



GenStat®

GenStat

统计软件用户手册

Roger Payne Darren Murray Simon Harding
David Baird Duncan Soutar & Peter Lane.

郑可锋 张浩 祝利莉 等 译
王磊 Xiang-ming Xu 校

Roger Payne Darren Murray Simon Harding
David Baird Duncan Soutar & Peter Lane.

GenStat 统计软件用户手册



图书在版编目 (CIP) 数据

GenStat 统计软件用户手册 / (英) 佩恩 (Payne, R.) 等著; 郑可锋等译. —北京: 中国农业科学技术出版社, 2009. 12

ISBN 978 - 7 - 5116 - 0081 - 3

I. G… II. ①佩…②郑… III. 统计分析 - 应用软件, GenStat - 手册
IV. C819 - 62

中国版本图书馆 CIP 数据核字 (2009) 第 225046 号

责任编辑 梅 红

责任校对 贾晓红

出 版 者 中国农业科学技术出版社

北京市中关村南大街 12 号 邮编: 100081

电 话 (010)82109704(发行部)(010)82106630(编辑室)

(010)82109703(读者服务部)

传 真 (010)82106636

网 址 <http://www.castp.cn>

经 销 者 新华书店北京发行所

印 刷 者 北京富泰印刷有限责任公司

开 本 889 mm × 1 194 mm 1/16

印 张 16.5

字 数 420 千字

版 次 2010 年 4 月第 1 版 2010 年 4 月第 1 次印刷

定 价 40.00 元

内容简介

本书是英国 VSNi 公司开发的 GenStat 统计软件的配套教材。基本内容包括 GenStat 的快速起步、数据输入、基础统计、电子表格、线性回归、方差分析、RMEL 混合模型分析、多元统计分析及时间序列等，其中包括近年来的一些较新进展。大部分统计方法都给出了 GenStat 软件的操作过程及输出结果的解读，便于广大科研人员的教学、自学和应用。

本书可作为农业科研机构科研人员、高等农业院校或综合性院校生物类各专业本科生、研究生的教材，也可供各领域需要进行数据分析处理的实际工作者自学参考。

声 明

本书的出版得到英国 VSNi 公司的许可，所采用的图、表属于英国 VSNi 公司所有，若文中出现错误，以英文原版为准。

特此声明。

序

欢迎使用 GenStat 统计软件用户手册。和世界上很多从事生物科学研究的科学家一样，你也选择了使用 GenStat 统计软件用于分析科研数据。

英国洛桑自 1843 年 John Bennet Lawes 开展小麦试验时开始农业研究，到目前为止，洛桑试验站是世界上运行时间最长的试验站。同时，洛桑试验站也率先在生物科学研究中应用统计方法。1919 年 Ronald Fisher 在研究小麦前期累积试验结果和后续试验结果时，他意识到在整个农业和生物研究中需要提高统计技术，故他和他的同事在 19 世纪 20 年代和 30 年代奠定了现代应用统计的基础。

当 Fisher 的继任者 Frank Yates 将一台 Elliot 401 计算机用于自己的统计工作时，洛桑试验站开始了统计计算工作。Fisher 将传统的统计研究延伸到现实中的生物统计，也促进了洛桑试验站在实践中更有效地应用现代统计软件。1968 年 John Nelder 接替 Yates 成为统计方面的负责人后，洛桑试验站开始开发 GenStat 统计软件。1985 年 Nelder 退休后，Roger Payne 继续领导 GenStat 的开发工作。

19 世纪 70 年代，GenStat 开始走出洛桑试验站。1979 年，世界上最古老的计算技术公司之一 Numerical Algorithms



Group (NAG) 开始负责 GenStat 的经营。近年来, VSNi 公司负责 GenStat 的开发和销售。VSNi 公司组建于 2000 年, 作为洛桑试验站和 NAG 旗下的子公司, 其整合了洛桑试验站的开发团队和 NAG 的商业经营团队, 促进了 GenStat 的开发和市场营销。然而, 开发团队作为洛桑试验站和 VSNi 公司连接者, 仍然与研究者保持着紧密的联系。所以, 用户不仅能从公司严格的质量控制中获益, 而且能感受到浓厚的研究氛围。

GenStat 的一个重要特点是, 其开发团队 (包括合作团队) 的成员本身就从事广泛的应用统计工作, 如方差分析、试验设计、线性模型、典型变量分析及最近开发的混合模型分析等。因此, GenStat 不仅能运行现有的统计模型, 而且还不断地发展和应用新的模型。所以我们相信, GenStat 不仅能满足生物统计方面的需求, 而且是解决其他应用统计问题的理想工具。

Roger Payne

Professor Roger Payne

Technical Director

November, 2009.



