

新世纪网络课程建设工程



经济数学基础

概率论与数理统计

李林曙 黎诣远 主编

高等教育出版社



经济数学基础

概率论与数理统计

李林曙 黎诣远 主编

高等教育出版社

内容提要

本书是与“新世纪网络课程建设工程”——经济数学基础网络课程相配套的文字教材,全书在编写过程中坚持“数学为体、经济为用”的原则.全书共分4册:包括网络课程学习指南、微积分、概率论与数理统计、线性代数等内容,涵盖了高等院校本专科经济管理类专业必要的数学基础.每册书配有学习光盘,可供学生课后使用,有条件的读者也可通过网络直接学习本课程.

本书可供全国各高等院校、广播电视大学、成人高校和职工大学经济管理类及相近各专业的学生使用.

图书在版编目(CIP)数据

经济数学基础.概率论与数理统计/李林曙,黎诣远

主编. —北京:高等教育出版社,2004.3

ISBN 7-04-013784-4

I. 经... II. ①李...②黎... III. ①经济数学—高等学校—教材②概率论—高等学校—教材③数理统计—高等学校—教材 IV. F224.0

中国版本图书馆 CIP 数据核字(2003)第 110929 号

出版发行 高等教育出版社

社 址 北京市西城区德外大街 4 号

邮政编码 100011

总 机 010-82028899

购书热线 010-64054588

免费咨询 800-810-0598

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

经 销 新华书店北京发行所

印 刷 北京中科印刷有限公司

开 本 787×1092 1/16

印 张 21.25

字 数 380 000

版 次 2004 年 3 月第 1 版

印 次 2004 年 3 月第 1 次印刷

定 价 38.80 元(含二张光盘)

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

序

经济数学,即在经济中应用的数学,是经济学与数学相互交叉的一个新的跨学科领域。

我国正处在社会主义初级阶段.这个阶段的社会主义建设是以经济建设为中心的.围绕这个中心,各门学科都发挥着自己的作用.经济学与数学结合起来,发挥着比它们各自的作用还要大的作用.

大家知道,数学作为一门主要的基础学科,有着极其广泛的应用.它的应用领域,首先是自然科学,进而到工程技术,再扩展到社会科学.在社会科学中,数学的首要应用领域,无疑是经济学.

经济学为了更好地对经济工作发挥指导作用,需要引入和运用数学,用数量分析来补充和发展质的分析,使这两种分析相互结合产生更大的威力.经济学中的均衡与优化等问题,以及经济工作中的计划、预测、评估、组织、控制等管理(包括决策)问题,都需要数学及其分支学科进行分析研究、计算求解.特别是电子计算机用于经济学和经济管理工作以来,数学模型的建立与运用,模拟实验方法的制订与实施,更离不开数学的帮助.借助于电子计算机,数学能成功地解决各类静态的和动态的、线性的和非线性的经济问题.所以,可以说,数学已成了经济学家和经济工作者的“良师益友”,或者说极其有用的重要工具.

当然,两个不同学科的交叉影响,决不会是单向的,往往总是双向的.经济问题的各种特殊性,如多维性、随机性、不确定性、模糊性、突发性、利益冲突性、信息非对称性等等,为数学的进一步发展提供了契机,促进新的数学方法的诞生.对策论的发展就是一个有说服力的例证.最近十多年来,对策论在经济学中所起的作用,正如许国志院士所说的那样,远比它在数学中的作用为大.

经济数学不同于数理经济学和数量经济学.数理经济学是用数学语言表述的或用数学方法分析的经济学.数量经济学是在经济理论质的分析的基础上,利用数学方法和计算技术,研究经济数量关系及其变化规律的学科.经济数学作为在经济中应用的数学,自然要以数学为主体,以其应用的对象即经济为客体.它与数理经济学不同,有强烈的实用性;也与数量经济学不同,属于应用数学.

经济数学既要提高,更要普及.鉴于普及的需要,我国曾在1991年由中国经济出版社出版过陈克式、陈开周、崔福荫三位同志主编的《经济数学辞典》,许国志院士和我还为该辞典写过序.现在,摆在我们面前的这本由李林曙、黎诣远同志主编的《经济数学基础》教材,是高等教育经济管理类各专业的基础课,为中央广播电视大学重点建设课程和全国电大共建课程多种媒体一体化教材的主教材之一,并被作为网络课程列入教育部启动的“新世纪网络课程建设工程”,其普及效果将会更大和更好.

