

对外经济贸易大学
国际工商管理学院 MBA 系列教材

数据、模型与决策

贾怀勤 主编

对外经济贸易大学出版社

对外经济贸易大学
国际工商管理学院 MBA 系列教材

数据、模型与决策

贾怀勤 主编

对外经济贸易大学出版社

(京)新登字 182 号

图书在版编目(CIP)数据

数据、模型与决策/贾怀勤主编. —北京:对外经济贸易大学出版社,2003
ISBN 7-81078-316-5

I. 数… II. 贾… III. 工商行政管理—决策模型 IV. F203.9

中国版本图书馆 CIP 数据核字(2003)第 122303 号

© 2004 年 对外经济贸易大学出版社出版发行

版权所有 翻印必究

数据、模型与决策

贾怀勤 主编

责任编辑:宋志红

对外经济贸易大学出版社

北京市朝阳区惠新东街 12 号 邮政编码:100029

网址:<http://www.uibep.com>

山东省莱芜市圣龙印务书刊有限责任公司印装 新华书店北京发行所发行

成品尺寸:185mm×230mm 34 印张 661 千字

2004 年 2 月北京第 1 版 2004 年 2 月第 1 次印刷

ISBN 7-81078-316-5/F·192

印数:0001—5000 册 定价:58.00 元

总 序

我国社会主义市场经济的迅速发展,对我国的工商管理教育提出了新的要求。如何更好地为社会培养具有扎实的理论基础,熟悉市场经济运行规则,了解现代工商管理活动的理念、规范、方法和技能,具有决策能力、创新意识和开拓精神的工商管理人才,是摆在我国管理教育界面前的一个紧迫问题。工商管理硕士(Master of Business Administration, MBA)教育是以美国为代表的西方国家普遍采用的培养高层次管理人才的主要方式,而我国从20世纪90年代初才正式开始MBA教育的试点。今天,以MBA教育为核心的各种形式、各种层次的工商管理教育和培训在我国方兴未艾,社会对工商管理知识和人才的需求仍在膨胀。

对外经济贸易大学于1982年成立国际企业管理系,在我国率先引进了美国工商管理教育的课程体系,研究生和本科生多数专业课程采用英文原版教材,并用英文授课。1988年,美国纽约州立大学董事会在对我校企业管理研究生专业的课程体系、教学质量以及学生质量进行严格评估的基础上,承认我校的企业管埋研究生相当于美国的工商管理硕士,并颁发了证书。我校也是我国第二批MBA教育的试点院校之一。多年来,我们走的是一条引进、消化、吸收的道路。在国际交流合作中,采用请进来和派出去的方式,培养了一批了解国际上MBA发展前沿、能够胜任MBA教学的骨干教师,在学科建设、教学和科研方面取得了丰硕成果。

作为国内较早开展工商管理教育的高校,我们有义务也有责任为我国的现代工商管理教育作出更大的贡献。为此,对外经济贸易大学国际工商管理学院经过认真策划,编写了这套“工商管理硕士课程系列教材”。我们希望通过这套教材将我们在多年教学和科研实践中的经验和成果展现给广大读者,为MBA学生和希望学习、了解MBA课程的企业管理人员以及其他人士提供最新的工商管理理念、方法和技能。

通过对我国多年工商管理教育的认真审视,结合我校工商管理教育的特色和课程体系结构,我们确定了20多个选题,涵盖了国家规定的MBA培养方案中的核心课程,同时也兼顾到课程体系的系统性与协调性、中国国情与国际可比性等问题。这套教材的编写者都是在我校多年从事MBA教学的教师,多数人有在国外长期学习和工作的经历,有着

丰富的 MBA 教学经验和资料的积累。在教材的编写过程中,他们本着理论联系实际这一 MBA 教学的基本要求,追求教材内容的实用性和可操作性,将他们多年在教学过程中积累和编写出的案例用于教材的编写,使教材的内容更加充实。他们还力求教材内容难易适度,图文并茂,可读性强。我相信,通过教材编著者的辛勤耕作和对外经济贸易大学出版社的大力支持,这套教材会以崭新的内容和形式呈现在广大读者面前。

