

计算机在经济统计 和文字档案中的应用

曹文君 编著



复旦大学出版社

计算机在经济统计和 文字档案中的应用

曹文君 编著 李廷燕 (审)

复旦大学出版社

内 容 提 要

本书由四部分组成：第一部分为计算机系统的基本知识，第二部分为 BASIC 语言，第三部分介绍常用经济统计的应用软件，第四部分为计算机文字处理知识，包括汉字信息处理、办公自动化、文献检索、机器翻译和辅助教育。通过对计算机知识的了解，介绍计算机在统计和文化领域中的应用。

本书是高等院校文科类专业、中等学校财经类专业的教材。但同样可作为具有中等文化以上的职工、管理人员、干部自修计算机知识的读本。

(沪)新登字202号 曹文君

计算机在经济统计和文字档案中的应用

曹文君 编著

复旦大学出版社出版

(上海国权路379号)

新华书店上海发行所发行 江苏大丰印刷二厂印刷

开本850×1168 1/32 印张9.875 字数280,000

1989年5月第1版

1992年11月第2次印刷

印数4,000—7,000

ISBN7—309—00953—3/T·65 定价：4.30元

序 言

在人类文化发展史上有过“石器文化”、“青铜器文化”和“铁器文化”几个阶段。纸和印刷术的发明,打字机、排铸机等文化工具的出现,都使人类社会进入到更文明的阶段。

本世纪中叶诞生的电子计算机,已经悄悄地进入了文化领域,成为建设现代文明必不可少的工具。请看:计算机编辑照排系统,计算机自动翻译系统,计算机情报检索系统,计算机辅助教学系统,计算机辅助作曲系统,计算机图书馆管理系统,计算机办公室自动化系统,电子词典,电子书籍等等枚不胜数。这就是有人称之为“计算机文化”的内容。

现代社会是一个信息化的社会,“知识爆炸”使人们无法再利用旧的处理工具和手段来对付日益增长的信息量。人们不得不借助计算机这个存贮量大、处理速度快的现代化工具来收集、处理、存贮、传递信息,从而使计算机成为现代社会的信息中心。作为文化工作的人们——作家、编辑、新闻记者、翻译家、教师等一切具有文化的人,终将不得不和计算机打交道,通过计算机网络上的终端向计算机发送信息,并从那儿得到信息和知识。因而,每个人都要学会使用计算机,不管你是科学家,工程师,还是文化工作者,这是现代社会发展的必然要求。

为了使从事文化工作的人们和在校(院)文科类学生能对计算机有一个正确的初步的了解,特编辑这本内容涉及到计算机基本知识、计算机社会应用知识和计算机文化三个方面的教材,希望它能对广大读者有较大的帮助。

由于本人才学浅薄,难免挂一漏万,或者存在错误的地方,敬请读者和这方面的行家批评指正。

何永保教授仔细地审阅了全稿,提出了不少宝贵意见,在此深表谢意。

编 者

于复旦大学

1987.3.

目 录

序言

第一篇 计算机系统

第一章 概论	1
§1 人脑与电脑	1
§2 计算机与人类社会	2
§3 计算机发展简史	4
§4 计算机与二进制	6
§5 计算机系统的构成	8
本章参考书习题	10
第二章 硬件	11
§1 冯·诺依曼型计算机的基本组成	11
§2 中央处理机	13
§3 存贮器	14
§4 总线与接口	16
§5 计算机局部网络	18
本章参考书习题	20
第三章 外部设备	22
§1 输入设备	22
§2 输出设备	24
§3 外存贮器	29
本章参考书习题	32
第四章 软件	33
§1 计算机软件的基本组成	33
§2 计算机语言与语言处理程序	35
§3 操作系统	39
§4 数据库与数据库管理系统	45
§5 通讯软件与网络软件	49

本章参考书习题	51
---------------	----

第二篇 BASIC 语言

第五章 基本概念	53
§1 BASIC 语言	53
§2 BASIC 程序结构	56
§3 BASIC 语言的基本符号	59
§4 常数	60
§5 变量与数组	61
§6 标准函数	63
§7 表达式	64
本章参考书习题	70
第六章 基本语句	72
§1 注释语句和终止语句	72
§2 赋值语句	73
§3 键盘输入语句	74
§4 读数语句和恢复语句	75
§5 输出语句和制表格式语句	77
习题	85
第七章 控制语句	88
§1 无条件转移语句	88
§2 暂停语句	88
§3 条件转移语句	89
§4 散转语句	90
§5 数组说明语句	91
§6 循环语句	92
§7 子程序	100
§8 通用程序的编写	102
习题	115
第八章 字符串与文件	119
§1 字符串变量与字符串数组	119
§2 字符串变量的赋值	120
§3 字符串操作	121