从内容看,教材不同于辞典,非用来查阅的工具书,是用来教学的课本,不仅有基础知识和运算方法,而且还有例题和习题,旨在使学生熟练掌握基本内容和提高解决实际问题的能力.由于这本教材的主要对象是高等成人业余教育的学生,其内容是针对他们入学基础的特点,适应他们今后学习经济管理课程和从事经济管理工作的需要,而设计与选定的,包括微积分、概率论与数

理统计、矩阵代数等.把经济融于数学.每章开始都有引子,用经济热点问题激发学生学习有关数学知识的兴趣,进而引导学生运用每章所学的数学通过分析引子提出的经济问题来加深对数学原理和方法的理解,以提高他们今后在经济管理工作中应用数学的能力.这本教材的另一个特点是充分注意到教学分层次循序渐进的要求,加强导学和助学,配有《跟我学经济数学》的辅助教材.

从形式看,《经济数学基础》既有文字教材,又有音像教材和网络课程,还有其他媒体如计算机辅助教学(CAI)课件、文具卡等,首次实现了同类教材的多种媒体一体化.由于主媒体与强化媒体有机配合,其他媒体有效补充,这就会有力地扩大经济数学的影响.随着我国1994年以来信息基础设施大规模建设的推进,远程教育如同远程医疗一样逐步成为现实.多媒体的宽带高速教育网正从根本上改变传统的教育方式.多媒体一体化教材和网络课程建设工程的出现,为网络化远程教育创造了条件,使师生们共同感受到信息时代在我国已经来到.

《经济数学基础》是全国电大高等财经科各专业的必修基础课,也是我国教育领导部门指定的11门财经课的核心课程之一,还是中央电大的一门重点课程,也是已建成的200多门新世纪网络课程中的一门重要课程.我衷心祝愿这套教材出版后产生为编写者和组织者所始料不及的深远而广泛的社会影响.

乌家培

1998年五一节,

改于2003年7月26日

前 言

这套《经济数学基础》教材是与教育部“新世纪网络课程建设工程”项目——“经济数学基础网络课程”配套的文字教材。

经济数学是指经济管理类专业所用的高等数学,这门课程与一般高等数学相比有其特殊性。因此学习本课程,首先需要正确认识经济与数学的关系。众所周知,任何事物都是质和量的统一体,没有无量的质,更没有无质的量。三百多年前,牛顿的《自然哲学的数学原理》和配第的《政治算术》,开辟了自然科学和社会科学数量化的时代。马克思认为:“一种科学只有在成功地运用了数学以后才算达到了完善的地步。”将数学用于经济学,可以深入揭示仅靠定性分析难以表达的现代经济错综复杂的相互关系及其变动趋势,可以提出经济决策的方向、力度和边界,可以预测这些决策的直接效果和间接效果。随着社会经济的快速发展,特别是社会主义市场经济的不断完善,加速提高经济效益和实现经济管理现代化的要求日益迫切,数量经济研究和定量分析越来越受到重视和加强,现实需要我们在经济管理和经济研究中卓有成效地运用数学方法实际问题,这些都对高等教育经济与管理学科各专业基础教学提出了更高的要求,也使经济数学课程成为高等教育经济与管理学科各专业学生的必修课之一。

这套《经济数学基础》教材坚持“数学为体”、“经济为用”的原则,根据课程教学大纲要求,包括一元函数微积分、二元函数微分学、概率论与数理统计和线性代数等内容,涵盖了高等教育经济与管理学科各专业必要的数学基础知识,并在内容选择和教学方法的改革上做了许多有益探索。首先,它立足于将经济有机地融于数学。每章开始都有一个短小精悍的“引子”,用当前经济生活中的热点问题激发学生学习有关数学知识的兴趣;在阐述内容时,尽可能以经济为例,使数学与经济不断结合;最后又以所学数学知识,回过头来分析和逐步解决“引子”提出的经济问题。这样,既帮助学生理解有关的数学原理和方法,又帮助学生了解它们在经济管理中的应用。专题讲座则通过若干问题的论述进一步说明经济学怎样运用数学分析和解决问题。其次,在编写过程中,贯彻“必需、够用”的指导思想,重视基本概念,重视基本运算技能的训练,重视培养学生运用数学方法解决实际问题的能力,而不拘泥于理论推导和繁琐的运算。在保证数学概念准确的基础上,在引例、解释和应用诸多方面力求更多联系与经济有关的问题。此外针对网络课程的特点,在文字叙述上力求深入浅出、通俗易懂、便于自学。

