

科研教学必备工具

唐启义 冯明光 著

# 实用统计分析 及其计算机处理平台

中国农业出版社

# 国家杰出青年科学基金资助



## 内 容 简 介

本书是一本简明实用的统计分析电子工具书，它从应用角度简要地阐述了经典统计学和现代生物数学，如方差分析、正交回归试验设计、品种区域试验分析(包括最新的 AMMI 技术)、多元分析等 100 多种实验数据统计分析和模型模拟方法原理及数据处理技术。全书由文字和磁盘两部分组成，共 6 篇 22 章，配以作者开发的计算机全屏幕交互式 DPS 数据处理软件系统压缩磁盘 1 张。

本书内容特点：一是注重专用性和通用性结合。即既有适用某一专业领域使用的统计分析技术，如生物测定；也有各行业领域通用的统计分析和建模技术。二是“傻瓜”性和灵活性结合。一般读者只需输入原始数据，即可进行各种繁冗的数据处理和分析；但对有经验和水平较高读者，在数据转换、模型模拟分析等方面又可灵活运用，挥洒自如。三是注重实用和可操作性，在试验数据编辑整理和分析技巧方面，均附有大量的应用实例及其在 DPS 系统中编辑、操作和分析的具体步骤，便于读者掌握。

本书可供从事自然科学和社会科学如农林牧渔、医药卫生、气象、水文地质、工程计算、市场调研、经济等领域的广大科研、教学、管理和技术推广人员使用，也可作为有关专业大学生和研究生学习生物统计课程的参考书及实践工具。书末附有用户联系卡，便于读者信息反馈和今后软件升级。

## 前 言

呈现在读者面前的《实用统计分析及其计算机处理平台》，既是一本以实用为目的的统计工具书，更是一张请柬。它邀请读者进入 DPS 数据处理系统中漫游，高复杂繁冗而又沉闷乏味的统计分析于轻松自如和乐趣之中。

现代科学研究，不管是自然科学还是社会科学，离不开统计分析。统计学的形成可以追溯到本世纪 20 年代。当时 Fisher 等以农业科学试验为对象，探讨试验资料的整理和统计分析方法。1922 年，Hayes 研究亲代与子代蛋白质含量的相关性，提出了相关分析的理论。翌年，Fisher 和 Maekenzie 研究肥料对马铃薯产量的影响，首次提出了方差分析和交互作用的概念。1925 年，Engedow 研究不同品种在不同地区和不同年份产量的表现时，进一步阐述了因素间的交互作用。Fisher 等正是在研究农业科学试验方法的基础上，建立起统计学的基本理论和方法。这些理论和方法目前已广泛应用于科学研究、工农业生产和现代经济管理。统计学的形成和发展始终同生物学研究和农业科学试验紧密相联，产生了生物统计学的分支，使数理统计分析技术如方差分析、相关分析、回归分析、列联表分析等在农业科学研究中广泛应用，极大地推动了农业和生物学的发展。

随着经典数理统计科学的发展，本世纪 40~50 年代，多元统计分析在 Pearson 及 Fisher 等人开拓性工作基础之上又有了坚实的理论基础。如关于多元正态分布的理论著述，Anderson (1958) 的《多元统计分析导论》等，标志着统计学进入到一个新的发展阶段。但是，多元统计分析的发展、完善和广泛应用，紧密地伴随着计算机和计算数学的发展。由于多元统计分析的计算工作量十分惊人，没有计算机是不可想象的。正是由于计算机技术的飞速发展，一些独特的多元统计分析技术，如主成分分析、典型相关（亦译为典范相关）、因子分析、对应分析、判别分析、聚类分析等等，在农林、医药、气象、地质、生物、心理、工业、经济、社会等许多研究和生产领域获得了广泛应用。目前，多元统计分析已是生物统计学中最活跃的分支学科之一。

传统农业研究以定性描述为主。Fisher 等创立的一套经典统计方法在定量研究方面迈出了一大步。但囿于其固有缺陷，试验结果的统计推断不能对有关数量关系从机理上作合理解释，因而其统计模型仍属经验范畴。因此，现代生物科学工作者们正力图使生理、生态、遗传、生物化学等学科理论与数学有机地结合，用模型来描述、解释自然现象和生物运动规律，最终解决科研和生产中的实际问题。通过计算机建立模型，如模拟作物生长发育过程或病虫害的发生变化过程，又用模型的模拟预测技术指导农业生产和病虫害防治等，是生物统计技术的研究热点之一。

本世纪 60~70 年代以来，一些新的统计理论和分析技术发展起来。如 60 年代 Zadeh 提出的模糊集合论已在各个领域应用十分广泛。实践证明，模糊数学方法在农业中用于病虫测报、种植区划、品种选育等方面，以及在图像识别、天气预报、地质地震、交通运输、

医疗诊断、信息控制、人工智能等领域的应用已初见成效，其发展前景广阔。作为预报和控制的有力工具，研究系统变化规律的时间序列分析技术目前正被广泛用于地质、石油、气象、商业、工程控制方面，在农业中也已用于作物病虫害的中长期预测预报，如李绍石等于 80 年代中期将这一新技术应用于湖南省水稻主要病虫害中长期预测预报，取得了巨大的经济效益。灰色系统是邓聚龙于 80 年代初提出的用于控制和预测的新技术，目前亦在我国农业、社会和经济等领域应用广泛，并已取得显著成就。

