

● 现代会计学系列教材 ● 姜灵敏 等编著

计算机统计



中南工业大学出版社

计 算 机 统 计

姜灵敏 朱顺泉 胡松强 编著

中南工业大学出版社

【湘】新登字 010 号

计 算 机 统 计

姜灵敏 朱顺泉 胡松强 编著

责任编辑：谢贵良

*

中南工业大学出版社出版发行

中南工业大学出版社印刷厂印装

新华书店总店北京发行所经销

*

开本：850×1168 1/32 印张：9.75 字数：250千字

1995年11月第1版 1995年11月第1次印刷

印数：0001—6000

*

ISBN 7-81020-800-4/TP·059

定价：12.00 元

本书如有印装质量问题，请直接与生产厂家联系解决

内 容 提 要

本书介绍了统计数据处理和计算机统计信息系统分析与设计,详细地阐述了统计业务中的各种平均数、指数、回归分析、相关分析、方差分析、统计图等各种实际问题的计算机实现,最后对两种流行的统计软件 SARP 和 SAS 作了简介。

本书内容丰富,实用性强,可作为大中专院校统计、会计、经济管理等专业的教材,也可作为从事统计工作人员的参考书。

前 言

以计算机为龙头的信息技术革命，正在迅速地改变着人们的生活方式和生活习惯。而高速信息公路的逐步开通，集成的信息和通讯系统逐步形成，使得迅速的处理数据、传输数据，缩短信息的流动时间，提高决策的效率和可靠性成为可能和现实。

统计是认识社会的有力武器，统计信息所反映的是国民经济和社会发展的总体情况，是国家制定社会经济发展战略的主要依据，是社会经济信息的主体。一般来说，现阶段我国的社会经济信息主要包括统计信息、会计信息、市场信息、决策信息及计划信息等。而统计信息是主体，因为离开统计信息，就谈不上认识社会经济运行规律，就无法进行正确决策、制定科学的计划。可以说，统计信息是其它社会经济信息的基础，而又综合、概括了其它社会经济信息。它是国家、企业和个人了解整个社会经济运行状况，参与社会经济活动，对复杂的社会经济运行进行全面、系统的定量监测和综合分析，准确描述、评价、预测社会经济发展，在全社会范围内合理开发、配置各种资源，有效调节和控制社会经济运行的根本保证。所以，统计工作现代化，是整个信息工作现代化的重点，在新的技术革命中占有极其重要的地位，实现统计工作现代化对提高决策效率、加速经济改革，对国家进行有效的宏观调控使国民经济健康迅速地发展都有着十分重要的意义。

本书全面介绍了统计数据处理过程，详细地介绍了统计业务中的各种平均数、指数、回归分析、相关分析、方差分析、统计图等各种实际问题的计算机实现方法和技术，最后对国内两种流行的统计软件 SARP 和 SAS 的功能和使用作了介绍。

本书内容丰富，实用性强，可作为大中专院校统计、会计、经济管理等经济类专业的教材，也可作为从事统计工作人员的参考书。

姜灵敏撰写了第1、2、6章，朱顺泉撰写了第3、4、5章，胡松强撰写了第7、8章。本书由姜灵敏任主编，朱顺泉任副主编。鉴于作者学识水平，不妥之处在所难免，敬请读者批评指正。

作者于长沙

1995.7

目 录

第 1 章 绪论	(1)
§ 1.1 计算机统计的意义与作用	(1)
§ 1.2 流行统计软件及统计软件包简介	(7)
第 2 章 计算机统计数据处理过程	(16)
§ 2.1 统计数据采集	(16)
§ 2.2 统计数据编码	(18)
§ 2.3 统计数据处理	(25)
§ 2.4 统计数据处理系统的系统分析	(34)
第 3 章 计算机平均数分析	(39)
§ 3.1 算术平均数	(40)
§ 3.2 调和平均数和几何平均数	(46)
§ 3.3 中位数及其应用	(57)
§ 3.4 众数及其应用	(63)
第 4 章 计算机比率	(68)
§ 4.1 统计比率	(68)
§ 4.2 发展速度和增长速度及其应用	(75)
§ 4.3 物价指数	(81)
§ 4.4 物量指数	(87)
第 5 章 计算机统计预测	(94)
§ 5.1 回归分析、相关分析和方差分析	(94)
§ 5.2 用 C 语言实现的统计预测程序	(104)
§ 5.3 统计分析的 FOXBASE 总控程序	(112)
§ 5.4 统计预测程序在股票分析中的应用	(114)
第 6 章 计算机统计图	(116)

