概率图(块状统计图)	(410)
附录 1	练习题	(412)
	STATISTICA/Win 基础习题	(412)
	统计分析(Analysis)习题	(413)
	图表分析(Graphs/Tables)习题	(419)
附录 2	STATISTICA/Win 数据文件的使用说明	(420)
附录 3	英汉词汇对照表	(421)
参考文献		(447)

第 1 章 STATISTICA/Win 基础

1.1 STATISTICA/Win 概述

计算机统计软件包或统计分析系统的各 Windows 版本, 诸如: SPSS for Windows(v5.0~v11.0), SAS for Windows(v6.10~v8.0), 以及 STATISTICA for Windows(v4.0~v6.0)等相继问世以来, 经多个国家地区的广大用户试用, 并与 DOS 版本比较, 由于其具有一看就懂, 一点就会的独特风格, 深受用户喜爱。

美国俄克拉荷马州(Oklahoma)StatSoft 软件公司继 1993 年推出 Windows 版本大型专业统计图表分析软件包 STATISTICA for Windows(v4.5)之后, 近年来又推出 v6.0 的 STATISTICA for Windows 版本, 它除保持了原有版本的诸多优点外, 还汇集了 SPSS, SAS, SYSTAT 与 BMDP for Windows 的长处, 增加了更多的统计分析与图表分析功能, 可以说是当今国际上最新的 Windows 版本统计分析与图表分析软件包之一。它的主要特点是:

(1) STATISTICA/Win 不但可以在 Windows 95/98/2000/ME/XP 下运行, 而且可在 Windows NT 的各种版本下运行。

(2) 运行速度比 SPSS/Win, SAS/Win, SYSTAT/Win 以及 BMDP/Win 更快。

(3) 菜单栏上详细的统计专业名词, 方便用户易于理解其功能, 除帮助功能外, 在对话框上都有简明扼要的解释, 并可配合中文使用, 增加可读性。

(4) 每一个统计或图形模块均能生成图标(按钮), 而且图标形象直观, 便于判别和理解, 用户可以根据不同的需要将自己常用的统计或图形模块生成用户自动任务按钮(Customize Auto Task Button), 减少利用菜单选择的时间。

(5) 有功能强大的数据管理系统, 可同时打开多个数据文件, 而且还可以在分析的过程中打开数据文件。

(6) 可分步进行分析, 用户可根据需要选择所需的统计或图表模块。

(7) 作图(制作图表并进行分析)速度比 SPSS/Win、SAS/Win 更快, 且能制作更多类型的统计图表, 对于图形的颜色, 线条的类型有更多的选择, 为用户制作多种类型的统计图表提供良好的环境, 还具有用户自定义统计图表的功能, 用户可以根据需要定义具有自己特色的统计图表。

(8) 输出结果简单明了, 对于具有显著性意义($P < 0.05$)的结果以不同颜色(例如红色)表示, 输出结果可以用指定文件或标准文本文件格式存盘, 便于资料的再开发, 还可以把图表、数据计算结果编辑成图文并茂的报告。

1.1.1 统计分析

STATISTICA/Win 统计分析(Analysis)有 15 个主命令, 130 多个子命令。

1. 基本统计分析/制表(Basic Statistics/Tables)

它包括描述性统计分析(Descriptive Statistics), 相关矩阵分析(Correlation Matrices), 独立

样本 t 检验(t-Test for Independent Samples),相依样本 t 检验(t-Test for Dependent Samples),细分与单因素方差分析(Breakdown & One-way ANOVA),频数分布分析(Frequency Tables),制表与标题(Tables and Banners),概率计算(Probability Calculators)与其他显著性检验(Other significance tests)等。

2. 非参数统计检验/分布拟合(Nonparametrics/Distrib)

(1) 非参数统计检验(Nonparametric Stats)。它有: 2×2 表(2×2 Table), $\chi^2/\nu^2/\phi^2$ 值, 麦克内马检验(McNemar), 费雪尔精确检验(Fisher Exact), 观察值与期望值的 χ^2 检验(Observed Versus Expected χ^2), 非参数相关分析(Correlations (Spearman, Kendall, Gamma)), 沃尔德-沃尔福维茨游程检验(Wald-Wolfowitz Runs Test), 曼-惠特尼 U 检验(Mann-Whitney U Tests), 柯尔莫哥罗夫-斯米尔诺夫两样本检验(Kolmogorov-Smirno Two-Samples Test), 克鲁斯凯-沃利斯方差分析(中位数检验)(Kruskal-Wallis ANOVA, Median Test), 符号检验(Sign Test), 威尔科克森配对检验(Wilcoxon Matched Pair Test), 弗里德曼方差分析与肯德尔一致性(Friedman ANOVA & Kendall's Concordance), 科克伦 Q 检验(Cochran Q Test)等。