生物统计学的发展源远流长，内容十分丰富。迄今，国内外关于统计学或生物统计学的专著和教科书已有多种版本见诸于世。如果按照一般统计学著述的内容和体裁，我们没有必要再推出一本教科书似的理论专著，让广大生物学科工作者在复杂的数学公式堆中艰难跋涉。《实用统计分析及其计算机处理平台》的编写宗旨强调实用性。作者在内容上刻意求新，形式上追求通俗易懂，在取材上尽可能全面，并注意反映当今国内外在试验设计和统计分析方面的最新研究进展。书中介绍的 100 多种统计分析方法，可以满足各种水平的统计分析需要，且都能在作者设计制作的 DPS 计算机数据处理系统中随意调用，操作自如。本着服务于读者、服务于用户的最高宗旨，作者对入选的 100 多种统计分析方法分门别类，每一类(种)统计分析技术按方法或原理简介和操作示例两个基本部分编写，全书共分六篇 22 章。同时，作者按书中篇目类别设计 DPS 系统的套叠下拉式主菜单和子菜单，将带给读者和用户极大的便利。下面分篇简单作一介绍，以引导读者高效率地使用书中的各种分析技术和 DPS 数据处理系统。

第一篇(共两章)主要介绍 DPS 系统的性能、特点和操作要领。DPS 数据处理系统是作者独立研究开发的大型通用多功能数据处理分析应用软件系统。该系统在 IBM PC 及其兼容机、直接写屏的软汉字操作系统或中文 Window 95 下运行。因此，在使用 DPS 系统之前，建议读者首先浏览本篇内容，熟悉该系统本身将使读者在数据处理和分析中事半功倍。

第二篇主要为各专业读者提供最基本的田间试验数据的处理分析技术。对于经常从事田间试验的读者，如研究土壤、施肥等农艺措施对作物产量的影响，方差分析和正交回归试验设计两章将会大有帮助。对于作物育种工作者，可能更加关心品种和环境的互作效应研究。对此，方差分析和品种比较区域试验两章中提供的分析技术将是必不可少的，功能卓越的 AMMI 模型技术可能更有帮助。对于从事动植物生态学研究的科技人员，抽样设计及群落参数统计分析两章介绍的技术将是得心应手的工具。第 9 章专门介绍生物测定的统计分析技术，既有经典的机率分析，也有 90 年代发展起来的时间—剂量—死亡率模型分析技术。总之，第二篇的编写充分考虑到经典试验统计的完整性和系统性，如提供了比较全面的试验设计和试验结果统计分析的内容，尤其对方差分析给予了较为详尽的介绍。因此，各个层次、各个领域的读者都可使用这些基本的统计分析方法。

第三篇以较大篇幅介绍多变量数值统计分析技术，并注意反映目前国内外的最新研究动态，将最新的技术提供给读者。近年来，多变量统计分析技术已广泛应用于各个领域的科学试验中，因为人们越来越多地将研究对象作为一个有机结合的整体即多元素系统(亚系统)对待，而不是分割开来，作为单独事物处理。“从树木到森林”，抓住系统运动变化的本质，舍弃次要因素。简化系统的结构，以探索系统运动变化的总趋势，为人类征服自然、

改造自然服务。由此，主成分分析、因子分析、对应分析、典型相关分析、聚类、判别分析以及马尔可夫链分析等多因素分析技术倍受自然和社会科学各领域科研人员的重视。

第四篇和第五篇向读者介绍相当灵活的通用数学模型的模拟分析和数值计算工具。因第二、三篇的分析方法和程式基本上都有完整的体系和要求，读者必须按部就班地照着去做；而利用第四、五篇的通用工具，读者将尽可能发挥自己的专长和展现自己的学识水平，实现思维活动的飞跃。例如，模拟作物生长发育过程或作物病虫害的发生过程，建立结合自己专业和学科特点的数学模型，而不是套用别人的现成模型。借助于 DPS 系统的强大功能，读者可在计算机上实现自己的思维活动，获得所想即所见，所见即所得的效果。

第六篇为读者提供三大类不属于经典统计学但近 20 余年来发展迅速且应用已经相当广泛的模型分析技术，即模糊数学分析、时间序列分析和灰色系统分析。只要读者有兴趣，DPS 系统可以帮助读者熟悉和掌握上述分析技术，并应用到研究实践中去。

综上所述，本书和 DPS 数据处理系统是密不可分的整体。作为通用的数据处理和统计分析的工作平台，作者在系统的设计方面注重专用性与通用性以及“傻瓜”性和灵活性的关系，以适应不同层次的读者需要。

**专用性和通用性** DPS 的有些功能是专为某些领域或完成某一类工作所专用的，如“生物测定”是专为从事化学或生物制剂的生物活性测定的人员而设计；种群空间分布型及群落参数统计分析、品种区域试验等章节都是服务于某个领域的科学工作者。相对来说，它们的用途比较专一。