§4 字符串函数	122
§5 文件概述	126
§6 文件标识和文件分类	127
§7 顺序文件的存取	128
§8 随机文件的存取	133
习题	148
第九章 上机知识	150
§1 键盘命令	150
§2 BASIC 源程序的输入和修改	151
§3 汉字输入	152
§4 BASIC 语言上机步骤	154
本章参考书	155
附录一 ASCII 字符代码集	155
附录二 BASIC 常见错误信息表	156

第三篇 常用社会经济统计与程序

第十章 数字特征	159
§1 次数分布	159
§2 集中量	164
§3 差异量	171
本章参考书习题	179
第十一章 社会与经济统计	182
§1 回归分析	182
§2 相关分析	185
§3 计算回归和相关的 BASIC 程序	186
§4 实例	187
§5 速度	190
§6 指数	197
本章参考书习题	206
第十二章 应用统计	209
§1 票数统计	209
§2 λ 值和 GAMMA 值的计算	219
习题	225

第四篇 计算机与文字处理

第十三章 汉字信息处理	227
§1 概述	227
§2 汉字库与汉卡	228
§3 汉字输入	229
§4 汉字输出	237
§5 汉字信息处理系统的组成	239
§6 汉字字处理软件	239
本章参考书	247
第十四章 办公自动化	248
§1 概述	248
§2 办公自动化系统的功能	251
§3 办公自动化系统的模型	254
§4 办公自动化系统中值得注意的几个问题	257
§5 办公自动化软件介绍	258
本章参考书	260
第十五章 文献检索	261
§1 概述	261
§2 词表和文献标引	265
§3 文献的存贮	268
§4 文献的检索	274
本章参考书	279
第十六章 机器翻译	280
§1 概述	280
§2 机器翻译系统的组成	284
§3 机器翻译的过程	286
§4 机助翻译	288
§5 我国的机器翻译	290
本章参考书	291
第十七章 计算机辅助教育	292
§1 概述	292
§2 计算机辅助教学	295

§ 3 计算机管理教学.....	297
§ 4 智能型计算机辅助教学.....	298
§ 5 PLATO 计算机辅助教学系统.....	299
本章参考书.....	300
全书内容索引	301

第一篇 计算机系统

第一章 概 论

§1 人脑与电脑

电子计算机也称为“电脑”。这是因为计算机具有人脑的某些功能,它能部分地替代人脑的工作。

为什么电脑能部分地替代人脑呢?因为它具备以下四项基本功能:

1) 控制功能; 2) 记忆功能; 3) 数值计算和逻辑判断功能; 4) 与外界进行信息交换的功能。这些功能使电脑具有像人脑一样思考问题的能力,从而能进行联想和学习,作普通的脑力劳动,甚至作富有创造性的脑力劳动。

电脑与人脑相比较,它在以下几个方面比人脑显得更加优越。

(1) 处理信息速度快:人脑要进行连续工作几年,甚至几十年才能解决的问题,电脑只要几个小时就能完成。电脑比人脑快几十万到几百万倍。

(2) 数值计算正确度高:对同一个数值计算问题,电脑计算的结果正确度能达到与计算精度要求几乎毫无差别的程度,而且算一遍与算一百遍的结果是一致的。人脑是无法达到这一点的。

(3) 记忆力强:用电脑附加的存储器记忆各种信息,可以保存几年至几十年,并且可以随时查看。人脑和电脑相比,记忆力差多了。

(4) 永不疲劳:电脑能长期连续工作,永不疲劳。而任何人,不管他精力多充沛,连续工作几个小时后,大脑总得要休息片刻。人脑和电脑相比,这是望尘莫及的。

当然话还得说回来,电脑也有不如人脑的短处。第一,电脑的记忆容量比人脑小,无论电脑的记忆容量有多么大,它总是有极限的,而人脑的记忆容量可以说是无限的。第二,电脑必须要在人脑的创造性劳动的帮助下才能具有脑力劳动的能力。因此,电脑永远不会代替人脑,机器人代替不了有血有肉的真正的人。