(2) 分布拟合(Distribution Fitting)。它既能对连续型分布(Continuous Distributions)进行拟合, 如正态分布(Normal), 均匀分布(Rectangular), 指数分布(Exponential), γ 分布(Gamma), 对数正态分布(Log-normal), χ^2 分布(Chi-square)与其他(Others)连续型分布; 又能对离散型分布(Discrete Distributions)进行拟合, 如二项分布(Binomial), 泊松分布(Poisson), 几何分布(Geometric)和伯努利分布(Bernoulli)。

3. 广义单因素方差分析与多因素方差分析(General ANOVA/MANOVA)

它能进行单因素与多因素方差分析(General Univariate and Multivariate ANOVA), 重复测量方差分析(Repeated Measures), 协方差分析(Covariates), 嵌套设计方差分析(Nested Design)以及后验比较检验(Post-hoc Tests)和对比分析(Contrasts)等。

4. 多重回归分析(Multiple Regression)

其中的线性回归分析提供了 3 种方式(Mode): 标准方式(Standard), 前向逐步法与后向逐步法(Stepwise Forward, Backward)。此外, 还有固定非线性回归分析(Fixed Nonlinear)(如: 多项式回归分析(Polynomial)), 岭回归分析(Ridge), 过原点回归分析(Through Origin)和残差分析(Residual Analysis)等。

5. 非线性估计法(Nonlinear Estimation)

它包括用户指定回归分析(User Specified Regression), 逻辑斯谛回归分析(Logistic Regression), 概率单位回归分析(Probit Regression), 指数增长回归分析(Exponential Growth Regression), 分段线性回归分析(Piecewise Linear Regression), 而所用的方法有最小二乘法(Least-Squares), 最大似然估计法(Maximum Likelihood Estimation), 用户自定义损失函数法(User-defined Lossfunctions)等。

6. 时间序列/预测预报分析(Time Series/Forecasting)

它包括时间序列探索(Time Series Exploratory), 模型预测预报(Modeling and Forecast-

ing), 相互转换(Interactive Transformations), 平滑法(Smoothing), 自相关分析(Autocorrelations), ARIMA 模型 Box-Jenkins X11 ARIMA(Census 11 X-11 法), 指数平滑法(Exponential Smoothing), 傅里叶分解(Fourier Decomposition)等。

7. 聚类分析(Cluster Analysis)

能对样本或变量进行聚类或分类。聚类方法有连接(树状聚类)(Joining(Tree Clustering)), 逐步聚类分析(K-Means Clustering)与双向聚类分析(Two-Way Joining)。

8. 因子分析(Factor Analysis)

有因子分析与主成分分析(Factor Analysis and Principal Components Analysis), 用户或回归定义公性成分分析(User or Regression Defined Communalities), 最大似然因子分析(Maximum-Likelihood Factor Analysis)和正交或斜交旋转分析(Orthogonal or Oblique Rotations)。

9. 典型分析(Canonical Analysis)

它包括典型分析(Canonical Analysis), 两组变量的多变量相关分析(Multivariate Correlation of Two List Variables), 典型载荷与因子系数分析(Canonical Loading and Factor Coefficients)。

10. 多维尺度分析(Multidimensional Scaling)

可分析距离(Analyze Distance)不一致性与相似性(Dissimilarities and Similarities)等。

11. 路径分析(SEPATH)

它包括全面选择一个结构方程模型(Structural Equation Modeling), 确认因子分析(Confirmatory Factor Analysis), 路径分析(Path Analysis), 蒙特卡罗模拟(Monte Carlo Simulation)和自助估计(Boot-Strap Estimation)等。

12. 信度/项目分析(Reliability/Item Analysis)

构造、分析测验表与调查表(Toconstruct and Analyze Tests and Questionnaires)。包括相关分析(Correlations), 克龙伯 α 分析(Cronbach Alpha), 项目-总计相关与多重相关分析(Item-Total Correlations and Multiple Correlations), 分半信度分析(Splithalf Reliability), 衰减校正(Attenuation Correction)等。

13. 判别分析(Discriminant Analysis)