但是，本书中绝大部分统计分析和建模技术是通用的，例如丰富的数学函数、交互方式的矩阵运算功能以及通用数学模型的模拟分析技术。DPS 处理数学公式的功能是非常独特的，有利于读者发挥潜在的想象力和创造力。同时，数学计算功能，特别是矩阵交互式处理功能，在当前微机普及的情况下，将有助于高等院校数理统计课程的教学，使繁冗无味的矩阵运算得以在计算机上精确而快速地进行，在一定程度上提高学生的学习积极性和教学质量。

**“傻瓜”性和灵活性** 在 DPS 平台上进行经典统计分析的过程十分简单，读者只需要编辑自己的实验数据，将它们定义成数据块，然后调用分析功能，按回车键就可在当前光标下面得到分析结果。有些统计分析方法需要读者输入参数，系统会自动给以提示，如用户事先没有相关的经验，系统会给出缺省值，用户只管按回车键，同样可以获得较优的分析结果。又如时间序列分析，系统以直观的图形方式让读者确定参数，只要读者稍有常识，无须牢记一大堆统计分析过程中参数辨识准则，就可成为高水平的建模能手。再如在系统聚类分析中，聚类结果一次成“像”，给出的是非常直观的聚类谱系图，等等。系统的这些功能，好比一只傻瓜型照相机，使用时只需按快门。

另一方面，DPS 系统在数据处理和模型模拟中又比较灵活，完全取决于读者的经验和知识水平。如数据行列的转换功能、自定义数学模型的拟合技术、逐步回归及逐步判别分析中  $F$  检验临界值的随意调整等等，使读者大有用武之地，挥洒自如。

应当指出，本书不是一本基础理论教材，而是一本不需要高深数理统计知识即可看懂的统计方法论专著。事实上，要把本书中每种统计方法的理论背景在这样一本书中完全交

待清楚是不太可能的，因为有些章节本身就是一本书的内容。作者的主要目的是要把读者从繁琐的统计分析中解放出来，既要用最适合、最先进的统计分析技术为自己的研究工作服务，又不被它们复杂的数学公式和运算所困扰。如果本书及其 DPS 数据处理系统，能使读者在数据处理和统计分析方面有所获益或帮助，作者的基本目的就达到了。

值此书出版之际，第一作者特别铭记业师李绍石研究员（湖南省农业厅）长期给予的关怀和培养，并感谢原农业部全国植保总站、湖南省农业厅植保站以及湖南省桃源县农业局对早期工作的支持和帮助。同时特别感谢胡国文研究员（中国水稻研究所）对作者多年的关怀和帮助。

本书的出版得到国家杰出青年科学基金资助，部分内容系国家“九五”攻关课题“水稻主要病虫害综合防治技术研究”（编号 96-005-01-01-06）项目的成果。本书在编写过程中，得到中国水稻研究所、浙江农业大学的支持，中国农业出版社在出版方面给予优先安排和合作，在此一并致谢。

最后，DPS 数据处理系统与国内外同类电脑统计分析软件相比，虽有独特之处，但毕竟是首次面世，加之作者本身专业学识的局限性，很可能存在种种不尽人意之处。对此，恳切希望广大读者提出宝贵的批评意见，尤其欢迎有典型意义的研究实例，使作者在对系统进行改进更新时参考，扩大系统的辐射面，以便更好地为读者服务。

作者

一九九七年五月于杭州

# 目 录

前言

第一篇 DPS 数据处理系统.....	1
第 1 章 DPS 系统概述.....	2
第 1 节 系统功能简介.....	2
第 2 节 DPS 系统的安装与使用.....	4
第 3 节 数据编辑处理.....	8
第 4 节 数据计算及其数学函数的定义.....	10
第 5 节 数据统计分析及其建模基本步骤.....	11
第 6 节 常用功能键的定义.....	12
第 7 节 其它.....	14
第 2 章 DPS 数据处理系统操作指南.....	15
第 1 节 数据编辑.....	15
第 2 节 数据的处理.....	16
第 3 节 数据统计的数学函数及其应用.....	18
第 4 节 常用统计分布函数.....	20
第 5 节 统计检验临界值函数.....	23
第 6 节 矩阵计算及其函数.....	23
第 7 节 数据作图和图形编辑.....	29
第二篇 试验统计与生态模型.....	34
第 3 章 样本基本特征参数与生物种群空间分布型.....	35
第 1 节 样本次数分布与基本特征参数计算.....	35
第 2 节 生物种群空间分布型的聚集度指标测定.....	37
第 3 节 生物种群空间分布型—频次分布检验.....	39
第 4 节 负二项分布公共 $k$ 值估计.....	43
第 4 章 方差分析.....	46
第 1 节 方差分析基本原理和步骤.....	46
第 2 节 单向分组完全随机区组设计.....	49
第 3 节 单向分组随机区组设计.....	51
第 4 节 组内分亚组的单向分组随机区组设计.....	53
第 5 节 双向分组(组内无重复)完全随机区组设计.....	55
第 6 节 双向分组完全随机试验设计.....	57
第 7 节 双向分组随机区组试验设计.....	62
第 8 节 双因素裂区试验设计.....	67
第 9 节 协方差分析设计.....	70