这套《经济数学基础》教材共分四个分册:《网络课程学习指南》、《微积分》、《概率论与数理统计》和《线性代数》。《网络课程学习指南》是指导刚刚跨进大学的同学如何进入经济数学的学习,如何利用“经济数学基础网络课程”进行自主学习,这一分册中还包括为同学拓展学习本课程相关内容专设的专题讲座的全部文稿。《微积分》、《概率论与数理统计》和《线性代数》三册均是主、辅合一的教材。每章的基本知识内容列前,“学习指导”列后。为了突出基本内容,每章内容的正文、例题、习题都以不同字体排印,重要概念标以黑体,每节都列出关键词、练习题。每章末都配有习题。值得一提的是,每章末还附有需要读者自己完成的小结,主要是通过一些简单的方法引导读者自己回顾本章主要概念、公式、定理和方法等,使读者在巩固所学知识的同时,逐步掌握

自主学习的方法和技能,学会学习.

本书编写分工如下:黎诣远:每章引子.各分册为:

《网络课程学习指南》分册.顾静相:预备知识;陈卫宏、胡新生:使用说明;黎诣远:专题讲座.

《微积分》分册.微分学部分:陈卫宏:第1章;赵坚:第2章;顾静相:第3章;周永胜:第4章.积分学部分:李林曙:第1、2章、第3章3.1~3.5节;陈卫宏:第3章3.6节.

《概率论与数理统计》分册,张旭红:第1、4、5章;冯泰:第2、3章.

《线性代数》分册,顾静相:第1章;赵坚:第2章;张旭红:第3章.

顾静相和冯泰作过初稿审阅,全书由李林曙和黎诣远总纂定稿.

需要特别指出的是,本套教材的编写是在高等教育出版社出版的《经济数学基础》(黎诣远主编、李林曙副主编)和《跟我学经济数学》(李林曙等编著)的基础上进行的,自始至终得到本课程强大的专家学者组的大力支持和直接参与,他们是:

经济数学基础学术咨询委员会:

乌家培 国务院学位委员会应用经济学科第二、三、四届评议组成员 国家信息中心专家委员会名誉主任 中国数量经济学会名誉理事长 中国信息经济学会名誉理事长 中国信息协会副会长

张恭庆 第三世界科学院院士 中国科学院院士 北京大学教授

李子奈 教育部经济学学科教学指导委员会委员 中国数量经济学会副理事长兼高等院校专业委员会主任 北京市经济学总会副会长 清华大学教授、经济系主任、中国经济研究中心主任

叶其孝 北京理工大学教授 全国大学生数学竞赛组委会副主任

胡显佑 中国人民大学教授 北京经济数学学会理事长

施光燕 大连理工大学教授 中央广播电视大学经济数学基础主讲

柳重堪 北京航空航天大学教授 中央广播电视大学高等数学主讲

经济数学基础教学设计顾问委员会:

任为民 中央广播电视大学教授 教育部现代远程教育专家委员会委员 教育部现代远程教育资源建设专家组成员

孙天正 中央广播电视大学教授

乌美娜 北京师范大学教授 教育部全国教师教育课程资源专家委员会委员 教育部现代远程教育工程资源建设基础项目专家

文 丽 北京大学教授 中央广播电视大学高等数学主讲

于 琛 人民教育出版社编审 课程教材研究所研究员

赵建华 北京气象学院大气视听研究所

教材审定专家:

施光燕 大连理工大学教授 中央广播电视大学经济数学基础主讲

叶其孝 北京理工大学教授 全国大学生数学竞赛组委会副主任

胡显佑 中国人民大学教授 北京经济数学学会理事长

孙天正 中央广播电视大学教授

柴全战 辽宁广播电视大学教授 教学处长

胡秀珍 天津广播电视大学副教授 理工处处长

张旭辉 南宁广播电视大学副教授

高等教育出版社的杨祥、张爱和、文小西、高尚华、郭思旭、郑洪深、胡凯飞、薛春玲等,为教材质量把关付出了辛勤劳动.