有人担心:电脑广泛使用后,代替了人们正在进行的许多工作,有一部分人会由此而没有工作做。其实这种担心是多余的。过去,没有交通工具时,人要外出,只能走路,人的活动范围受到很大的限制。有了汽车、火车、轮船、飞机后,人的脚并没有因此而退化,相反,人们的足迹将离家更远。电脑广泛使用,代替部分人脑所作的脑力劳动;人的工作并没有减少,人们用更多的时间研究新问题,向科学的深度和广度进军,创造出更新更美的能丰富人们生活的杰作。未来的社会对人脑的要求更高了,人们都要有电脑知识,否则就难以工作和学习。

对于一个刚接触电脑的初学者来说,往往会被五花八门的电脑弄得眼花缭乱,又会对那些高深莫测的电脑技术知识望而生畏,似乎电脑不可捉摸。其实不然,在第一篇的四章内容里,将首先帮助你对于千变万化的电脑理出一个头绪,使初学者对电脑建立一个初步的正确的整体认识。然后在以后的篇章中逐步介绍电脑的技术和应用方面的知识。

§ 2 计算机与人类社会

计算机自从 1946 年降生到社会上来,由于它有神奇的功能,短短四十年,已经从“丑小鸭”成长为八十年代的风云人物。如果说十八世纪开始的第一次产业革命给社会这个躯体添上“肌肉”,那么作为第二次产业革命的产儿,计算机给社会添上“头脑”。人脑在电脑的辅助下,变得更加聪明,更富有创造性,整个社会随之发生了翻天覆地的变化。

计算机能以每秒亿万次的运算速度,协助科学家计算宇宙航行的轨迹,帮助工程师绘制新的蓝图。人类的足迹已经到达月球,但这不是遨游太空的最终目标,以后将会走得更远更远。

带有电脑的机器人已经开始在工厂里代替工人完成一些艰苦重复的体力劳动，全部由计算机控制的无人车间正在逐日增加数目。由电脑设计出来的服装以款式新，穿着适合得体而受到人们的青睐。

计算机正在使办公室工作发生根本的变化，工作人员不再为整理、保存大量的资料而发愁，利用像塑料唱片那样大小的一片片磁盘，将大量的文件、档案、资料、信件等存放起来。等到需要的时候，又只需按一下键钮，所需的资料可以在几秒钟内在屏幕上重新出现，必要时可以由电脑驱动复印机复制复印件。

由计算机制作出来的电子词典体积小，并且可以发音，帮助你校正单词的读音；电子书籍不但可以阅读并且可以听，其体积之小，全部辞海可以放进一个火柴盒子里。另外，计算机可以辨认笔迹，能鉴别文章和图画的真伪。计算机还可以用来辅助创作，由电脑写出来的古诗比人脑写出来的毫不逊色，诗意盎然，用词正确，韵和对仗也无可指责之处。

一种将一国文字译成另一国文字的计算机翻译系统已经陆续登场。可以想象，这样的时期一定能到来：放在口袋里的微电脑翻译机，听到英语后，能立即用中文将意思告诉你。在这样的世界里，人们不必为语言上的隔阂而担心了。国际交往将发生何等巨大变化。

信息化的现代社会每年全世界要出版近十万种杂志，五十万种新书，上万种文摘，通过计算机的帮助，研究人员能够在七秒钟内，从这些浩如烟海的文献中找到所需要的资料。并且当你需要文献的话，可以在三秒钟内将内容拷贝到微缩胶片上。

计算机编辑照排系统的诞生，大大提高了报刊和书籍的出版效率，编印一期几十个版面，发行量为几百万份的报纸的周期仅需短短一个小时，而出版几十万字的一本书亦仅需几十天乃至几天时间。

在学校里，计算机将成为师生们的良师益友。它既能充当学生们的游戏，玩耍的伙伴，又是学生们最具有耐心的老师，帮助学生们学习。它既可以给学生们上课，又能出习题，批改作业和试卷，并且还能登录成绩和作出各种各样的教学统计。

计算机已经悄悄地使千家万户的传统生活发生变化,录音机、洗衣机、电冰箱、电视机放入电脑后,功能大大加强了。“个人电脑”在家庭里不但能够训练、培养儿童的智力,还能够帮助主妇定时做饭,开关各种电器,为花草浇水,它可以当警卫员,看大门、报火警、防窃盗。它还可以帮助主人,安排每天工作的日程。

计算机的用途真是不胜枚举,它不仅使工矿企业实现自动化,而且使办公室实现自动化,现在和将来还要使家庭实现自动化。可以想象要不了多久,计算机会成为人类不可缺少的伴侣,它不仅是干活的帮手,而且又是你的良师益友。到了那时,就像你现在离不开水、电和煤气一样,在生活中将离不开忠实为你服务的计算机。