第5章 正交回归试验设计及其分析.....	77
第1节 二次正交旋转组合设计.....	77
第2节 二次正交旋转组合设计试验结果分析.....	80
第3节 二次通用旋转组合设计.....	85
第4节 二次通用旋转组合设计试验结果分析.....	87
第6章 品种区域试验的统计分析.....	91
第1节 一年多点区域试验的统计分析.....	91
第2节 多年多点品种区域试验的统计分析.....	95
第3节 品种区域试验 AMMI 模型分析.....	98
第7章 群落参数统计分析.....	108
第1节 二项分布系数计算及排序.....	108
第2节 距离系数计算及排序.....	111
第3节 种群丰富度计算.....	115
第4节 对数序列参数估计.....	116
第5节 对数正态分布模型参数估计.....	118
第6节 群落多样性指数计算.....	119
第7节 生态位宽度指数计算.....	121
第8节 生态位重叠指数计算.....	123
第8章 抽样设计.....	125
第1节 分层随机抽样.....	125
第2节 序贯抽样.....	127
第3节 标记—回捕 Petersen 估计.....	131
第4节 标记—回捕 Schnabel 估计.....	131
第5节 标记—回捕 Jolly-Seber 估计.....	132
第9章 生物测定.....	135
第1节 机率分析.....	135
第2节 时间—剂量—死亡率模型分析.....	139
第三篇 多元统计分析.....	145
第10章 回归分析.....	146
第1节 线性回归和逐步回归分析.....	146
第2节 双重筛选逐步回归.....	160
第3节 二次多项式回归分析.....	164
第4节 岭回归分析.....	167
第5节 趋势面分析.....	171
第11章 聚类分析.....	174
第1节 系统聚类分析.....	174
第2节 动态聚类分析.....	185



第3节 有序样本的分类.....	189
第4节 非线性映射分析.....	191
第12章 判别分析.....	195
第1节 两组判别.....	195
第2节 逐步判别分析.....	198
第13章 多因素分析.....	206
第1节 主成分分析.....	206
第2节 因子分析.....	211
第3节 对应分析.....	224
第4节 典型相关分析.....	231
第14章 概率统计模型.....	240
第1节 连续数据序列分级.....	240
第2节 马尔可夫链.....	241
第3节 多元时空序列马尔可夫链分析.....	244
第4节 加权列联表分析.....	248
第5节 多因子综合相关分析.....	250
第四篇 数学模型模拟分析.....	253
第15章 模型参数估计.....	254
第1节 单因变量非线性模型的参数估计.....	254
第2节 二项分布模型的参数估计.....	272
第3节 多因变量联立方程的参数估计.....	279
第16章 数学模型模拟与优化.....	283
第1节 模型模拟分析.....	283
第2节 模型参数灵敏度分析.....	288
第3节 模型优化.....	290
第4节 线性规划.....	292
第五篇 常用数值分析.....	297
第17章 方程求解.....	298
第1节 矩阵求逆.....	298
第2节 求解线性方程组.....	299
第3节 非线性方程组求解.....	300
第4节 实系数多项式求根.....	302
第18章 特征值和特征向量.....	305
第1节 雅可比法求实对称矩阵特征值和特.....	305
第2节 QR法求实矩阵特征值和特征向量.....	307
第19章 微积分数值计算.....	311
第1节 定积分.....	311

第2节 多重积分.....	312
第3节 数值微分.....	314
第4节 微分方程(组)初值求解.....	315
<b>第六篇 其它统计方法.....</b>	<b>319</b>
<b>第20章 模糊数学方法.....</b>	<b>320</b>
第1节 模糊聚类分析.....	320
第2节 模糊模式识别.....	327
第3节 模糊相似优先比方法.....	330
第4节 模糊综合评判.....	333
第5节 模糊关系方程求解.....	334
第6节 综合评判的逆问题.....	336
<b>第21章 时间序列分析.....</b>	<b>339</b>
第1节 时间序列数据基本统计特性分析.....	340
第2节 季节性水平模型.....	346
第3节 季节性交乘趋势模型.....	350
第4节 季节性叠加趋势模型.....	353
第5节 时间序列方差分析外推法.....	357
第6节 差分白回归移动平均模型.....	360
第7节 多变量时间序列 CAR 模型.....	375
<b>第22章 灰色系统分析.....</b>	<b>381</b>
第1节 关联度分析.....	381
第2节 灰色动态(GM)建模基本原理.....	388
第3节 灰色数列 GM (1, 1)模型.....	391
第4节 灰色数列 GM (2, 1)模型.....	395
第5节 灰色数列 GM (1, N)模型.....	397
参考文献.....	401
读者联系卡.....	408

# 第一篇 DPS 数据处理系统

## ( Data Processing System )

DPS 数据处理系统, 英文名称为 Data Processing System, 取首字母缩写为 DPS。该系统采用全屏幕直接写屏方式设计编制, 配有多级下拉式菜单, 用户使用时整个屏幕犹如一张工作平台, 随意调整, 操作自如, 故形象地称其为 DPS 数据处理工作平台, 简称 DPS 平台。