它包括逐步判别函数分析(前向法、后向法或集团分析)(Stepwise Discriminant Function Analysis(Forward, Backward, or in Block Analysis), 典型分析与得分分析(Canonical Analysis and Scores), 分类统计分析(Classification Statistics)等。

14. 对数线性模型分析(Log-linear-analysis)

它包括多向列联表的对数线性分析(Multiway Crosstabulations), 选择一个列联表的定义格式(A Selection of Table Definition Formats), 分析原始数据或汇总表(Analyze Raw Data or Aggregated Tables)等。

15. 生存分析(Survival Analysis)

它包括生存时间的寿命表与分布(Life Tables and Distributions),卡卜兰-迈尔乘积极限法(Kaplan & Meier Product-Limit Method),两样本的比较(Comparing Two Samples),多样本的比较(Comparing Multiple Samples),回归模型(Regression Models)。而回归模型提供了5种方法:即比例风险(Cox)回归分析(Proportional Hazard(Cox) Regression),具有时间协变量的比例风险(Cox)回归分析(Prop. Hazard(Cox)/w Time-dep. Covariats),指数回归分析(Exponential Regression),对数正态回归分析(Lognormal Regression)和正态回归分析(Normal Regression)。

1.1.2 图表分析

STATISTICA /Win 在统计分析模块(Analysis)的13个主命令及其各级子命令中,都配有丰富齐全的图表功能选项,供用户随时选用,用户还可利用中间结果或最后结果制作图表。此外,专用的图表模块(Graphs)提供了13个主命令(13种类型)以及80多个丰富的子命令。

1. 快速统计图(Quick Stats Graphs)

它可以按性质制作二维、三维直方图(2D,3D Histogram),箱形-触须线图(Box-Whisker),概率图(Probability Plot),还能制作散点图(Scatterplot),矩阵散点图或散布图(Matrix Scatterplot)等10多种类型的图形。

2. 二维统计图(Stats 2D Graphs)

它可以制作直方图(Histograms),散点图或散布图(Scatterplots),箱形图(Box plots),正态概率图(Normal Probability Plots),分位数-分位数图(Quartile-Quartile Plots),概率-概率图(Probability-Probability Plots),全距图(Range Plots),条形/柱状图(Bar/Column Plots),变量/个案轮廓线图(Line Plots Variables/Cases Profiles),序列/成堆图(Sequential/Stacked),饼分图(Pie Charts),缺失/全距数据图(Missing/Range Data Plots)与常用函数图(Custom Function Plots)等16种以上的统计图表。

3. 三维序列统计图(Stats 3D Sequential Graphs)

它可以制作原始数据图(Raw DataPlots),二元直方图(Bivariate Histograms),全距图(Range Plots)和箱形图(Box Plots)等三维序列图。

4. 三维 XYZ 统计图(Stats 3D XYZ Graphs)

它可以制作散点图(Scatterplots),曲面图(Surface Plots),轮廓线(等高线)图(Contour Plots),轨迹线图(Trace Plots),三元图(Ternary Plots),分类 XYZ 图(Categorized XYZ Plots),分类三元图(Categorized Ternary Plots)和常用函数图(Custom Function Plots)等。

5. 统计矩阵图(Stats Matrix Graphs)

它提供6种作图类型(Graphs Type),包括正方形/长方形散点阵图(Square Scatter Matrix),正方形/长方形曲线阵图(Square Line Matrix),正方形/长方形柱状阵图(Square Column Matrix)。此外,还能制作线性图(Linear),对数图(Logarithmic),指数图(Exponential),条样图(Spline),多项式图(Polynomial),最小二乘图(Least Squares),负指数/加权图(Neg Exp/

Wght, 即 Negative Exponential/Weight)和常用函数图(Custom Function)等。

6. 统计画像图(Stats Icon Graphs)

图型(Graph Type)有:切尔洛夫脸形图(Chernoff Faces),星形图(Stars),日射图(Sun Rays),多边形图(Polygons),饼分图(Pies),柱状图(Columns),线图(Lines)和轮廓图(Profiles)等 8 种。

7. 统计分类图(Stats Categorized Graphs)

它可以制作直方图(Histograms),散点图(Scatterplots),正态概率图(Normal Probability Plots),分位数-分位数图(Quartile-Quartile Plots),概率-概率图(Probability-Probability Plots),线图(Line Plots),箱形图(Box Plots),饼分图(Pie Charts),缺失值/全距数据图(Missing/Range Data Plots),三维 XYZ 图(3D XYZ Plots)和三维三元图(3D Ternary Plots)等。