前言

随着现代信息技术的发展，在农业和生物科学研究中利用统计软件进行统计分析越来越普遍。统计软件不仅为相关科研工作者消除了大量数据处理的烦恼，同时可以促进其对统计理论和方法的深入理解，不断提高人们应用统计学的能力，统计软件已成为农业和生物科学研究工作中最有力且不可缺少的工具。GenStat 统计软件是一款强大而灵活的统计软件，它属于完全交互式系统，具有先进的图形化工具和友好的图形化用户界面，还有强大的统计学程序编制功能。而且，GenStat 具有悠久的历史，经过不断更新发展，它始终能活跃在生物统计学技术的最前沿。

GenStat 统计软件有比较详细的原版使用手册，但对于大多数中国读者而言，阅读这些英文材料无疑会耗费较多的时间。因此，如何让中国读者快速熟悉和使用 GenStat 统计软件，并能解决农业和生物科学研究中实际问题，是编译《GenStat 统计软件用户手册》一书的初衷。因译者水平有限，本书的翻译出版只是一次初步的尝试，但愿它能起到抛砖引玉之作用。

本书以 GenStat 7.0 统计软件包作为实现复杂统计计算的工具，从而省去了大量的篇幅着重介绍各种试验设计方法、各种统计分析方法及其适用条件、结合具体问



题正确选用统计方法的技术以及对计算结果的正确解释和应用。在一切从实际出发的思想指导下，经过合理调整结构及书写形式，貌似复杂的统计问题被化繁为简，更实用更方便。

全书分为 3 部分共 11 章。第 1 章部分由第 1 章至第 4 章组成，主要介绍 GenStat 软件的基础知识，包括 GenStat 的快速起步、数据输入、基础统计、电子表格。第 2 部分由第 5 章至第 9 章组成，着重介绍 GenStat 在线性回归、方差分析、RMEL 混合模型分析、多元统计分析及时间序列方面的应用。第 3 部分由第 10 章和第 11 章组成，分别介绍了 GenStat 其他命令和其他统计方法。

整体编译工作安排如下：第 1 章至第 4 章由郑可锋、张浩、祝利莉同志编译，第 5 章至第 9 章由郑可锋、张浩、张小斌和姚旭国同志编译，第 10 章和第 11 章分别由胡为群和叶少挺同志编译，郑可锋同志负责全书编译工作的统筹和安排。

本书可用作生物类研究生、本科生、大中专生的统计学教材，亦可作为涉农高等院校和科研机构的教师、学者、科技人员、管理者等学习和应用统计方法的参考书，还可作为用 GenStat 软件解决统计问题的实用手册。

本书的问世，与英国 VSNi 公司和浙江省农业科学院在经济上给予的大力扶持息息相关，与中国农业科学技术出版社的热情关心和帮助分不开。中国水稻研究所的王磊博士对本书内容提出了很多宝贵的建议。英国洛桑试验站的 Xiang-ming Xu 博士为本书的初稿作了大量认真而又细致的校对工作。在此，一并表示衷心的感谢！