在此一并向他们表示衷心感谢.

本套教材的各模块及其组合可供各类高等院校经济与管理学科各专业根据其教学需要自行选用.对于广大经济管理工作者来说,本书在充实数学知识和掌握定量分析方法上也大有裨益.

毕竟,在我国开展网上教学还是刚刚起步,如何处理网络课程与文字教材之间的关系,我们还缺乏经验,加之作者水平所限,书中难免有不当之处,敬请使用本教材的师生和其他读者,毫无保留地提出批评和建议,以期及时修正.

编 者

2003年9月于北京

目 录

第 1 章 数据处理	2
1.1 总体和样本	3
1.2 重要的特征数(1)——平均数	4
1.3 重要的特征数(2)——方差和标准差	12
1.4 频数分布表和频数直方图	18
1.5 频率直方图	25
习题 1	28
跟我学	30
第 2 章 随机事件与概率	51
2.1 随机事件与概率	51
2.2 事件的关系与运算	55
2.3 古典概型与概率的性质	60
2.4 概率加法公式	65
2.5 条件概率与概率乘法公式	69
2.6 事件的独立性	75
2.7 全概率公式	78
习题 2	81
跟我学	84
第 3 章 随机变量与数字特征	110
3.1 随机变量概念	110
3.2 离散型随机变量	114
3.3 连续型随机变量	119
3.4 正态分布	123
3.5 分布函数与函数的分布	131
3.6 数学期望	137
3.7 方差	143
3.8 n 维随机变量	148
习题 3	168
跟我学	171
第 4 章 参数估计	216
4.1 总体、样本、统计量	216
4.2 抽样分布	218
4.3 参数的点估计	225
4.4 区间估计	233
习题 4	239
跟我学	241

第 5 章 假设检验	259
5.1 假设检验问题	259
5.2 正态总体的假设检验	264
5.3 一元线性回归分析	271
习题 5	278
跟我学	280
练习与习题答案	300
附表 1 标准正态分布数值表	312
附表 2 t 分布表	313
附表 3 χ^2 分布的上侧临界值表	315
附表 4 F 分布上侧分位数表	318
参考文献	328

引子

平均数“骗人”

2000年,全世界总人口60亿平均每平方公里45人,可是,新加坡超过6000人,蒙古则不到2人.

2000年,我国居民储蓄存款余额64332亿元,平均每人5082元.可是,全国农户年纯收入超过5000元者仅占7.45%,实际上超过80%的储蓄集中在不到20%的人手里.

2000年,我国许多省市上市公司的平均利润额都是上升的,可是,除了一两个盈利大户外,这些省市绝大多数上市公司的利润额都是下降的,甚至是亏损的.

据此,有人认为:“平均数最能骗人!”

实际上,平均数又最能服人:

2000年,我国国内生产总值89404亿元,早已进入世界前10名,可是,人均只有7078元,排在100名以外.

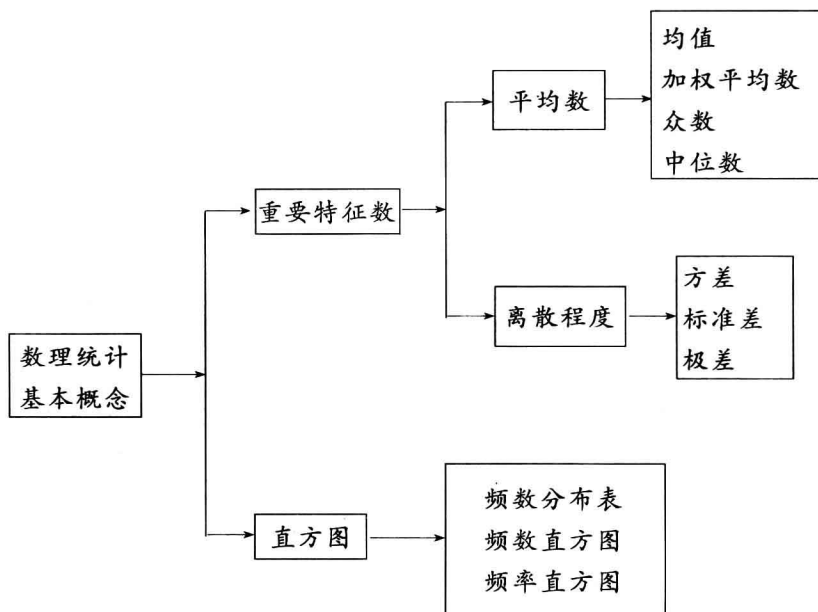
2000年,我国钢产量1.285亿吨,居世界第一,可是,人均产量不到100公斤,只是我们邻居韩、日两国的1/8左右.