DPS 平台是我们设计研制的通用多功能数理统计和数学模型处理软件系统。它将文字处理和数值计算、统计分析、模型模拟以及画线制表等功能融为一体。因此, DPS 系统主要是作为数据处理和分析工具而面向广大用户, 同时它又保留了 WPS 文字处理器的一般功能。作为数据处理和分析工具, DPS 系统兼有如 Lotus 123、Excel 等流行电子表格软件系统和若干专业统计分析软件系统的功能。与流行的电子表格系统比较, DPS 平台显得更为灵活, 行和列由系统辨认, 但用户面对的屏幕则是无形的, 操作时不必拘泥于一般电子表格系统的行列规定, 可自由发挥, 任意移动。与国外同类专业统计分析软件系统 (如 SAS、STAT、STATISTICA 等) 相比, DPS 系统是独特的, 在统计分析和模型模拟方面功能齐全, 易于掌握, 尤其是对广大中国用户, 其工作界面友好, 只需熟悉它的一般操作规则就可灵活应用。

DPS 平台既可在 IBM PC/XT AT 微机及其兼容机、直接写屏的软汉字操作系统下运行, 亦可在当前流行的中文 Windows 95 环境下运行。系统集数据全屏幕编辑制表、试验设计及统计分析、多元分析、数值计算以及建立各种数学模型等多项功能为一体, 可广泛适用于教学、科研和生产各个领域。不管是青年学生、还是高级科研人员, 不管是计算机应用的初学者, 还是经验丰富的计算机应用专家, 用户都可以在本系统中找到自己感兴趣或有用的部分。

本篇介绍 DPS 平台的主要性能和技术特点。

# 第 1 章 DPS 系统概述

## 第 1 节 系统功能简介

### 1. 独特优良的人机用户界面

一旦开机进入 DPS 平台系统，用户的整个数据处理过程，无论是计算还是数学公式分析都将处于全屏幕编辑状态。可把整个系统视为一张稿纸，键盘和鼠标作为笔。用户可以任意地在上边写、擦、画表和进行数值计算、统计分析、建立数学模型等，计算结果也都返回于编辑状态下。用户需要做的事情主要是输入编辑数据、定义数据块和在菜单方式下调用统计计算功能。当执行统计分析、建立数学模型等高级功能的调用时，只需要按一下 F10 键进入下拉式菜单，在菜单方式下选择执行相应的功能。当相应功能执行完毕后，计算结果存放在编辑器里(当前的光标下面)，系统又处于编辑状态。因此用户的任何工作都能在编辑方式下完成。

当按 F10 键进入下拉式菜单操作时，菜单上部是主栏条形菜单，它总是看得见的，并显示在屏幕的顶行。通过移动左右光标键选择主菜单的功能。如从主菜单上选择一项后，将有两种情况可能出现：一是直接转去执行某一项操作程序；一是在主菜单上作出一项选择后，一个下拉式子菜单就会立即弹出。在子菜单中，使用上下光标键使某项选择变为高亮度，然后再按回车键就可激活该选择项。下拉式子菜单又可产生次一级的下拉式子菜单。当选定相应功能后，按下回车键即可执行。

性能良好的人机用户接口是 DPS 系统特有的操作界面，它是目前国内外同类软件所不具备的。

### 2. 功能完备的数据编辑数字计算器

目前日常数据编辑处理大都借助于字处理器或电子表格、数据库管理系统，它们有的有方便的全屏幕处理性能但无计算功能、有的有计算功能但缺少方便的全屏幕交互处理界面。本编辑器是针对数据处理的特点而开发的数据处理集成软件，它将字处理、数据加工和画线制表融为一体。它既是一个文本编辑器，但通过按 Alt+S 切换，立即又是一个全屏幕交互处理的数字计算器。这时可以在编辑器的任何地方随时写入算式或查询有关函数，进行计算或查询函数值。如键入 “ $4.5^3$ ” 后再键入 “=” 即可得到得数 “91.125”；如键入 “f<sub>test</sub>(5, 20, 0.05)” 后再键入 “=” 即可得到方差分析中大均方自由度为 5、小均方自由度为 20、显著水平为 0.05 时 F 检验值临界值 2.7109。

用本系统编辑的数据无任何隐含字符，故可用于源程序的编辑。其它数据处理系统输出的 ASCII 码文本文件，在本系统下都能处理。数据编辑器功能齐全，简单易学。任何人都能够在很短的时间内掌握本系统的操作方法。

### 3. 接近自然语言的数学表达方式

DPS 系统在统计分析、数学模型模拟以及数学计算等方面，其数学表达式非常接近自然形式(甚至公式中的乘号都可省略)。例如我们输入编辑了两列数，将这两列数据定义成数据矩阵后，系统即默认数据矩阵块中第 1 列变量为  $x_1$ ，第 2 列为  $x_2$ ，然后对这两列变量可进行加、减、乘、除等各种运算。若第 1 列变量取以 10 为底的对数后加上第 2 列变量只要在计算状态下输入“ $\lg(x_1)+x_2=$ ”即可。若第 1 列为因变量，第 2 列为自变量，要建立两变量之间的 Logistic 方程，那么有三个参数需求出。在本系统中，待求参数用  $c_1, c_2, \dots, c_m$  (若有  $m$  个参数的话)表示，这时只要在编辑状态下定义 Logistic 方程的公式“ $x_1=c_1/(1+\exp(c_2+c_3x_2))$ ”，再选择建立模型的菜单功能项即可得到参数精估计值、参数标准误差、 $t$  测验值和置信度等统计参数。

