

微型计算机 在统计及经济预测中的应用

尹建生 编著

WXJSJ
ZTJJJJYCZDYY

地震出版社

7-222
10

微型计算机 在统计及经济预测中的应用

尹建生 编著



地震出版社

内 容 提 要

本书着重介绍微型机在统计与经济预测中的应用，目的是使广大管理工作者在实际工作中掌握和应用微型机。书中简单明了地论述了目前国内外在统计与经济预测中常用的几十种方法，书中所讲的“统计与经济预测常用方法软件包”是将上述方法与微型计算机技术结合起来，通过八个程序系统，大量的实例，系统地介绍了微型机在原始资料处理、统计检验，各种时间序列的预测及相关与回归分析中的应用。本书还系统地介绍了微型机数据库dBASE II 及一些简单的程序设计方法，并在本书的最后一章给出了该软件包的使用说明。

本书可供统计人员、计算机软件人员及经济工作者使用，也可供各行各业管理工作者及大专院校师生参考。

2549/23

微 型 计 算 机
在统计及经济预测中的应用
尹建生 编著
责任编辑： 张晓梅

地 震 出 版 社 出 版

北京复兴路63号
山东电子工业印刷厂印刷
新华书店北京发行所发行
全国各地新华书店经售

787×1092 1/16 15.25印张 386千字

1987年4月第一版 1987年4月第一次印刷

印数 00001—10000

统一书号：17180·387 定价：4.00元

序

实现管理现代化,计算机、特别是微型计算机的应用显得愈来愈重要。这是因为,现代社会已进入信息化的时代,不仅信息量大,而且要求迅速、准确,及时地加工处理,以便为决策提供依据。在这种形势下,靠手工处理的办法已不能适应实际的需要。微型机具有数据处理的功能,其快速性、准确性,非凡的记忆能力和数据的存储能力是人的脑力和体力所无法比拟的。因而微型机的应用,为解决现代管理中这一矛盾提供了新手段。

近几年来,我国推广应用微型机,实现办公室自动化方面已取得不少进展。过去靠手工处理或难以用手工处理的工作,由计算机代替,极大地提高了工作效率和工作质量。计算机在国家宏观决策和管理方面的应用也初见成效。目前,国家计委和国家统计局已分别建立了综合、能源、农业、人口、人材和骨干企业基本情况等数据库,建立了经济发展预测、政策模拟的经济数学模型,为政策和计划的制定提供了依据。同时,也可看到,计算机的应用,能给社会带来巨大的经济效益。如石化总公司采用计算机改进生产和管理,1985年一年中就取得经济效益2.6亿元;国家物资局通过计算机信息系统对轴承产品进行管理,使库存量下降30%,压缩占用资金二亿多元;又如京棉三厂,1982年下达利润指标,经人工预算认为难以完成,而由于应用了微型机,采用目标规划方法,建立自动配棉的软件系统,利用微型机实现优化所得结果比下达指标多得利润134.4万元,与人算相比可多得利润253.5万元。

毫无疑问,我国在推广应用计算机方面已取得突破性的发展。但应当承认,从总体来看,在我国,计算机的应用还急待进一步提高。从我国目前拥有的计算机数量和它们可能发挥的潜在作用来说,还是不相称的。这不仅与我国目前的管理水平较低,管理人员对计算机应用的基本知识缺乏应有的了解有关,也与国内应用软件质量不高有关。因此,在实现管理现代化过程中,为了不断扩大计算机的应用领域,充分发挥计算机的潜在作用,在不断提高我国管理人员的素质和管理水平的同时,应当进行计算机、特别是微型机应用知识、现代管理知识的普及教育和基础训练。同时,还应当集中力量,结合我国的国情,有系统地开发高水平、高质量的应用软件,用到各行各业的管理中去。

现出版的《微型计算机在统计及经济预测中的应用》这本书,结合大量实例,介绍了几十种常用的方法,并和微型机计算技术相结合,建立了“统计与经济预测常用方法软件包”,只要实际工作者具有应用微型机的基本知识和掌握操作方法,就可以在实际工作中应用。因此,根据目前我国实际情况,该书具有一定的实用价值。同时,也可作为对管理人员进行微型机应用知识普及教育和基本训练的很好的参考书。

该书的作者是一位较年轻的科技工作者,勇于实践和探索,该书实际上是作者实践经验的总结。考虑到作者的阅历有限,所著之书难免有不足之处。笔者期望该书作者继续努力进取,为开发符合我国国情的高水平、高质量的应用软件作出出色的成绩,为实现我国管理现代化作出创新的贡献!