我们编写这套 MBA 教材的另一个目的是积极寻求与国内兄弟院校的合作与交流。我们恳请国内同行和此套教材的使用者提出建议和批评,以便我们改进。

对外经济贸易大学副校长

徐子健

2000 年 6 月 30 日

前 言

本书是对外经济贸易大学工商管理学院策划编写的工商管理硕士(MBA)课程系列教材之一,体现贾怀勤教授和他的同事们对MBA教育的一点贡献。

把《应用统计学》和《运筹学》合为一门《数据、模型与决策》,是20世纪90年代北美各大学商学院在数量方法课程方面的一次重大变革,其目的是使数量方法课程更贴近工商管理的实践应用。在中国,自1998年全国MBA指导委员会主干课程研讨会提出把“应用统计学”和“运筹学”合并开设“数据、模型与决策”的建议以来,各高校的商学院都开始了包括教材编写在内的课程建设。

贾怀勤教授在本课程建设之初,即制定了明确的建设规划,将建设工作的重点放在课程内容体系的创新和教学方法的创新上。在本课程建设的实践中,创新主要体现在如下几个方面:

首先,本课程以工商管理中所接触到的数据如何通过模型以用于决策为导向,从原有两门课程各自的知识体系中跳了出来,重新建立起认知体系。

其次,本课程的知识展开不是沿固有的“概念——定理——公式——算例——应用”的模式,而是本着贴近应用的原则,以案例切入,提出问题,展开讨论,引出解决方法,归纳原理,提升认识,举一反三,解决问题。

再次,本课程所用案例,皆由工商管理实践中取得,或涉及生产运作和质量管理,或涉及财务管理和投资决策,或涉及营销调研和广告,或设计人力管理和企业组织,等等。

最后,本课程的模型和各种管理问题的讨论,都从电子表格数据出发,在电子表格平台上展开,运用EXCEL强大而便捷函数运算功能,实现数据处理、建模、求解,并由输出结果导出有用的结论。

这是一个逐步积累的过程,而非突变。这当中已经有了一定的经验总结,反映在贾怀勤教授的两篇教学研究论文中:《CAI在MBA教育中的应用》(1999)^①和《〈数据、模型与

^① 载入全国现代教育技术研究会《现代教育技术研究与应用》论文集,高等教育出版社,1999.12。

决策》教学与案例运用》(2001)①。前一篇主要总结如何用 EXCEL 做数据计算,后一篇集中总结举行案例作业和课堂讨论以至进行讲评的经验。经过几年的教学活动,我们的教学思路、教学方法、教学资料 and 教学辅助手段都渐趋成熟,教材也就瓜熟蒂落可以付梓出版了。

《数据、模型与决策》这部教材,在内容安排上符合全国 MBA 指导委员会决定设立“数据、模型与决策”课程时提出的“数理统计的内容约占三分之二,运筹学的内容约占三分之一”②的指导思想,同时又以数据——模型——决策为主线,将各部分内容有机地串起来。在数据处理和计算手段上突出 EXCEL 对于各种分析方法的实现,相关的 EXCEL 运行展示,视所用软件工具在 EXCEL 中的编挂方式,或留在各章正文之后,或渗入到各章正文之中。本教材一改以往每章后编有容易“对号入座”的计算题的惯例,只在书后提供 46 个案例,由教师根据不同教学内容和进度的需要选用,既可以安排学生做与某部分教学内容联系较为直接的分析,也可以做综合性分析。

在课程建设中,重点参考了以下几部美国教材:

Charles F. Tonini, Warren H. Hausman and Howard Bierman, Jr., *Quantitative Analysis for Management*, McGraw-Hill, 1997. (斯坦福大学教授编著,机械工业出版社译本,1999 年)。

Frederick S. Hiller, Mark S. Hiller and Gerald J. Lieberman, *Introduction to Management Science*, McGraw-Hill, 1999. (斯坦福大学教授编著,中国财政经济出版社译本,2001 年)。

Dimitris Bertsimas, Robert M. Freund, *Data, Model & Decision: The Fundamentals of Management Science*, Thomson Duxbury, 2000. (麻省理工学院斯隆管理学院教授编著,中信出版社译本,2002 年)。