### 4. 完整实用的多功能统计分析功能

DPS 系统的统计分析功能齐全，主要包括方差(协方差)分析、品种区试分析、正交回归实验设计及结果分析、一元及多元回归分析及逐步回归、二次多项式(逐步)回归。典型相关分析、聚类分析、非线性映射、有序样本最优分割。两类判别和多类(逐步)判别分析、时间序列分析、主成分分析、因子分析、对应分析、线性规划、马尔可夫链分析及空间分布型和多样性指数计算等 100 多种方法。

### 5. 用途广泛的数学模型建模工具

DPS 平台除具有功能齐全的统计分析功能外，还为用户提供了通用的试验数据分析和建立数学模型的工具。用户可根据自己的需要，建立符合自己要求和目标而定义的模型。拟建模型可以是单因变量的线性回归模型或非线性经验模型，也可是由多个因变量建立的联立线性回归模型或非线性模型方程组。这些线性、非线性模型又可以是差分模型，或是对各个变量进行加权处理后的模型。

此外，DPS 系统还可对模型进行模拟分析。例如，数学模型中各个参数的变化速率分析(即农业技术经济学中的边际分析)、求解模型的最大值、最小值(即建立模型后对模型模拟寻优)以及线性规划等，DPS 系统均有相应的功能模块代用户之劳。

用 DPS 系统建立和分析数学模型，将使复杂工作变得十分容易，用户只需要按系统规则定义模型就行了。DPS 系统在数学模型处理方面实现了“所想即所见，所见即所得”。

### 6. 简便的数值计算

方程求解包括求解线性规划或非线性方程组(如无约束非线性规划)。

DPS 系统对矩阵的处理采用交互方式，用函数形式表达矩阵计算功能，包括两个矩阵相加、相减、相乘(左乘或右乘)、矩阵求逆、矩阵转置、求矩阵特征值及特征向量等繁琐的计算工作。

其它数学分析和计算功能包括实系数多项式求根、微分方程求导数值、定积分、多重积分数值的计算以及微分方程的初值求解。

数学计算功能，特别是矩阵交互式处理功能使得在当前微机普及情况下，农林高等院校的数理统计分析教学工作中遇到的繁冗无味的矩阵计算，若用本系统作为辅助教学工具，这些计算可以精确而快速地进行，可在很大程度上提高教学质量。

### 7. 丰富多样的数学函数

DPS 系统定义的数学函数除包括一般常用数学函数,如 sqrt(求平方根)、exp(以 e 为底的幂)、lg(以 10 为底的常用对数)、ln(以 e 为底的自然对数)外,还提供数理统计分析中常用的统计函数。主要包括以下类群。

数理统计检验临界值,包括 Fisher 检验临界值、Student's *t* 检验临界值、卡方 ( $\chi^2$ ) 检验临界值、相关系数检验临界值,等等。

数理统计置信度,如 Fisher 检验置信度、Student's *t* 检验置信度、卡方检验置信度以及相关系数检验置信度等;及概率分布函数,如 *F* 分布的分布函数、卡方分布的分布函数、*t* 分布的分布函数、二项分布的分布函数及泊松分布的分布函数,等等。

变量处理函数,包括插值处理函数(spln)、变量自动分级函数(scale)以及计算基本统计量函数。如求某变量平均值(mean)、标准差(std)、方差(var)、平方和(ssqr)、两变量的乘积和(smul)、偏度系数(skew)和峰态系数(kurt)等。

此外, DPS 系统还提供了计算生物群落多样性指数计算等的多种参数计算功能。

### 8. 直观捷达的数据作图辅助分析系统

在数据资料比较及动态分析中,图形以其直观的表达深受人们欢迎,但一般图形处理速度慢,费时费力。DPS 系统的另一突出特点是将数据分析和数据作图有机地结合起来,有些分析结果系统除以文字形式输出外,还自动以图形方式给出(如聚类谱系图)。当然亦可对分析数据作图,如拟合曲线图、*x-y* 折线图等等。

### 9. 灵活方便的表格处理技术

制表是数据编辑输出过程中费时费力的工作,本系统为方便用户,提供了 3 种数据制表方法,包括:① 数据套表对定义的数据矩阵块的文字、数据自动制表连线、形成表格;② 自动制表在当前行下可按用户的要求生成一个具任意多层表头的空表,供填写数据用;③ 交互制表在交互制表状态下,系统可进入图形状态,并将屏幕划分为若干行和列,在屏幕上移动光标键进行制表。

### 10. 单键触发、内容详尽的联机求助

DPS 系统提供了单键触发、内容详尽的联机求助功能。在本系统中,用户可以在任何时刻(除正在执行统计分析功能时外)按 F1 键就能得到相应的解释、说明。在编辑状态下,求助系统详述了 DPS 系统的功能设置、编辑操作以及进行统计分析的步骤等内容。在菜单状态下,求助系统将动态地详述光标所处的当前菜单项目的功能、数据准备方式、统计分析步骤和将得到的分析结果(有的还附有实例),并指明注意事项。