中国管理现代化研究会

金良超 周子康

1986.6.10.北京

前 言

随着现代科学技术的迅猛发展，人类社会开始进入到一个利用和开发信息资源的信息社会。信息数量之大、范围之广、变化之快，致使信息管理已成为一门科学。这门科学中包括了计算机的应用。由此，计算机这一处理和开发信息的工具更加引起人们的重视。特别是近几年，微型机以其成本低、体积小、功能强、易掌握等特点，进入到社会的各个领域，它与其它一些新技术相结合，形成了强大的生产力，发挥着越来越重要的作用。

本书从广大经济管理工作者的实际出发，详细介绍了微型机在统计及经济预测中的应用。

从具体分析解剖一个微型机“统计与经济预测常用方法软件包”入手，在计算方法、程序设计及实际应用上结合实例加以详细论述。读者在此基础上，只需对微型机的操作稍加学习，便可在实际工作中加以应用。为了使读者能够顺利地应用本软件包进行经济预测，本书系统地介绍了dBASE II命令及软件包的使用说明，对于各种统计与预测方法的直观原理也作了一些介绍。

本书所介绍的“统计与经济预测常用方法软件包”业经中国工商银行总行及有关专家的鉴定，认为该软件包达到了原设计标准，在统计与经济预测工作中有较高的实用价值。

本书的写作过程中，中国管理现代化研究会给予了大力支持，李茜同志作了大量的整理及图、表的绘制工作，金良超、李德宣、崔德海同志审阅了本书，并得到了伊世忠、张习武、史凤歧、何建文、李静宜、李栓柱等同志的帮助，中国工商银行总行、燕山计算机应用中心、北京市统计局、北京公交第一汽车公司、国际旅行社北京分社、延吉玻璃厂等有关单位给予了大力支持和热情帮助，在此深表感谢。

由于作者水平有限，本书中难免有不妥和谬误之处，敬请读者批评指正。

作 者

一九八六年四月于北京

目 录

第一章 概述	(1)
第一节 微型机在统计及经济预测中的应用.....	(1)
第二节 微型机应用程序设计简介.....	(2)
第二章 统计资料处理程序系统	(5)
第一节 原始资料的次数分配程序.....	(5)
第二节 原始资料的分析处理程序.....	(7)
第三节 dBASE II 基本性能指标及建立数据库命令.....	(10)
第四节 程序应用实例.....	(14)
第三章 概率统计检验程序系统	(17)
第一节 基本概念及符号约定.....	(17)
第二节 U检验法程序.....	(18)
第三节 T检验法程序.....	(20)
第四节 χ^2 检验和F检验程序.....	(23)
第五节 符号检验程序.....	(24)
第六节 dBASE II 的命令文件.....	(25)
第七节 程序应用实例.....	(29)
第四章 产销计划预测分析程序系统	(31)
第一节 产销计划预测分析程序.....	(31)
第二节 dBASE II 的运算符及函数.....	(34)
第三节 程序应用实例.....	(37)
第五章 简单时间序列预测程序系统	(39)
第一节 统计分析程序.....	(39)
第二节 简单时间序列预测程序.....	(40)
第三节 dBASE II 的显示及打印命令.....	(43)
第四节 程序应用实例.....	(46)
第六章 长期趋势预测程序系统	(53)
第一节 加权外推法预测程序.....	(53)
第二节 直线模型预测程序.....	(54)