§ 3 计算机发展简史^[1]

早在公元前 4000~3000 年出现的算盘,可以算得上是现代计算机的祖先,当时算盘是个用小珠和线做成的计算装置,小珠在每根线上的位置决定了数字的值,用小珠在线上的移动来进行加、减、乘、除的算术运算。

公元 1642 年法国科学家帕斯卡研制成功旋转字轮式机械计算器。以后随着时间的推移,从字轮式计算器的基础上,改进成为台式机械计算器。人们使用机械计算器进行数值计算的时期大约持续了三百多年,1944 年美国哈佛大学成功地研制出 mark I 自动顺序控制机械数字计算机,它实行除法需用约 60 秒时间,而现代计算机实行除法少于 1 微秒。

世界上第一台电子计算机于 1946 年由美国宾夕法尼亚大学的 J. PECKERT 和 J. W. MAUCULY 领导的一个工程和数学小组建造成功。它被取名叫 ENIAC。全机用了 18,000 个电子管,1,500 个继电器,由于机器的元器件全部采用电子管,所以称为第一代的电子计算机。该机运算速度为 5,000 次/秒,耗电 150 瓩,占地 1,800 平方米,重 36 吨。它使用于陆军炮兵部队用来计算火炮弹道及弹落点。它作为第一台电

子计算机一直使用到 1958 年。

1958 年第一台全晶体管化的电子计算机 NCR304 交货使用。这是一台商用计算机，用于美国国家现金收入情况统计。该机由于采用以晶体管为元器件，称为第二代的电子计算机。由于晶体管体积小，重量轻，耗电省，寿命长，速度快，故障少，使计算机技术有了重大的突破。与第一代电子计算机相比，体积缩小到几十分之一，运算速度提高了几十倍，达到每秒几十万次。

1965 年以后电子工业的发展出现了先进的集成电路，它是在仅有几平方毫米的单晶片上，集中做成含有几十至几百个晶体管元件的电路，由集成元件组成的电子计算机称为第三代电子计算机。1964 年美国 IBM 公司首先宣布该公司的 360 系列机器采用了混合集成电路，1970 年美国 IBM 公司宣布新的系列机 370 全部采用集成电路。集成化的计算机体积更小，重量更轻，耗电更省，速度更快，达到每秒几百万次，它的价格随之更低廉，便于各行各业推广使用，从此以后进入了电子计算机大发展的新时期。

1970 年以后，集成技术发展很快，在一个芯片上可做出含有几千到几万个电子元件，称之为大规模集成电路 (LSI)，而集成十万个电子元件以上的叫超大规模集成电路 (VLSI)。用 LSI 或者 VLSI 电路组成的计算机称为第四代电子计算机。由于元器件的微型化，出现小型化和微型化的计算机。这些计算机整个电路组成只有几块电路板子或者只有一块电路板。前者称为微型机，后者称为单板机。最小的计算机只由一个芯片电路组成，称为单片机。近几年最流行的是微型机，该机有体积小，功能全，价格低等优点，其大小只占半个写字台，每台机器价格为几百~几千美元，因而个人有条件买得起，所以又称为个人计算机。

目前已经有人在着手研制用超大规模集成电路组成的具有智能的第五代计算机。这种计算机以新的理论和新的技术为基础，具有并行处理(一个问题被分解为多个子问题同时处理)功能，容错(在发生故障情况下继续工作)功能。整个计算机由推理系统，知识库系统及智能接口系统组成。主机将由 1000~10000 个处理机构成，主存容量达 1000~

10000 兆字节，处理速度为 10~100 亿次/秒。能用自然语言、声音、图像和图形输入计算机并和计算机会话。它将具有近似人脑的功能，能推断，能联想和学习，是一种全新的智能计算机。

六十年代以后，电子计算机技术的发展是“迅猛的”。这不仅反映在计算机台数的增长，同时也反映在机器的质量和用途上有飞跃的发展。自从 1946 年 ENIAC 问世以后，到 1950 年才只有 10 台计算机。但到了 1970 年全世界已有计算机总数达到 10 万台。到 1980 年全世界计算机台数突破 1000 万台。