8. 多重图形版面(Multiple Graph Layouts)

可以是奇才(Wizard)或模板(Templates)形式。

9. 空白图形(Blank Graphs)

可以是预设(默认)比例(Default Proportion)或页/打印比例(Page/printer Proportion)。

10. 常用图形(Custom Graphs)

有 5 种类型:二维图(2D Graphs),三维序列图(3D Sequential Graphs),三维 XYZ 图(3D XYZ Graphs),矩阵图(Matrix Plots),画像图(Icon Plots)以及三元图(Ternary Graphs)。

11. 块状统计图(Block Stats Graphs)

它可以制作纵/横直方图(Histograms of Columns/Rows),纵/横均数箱形图(Box Plot/Mean of Columns/Rows),纵/横中位数箱形图(Box Plot/Medians of Columns/Rows),纵/横正态概率图(Normal Prob. Plots of Columns/Rows)。

12. 用户自定义统计图(Stats User-defined Graphs)

可以通过帮助(Help)完成。

13. 图形数据连接(Graph Data Links)

1.1.3 STATISTICA/Win 的运行环境

STATISTICA/Win v6.0 能在 Windows 95/98/2000/ME/XP 或 Windows NT 3.51 以及更高版本下进行运算。它推荐的基本设置要求如下。

(1) 一个能运行 Windows 95/98/2000 操作系统的 PC 机,本书的实例是在 Windows 98 操作系统支持下运行的。

(2) 至少需要 16MB 内存。

(3) S-VGA 显示器以及装有与 Windows 98/2000 相兼容的高级图形适配卡。

(4) 整个 STATISTICA/Win 约占据 20MB 硬盘空间。

- (5) 双速或四速以上 CD-ROM 驱动器及至少有一个 3.5 英寸的软驱。
- (6) 支持 Windows 98 或 Windows 2000 的图形打印机一台。
- (7) 鼠标(Mouse)。

1.1.4 STATISTICA/Win 的启动

STATISTICA/Win 的启动可以按如下 3 种方法进行。

方法一 开机后,启动 Windows 95/98/2000,在桌面上如图 1-1 所示的快捷图标处双击鼠标左键,立即显示 STATISTICA /Win 的 STATISTICA Module Switcher (STATISTICA 模型切换)对话框(见图 1-2)。



图 1-1 STATISTICA/Win 图标

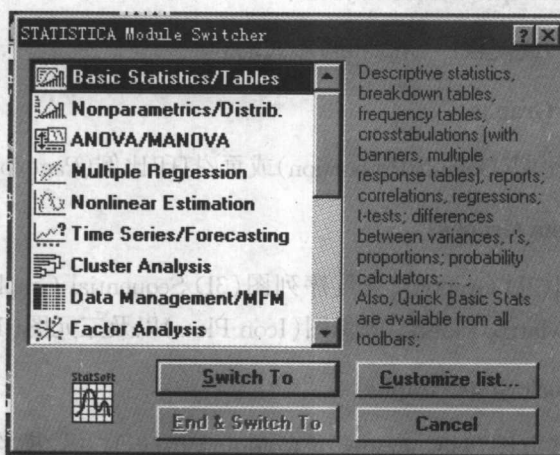


图 1-2 STATISTICA Module Switcher(STATISTICA 模型切换)对话框

方法二 开机后,启动 Windows 95/98/2000,单击【开始】⇒【程序】⇒(双击)【STATISTICA】⇒【STATISTICA】,也立即显示 STATISTICA /Win 的 STATISTICA Module Switcher 对话框(见图 1-2)。

方法三 开机后,启动 Windows 95/98/2000,双击桌面上“我的电脑”快捷图标(图 1-3),然后选用地址 D:\ ⇒ Stat,最后,双击“StatSoft 软件”图标(见图 1-4),也会显示 STATISTICA/Win 的 STATISTICA Module Switcher 对话框(见图 1-2)。



图 1-3 “我的电脑”图标



图 1-4 StatSoft 图标

STATISTICA Module Switcher(STATISTICA 模型切换)对话框(见图 1-2)有 3 个区:

- (1) 左侧(第一区)提供了本统计图表软件可选用的模型(包含数据管理、统计分析 with 图表制作等模块)。当前选用的模块是 Basic Statistics/Table(基本统计分析/制表)。
- (2) 右侧(第二区)显示当前模块的简要功能说明。本例说明“基本统计分析/制表”(Basic

Statistics/Table), 它包含有:

- ① Descriptive statistics(描述性统计分析);
- ② Breakdown tables(细分表);
- ③ Frequency tables(频数表);
- ④ Crosstabulations(with banners,multiple response tables)(带标题的多重反应列联表);
- ⑤ Reports(统计报表);
- ⑥ Correlations(相关分析);
- ⑦ Regressions(回归分析);
- ⑧ t-tests(t 检验);
- ⑨ Differences between variances(方差之间的变异分析);
- ⑩ Probability calculators(概率计算)等。

(3)下部(第三区)有4个选项:Switch To(切换),Customize list(显示常用模型清单),End & Switch To(结束并切换),Cancel(删除当前对话框)。

1.1.5 STATISTICA/Win 的主画面

当启动 STATISTICA/Win 以后,会显示如图 1-5 所示的 STATISTICA/Win 的主画面。

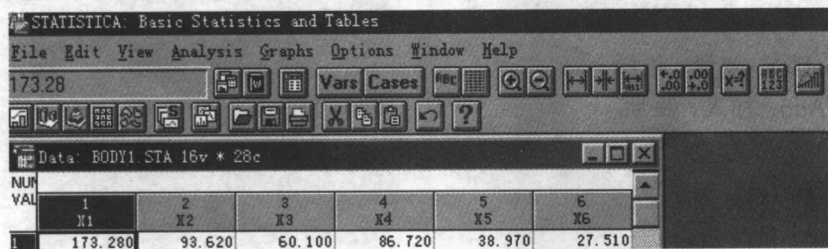


图 1-5 STATISTICA/Win 的主画面

STATISTICA/Win 主画面(图 1-5)的第一行是 StatSoft 软件标题。STATISTICA: Basic Statistics and Tables, 表明当前已进入 STATISTICA/Win 软件的“基本统计分析与制表”模块。右上角有  图标,依次是:缩小、放大、关闭/激活的图标。

STATISTICA /Win 主画面(见图 1-5)的第二行是菜单栏。它由 8 个下拉式菜单组成,它们是:

1. File(文件操作)

可以实现建立、导入、打开、保存、输出、打印等操作,有如下选项。

- 📁 New Data, 建立新的数据文件。
- 📁 Import Data, 导入数据文件。
 - ☆ Quick, 快速导入数据文件。
 - ☆ ODBC, 打开数据库连接。
- 📁 Open Data, 打开数据文件。
- 📁 Open Other, 打开其他文件。
 - ☆ New Data, 新的数据文件。
 - ☆ Date File, 日期文件。
 - ☆ Graph File, 图形文件。

- ☆ Scrollsheet File, 滚动工作单文件。
- ☆ Text/Output File, 文本/输出结果文件。
- ☆ STATISTICA BASIC Program, STATISTICA BASIC 程序文件。
- ☆ Command Language(SCL), 命令语言文件(SCL)。

- ↘ Save, 储存当前文件。
- ↘ Save As, 将当前文件另存为其他格式类型的文件。
- ↘ Export Data, 导出数据文件。
- ↘ Convert to Scrollsheet, 转变为滚动工作单文件。
- ↘ Print/Eject Current Pages, 打印/跳过当前页。
- ↘ Print, 打印。
- ↘ Print Files, 打印文件。
- ↘ Page/Output Setup, 页面/输出格式设置。
- ↘ Printer Setup, 打印机设置。

2. Edit(编辑)

可以对变量、个案进行添加、修改、复制、粘贴、删除与清除等操作, 有如下选项。

- ↘ Cell, 单元格。
- ↘ Variables, 变量, 指标。
 - ☆ Add, 加入, 添加。
 - ☆ Move, 移动。
 - ☆ Copy, 复制。
 - ☆ Delete, 删除。
 - ☆ Current Specs, 当前设定值。
 - ☆ All Specs, 所有设定(置)值。
 - ☆ Text Values, 文本值。
 - ☆ Date Values, 日期值。
 - ☆ Recalculate, 重新计算。
 - ☆ Shift(lag), 转变(滞后), 改变位置(滞后)。
 - ☆ Rank, 编秩。
 - ☆ Recode, 重新编码。
- ↘ Cases, 个案。
 - ☆ Add, 加入。
 - ☆ Move, 移动。
 - ☆ Copy, 复制。
 - ☆ Delete, 删除。
 - ☆ Names, 命名。
- ↘ Workbook/Headers, 工作表/抬头。
- ↘ Undo, 取消前一次操作。
- ↘ Cut, 剪切。
- ↘ Copy, 复制。