最后，由于我们水平有限，编译时间仓促，难免有错漏之处，敬请读者批评指正，以便我们不断改进。

译者

二〇一〇年三月

目 录

第 1 部分

第 1 章 快速起步	1
1.1 什么是 GenStat?	1
1.2 使用菜单	2
1.3 执行命令	9
1.4 执行程序.....	12
1.5 窗口界面.....	13
1.6 练习.....	16
第 2 章 数据输入、计算和显示	17
2.1 GenStat 数据结构	17
2.2 数据输入.....	18
2.3 从 ASCII 文件中读取数据	23
2.4 显示数据.....	26
2.5 转换数据结构.....	28
2.6 保存数据到文件.....	29
2.7 计算.....	31
2.8 GenStat 图形向导	38
2.9 数据输入、计算以及显示命令.....	42
2.10 练习	49
第 3 章 基础统计	52
3.1 比较两个样本.....	52
3.2 分类数据摘要.....	58
3.3 分组数据摘要.....	59
3.4 分类变量间的关联.....	60
3.5 将数据转移至其他应用程序.....	63
3.6 基础统计命令.....	66
3.7 练习	67
第 4 章 GenStat 电子表格	69
4.1 在电子表格中输入数据.....	69
4.2 数据检验.....	73
4.3 插入和删除行或列.....	76
4.4 定义数据值子集.....	77



4.5	排列数据	81
4.6	数据处理	83
4.7	书签和注释	87
4.8	动态数据交换 (DDE)	89
4.9	读写数据库数据	92
4.10	其他功能	99
4.11	命令	99
4.12	练习	102

第2部分

第5章	线性回归	104
5.1	拟合直线	104
5.2	多元线性回归	109
5.3	逐步回归和所有子集回归	114
5.4	分组数据回归	118
5.5	拟合曲线和多项式	125
5.6	广义线性模型	132
5.7	回归命令	134
5.8	其他功能	135
5.9	练习	135
第6章	方差分析	140
6.1	单向方差分析	140
6.2	双向方差分析	141
6.3	随机区组设计	144
6.4	拟合对比	147
6.5	设计实验	152
6.6	模型公式语法	157
6.7	不平衡设计	159
6.8	裂区设计	164
6.9	方差分析命令	167
6.10	其他功能	170
6.11	练习	170
第7章	REML 混合模型分析	175
7.1	线性混合模型：裂区设计	175
7.2	线性混合模型：非正交数据	179
7.3	空间分析	185
7.4	重复测量	189
7.5	REML 分析命令	194
7.6	其他功能	195



7.7 练习	195
第8章 多元分析	196
8.1 主成分分析	196
8.2 典型变量分析	201
8.3 多维尺度	204
8.4 分层聚类分析	207
8.5 非分层聚类分析	210
8.6 多元方差分析	216
8.7 分类树	218
8.8 其他功能	225
8.9 练习	225
第9章 时间序列	226
9.1 时间序列研究	226
9.2 ARIMA 模型拟合	227
9.3 时间序列命令	230
9.4 练习	231

第3部分

第10章 更多命令及其用法	232
10.1 使用命令	232
10.2 重复命令序列	237
10.3 GenStat 命令语法	238
10.4 缩写规则	240
10.5 重复语句	241
10.6 压缩列表	242
10.7 后缀标识符和指针	243
10.8 未命名数据结构	244
10.9 练习	245
第11章 其他统计方法	247
11.1 六西格玛	247
11.2 调查数据	247
11.3 地质统计学	248
11.4 生存分析	248
11.5 重复测量	249
11.6 多重实验	250
主要参考文献	251



第1章 快速起步

1.1 什么是 GenStat?

GenStat 是一个用于总结、显示和分析数据的综合统计系统。使用计算机分析数据可以节约大量时间和精力,但是如何让计算机做到这一点是一个让人头疼的问题。使用诸如 Fortran、C++ 之类的一般编程语言可以用于处理数学和人机交互的细节,但是一个普通的分析方法就需要很长的程序来实现。专业的统计包可以提供一个易于操作的环境,不过菜单中往往只有寥寥数项可以实现一个标准的分析,当它所面对的是非标准的分析时,往往很难甚至不可能改变。

Windows 版的 GenStat 使得鱼与熊掌可以兼得,它使用由数个窗口和菜单项组成的 Windows 标准界面实现了标准分析,结合了编程语言的灵活性和菜单驱动程序包的操作简单性。菜单使用 GenStat 的高级统计编程语言自动生成执行所选项的命令,同时,用户也可以轻松地使用命令语言来构建自定义的非标准分析。

GenStat 可以是实现下列功能:

- 处理用 GenStat 自带的表输入或者从别的计算机文件导入的数据;
- 使用柱状图、箱线图、散布图、线图、格图、轮廓图和三维表面图等图表展示数据;
- 使用包含大量数学和统计学方法的常规计算工具转化数据;
- 使用线性或非线性回归模型、广义线性模型、广义相加模型、广义混合线性模型、多层广义线性模型对参数关系进行建模;
- 设计和分析从单因素方差分析到复杂多层设计的实验,后者可以通过使用平衡方差分析或者约束性最大似然法(其中包括关联结构建模)来实现;
- 使用典型变量分析、主成分分析、主坐标分析、对应分析、偏最小二乘回归、分类树分析和聚类分析等多变量分析技术确定数据模式;
- 分析来自整群采样和非结构化采样的结果;
- 绘制控制图,打印 Pareto 图,计算容量统计;
- 使用 Box-Jenkins 模型或光谱分析法分析时间序列;
- 通过变量分析、使用 antedependence 结构或对时间相关性建模分析重复测量;
- 使用克里格法或空间点过程分析空间模式。

这些技术被广泛应用于包括农业、生态学、基因学、医学研究在内的生物学领域和工业研究、质量控制以及经济和社会调查。可以说在所有应用统计学的科学研究、商业运行、政府管理和教育领域,这些技术都无处不在。

本手册所介绍的 GenStat 是针对使用 Windows95、98、NT、2000 和 XP 系统的个人电脑推出的第七版。Windows XP 下的最低推荐配置如下:奔腾处理器、32 兆内存。GenStat7.1 版本菜单是用底层命令语言编写的,该语言也可以用于编写非标准分析,适用于包括工作站和大型机版本在内的所有 GenStat 系统。本书第二部分详细说明 GenStat 的统计学内容,包括介绍基础方法、解释输出和描述相应的 GenStat 命令。除此之外,GenStat 参考手册更是对每一条命令做了描述。这些文档可



以从帮助菜单以 PDF 格式打开（图 1.4）。

本入门手册指导用户如何使用 GenStat 完成简单的统计任务。每章末尾都配有小习题，可用于提高你对练习使用 GenStat 的兴趣。习题和示范样例中使用的数据文件是随本软件提供的，位于 Gen7ed \ Introduction 文件夹下。

声明：Microsoft、Windows、Windows NT、Windows 2000、Windows XP、Excel、Explorer、Word 和 Access 是微软公司的商标和注册商标。

1.2 使用菜单



GenStat 7th
Edition

图 1.1

在使用 Windows 操作系统的电脑上，你可以从程序菜单中选择 GenStat 图标（图 1.1）或者双击桌面上的该图标打开 GenStat。

电脑上会开始运行两个程序，一个是 GenStat 客户端，一个是 GenStat 服务器。前者用于控制 GenStat 的 Windows 界面，收集信息并发送给 GenStat 服务器。GenStat 服务器在后台运行，执行计算功能。图 1.2 是 GenStat 在 Windows XP 下启动后的界面，基本与 Windows 95、98、NT4.0、2000 下的界面一致。客户端图标显示在任务栏中，服务器的图标则显示在右侧的托盘中。