§ 6.1	计算机统计制图概述	(116)
§ 6.2	条形图	(118)
§ 6.3	圆形图	(130)
§ 6.4	曲线图	(137)
§ 6.5	图形打印	(140)
第7章	SARP 统计软件	(146)
§ 7.1	系统功能概述	(146)
§ 7.2	系统文件的组成结构和作用	(147)
§ 7.3	SARP 系统的运行环境及安装过程	(151)
§ 7.4	系统基本约定及用户描述环境	(154)
§ 7.5	SARP 系统加载	(164)
§ 7.6	专业字典、词典	(167)
§ 7.7	数据表描述	(178)
§ 7.8	数据输入	(184)
§ 7.9	数据检索	(191)
§ 7.10	审核和计算	(192)
§ 7.11	汇总	(200)
§ 7.12	制表描述及操作	(204)
§ 7.13	数据的备份	(218)
§ 7.14	批处理作业	(223)
§ 7.15	系统维护	(225)
附录 A	错误返回码	(226)
附录 B	各主要模块功能键一览表	(231)
第8章	SAS 统计软件包	(235)
§ 8.1	SAS 统计软件概述	(235)
§ 8.2	SAS 显示管理系统	(241)
§ 8.3	数据步及其语句	(250)
§ 8.4	过程步语句	(282)

第 1 章 绪 论

§ 1.1 计算机统计的意义与作用

计算机统计就是统计工作计算机化，是统计工作现代化的核心组成部分。发达工业国家早在 50 年代就把计算机用到了统计工作中。1951 年，美国普查局就使用计算机完成了 1950 年人口与居住普查的相当一部分数据处理工作。1960 年美国人口普查的汇总工作，仅利用计算机输入设备这一项，就节省了 500 万美元，并缩短时间 7 个月；1970 年的人口普查在这一项上节省的费用达到了 1000 万美元。计算机早已成为美国各统计机构工作人员的必不可少的工具，除了利用计算机进行数据汇总、整理、制表、打印等简单的数据处理之外，还使用计算机进行统计分析，预测及辅助决策等工作。

80 年代初我国的统计工作现代化就提到了议事日程，1984 年 1 月 6 日，国务院《关于加强统计工作的决定》指出：“统计工作现代化，就是运用先进的统计科学和现代计算技术，来改革和完善我国的统计工作，进一步做到数字准确、资料丰富、信息灵通，以适应社会主义现代化建设的需要。”我国的统计工作“必须逐步实现统计指标体系完整化，统计分类标准化，统计调查工作科学化，统计基础工作规范化，统计计算和数据传输技术现代化，统计服务工作优质化。”并指出：“电子计算机的应用是统计工作的一项重大改革，要充分发挥现有电子计算机和传输设备的作用，争取在几年内县以上各级统计局要配备微处理机，逐步建立健全现

代化的统计体系。”国务院批准的《统计法实施细则》进一步明确规定：“国家有计划地用现代化信息技术装备各级人民政府统计机构，建立健全国家统计信息自动化系统”。根据国务院的决定和规定，国家统计局在1986年制定了“国家统计信息自动化系统总体规划”，在这个规划的指导下，经过十年的努力，统计信息自动化建设已经取得了重大进展，国家统计信息自动化系统已初具规模。

自1986年以来，相继建立了国家统计局计算中心和30个省、自治区、直辖市统计局计算中心。到1989年，全国318个地（市）成立了计算站，绝大多数县统计局建立了计算机室。据统计，目前全国统计系统从事专职计算机技术和管理的人员约4000人，40岁以下的统计专业人员都已初步掌握了计算机的使用技术。全系统配备的微型计算机已逾万台。统计系统的常规报表处理已实现计算机化；已具备承担大型数据处理工程的能力，如1987年的全国投入产出普查、1990年第四次人口普查、1994年第三产业普查等。

以计算机为龙头的信息革命，正在迅速地改变人们的生活方式和生活习惯。而高速信息公路的逐步开通，集成的信息和通讯系统逐步形成，使得迅速的处理数据、传输数据，缩短信息的流动时间，提高决策的效率和可靠性成为可能和现实。

统计是认识社会的有力武器，统计信息所反映的是国民经济和社会发展的总体情况，是国家制定社会经济发展战略的主要依据，是社会经济信息的主体。一般来说，现阶段我国的社会经济信息主要包括统计信息、会计信息、市场信息、决策信息及计划信息等。而统计信息是主体，因为离开统计信息，就谈不上认识社会经济运行规律，就无法进行正确决策、制定科学的计划。可以说，统计信息是其它社会经济信息的基础，而又综合、概括了其它社会经济信息。它是国家、企业和个人了解整个社会经济运行状况，参与社会经济活动，对复杂的社会经济运行进行全面、系统的定量监测和综合分析，准确描述、评价、预测社会经济发展，