可见,根据经济分析的不同目的,可以使用不同的数据指标:说明一国经济实力,用国民财富总量为宜;说明一国发展阶段,用人均国民总收入为宜;说明一国发展速度,用国内生产总值增长率为宜;说明一国经济效益,用人均劳动生产率为宜;说明一国收入分配,以基尼系数为宜.

为此,要全面描述一个总体的各种特点,必须使用不同方法处理不同数据.平均数只是从平均的意义上描述一个总体的状态,就看你使用是否得当.数据指标没有好坏,使用方法却分优劣.

第1章 数据处理

本章知识结构



7

在实际工作中,为了解某方面的情况,或对某些问题作出判断,经常要搜集许多数据.这些数据表面看起来参差不齐,大小不一,但经过一系列的统计方法处理后,就会呈现出一定的统计规律性,为我们认识 and 解决问题提供依据.

数据处理就是对搜集到的数据进行归纳、整理和简化的过程,基本方法有三种:

1. 计算 计算统计数据的一些重要特征数,如均值、方差等;
2. 列表 对数据分组列表,找出它们的统计规律性,如频数分布表等;
3. 作图 给数据以直观的表现形式,如频数分布直方图、频率分布直方图及频率密度曲线等.

对统计数据进行上述处理后,根据数据所反映出来的规律性,就可以对总体进行粗略的定性分析.

下面先介绍统计中常用的两个概念——总体和样本.

1.1 总体和样本



3.1.1

为了说明总体和样本的概念,先看下面几个例子.

例 1 要了解 A 城市居民 1997 年的年收入情况,一般不会花费很多人力物力去一一调查,而是采取抽样调查的方法,即抽查该城市一小部分居民的收入情况.例如抽取 1 000 个人,统计他们 1997 年的年收入,由此推断该城市居民的年收入状况.

例 2 为检验一批钢筋的质量是否合格,从中任意抽取 2 根钢筋,进行拉力和冷弯试验.

例 3 测试某品牌电视机的开箱合格率,从一批产品中任意抽取 3 台开箱测试.

上面这些例子都有一个共同的特点,就是为了研究某个对象的性质,不是一一研究对象包含的所有个体,而是只研究其中的一部分.通过对这部分个体的研究,推断对象全体的性质.这就引出了总体和样本的概念.

我们将所研究对象某一数量指标的全体称为**总体**,而组成总体的基本单位称为**个体**.从总体中抽取出来的个体称为**样品**,若干个样品组成的集合称为**样本**,一个样本中所含样品的个数称为**样本容量**(或大小).由 n 个样品组成的样本用 $x_1, x_2, \dots, x_n$ 表示.

例 1 中 A 城市居民 1997 年的年收入,是我们研究的对象,就是总体.城市中每一个居民 1997 年的年收入,就是一个个体.从总体中随机抽取(所谓随机抽取,就是指每个个体被抽取的机会都是相等的)出来的一个城市居民的年收入,就是一个样品.所抽取出来的 1 000 个城市居民的年收入,就组成一个样本,这个样本的容量是 1 000.

例 2 中要检验的那批钢筋的质量就构成一个总体,其中每一根钢筋的质量就是个体.从总体中随机抽取出来的一根钢筋的质量就是样品,所抽取出来的 2 根钢筋的质量就组成了一个样本,样本的容量是 2.



1. 例 3 中的总体和样本各是什么?

由于样品所表示的是某些特性的数量指标(如某人的年收入)以及样品抽取的随机性,因此样品是一个变量,我们所看到的都是样品的取值.如例 1,抽取样品的时候,并不知道这个人的年收入是多少(也就是个变量),抽取出来后,知道这个人的年收入是 9 600 元(样品的取值).我们将样品的取值称为**样品值**,样本的取值称为**样本值**,也称为**样本数据**.本书后面涉及的数据指的都是样本数据.