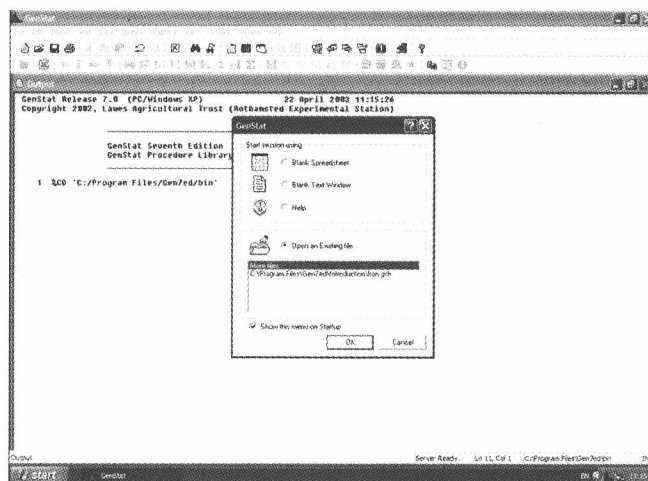


图 1.2

图 1.2 显示客户端使用标准 Windows 界面，菜单和工具栏在窗口顶部，状态栏位于底部。工具栏有两个，较低的一个用于显示 GenStat 表，有关该对象的细节请见第 4 章。客户端可生成数个子窗口。系统初始化时启动会话菜单，在该菜单下你可以打开一个新的 GenStat 表或者文本框，这样你可以访问 GenStat 帮助系统，也可以打开先前会话中使用的文件。如果你不想在以后打开 GenStat 时出现该菜单，你可以反选“启动时显示该菜单”项，系统就只会打开一个输出窗口。初始状态下，该窗口只显示系统的版本信息，之后将显示用户操作的输出。窗口标题栏高亮显示，表明该窗口是当前窗口，鼠标指针会在该窗口里跳动。状态栏的左边也会显示当前窗口的名称，右边则显示 GenStat 服务器状态、鼠标位置、当前工作路径以及当前窗口模式（插入和覆盖）。我们可以点击菜单栏里的 Windows 选项察看当前可用的窗口（图 1.3）。如 Input Log 记录了客户端向服务器发送的用户请求的命令。

你可以通过使用一般的 Windows 方法来改变 GenStat 窗口的外观，例如点击窗口右上角的盒状



图标可使 GenStat 窗口填满整个屏幕，也可以自由改变每个子窗口的大小来合理利用空间，例如你可以点击 Input Log 窗口，再改变 Output 窗口使得它们可以共享空余的空间。

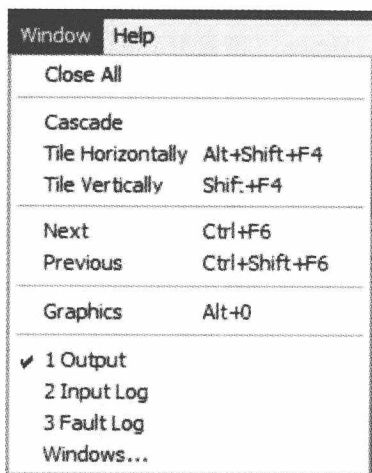


图 1.3

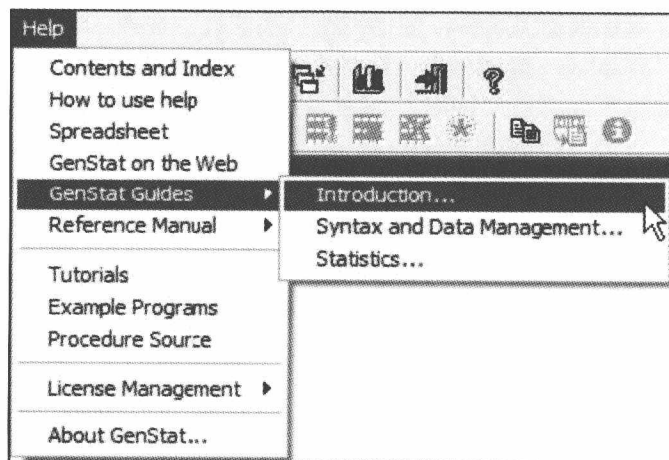


图 1.4

菜单中的许多选项都是 Windows 标准程序选项。1.5 节将简要介绍这些选项，本书的其他章节则着重介绍那些 GenStat 特有的完成统计分析和展示的选项。不过在开始这些之前，我们先来熟悉一下帮助菜单所提供的在线帮助功能，帮助是菜单栏上最右侧的下拉菜单，如图 1.4 所示。通过该菜单我们可以访问一些有用的资源（点击“Introduction”即可得到本书的 PDF 版本）。