S. Christian Albright, Wayne L. Winston and Christopher Zappe, *Data Analysis & Decision Making with Microsoft Excel*, Thomson Brooks/Cole, 2003. (印地安那大学教授编著,无中译本)。

此外,我们还观摩了美国马里兰大学 G. Anandalingam 教授在我校中美管理学院“Data Analysis and Decision Modeling”课程的授课。

以上所提的教材参考和课程观摩对本教材的编写无疑是很有启发的。在此我们谨向上述教授致谢。

① 对外经济贸易大学《高等教育问题研究》,2001 年第 2 期。

② 复旦大学管理学院郑祖康:《〈数据、模型与决策〉教学大纲》。

本教材正文后 46 个案例及正文中的例题,少数取材于公开出版的书刊,多数是编著者和 MBA 研究生调研所获。编著者衷心感谢案例资料提供者对于编写这部教材的贡献。

第十一章至第十三章由杨恭辅副教授执笔,第八章由王玉荣副教授执笔,第一至八章的附录——关于 EXCEL 的使用由唐振宇执笔,贾怀勤教授全面策划、总纂,并负责除以上提及的其他部分的编写。

编著者

2003 年 9 月于惠园

目 录

第一章 数据的意义、分类和来源	(1)
第一节 数据、模型与决策的联系	(1)
第二节 数据的分类	(2)
第三节 数据的获取途径和方法	(6)
第四节 数据获取的组织方式	(8)
EXCEL 应用一	(10)
第二章 数据特征的展示和概括	(17)
第一节 原始数据的有序化:频数分布数列	(17)
第二节 数据特征的图形展示	(28)
第三节 数据趋势的量数	(37)
第四节 数据特征的探索工具	(48)
EXCEL 应用二	(53)
第三章 数据推理的理论基础:概率与概率分布	(68)
第一节 概率	(68)
第二节 常用的离散型概率分布	(84)
第三节 典型的连续型概率分布	(92)
第四节 随机数据的生成和模拟决策	(100)
EXCEL 应用三	(109)
第四章 数据推理的演绎与归纳:从抽样分布到参数估计	(113)
第一节 抽样分布的概念	(114)
第二节 平均数的抽样分布(大样本):中心极限定理	(120)
第三节 平均数的抽样分布(小样本):学生氏 t 分布	(122)
第四节 参数估计的原理	(125)
第五节 区间估计的应用	(129)
第六节 必要样本容量的确定	(135)

EXCEL 应用四	(137)
第五章 关于总体数据特征的比较:假设检验	(140)
第一节 假设检验的基本理念	(140)
第二节 关于总体平均数的假设检验	(145)
第三节 关于总体比率的假设检验	(157)
第四节 关于总体方差的估计和假设检验	(160)
第五节 两总体方差间差异的估计和假设检验	(165)
EXCEL 应用五	(168)
第六章 涉及数据分析的其他假设检验问题	(174)
第一节 多总体平均数齐一性检验:方差分析	(174)
第二节 交叉分析中的独立性检验	(184)
第三节 数据分布的拟合检验	(189)
第四节 数据的随机性检验	(194)
EXCEL 应用六	(195)
第七章 回归模型应用:由解释数据推测响应数据	(201)
第一节 简单相关和回归	(201)
第二节 简单回归中的统计推断	(215)
第三节 多元回归	(218)
第四节 回归建模中的专门技术	(235)
EXCEL 应用七	(248)
第八章 时间数列模型应用:由历史数据预测未来数据	(254)
第一节 时间数列研究的朴素方法	(254)
第二节 时间数列的解析研究方法	(261)
第三节 时间数列的预测	(280)
EXCEL 应用八	(292)
第九章 最优化决策(I):线性规划的基本模型及其应用	(295)
第一节 线性规划的一般模型	(295)