在全社会范围内合理开发、配置各种资源，有效调节和控制社会经济运行的根本保证。

所以，统计工作现代化，是整个信息工作现代化的重点，在新的技术革命中占有极其重要的地位。实现统计工作现代化对提高决策效率、加速经济改革，提高宏观调控的有效性、健康迅速地发展国民经济都有着十分重要的意义。

目前，我国由于国民经济指标体系还不够完整，统计标准化工作还没跟上，特别是统计信息的搜集、处理、传输的技术水平还不高，所以统计系统搜集和提供的社会经济信息远远满足不了社会经济发展的需要。我国目前统计信息的处理是“金字塔”式，城乡居民和基层单位提供的巨量信息，构成了“金字塔”庞大的底座，而这些信息经过层层加工处理后，报到国家统计局的信息比原始信息量少了90%以上，构成了“金字塔”尖细的塔顶。这一现象严重地影响着国家对社会经济发展的决策，影响和制约着经济的高速发展。

为了从根本上解决上述现象，必须使统计工作现代化，采用以电子计算机和网络通讯为核心的现代信息技术，对大量的统计信息进行搜集、加工、传输、计算和出版，提高信息处理的全面性和时效性。电子计算机具有运算速度快、存储容量大、逻辑功能强、自动化程度高、保存信息的时间长等特点，而网络通讯技术可使信息高速传递、共享。在统计工作中，只有广泛地运用电子计算机和现代信息传输技术，才能完成各项复杂的信息搜集、储存和计算任务，才能大大缩短信息传递的周期。统计工作中运用计算机的目的，不仅仅是为了简单地替代统计人员的手工劳动，更主要是通过电子计算机系统准确、及时地整理大量的、详细的基础数据，在这一基础上进行各种复杂的分类、分组，并运用数理统计学和其它科学方法，去完成人们手工计算难以做到的大量数据的搜集、整理以及统计分析和统计预测工作。电子计算机和现代传输技术的应用，统计自动化系统的建立，能极大地提高统

计工作的水平和质量，更加有效地发挥统计工作在我国经济建设中的作用。

在统计工作中广泛运用电子计算机和现代信息传输技术，具有如下优越性：

1. 丰富统计内容，扩大信息量

我国统计工作由于信息技术水平还不高，致使基层提供的大量丰富的资料未能得到整理应用，白白地浪费了，形成了信息处理的“金字塔”式。以工业统计为例，用手工计算的方法，一般只能把各个工业企业的成本、利润等经济效益指标加在一起，得出全市（或全行业）的总数上报，上一级得到的只是全市（或全行业）成本利润的数字，而失去了各个企业的经济效益的信息，失去了哪个企业效益好、哪个企业效益差的信息；如果一个省把各市的这些数加在一起上报，那么省里得到的仅是一些总数，而失去了经济效益数值在各市分布情况的信息，这就是落后的手工层层汇总的缺陷。正因为如此，越往上信息越少，信息的损失越严重。那么是否可以将基层的所有数据尽数上报，而不是上报汇总后的数据呢？这样的话，巨量的数据传输、存储、处理用手工方法是无法完成的。

应用电子计算机进行统计数据处理就大不相同了，它可以改变手工方式逐级汇总只得到少量数据的状况，实行“超级汇总”得到大量数据。用计算机处理统计数据，不仅可以扩大统计的规模、增加统计内容、提高汇总速度、扩展统计信息的内容，而且还可以提高统计信息的利用率，采用现代科学方法产生出大量新的信息。利用计算机巨大的存储容量，还可以收集、存储大量的历史信息和国际信息。这样，就可以利用这些信息进行各种统计分析、统计预测、统计决策等方面的分析研究，从而产生出大量的再生信息。这是手工方式不可能做到的。所以，使用电子计算机以后，不仅避免了手工层层汇总造成的信息损失，而且通过对基本信息综合、科学地分析（复杂的交叉分组、数理统计分析、各种数

学模型分析等)产生出比基本统计信息多得多的信息,这样一来,一方面扩大了统计调查信息量,另一方面越向上汇总,处理产生的信息量就越多。完整、详细、全面的统计信息对于国家制定国民经济和社会发展决策将是极为有益的。