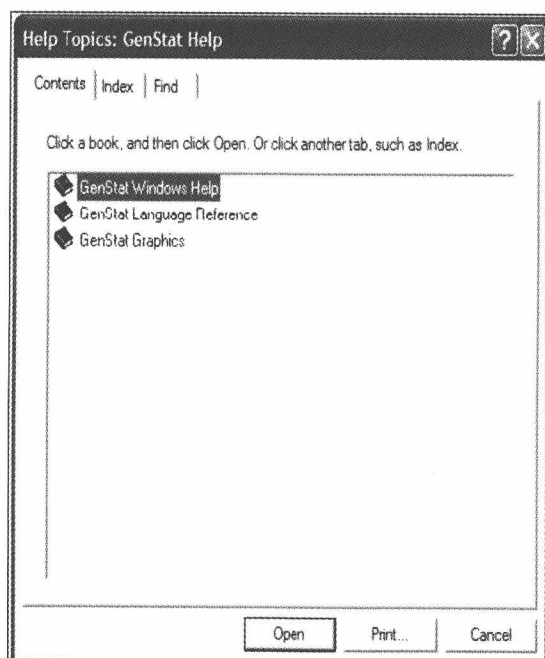


图 1.5

点击“Contents and Index”，你将进入 GenStat 的主要帮助系统。初始界面如图 1.5 所示，



“Contents” 标签下可以看到三本“书”：《GenStat Windows Help》提供 Windows 版 GenStat 及其菜单的相关信息。《GenStat Language Reference》描述 GenStat 命令语言和它所提供方法的细节。《GenStat Graphics》提供辅助 GenStat 绘图的子系统信息。

你可以根据 Windows 惯例浏览这些信息，选择一本书，点击“Open”键打开，直到打开某页；然后“Display”键会出现，你就可以阅读该页的内容了。

你也可以选择“Index”标签通过索引来浏览主题。在图 1.6 中，我们选中了有关“t-test”的条目。点击“Display”会产生一个列表，你可以通过它选择显示相关的主题。你可以像往常一样使用独立的帮助区保存书签，通过热连接访问其他区域等。你可以点击窗口右上角的按钮或者“Exit”和“Cancel”键退出帮助系统。

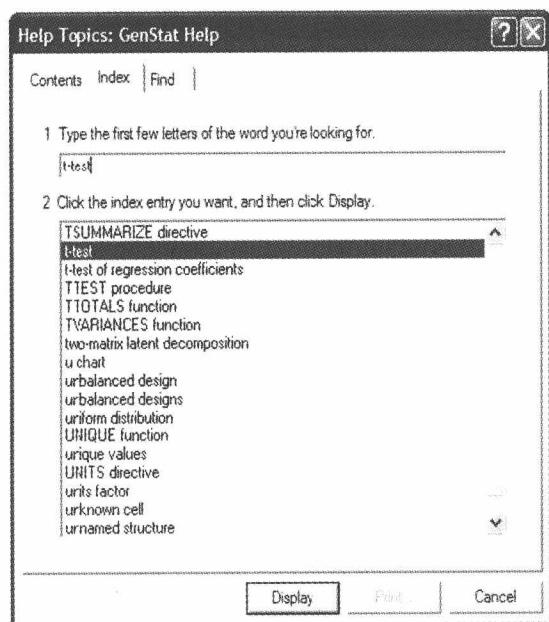


图 1.6

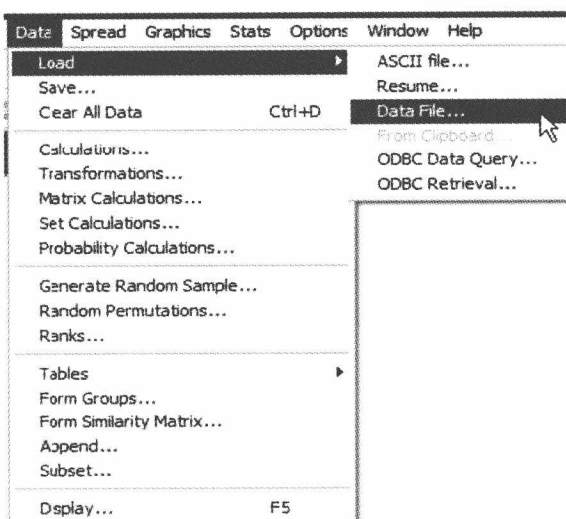


图 1.7

通过点击帮助菜单第七行“Tutorials”，你可以观看视频或运行互动的练习，从而更好的了解 GenStat。

通常打开 GenStat 要做的第一件事就是访问数据。我们将在之后的章节展示直接从屏幕输入和从多种文件中导入数据的方法，在本节中我们只介绍最简单的方法，即从先前的 GenStat 会话导入其中的数据。点击菜单栏中的“Data”，该下拉菜单的选项如图 1.7 所示，选择“Load”然后点击“Data File”，此时将弹出“Select Input File”对话框，你就可以选择一个保存有先前数据的文件了。