在实际问题中,总体所含个体的数目可以是有限的,也可以是无限

的. 在研究总体的性质时, 往往由于种种原因, 诸如总体里个体的数目无限, 或个体数目有限但数目太大, 以及试验带有破坏性或试验费用昂贵等等, 使得我们不能对总体的全部个体一一研究, 只能从总体中抽取一部分个体(即样本)进行观测, 通过研究所获得的样本值去推断总体的情况. 可见, 样本在统计分析中具有非常重要的意义.

正是由于很多总体的性质都要通过样本来推断, 因此样本的抽取非常重要. 限于篇幅, 关于样本的抽取方法本书不多涉及, 有兴趣的读者请参考有关书籍.

本节关键词

总体 个体 样品 样本 样本容量 样品值 样本值

练习 1.1

说出下列问题里的总体、个体、样品、样本、样本容量.

1. 为了解某城市居民 1997 年的年消费支出情况, 随机调查了 1 000 户居民的年支出费用.
2. 某日光灯生产厂要了解一批日光灯的质量, 从中抽出 100 只灯管进行测试.
3. 某盐业公司用自动包装机装食盐, 为了解机器的生产状态是否正常, 从一批产品中抽取 100 袋食盐进行重量检测.
4. 调查 C 城市家庭主要家用电器(彩色电视机、冰箱、收录机、洗衣机)的拥有量, 随机访问了 200 户家庭.

1.2 重要的特征数(1)——平均数

面对收集到的一批样本数据, 如何归纳、整理、分析它们, 以推断总体的性质呢? 计算样本数据的特征数是一个很重要的方法. 一般将能够反映统计数据主要特征的数, 称为**统计数据的特征数**(简称**特征数**). 实际上, 数理统计的核心内容就是通过数据的特征数来对总体的数字特征作出估计, 以及对总体的其他特性进行分析和推断.

数据分析中最常用的特征数可分为两类: 一类是表示数据总体水平的数, 包括均值、加权平均数、几何平均数、中位数和众数等, 它们统称为**平均数**; 另一类是表示数据分散程度的数, 常用的有方差、标准差、极差和变异系数等.

1.2.1 均值

定义 1.1 给定一组数据 $x_1, x_2, \dots, x_n$, 称

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.2.1)$$

为数据 $x_1, x_2, \dots, x_n$ 的均值^①.

均值就是通常所说的算术平均数, 是反映数据整体水平的特征数. 实际问题中, 经常用样本的均值来估计总体的均值, 或用均值代表总体水平, 对不同的总体进行比较.

例 1 商场销售一种新产品, 统计前 5 天的销售量, 分别为 (单位: 个):

38 42 36 45 39

求这个新产品平均每天的销售量是多少?

解 根据均值的计算公式得

$$\begin{aligned} \bar{x} &= \frac{1}{5}(38 + 42 + 36 + 45 + 39) \\ &= \frac{1}{5} \times 200 = 40 (\text{个}) \end{aligned}$$

即这种新产品平均每天销售 40 个.

根据均值的计算公式, 可以推出关于均值的两个性质.

性质 1

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \quad (1.2.2)$$

! 数据 $x_1, x_2, \dots, x_n$ 的偏差 $x_i - \bar{x}$ 之和等于 0.

证明

因为

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

即

$$\begin{aligned} n\bar{x} &= \sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n \\ &= (x_1 - \bar{x}) + (x_2 - \bar{x}) + \dots + (x_n - \bar{x}) + n\bar{x} \end{aligned}$$

所以

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

性质 2 任给一个常数 c , 总有

$$\sum_{i=1}^n (x_i - c)^2 \geq \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.2.3)$$

等号仅在 $c = \bar{x}$ 时成立.

^① Σ 表示对数据求和, 如 $x_1 + x_2 + \dots + x_n$ 就可以表示成 $\sum_{i=1}^n x_i$, 即 $\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$.

这个性质的证明留作习题.

在实际工作中,均值对于作出判断、进行预测和制定计划都有十分重要的意义.请看下面的例子.

例2 某商场统计了1995—1997连续三年羊毛衫的销售情况,见表1-1.根据统计资料,分别计算三年来同月的平均销售量.