计算机的发展情况可由表 1-1 看出。

表 1-1 电子计算机各代的划分

计算机年代	起迄年代	代表机种	元器件
1	1946—1957	ENIAC	电子管
2	1958—1964	NCR304	晶体管
3	1965—1970	IBM370	集成电路
4	1970—	IBMPC	大规模集成电路
5	1990—		超大规模集成电路

§ 4 计算机与二进制^[1]

在电子计算机中使用二进制数码作为运算和表示信息的基本单位。

1. 二进制数

我们在日常生活和生产劳动中，经常碰到的数都是十进制的，它有“0, 1, 2, 3, 4, 5, 6, 7, 8, 9”十个数字和“.”一个小数点组成。从低位到高位是逢十进位的。

二进制数制只有“0, 1”二个数字和“.”一个小数点组成。从低位到高位是逢二进位的。

表 1-2 列出了十个数的十进制数和二进制数的对照表。

表 1-2 十进制数和二进制数

十进制	0	1	2	3	4	5	6	7	8	9
二进制	0	1	10	11	100	101	110	111	1000	1001

那么计算机为什么要采用二进制数制呢？这是因为二进制比十进制有以下几处优点。

第一，由于二进制只使用“0”和“1”两个不同的数字符号，所以任何具有两个不同稳定状态的元件都可用来表示数的每一位。而制造具有两个稳定状态的元件要比制造多个稳定状态（如十个稳定状态）的元件容易得多。

第二，从运算操作的角度来说，二进制也是最简单方便的一种计数制。二进制的加法和乘法公式只有四条。

$$0+0=0 \quad 0 \times 0=0$$

$$0+1=1 \quad 0 \times 1=0$$

$$1+0=1 \quad 1 \times 0=0$$

$$1+1=10 \quad 1 \times 1=1$$

而十进制的加法和乘法公式各有 100 条。很显然，二进制的运算规则比十进制运算规则简单得多。

第三，采用二进制表示数可以节省存贮设备。例如：我们要表示 0~99 这 100 个数码。用十进制表示需要二位数，每位数需十个元件来获得十个稳定状态，共需 $2 \times 10 = 20$ 个元件。而用二进制表示需要七位数，每位数需要二个元件来获得二个稳定状态，共需 $7 \times 2 = 14$ 个元件。可见，用二进制表示时所需设备量比十进制少。

第四，由于二进制中只有“0”和“1”两个不同符号，因此可以使用逻辑代数这一工具来分析和综合电子计算机中有关的逻辑线路，设计出简单合理和高效率的计算机。

2. 二进制位、字节、字长

在计算机中一位二进制信息是最小的信息量，称为“比特”(bit)。也叫“二进制位”，简称“位”。

在计算机进行运算和传递数据的过程中,以一定个数的二进制信息为组,作为机器处理信息过程中的基本信息单位叫“字节”(Byte),通常一个字节为 8bits。由图 1-1 所示。

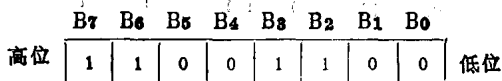


图 1-1 8 位(bit)组成的一个字节(Byte)

计算机的字长(word)是指内存贮器的存贮单元的宽度(位数)或者数据总线的宽度。一般字长为二倍或四倍字节数。即一个字长为 16bits 或 32bits。图 1-2 所示为由二个字节组成的一个字长。

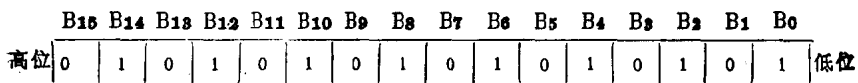


图 1-2 2 个字节(Bytes)组成一个字长(word)

3. 代码,指令,程序

(1) 代码是表示数据或指令的符号。用二进制数为符号来表示的叫二进制代码。计算机只直接认识二进制代码。

(2) 指令是规定计算机的操作类型,操作数的值及其存贮器的地址的一组编码字符。操作码指定计算机进行什么操作(如:加、减、乘、除、输入、输出等)。地址是指操作数所在地址和操作结果送往的地址。

(3) 程序是实现特定的操作目标而编制的有次序的一组指令的集合。程序是计算机能自动地完成处理信息任务的依据。

程序由指令组成,指令由代码组成,代码由二进制数表示。所以二进制数为计算机具有智能的基础。

§ 5 计算机系统的构成^[3]

一个典型的计算机系统包括硬件和软件两大部分。硬件部分主要有主机,系统总线,输入输出接口,外部设备等组成。图 1-3 为硬件部分组成框图。

软件部分主要由系统软件和应用软件组成。表 1-3 所示各种软件