该对话框开始于当前工作目录，如果 GenStat 是首次运行，当前工作目录就是包含 GenStat 可执行文件的目录或文件夹。你可以点击当前所处目录右侧的上一级目录图标到保存样例和练习数据的文件夹下，然后点击文件夹“Introduction”。你可以通过点击菜单左下角的“Set as Working Directory”按钮将其设为工作目录。在图 1.8 中，我们在文件类型中选择了 Genstat Spreadsheet File (*.gsh) 并选中了 Iron.gsh 文件。最后你就可以点击“Open”导入数据了。

GenStat 从该文件中导入数据，然后再输出窗口显示如下信息：

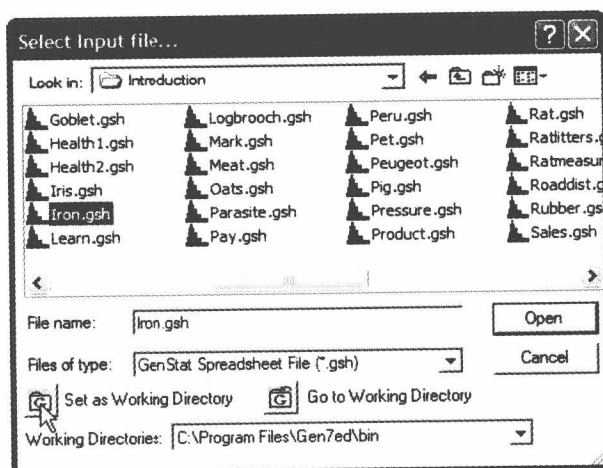


图 1.8

Identifier	Values	Missing	Levels	
sample	136	0	12	
Identifier	Values	Missing	Levels	
Site	136	0	6	
Identifier	Minimum	Mean	Maximum	Values
FE	200. 6	246. 6	308. 2	136
Identifier	Minimum	Mean	Maximum	Values
weight	11. 36	12. 42	13. 00	136

该文件中共有四组信息和数据结构。其中两组只保存了一些数字，名为 Fe 的结构保存测量的铁浓度（PPM），weight 保存被分析的沙子的质量。这些结构的值与整个研究中的 136 个沙子样本相对应。另外两组结构对样本进行了分类，site 的每个值记录做出分析的实验室编号，sample 记录交由实验室分析的沙子样本编号。请注意分类性结构与单纯的数字列是不同的。根据你所使用的 GenStat 不同的设置，Output 窗口可能会显示你用于导入数据的命令。你可以使用图 1.30 中的“Options”菜单来关闭命令显示。

你也可以通过“Data”菜单得到当前保存在 GenStat 中的数据列表，点击“Display”产生“Data Display”对话框（图 1.9）。你可以从菜单的左边选择你需要的数据结构显示在右边：在图 1.9 中，我们选中的是“All Data”。

可以看到，GenStat 中保存数字的数据结构称做变量，而分类性结构被称为因素。GenStat 提供一系列数据结构以满足不同数据类型的需求，这两种是最常见的。点击“Options”按钮，下拉菜单中的其他按钮可以实现删除、重命名和显示（比如在“Output”窗口中打印）右边选中的结构。你还可以拖曳列表中的结构到大多数的 GenStat 菜单中。点击“Close”会关闭该菜单。

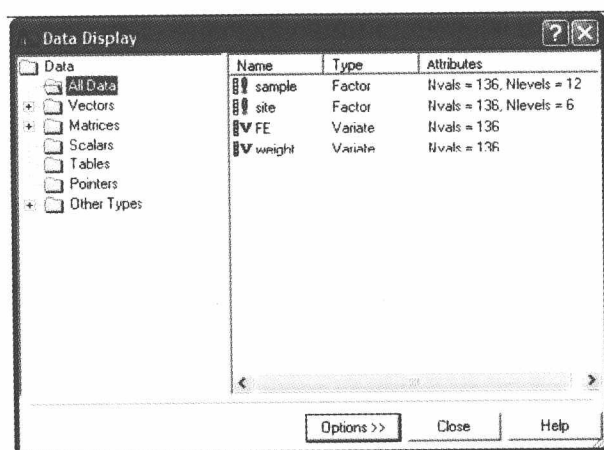


图 1.9