第三节	曲线模型预测程序	(59)
第四节	修正指数曲线预测程序	(61)
第五节	罗吉斯曲线预测程序	(65)
第六节	长期趋势法预测程序的选择	(67)
第七节	dBASE II 的内存变量	(68)
第八节	程序应用实例	(70)
第七章	季节波动预测分析程序系统	(73)
第一节	同季平均法预测程序	(73)
第二节	平均数趋势整理法预测程序	(74)
第三节	季节比例法预测程序	(77)
第四节	环比法预测程序	(78)
第五节	dBASE II 的求和、排序及SET命令	(80)
第六节	程序应用实例	(84)
第八章	短近期预测程序系统	(87)
第一节	移动平均法预测程序	(87)
第二节	加权移动平均法预测程序	(95)
第三节	指数修匀法预测程序	(98)
第四节	短近期预测方法的选择	(103)
第五节	dBASE II 的屏幕编辑功能	(103)
第六节	程序应用实例	(104)
第九章	相关与回归预测程序系统	(107)
第一节	相关与回归预测程序	(107)
第二节	计算相关系数程序	(108)
第三节	直线回归预测程序	(110)
第四节	曲线回归预测程序	(113)
第五节	自身回归预测程序	(120)
第六节	二元线性回归预测程序	(123)
第七节	三元线性回归预测程序	(130)
第八节	预测回归模型的选择	(136)
第九节	应用dBASE II 程序的准备工作	(137)
第十节	程序应用实例	(139)
第十章	“统计与经济预测常用方法软件包” (简称“软件包”)	

使用说明	(144)
第一节 “统计与经济预测常用方法软件包”的基本介绍.....	(144)
第二节 进入“统计与经济预测常用方法软件包”程序系统.....	(147)
第三节 统计资料分析处理程序系统的使用.....	(148)
第四节 概率统计检验程序系统的使用.....	(155)
第五节 预测分析程序系统的使用.....	(164)
第六节 简单时间序列预测程序系统的使用.....	(168)
第七节 长期趋势预测程序系统的使用.....	(171)
第八节 季节波动预测程序系统的使用.....	(182)
第九节 短期趋势预测程序系统的使用.....	(192)
第十节 相关与回归预测程序系统的使用.....	(204)
第十一节 退出程序系统.....	(225)
附录一	
dBASE II 命令索引表.....	(226)
附录二	
表1 标准正态分布表.....	(228)
表2 t分布表.....	(230)
表3 X^2 分布表.....	(231)
表4 F分布表.....	(232)
表5 符号检验表.....	(235)
参考文献	(236)

第一章 概 述

近几年来,随着微型机本身硬、软件功能的加强及配有汉字的微型机数据库 dBASE II 的出现,使我国微型机的普及应用出现了一个崭新的局面。目前,微型机已广泛地应用在办公室管理自动化等方面,并利用微型机数据库 dBASE II 建立了诸如人事档案管理、物资管理、情报检索及工资发放等应用软件系统,为人们的日常工作提供了方便。但是,微型机仅仅应用于这些简单的查询、检索、存贮等工作,还是远远不能满足实际工作的需要。如何用微型机开发利用信息资源,如进行统计分析 & 经济预测等工作,已成为急待解决的实际问题。

第一节 微型机在统计及经济预测中的应用

一、可行性分析

开展统计与经济预测工作离不开对大量数据资料的分析处理。从以往的实践中,尽管人们对大、中型计算机进行数据处理的功能是深信不疑的,然而大、中型计算机的应用范围却是相当有限的。随着近几年微型机的迅速普及,人们越来越多地依靠微型机开展各方面的工作。实践证明,微型机是能够完成对一般中、小型数据库的存贮及对数据进行各种分析处理的。其原因有以下两点:

1. 从微型机的硬件配置来分析 由于微电子技术的迅速发展,从七十年代至今,微型机更新了四代,硬件设备能力与中、小型计算机的差距越来越小。以机器的内存为例,八位机的内存一般是在 32K—128K 之间,而我国的一些小型计算机如: TQ-15、DJS-130 等,其内存只在 32K—64K 之间。目前国内常用的十六位微型机的存贮容量如 IBM-PC/XT、ZD-3100 等,其内存一般在 256—640K 之间,并配有一台或两台软盘驱动器及一台温氏硬盘。由此可见,微型机的外存容量已相当可观,其硬件本身的配置功能是可以完成中、小型数据库的存贮及开展数据的分析处理工作的。