## 第 2 节 DPS 系统的安装与使用

DPS 平台系统可在 DOS 3.0 以上的操作系统、直接写屏汉字环境,或当前流行的中文 Windows 95 环境下运行。系统运行时需要 400kb 以上的可用内存,还需 4 Mb 以上硬盘空间,高分辨率彩显(EGA, VGA)。

### 1. 系统安装

DPS 平台系统全部程序以压缩形式存放在 1 张 1.44 Mb 软盘上,必须安装到硬盘后方可使用。安装过程如下:

(1) 准备。启动时按下 F5 键，以确保内存中无与安装程序冲突的其它常驻内存程序。

(2) 安装。进入 DOS 环境后，将系统盘插入软盘驱动器中，键入驱动器名称 A:或 B:，按回车进入所在驱动器，再键入 INSTALL C:\DPS 后回车，或直接键入 X:\INSTALL C:\DPS 按回车。此处 X 指任意一个驱动器。

## 2. 系统的使用和退出

(1) 进入 C 盘的 DPS 子目录，即在 DOS 提示符下键入 CD C:\DPS，并回车。

(2) 进入系统，有两种方式：1) 键入 DPS <文件名>，再回车，系统将直接进入编辑状态；2) 键入 DPS，不带文件名就直接回车，这时就会进入数据文件名输入、选择用户界面，界面左上方是数据文件名输入区，下部是文件目录浏览显示区，右上部为帮助、提示信息，右下部则显示当前文件部分内容。

在左上方的数据文件名输入区内输入一个合法的文件名，或按 Tab 键将光标转换到文件目录浏览区，通过光标键的移动，选择数据文件。这时，屏幕的右下角将动态地显示文件目录浏览区内当前光标所在文件的部分数据内容。在文件目录区，可通过移动光标来选择已经存在的数据文件，按回车键后就进入编辑状态。

按 Tab 键可使光标在数据文件名输入区和文件名目录浏览区两者之间进行切换。

在数据文件名输入区中，可在输入的文件名中含有通配符“\*”，回车后在数据文件目录显示区显示所定义的一类数据文件名，再在数据文件目录显示浏览区中，用移动光标的方法来选择数据文件名。

进入编辑状态后，系统会出现如图 1-1 的用户界面。

用户界面是一个全屏幕数据编辑处理平台，其上部第 1 行为状态行，在这一行显示当前正在处理的数据文件名，当前光标所处的位置，系统当前所处的状态（编辑或计算，插入或改写等）。最后显示提示按 F10 进入菜单操作。在数据编辑器中，你可完成各种数据

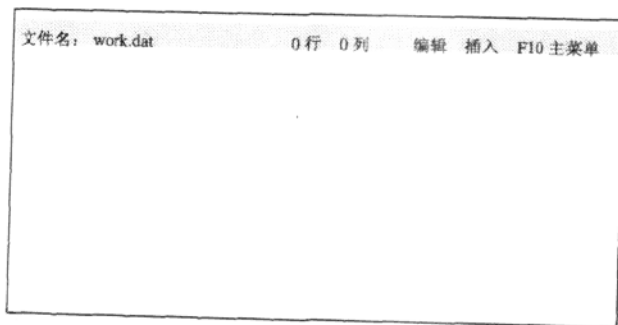


图 1-1 数据编辑处理器用户界面

编辑和计算操作。

欲退出 DPS 系统，按 Alt+X，或在菜单方式下退出。退出时，系统将询问用户正在编辑的文件是否存盘，这时键入 Y 则保存当前正在编辑的文件，再退出。如果键入 N 则不保存当前正在编辑的文件，并退出。如果键入 ESC 键，系统将返回原来状态。

(3) 欲了解如何使用菜单，请继续阅读下面的介绍。

### 3. 菜单操作方法

在编辑状态下,若用户需要调用菜单中功能,如已经将待分析的数据编辑好并定义成数据矩阵块,只需按 F10 键,系统即可弹出下拉式菜单(如图 1-2)操作。

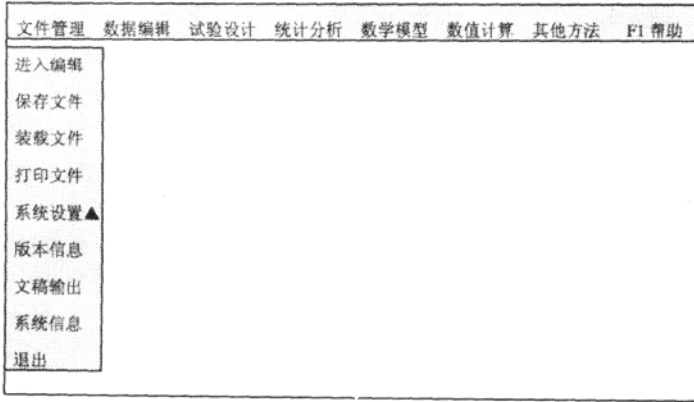


图 1-2 DPS 平台下拉式菜单的主菜单

图 1-2 下拉式菜单的顶部是主菜单。下拉式菜单的优点是使用户在系统中不迷失方向。每个新的菜单层都在上一个选择的下面显示,因此总是知道当前所处的位置和怎样选择下一级功能。下拉式菜单的另一好处是菜单并不完全掩盖屏幕。当进行选择时,新的菜单不断下拉出来后又消失,所以用户总可以看到在菜单下面的原有部分屏幕。