2. 提高统计数字的准确性

保证统计数字的准确性,是统计工作的根本要求。但是,采用手工方式逐级汇总,由于综合统计部门的统计人员忙于过录、汇总,没有时间对基层表进行严格的仔细审查,因此,基层表中许多数字质量方面的问题,往往被掩盖起来了,汇总中重复遗漏问题也难予发现。另外,在逐级手工过录、汇总、计算、复写的过程中又难免发生新的错误。所以在手工方式进行统计数据处理时,不能保证统计数字的高质量。而利用电子计算机进行统计数据处理,不仅可以避免手工汇总时转录、摘抄、计算复写工作中的技术性差错,而且还可以根据人们事先规定的逻辑检查规则,对输入的原始数据进行逐行计算检查和逻辑检查,可对表与表之间、表格中项目与项目之间、指标与指标之间的数据进行分析、比较、筛选、检查,判断其是否正确可靠,并把检查出来的错误一一记录下来,也可以打印出来,供统计人员查对改正,从而大大降低调查登记、填报、编码、录入各阶段所引入的差错,保证基础统计资料的准确性。

在统计分析、统计预测时,由于方法复杂、计算量大,手工处理发生错误的可能性是很大的,而利用计算机进行这些工作,只要原始数据正确,编码和录入工作准确无误,处理方法科学,设计的程序合理,就一定能保证计算结果正确。

3. 处理速度快,提高了统计信息的时效性

信息的一个极其重要的特点就是具有时效性。应用电子计算机进行统计信息处理,对于大量信息不仅可以缩短汇总处理和信息传递的时间,而且还可以及时地进行统计分析,为领导机关制定计划、措施和科研人员进行科学研究提供重要依据。如果建成

国家统计信息自动化系统,从信息的搜集、整理、存储、传输、计算到出版都采用现代信息技术,那么就可以大大缩短调查周期、处理加工的周期、信息传输的周期以及信息服务的周期,就能实现信息灵通的目的,更加充分地发挥统计工作的职能作用。

4. 提高工作效率,延伸人的智能

在统计工作中应用了计算机以后,节省了人力,提高了工作效率。用计算机处理和传输统计信息比手工要快几十倍,甚至几百倍。电子计算机能够延伸人们的智能,并且不会象人那样会发生疲劳,受情绪的影响,它可以超越人体机能的限制,去完成人们力不能及的工作。从这种意义上说,电子计算机能极大地扩大和加强人们的脑力劳动。

5. 提高了统计信息的共享性

随着“高速信息公路”的开通,局域网、广域网和综合数字网的应用,统计信息除了汇集到国家有关部门作为决策的主要依据外,还可以广泛地服务于社会和大众,提高了统计信息的效益和共享性。企业和个人在非常方便地使用统计信息的情况下,可以及时地对经济形势和社会发展情况进行分析、作出判断和决策,同时也可通过网络及时地反馈信息,以利国家及时进行宏观调控,使社会经济的发展进入良性循环,从而促进经济高速健康地发展。

6. 提高统计信息的利用率,开拓服务领域

利用电子计算机可以把完整的、动态的、连续的统计资料存入数据库,长期保存,根据国民经济和社会发展的需要,可以随时根据要求重新加工整理,从而大大提高了统计信息的利用率,广泛地开拓服务领域。手工方式管理统计信息,人们查找一个数据,或计算一个数据,都是相当慢的,再加上数据关系的复杂性,难以把大量统计数据构成一个整体,进行综合利用。利用电子计算机管理统计信息,完全可以把大量的统计信息有机地构成一个整体。这样,就可以利用现代数学方法完善统计分析和预测,增加分析项目,扩大统计内容,从而使那些在手工方式操作的情况下,

可望而不可及的大量统计汇总、统计分析工作变成现实。

在统计工作中运用计算机，不仅是统计计算技术方面的一场革命，而且对原来建立在手工计算、逐级汇总基础上的调查统计报表，在调查内容、工作组织和统计基础工作等方面，也会产生一系列的深刻变革；同时为建立基层企业一套表，适应各级各部门的需要，提供了可能性。电子计算机的应用，为统计资料的超级汇总提供了技术基础，而统计资料的超级汇总，不仅可以加工整理丰富的统计资料，为实现统计服务的优质化创造了条件，而且对统计指标体系的完整化、统计分类的标准化、统计调查工作的科学化以及统计基础工作的规范化，都将起很大的推动作用。

§ 1.2 流行统计软件及统计软件包简介

目前我国除国家统计局和省一级统计局的信息中心配备了大型或中型计算机以外，大多数统计部门使用的是微机，总数有近百万台。在微机上使用的统计软件有很多，在实际工作中应用得比较普遍的主要有 Lotus 1-2-3、TPL、CENTS、DET、SAS、SPSS、TSP、SP、SARP 等若干种，为帮助读者了解选择和使用这些软件，我们在此作一简单介绍，具体的应用技术和操作方法请参阅相应的软件说明书。