2. 从微型机的系统软件上分析 目前,大、中型计算机上配有的各种高级语言,如: BASIC 语言、FORTRAN 语言、COBOL 语言等,微型机上也都配有,大、中型计算机上配有数据库系统,微型机上则配有微型机数据库,如目前国内外比较流行的 dBASE II 等,具备了大型机数据库的基本功能,因此可以在微型机上建立一些中、小型的数据库。

虽然微型机能够完成中、小型数据库的存贮及数据的分析处理,但并不是说微型机能代替大中型计算机。因为用微型机进行数据处理的速度、精度及其它一些功能还远远低于大、中型计算机。然而微型机之所以能够广泛地应用于工矿企业、机关等国民经济的各个部门,是因为它具有价廉、灵活、操作简便、易于推广普及等特点。

二、在统计及经济预测中的重要作用

统计工作要实现现代化,就要运用先进的统计科学和现代化的计算技术,而运用计算机

技术,不但能使人们从大量繁杂的手工处理数据的工作中解脱出来,而且还可以大大提高对统计数据的信息利用率。

在今天的信息社会里,预测分析研究工作越来越引起经济部门的重视。对经济工作者及管理决策部门,要想做到纵观全局,了解现状,预测未来,就需要对大量的信息资料进行分析处理。然而,今天要想通过传统的手工方法来获得及时准确的信息是十分困难的。因此,只有发挥计算机的强大功能,才能获得具有高度逻辑性、预见性的信息。

应用微型机开展统计及经济预测工作,首先要明确的是,就计算机本身而言是不会思考、判断、分析处理问题的,它只能根据人们的安排(编写成的计算机程序)来做某些固定模式的判断,然后根据判断进行相应的操作。因此,它最适用于进行大量重复性的信息操作。

运用微型机进行数据处理的步骤(程序)与手工方法基本一致,都是从原始数据即尚未作评价的各种事实(信息)的处理开始,以产生新的数据(新的信息)为结束。其步骤包括对数据的收集、记录、分类、排序、存贮、计算加工、传输,制表和递交等处理过程。根据这些过程的要求,编写出与之相应的程序,由计算机直接进行数据处理。

微型机在统计工作中可以及时、准确地将有关统计资料的参量如均值、方差、标准差、偏态等迅速地提供给人们,特别是对一些数据量较大,难于用手工完成的统计工作,微型机更体现出了它的优越性。

微型机在统计中的应用,还表现在利用经过分析处理的统计资料,进行统计检验、概率分析,这在统计分析中是十分重要的。通过对某种产品的实验研究,或对某种现象的抽样调查,有助于在实际工作中避免主观性和盲目性,并最终选择出一种最佳方案。以前人们往往不愿意或很少做这方面的检验工作,其主要原因是因为数据量太大、太复杂。而应用微型机后,实现上述检验却是相当容易的。这些在本书后面各章中将做详细论述。

统计与预测的区别在于统计是侧重于对客观事物现象的事后写实和研究,而预测是侧重于对客观事物现象的未来进行推测和研究。即预测是按照客观事物的发展规律,在掌握大量信息的基础上进行分析判断,从而提出对未来的推测。如在商品生产中,生产管理者根据商品历年销售量的统计,可以获得该商品历年的递增(递减)值,销售量的最大、最小值及均数和获利总额等等许多数据,这是统计工作。如果再根据最近市场需要和发展趋势的调查,进行数据加工处理,就能得到有关该商品的市场预测信息。生产管理者通过对这些信息进行分析,从而作出对该商品发展趋势的评价,这就是预测。

微型机在经济预测中的应用,主要是以经济预测理论为依据,将常用的一些经济预测方法的数学模型,应用计算机语言编写成程序,在微型机上应用。例如:各种外推预测法、时间序列的长、中、短期趋势预测法及相关与回归的分析预测等。

微型机在经济预测中的应用,对人们充分利用信息资源,提高预测的精度、可信度和及时性等方面具有积极意义。目前,已有许多部门和企业正在开展这方面的工作。

第二节 微型机应用程序设计简介

一、dBASE II

在微型机应用程序开发中使用的计算机语言有几十种,国内常见的有: BASIC语言、

COBOL语言、FORTRAN语言、PASCAL语言、C语言及dBASE数据库语言。

dBASE II是一个功能强、实用价值高的微型机数据库管理系统。在国内微型机上应用较为普及。利用它可以进行各种数据库的建立和数据库检索。如人事档案、资料管理等,但是一般利用dBASE II对已建立的数据库进行各种综合统计及开展对经济预测等方面的工作还比较少。从dBASE II的功能来看,对大量的数据进行处理分析正是其特长。dBASE II的主要功能如下:

1. 在语言上采用英语化结构,每条命令都是以动词开头,后面跟有若干子句或短语,因此,使用起来就显得灵活方便。具有会话性,易学、易记、简单明了。

2. 具有较强的全屏幕操作编辑功能,可以灵活地对数据库中的记录进行增、删、插、改,并有对数据库的复制、连接等功能。

3. 在数据库组织中能排序和索引,因此可以对数据进行快速定位及查询、检索。

4. 具有灵活的报表打印及统计、求和等功能。

5. 可与各种高级语言交换数据,而且自身具有的程序编辑命令要比其它高级语言简单实用。其程序编写速度比COBOL等其它高级语言快6—10(倍)。¹⁾

6. 具有很强的通用性,可以在CP/M、MS/DOS、CDOS、M/CM等任一操作系统支持下运行,也可在IBM、ZD、长城0520系列等不同种类的微型机上运行。目前,在dBASE II的基础上,汉字dBASE III已开始应用,功能有了进一步提高。当然dBASE II作为微型机数据库还具有一定的局限性,一般只适合于中、小型数据库管理。

二、微型机应用程序设计中的PAD图

所谓PAD图是由二维空间和过程结构组成,即将程序的处理过程和处理时序作为两个坐标轴。一个是时间(序)轴,一个是过程(层次)轴,使整个程序设计层次清晰、直观。

PAD的示意图如图1-1所示²⁾:

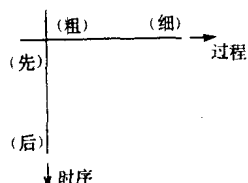
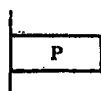


图1-1

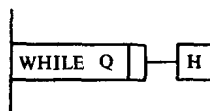
下面给出PAD图的基本图示

1. 处理表示符



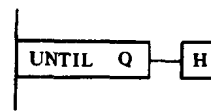
2. 循环表示符

a. WHILE



当条件Q成立时做H操作

b. UNTIL



重复H操作直到条件Q成立

1) 《微型机数据库》, 徐洁磐等编著。

2) 《汉字dBASE II实用技术》, 王电编。

c. DO



变量 I 由m增至n的过程中重复做H操作

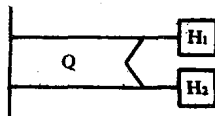
d. DOWN TO



变量I由m减至n的过程中重复做H操作

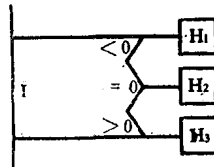
3. 条件表示符

a. IF.....THEN.....ELSE



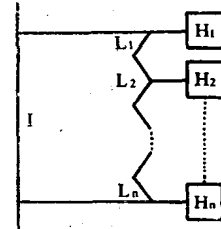
如果条件Q成立就做H₁操作，
否则，就做H₂操作

b. 算术IF型



如果 $I > 0$ 就做H₁操作
 $I = 0$ 就做H₂操作
 $I < 0$ 就做H₃操作

c. 计算GO TO型



当 $I = L_1$ 时做H₁操作
 $I = L_2$ 时做H₂操作
...
 $I = L_n$ 时做H_n操作

图1-2 PAD图的基本图示

最后，将微型机应用程序的设计过程归纳如下：

(1) 确立二次开发软件的选题。即根据自己的工作需求，提出有关应用软件二次开发的题目。

(2) 技术准备工作。掌握准备采用的计算机语言，如dBASE II等。

(3) 可行性分析。即分析机器本身的硬、软件功能，能否满足对应用软件设计的要求。

(4) 画出程序的PAD图。

(5) 编制程序。即在PAD图形成的基础上，编写程序。

(6) 程序调试。调试包括对程序的运行、总体布局的优化、改进等工作。

(7) 应用软件说明书的撰写。软件说明书一般包括以下几个内容：①对程序设计可行性的分析，②程序的技术说明书。③程序的操作使用说明书。④程序的调试验收报告。⑤程序的PAD图和全部系统清单等。