菜单操作时,用户只需按光标键,左右、上下移动就可选择到所需要的功能。

在下拉式菜单中,二级菜单中标有▲的项目表示其下面还有一级菜单,如图 1-2 中的【系统设置▲】就是个二级菜单,它下面又包含 7 个功能项(见表 1-1)。当光标移动到该菜单项并按回车键后,就可弹出它下一级菜单。整个系统的菜单功能见表 1-1。

### 4. 数据编辑格式及数据块的定义

数据编辑器编辑的数据文件为文本文件,一般文本可自由编辑,但需要进行统计分析的数据,各个数据之间至少要用一个空格或逗号隔开。

对于待统计分析的数据,必须先按^KB(即按下 Ctrl 键不松开按 K,松开 Ctrl 键后再按 B,简记为^KB),定义数据块首;按^KK(即按下 Ctrl 键不松开按 K,松开 Ctrl 键后又按 K)定义数据块块尾。由块首和块尾所标记的数据在屏幕上呈红底白字显示(在本书中,标记数据块的背景用阴影显示),同时只有标记后的数据才能进行统计分析。

如果使用鼠标,数据块的定义方法是将鼠标位置移到数据块的块首,按下左按钮并拖动鼠标到数据块的块尾位置,然后松开按钮,即可完成数据块的定义。

在标记的数据块中,数据格式是一行为一个记录(样本),每个记录有多个变量(即当定义数据块之后),各个变量因子可用  $x_1, x_2, \dots, x_m$  表示,参与运算。图 1-3 即为一标记的数据块,它有 7 个变量,8 个样本。



表 1-1 DPS 系统下拉式菜单项目一览表(有▲为二级子菜单)

文件管理	数据编辑	试验设计	统计分析	数学模型	数学计算	其它方法
进入编辑	数据块操作▲	数据分布▲	回归分析▲	管内变温模型参数估计▲	方程求解▲	模糊数学▲
保存文件	定义块首	次数分布	单因变量逐步回归	麦夸特法	线性方程	模糊聚类分析
装载文件	定义块尾	空间分布聚类判别	多重筛选逐步回归	加速单纯形法	非线性方程	模糊识别
打印文件	拷贝数据块	空间分布频次判定	二次多项式回归	二项分布拟合参数估计	多式求根	判别[相似优先
系统设置▲	打印数据块	负二项分布公共 K 值	岭回归分析	联立方程参数估计▲	线性块级▲	模糊综合评价
是否划线	读文本文件	方差分析▲	趋势值分析	麦夸特法	矩阵计算▲	模糊方程求解
数据文件目录	块中数据存盘	单向分组完全随机设计	病态数据回归分析	加速单纯形法	列计算	综合评价逆问题
版面宽度	取消数据块	单向分组随机区组设计	聚类分析▲	模糊识别	行计算	时间序列▲
保留小数位数	删除数据块	单向分组组内分亚组	系统聚类	参数交叉速率分析	相关系数普查	数据稳定性分析
计算状态开关	公式块操作▲	双向分组完全随机设计	动态聚类	求模型极大值	实对称矩阵特征	季节水平模型
数据项宽度	定义块首	双向分组有重复完全随机	有样本本最优判别	求模型最小值	一般实矩阵特征	交联趋势模型
是否缩进	定义块尾	双向分组有重复随机区组	非线性映射	线性规划	矩阵求逆	迭代的模型
版本信息	计算值及函数	双因素区组设计分析	判别分析▲		微分▲	方差分析外推
文档输出	排版	协方差分析	二组判别		定积分	ARMA 模型
系统有关信息	矩阵转置	正交回归试验▲	多类逐步判别		多重积分	多变量自回归模型
退出	查找片替换	正交旋转组合设计方案	多因素分析▲		导数(偏导数值	
	表格处理▲	处理正交旋转试验结果	对应分析		常微分方程组	
	自动缩放	二次通用组合设计方案	主成分分析			
	数据套表	处理二次通用试验结果	因子分析			
	交互列表	群集统计分析▲	典型相关			
	数据列表	二项分布相似系数排序	模糊统计模型▲			
	压缩表宽度	距离系数及其排序	马尔可夫链			
	扩展表宽度	种群丰富度	加权关联表			
	数据排序	多样性指数	多因子综合相关			
	图形编辑	生态位宽度	多因子马尔可夫概			
		生态位重叠				
		对数序列参数估计				
		群集对数正态分布				
		品种区域试验▲				
		一年多点				
		多年多点				
		品种试验 AMMI 模型				
		抽样技术▲				
		分层随机抽样				
		序贯抽样				
		标记 PETERSEN 估计				
		标记 SCHNABEL 估计				
		标记 JOLLY 估计				
		生物测定▲				
		机值分析				
		时间序列死亡率模型				