解 三年里1月份的平均销售量是

$$\frac{1}{3}(3\ 635 + 3\ 827 + 4\ 265) = \frac{1}{3} \times 11\ 727 = 3\ 909 \text{ (件)}$$

2月份的平均销售量是

$$\frac{1}{3}(3\ 298 + 3\ 478 + 3\ 721) = \frac{1}{3} \times 10\ 497 = 3\ 499 \text{ (件)}$$

同理,依次算出12个月的月平均销售量,见表1-1的最后一列.从表中看到每年12月和1月的销售量最多,6月和7月的销售量最少.掌握了这一变化规律,商场可以合理地组织货源,有效地利用资金.

表1-1 羊毛衫销售量统计表

单位:件

	1995 年	1996 年	1997 年	三年合计	月平均数
1 月	3 635	3 827	4 265	11 727	3 909
2 月	3 298	3 478	3 721	10 497	3 499
3 月	2 489	2 586	2 460	7 535	2 511
4 月	1 534	1 657	1 706	4 897	1 632
5 月	692	687	621	2 000	667
6 月	297	265	303	865	288
7 月	152	168	148	468	156
8 月	465	504	522	1 491	497
9 月	1 896	2 085	2 011	5 992	1 999
10 月	2 386	2 648	2 539	1 589	2 524
11 月	2 789	2 674	2 844	8 307	2 769
12 月	3 464	3 634	3 960	11 058	3 686
全年	23 097	24 213	25 100	72 410	24 137

? 2. 每年的月平均销售量如何计算? 三年的月平均销售量如何计算?

均值是平均数中最基本的形式,是根据全部统计数据科学地抽象出来的一种平均指标.它具有计算简单、代表性强的特点,是研究社会经济现象中的数量关系时最常用的指标,在统计分析中有着广泛的应用.

1.2.2 加权平均数

计算一组数据的均值时,若考虑数据中各数据的出现次数,或权衡数

据的作用程度,则需引入加权平均数的概念.先看一个例子.

例 3 某小学一年级二班共有学生 40 名,期末数学考试的成绩见表 1-2,计算该班学生的数学平均成绩.

 例 2

表 1-2 某小学一年级二班期末数学考试成绩

单位:分

100	100	98	99	100	98	99	96	95	99
98	100	99	98	95	100	100	99	94	100
100	99	98	100	97	96	90	100	100	96
99	100	98	100	94	100	99	96	94	97

该班学生的数学平均成绩可以按公式(1.2.1)

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

计算,但这样计算是很麻烦的.我们采用一种简便的办法,即对这组数据归纳整理,将相同的数据划分在一起,列出表 1-3.

表 1-3 某小学一年级二班数学考试成绩归纳表

成绩/分	100	99	98	97	96	95	94	90
个数	14	8	6	2	4	2	3	1

于是该班学生的数学平均成绩为

$$\begin{aligned}\bar{x} &= \frac{1}{40}(100 \times 14 + 99 \times 8 + 98 \times 6 + 97 \times 2 \\ &\quad + 96 \times 4 + 95 \times 2 + 94 \times 3 + 90 \times 1) \\ &= \frac{1}{40} \times 3\,920 = 98(\text{分})\end{aligned}$$

为了使这个式子的统计意义更加清楚,将它改写成以下形式:

$$\begin{aligned}\bar{x} &= 100 \times \frac{14}{40} + 99 \times \frac{8}{40} + 98 \times \frac{6}{40} + 97 \times \frac{2}{40} \\ &\quad + 96 \times \frac{4}{40} + 95 \times \frac{2}{40} + 94 \times \frac{3}{40} + 90 \times \frac{1}{40}\end{aligned}$$

上式可认为是 8 个数 100, 99, 98, 97, 96, 95, 94, 90 分别乘以 $\frac{14}{40}, \frac{8}{40}, \frac{6}{40}, \frac{2}{40}, \frac{4}{40}, \frac{2}{40}, \frac{3}{40}, \frac{1}{40}$ 之后的和,这就是我们下面要讨论的“加权平均数”.

 这 8 个数恰是它们各自在全部 40 个数据中所占的“比重”.

一般地,求 $x_1, x_2, \cdots, x_n$ 的均值时,如果 $x_i (i = 1, 2, \cdots, n)$ 中只有不同的 k 个值 $a_1, a_2, \cdots, a_k$ 出现,并且 a_j 出现 n_j 次, $j = 1, 2, \cdots, k, (\sum_{j=1}^k n_j = n)$,