第二章 统计资料处理程序系统

统计是处理数据的重要方法和手段。是专门研究数量资料、从事数量的搜集、整理和分析等数学领域中的一门学科，它有自己的符号、术语、定理和计算方法。

在统计中，描述统计是十分重要的。对于原始资料的分组，对样本平均数、中位数、及众数的计算以及方差、标准差、全距等计算都是描述性统计。

本章将主要介绍如何利用微型机解决上述问题，并结合实例加以说明。

第一节 原始资料的次数分配程序

一、传统的原始资料次数分配方法

传统的原始资料整理法，首先是为原始数据编制次数表，把原始数据按一定序列分成几组或几类，并分别记录各组事件的次数的系统材料¹⁾。

以某工厂的人事部门想分析该厂30个技术工人的月工资情况为例，首先要搜集这些人的工资资料，如表2-1所示。30个技术工人未经整理的月工资数值，称为原始资料²⁾。

表2-1 30个技术工人月工资数值					
工人序号	每月工资 (元)	工人序号	每月工资 (元)	工人序号	每月工资 (元)
一	106	十一	119	二十一	96
二	84	十二	87	二十二	105
三	110	十三	118	二十三	107
四	91	十四	97	二十四	121
五	109	十五	103	二十五	105
六	91	十六	106	二十六	95
七	111	十七	85	二十七	106
八	107	十八	106	二十八	128
九	99	十九	101	二十九	111
十	94	二十	105	三十	101

对于表2-1的原始资料，要分成两步才能把它组织成一种合乎逻辑的形式。

第一步，需要给数据安排序列，这种序列是把数值，按从小至大(或从大至小)的次序排列。从表2-1中可见，最低工资(二号，84元)和最高工资(二十八号，128元)，并从84元到128元列出每一数值的序列。

第二步，是按传统的划记法，把所有工人的工资划记在序列中，如表2-2所示。

从表2-2中可见，月工资84元者出现一次，91元者出现两次，105元者出现三次，106元者出现四次等。

对于仅仅限于30个技术工人的工资序列情况，尚可按照这种方法进行次序排列。但是，对于成千上万的数据资料，如果还用这种方法进行处理就很困难了。为此，必须借助于现代化的工具——计算机，才能从传统的手工方法的束缚下解放出来。

二、原始资料次数分配程序

1)、2)《工商业和经济学用统计方法技术》第一册。罗·第·梅森著，中国人民大学统计研究室译。

原始资料次数分配程序的设计过程同第一章第二节“程序设计简介”中的方法是一致的,需要经过调研、可行性分析、系统分析、画出程序框图,最后是程序的编制。

表2-2 序列与划记

工 资 (元)	划 记	工 资 (元)	划 记	工 资 (元)	划 记
84	/	99	/	114	
85	/	100		115	
86		101	//	116	
87	/	102		117	
88		103	/	118	/
89		104		119	/
90		105	///	120	
91	//	106	////	121	/
92		107	//	122	
93		108		123	
94	/	109	/	124	
95	/	110	/	125	
96	/	111	//	126	
97	/	112		127	
98		113		128	/

1. 程序设计的目的 利用计算机处理原始资料的目的与手工方法基本上是一致的,都是将原始资料从杂乱无章的状态下,处理成为按着一定序列排序的数据资料,经过计算机处理的这些资料,可以向人们提供建议性次数分组表,并同时给出分组后的最大值、最小值、全距、组距、组数等简单的统计分析值。

2. 程序的框图设计 在经过调研、可行性分析、系统分析之后,将程序的框图画出来(见图

2-1)。框图设计是编写程序的重要一步,程序设计中技巧的高低和功能的强弱,多取决于框图的设计水平。从图2-1中,可以看到,原始资料次数分配程序基本上由四个模块组成。即主程序模块、建库模块、计算模块、输出模块。有关框图中的具体内容,我们将在本章第四节“程序设计实例”中做详细介绍。

3. 程序运算的执行过程 使用者首先要将原始资料装入程序的数据库中,并根据实际情况给出“建议组数”,然后程序开始执行计算,对原始资料进行排序和分组。最终结果即可在屏幕上显示出来,也可按表格形式打印出来。以表2-1为例,结果见表2-3、2-4、2-5。本程序的详细操作步骤见第十章的使用操作说明。