1.2.1 LOTUS 1-2-3

LOTUS 1-2-3 是美国麻省 LOTUS 公司 1983 年 1 月推出的集成软件，它把电子数据表处理、数据库管理系统和图形处理程序三者有机地结合在一起。1-2-3 就是依次表示该软件的三种功能。该软件表格处理能力很强，能在屏幕上直观地显示数据表格（通过窗口移动可显示大到 8192×256 项的表格），操作者可以在屏幕上随意进行数据的录入、删改、移动、复制、保护等，并用其很强的计算能力进行数学和统计运算。该软件的数据库管理功能主要用于实现 LOTUS 1-2-3 数据与其它软件如 TSP、

FOXBASE、SC3 等的数据的转换，并完成文件的复制、删除、更名、备份、目录排序。统计图绘制功能使用户可在屏幕上直接绘制多种多样的统计图形（如直线图、圆饼图、折线图），并可实现统计图形的打印输出。

LOTUS 1-2-3 有单命令和批处理（宏）两种方式，采用菜单操作，简单易学。该软件在统计汇总、分析和预测方面十分有用，目前已有许多统计专业人员熟练地掌握了它的使用方法，广泛地用于简单报表的汇总和统计分析。

1.2.2 TPL 制表软件包

TPL (Table Producing Language) 制表软件包是美国劳工部统计局研制的一个通用报表生成系统。TPL 是一个用于统计数据交叉制表的比较灵活、而且功能比较齐全的制表软件。TPL 能适应各种统计报表的要求，比较方便地处理大容量的统计数据，生成用户所需要的报表。

TPL 主要是面向非计算机专业用户的。因此，不需要用户了解太多的有关计算机方面的知识，不需要专门为制表研制程序。TPL 本身有一套简捷方便的语言，通过使用 TPL 语言，就可以调用 TPL 的各种功能，达到用户的要求。熟悉统计制表的用户大概都了解制表流程，首先，对数据文件进行详细的了解和描述；其后，根据汇总表的要求对数据进行加工；最后，根据用户所需要的表的格式和内容来选择参加制表的数据项和输出格式。针对这个流程，TPL 软件包相应地分成三部份，即编码库，制表需求，制表格式。在此强调一点，即 TPL 只是一个制表软件，没有任何数据录入/编辑功能，它对用户的要求是必须拿到净化的数据文件（检查/编辑好的数据文件）。

TPL 软件对数据的要求不是很严格的，数据记录格式可以是等长记录，也可以是不等长记录，但必须保证数据文件是文本文件，数据可以从磁带、软盘、硬盘等介质中读取出来，数据文件的结构可以是分层文件，也可以是非分层文件。所谓分层文件，就

是指多记录的标识项。分层文件最典型的例子便是人口普查中的数据文件，它分为户记录和人记录，当我们用到该数据文件时，必须明确分开户记录与人记录，因为相同位置的数据项所代表的内容很可能不一样，否则，结果会发生混乱。TPL 每次只能处理一个数据文件，这一点一定要引起注意。

编码库用来对数据文件进行详细描述，用户在编码库中可以定义数据项的名称类型，长度/显示的方式及最后出表所需要的打印标志。在编写编码库之前，一定要先熟悉数据文件的物理特性及内容，只有这样，才能使所编写的编码库更适合你出表的要求，编写编码库是一项繁重的工作，但命令比较简单，而且编码库一编写好，用户就可以多次调用它，而无须再一次编写。

编码库定义好之后，必须经过一个专门的程序检查语法，并且把它转换成下一项工作专用的内部格式。

制表需求就是表生成这一部分，在制表要求中，用户可以对数据进行选择，重新分组，计算生成一个新变量等，并根据结果表的要求，选择参加制表的数据项。制表要求的运行必须依赖于编码库的存在并通过语法检查。

TPL 软件有了上面所说的编码库和制表需求这两部分，就可以运行并生成报表，但这个表的格式都是机器内部定义的，如果用户想改变表的输出格式，就需要专门来编写制表格式这一部分程序。制表格式就是根据用户的需要，改变表的输出格式，如表的标题、脚注、列注以及表中列的宽度等等。

如果编写了编码库，根据制表需求运行 TPL，就可以生成一套使你满意的报表。

1.2.3 CENTS 制表软件包

1970 年世界人口普查期间，为了改善数据处理能力，美国普查局国际统计项目中心 ISPC 用 IBM 汇编语言开发了一个通用的制表软件 CENTS。CENTS 是英文 CENsus Tabulation System 的缩写，意为“普查制表系统”。以后，又用 COBOL 语言改写为 CO-