4. 计算结果分析 表2-3中的数据是未经处理的原始资料。表2-4中的数据是经过计算机处理后的排序资料。表2-5是经过处理后的原始数据次数分配最终结果资料。其中,全距是指数据的最大值与最小值之间的距离,组距是指两组之间的间隔距离。

值得注意的是:原始资料的次数分配中,组距的宽度应尽可能相等,而且要避免使用开口组,例如:工资80元以上或100元以下等。最后在选组距时应使组织次数分配的结果仅有一个集中处(或单峰)。如果出现双峰分配现象,往往说明该组的组距还应扩大,或者编制

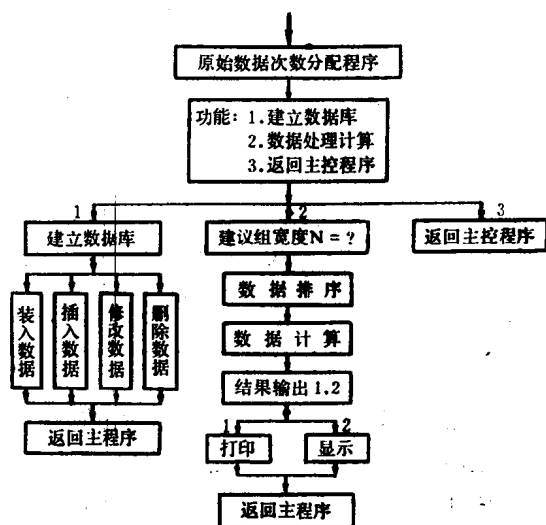


图2-1

次数分配的资料是异质的，原因也许来自两个不同的总体¹⁾。

表 2-3 原始数据次数分配
程序原始数据表

序 号	数 据
1	106.00
2	84.00
3	110.00
4	91.00
5	109.00
6	91.00
7	111.00
8	107.00
9	99.00
10	94.00
11	119.00
12	87.00
13	118.00
14	97.00
15	103.00
16	106.00
17	85.00
18	106.00
19	101.00
20	105.00
21	96.00
22	105.00
23	107.00
24	121.00
25	105.00
26	95.00
27	106.00
28	128.00
29	111.00
30	101.00

表 2-4 原始数据次数分配
程序数据整理表

序 号	数 据
2	84.00
17	85.00
12	87.00
4	91.00
6	91.00
10	94.00
26	95.00
21	96.00
14	97.00
9	99.00
19	101.00
30	101.00
15	103.00
20	105.00
22	105.00
25	105.00
1	106.00
16	106.00
18	106.00
27	106.00
8	107.00
23	107.00
5	109.00
3	110.00
7	111.00
29	111.00
13	118.00
11	119.00
24	121.00
28	128.00

第二节 原始资料的分析处理程序

上一节中我们所介绍的原始资料次数分配程序，仅仅是对原始资料做了最简单的初步处理。如果要想通过对原始资料的分析处理达到探求未知，做出推论的目的，那么还需对原始资料做进一步的研究。原始资料的分析处理程序，就是利用微型机对原始资料做进一步的分析处理工作。

本节在介绍程序之前，有必要对有关统计学中的计算公式及定义加以概述，使读者对利用微型机处理统计资料的过程有较为全面的了解。

一、统计资料计量方法的介绍

1. 集中趋势的测定 在统计学中对于测定中心值最常见的计量数字是算术平均数、中位数及众数。

算术平均数：算术平均数等于观察值总和除以观察值的项数。即：

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} \text{ 或 } \bar{X} = \frac{\sum X}{n}$$

式中： $\bar{X}$ ：算术平均值； X ：观察值； n ：项数。

中位数：中位数是位于样本分配中间的一个数值，其样本中大于或小于这个数值的两个数相同。其公式为：

$$\text{中位数} = b_L + \frac{\frac{n}{2} - F_{M-1}}{f_M} \cdot C \quad (2)$$

1)《工商业和经济学用统计方法技术》第一册，罗·第·梅森著，中国人民大学统计研究室译。

2)公式中假设的“分组资料”是“均匀分布”。

表 2-5

原始数据次数分配程序次数分配表

序 号	数 据	组 一	组 二	组 三	组 四	组 五
2	84.00	84.00	0.00	0.00	0.00	0.00
17	85.00	85.00	0.00	0.00	0.00	0.00
12	87.00	87.00	0.00	0.00	0.00	0.00
4	91.00	91.00	0.00	0.00	0.00	0.00
6	91.00	91.00	0.00	0.00	0.00	0.00
10	94.00	0.00	94.00	0.00	0.00	0.00
26	95.00	0.00	95.00	0.00	0.00	0.00
21	96.00	0.00	96.00	0.00	0.00	0.00
14	97.00	0.00	97.00	0.00	0.00	0.00
9	99.00	0.00	99.00	0.00	0.00	0.00
19	101.00	0.00	101.00	0.00	0.00	0.00
30	101.00	0.00	101.00	0.00	0.00	0.00
15	103.00	0.00	0.00	103.00	0.00	0.00
20	105.00	0.00	0.00	105.00	0.00	0.00
22	105.00	0.00	0.00	105.00	0.00	0.00
25	105.00	0.00	0.00	105.00	0.00	0.00
1	106.00	0.00	0.00	106.00	0.00	0.00
16	106.00	0.00	0.00	106.00	0.00	0.00
18	106.00	0.00	0.00	106.00	0.00	0.00
27	106.00	0.00	0.00	106.00	0.00	0.00
8	107.00	0.00	0.00	107.00	0.00	0.00
23	107.00	0.00	0.00	107.00	0.00	0.00
5	109.00	0.00	0.00	109.00	0.00	0.00
3	110.00	0.00	0.00	110.00	0.00	0.00
7	111.00	0.00	0.00	0.00	111.00	0.00
29	111.00	0.00	0.00	0.00	111.00	0.00
13	118.00	0.00	0.00	0.00	118.00	0.00
11	119.00	0.00	0.00	0.00	119.00	0.00
24	121.00	0.00	0.00	0.00	0.00	121.00
28	128.00	0.00	0.00	0.00	0.00	128.00

式中： b_L ：中位数组的下限；

n ：观察值的项数；

f_M ：中位数组中观察值的项数；

F_{M-1} ：在中位数组前边的第 $(m-1)$ 组的观察值的项数；

C ：宽度为 C 单位的一个组距。

众数：是在样本中出现次数最多的一个观察值。

算术平均数、中位数、众数三者之间的近似估计公式：

众数 = 算术平均数 - 3(算术平均数 - 中位数)

算术平均数 = $\frac{3(\text{中位数}) - \text{众数}}{2}$

中位数 = $\frac{2(\text{算术平均数}) + \text{众数}}{3}$

2. 离势数量的测定 在统计中对于测定离势的常用计量数字是全距、平均差、方差和

标准差等。

全距：是一个样本观察值的两个极端数值之差。

全距 = 样本最大观察值 - 样本最小观察值

绝对平均差：是每一项观察值与样本平均数的绝对差异的算术平均数。即：

$$M \cdot A \cdot D = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

式中：M·A·D：平均绝对差；

$|X_i - \bar{X}|$ ：($X_i - \bar{X}$)的绝对值。

方差：是距算术平均数的各离差平方的算术平均的数字。其公式：

未分组时：

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

分成K组时：

$$S^2 = \frac{1}{n-1} \sum_{i=1}^k (X_i - \bar{X})^2 F_i$$

式中： X_i = 第i组的组中值；

F_i = 第i组观察值的项数；

$\bar{X}$ = 样本平均数。

标准差：标准差是方差的平方根。其公式：

未分组时：

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

分成K组时：

$$S = \sqrt{\frac{\sum_{i=1}^k (X_i - \bar{X})^2 F_i}{n-1}}$$

3. 偏态 是测定一个次数分配系不对称的偏态程度。一般说来，统计中用集中趋势、离势数量和偏态的测定，基本可以概括地描述一个资料的次数分配性质。

偏态系数测定公式：

$$S_K = \frac{3(\bar{X} - \text{中位数})}{S}$$

其中： $\bar{X}$ ：算术平均数；

S：样本的标准差。

用图 2-2 表示如下：