第二节 线性规划求解和敏感性分析	(299)
第三节 线性规划的应用	(315)
第十章 最优化决策(Ⅱ):线性规划的特殊模型和扩展模型的应用	(322)
第一节 运输问题	(322)
第二节 指派问题	(327)
第三节 进度计划问题	(331)
第四节 整数规划	(334)
第五节 非线性最优化	(338)
第十一章 统计决策(Ⅰ):先验分析	(347)
第一节 统计决策及其要素	(347)
第二节 不确定型决策方法	(348)
第三节 风险型决策方法	(355)
EXCEL 应用九	(363)
第十二章 统计决策(Ⅱ):后验分析与预后验分析	(366)
第一节 使用附加信息的决策:后验分析	(366)
第二节 附加信息的价值:预后验分析	(374)
第十三章 统计决策(Ⅲ):效用评价标准和连续型自然状态的决策	(389)
第一节 以效用为损益值的决策	(389)
第二节 连续型自然状态的决策	(397)
案例	(410)
1. 北京市场上饼干价格与销售量	(410)
2. 滨河市企业职工基本数据特征	(414)
3. 航空客机配餐管理	(416)
4. 中央银行城镇储户抽样调查	(417)
5. 国产与国外原产助听器集成线路质量测试比较	(418)
6. 美国用户汽车保有量与其社会经济特征的关系	(420)
7. 对日红小豆出口市场调研	(422)

8. 泛世通防滑轮胎促销	(423)
9. 捷利物流油罐车购置	(427)
10. 德士古新炼油厂的选址	(428)
11. 鹏腾制衣有限公司赊销决策	(431)
12. 轻轨路基预制板质量管理	(433)
13. 中储总公司 2002 年上半年仓储业务	(436)
14. 影响钢材销售的因素分析	(438)
15. CHW 饭店员工意见调查	(440)
16. CCTV-3《艺术人生》改版收视调查	(442)
17. 新可乐的开发与味感测试	(443)
18. 固定电话和移动电话用户数量	(446)
19. 上市公司股票数据	(448)
20. 洗涤用品的满意度调查	(451)
21. 北京市 2001 年土地出让价格	(456)
22. 北京铁路局客货运形势趋好	(461)
23. 离合器片的使用寿命检测	(464)
24. R130 挖掘机的销售价格	(465)
25. 华电集团钢结构事业部的项目分配	(466)
26. 山海旅游公司的旅行车购置规划	(467)
27. 燕京电机公司的汽轮机订单承接能力	(468)
28. 高压线圈铝箔厚度检验	(469)
29. 芭芭娃娃增值分析	(471)
30. 鸿基桥梁厂的轨枕外销	(472)
31. 英国建筑业营业额的季节变化	(474)
32. 企业规模与税收偷逃	(475)
33. 防洪堤建设方案决策	(477)
34. 统一食品有限公司	(479)
35. 消费者调查有限公司	(481)
36. 质量联盟有限公司	(483)
37. 长城重型机械公司	(485)
38. 天广线进口日本电缆质量索赔	(486)
39. 戈龙公司的奖励计划	(487)

40. 电力公司应收电费账户的抽样审计·····	(494)
41. 大理石石材光洁度标准的检验·····	(496)
42. 运用本德福法则选取审计账户·····	(498)
43. 电台调试工人的技术培训·····	(500)
44. 三家出口供货商的选取·····	(502)
45. 中国邮政业务单式集中采购的配货方案·····	(504)
46. 泰民瓶装杀蟑粉剂是否不足秤? ·····	(508)
47. 西单商场 18 个服装品牌的销售因素 ·····	(509)
48. 直升机旋转翼梁制造工艺的选择·····	(510)
统计数表 ·····	(511)
附表 1 二项分布累积概率表·····	(511)
附表 2 普哇松分布累积概率表·····	(514)
附表 3 标准正态分布·····	(519)
附表 4 随机数字 4 000 个 ·····	(521)
附表 5 t 分布 ·····	(525)
附表 6 χ^2 分布分位数 $\chi^2_{1-\alpha}(n)$ 表 ·····	(527)
附表 7 F 分布:显著水准 1% ($\alpha=0.01$) 的临界值 ·····	(529)
附表 8 F 分布:显著水准 5% ($\alpha=0.05$) 的临界值 ·····	(530)

第一章 数据的意义、 分类和来源

第一节 数据、模型与决策的联系

决策,是指从思维做出决定的过程。从管理学的角度说,决策可以表述为:为了达到某一预定目标,在掌握充分、必要的数据的前提下,按照一定的价值评判标准,运用逻辑和数学推理的方法,对几种可能采取的方案做出合理的选择。在企业管理中,无论是上层、中层,还是基层,无论是生产运作、营销,还是人力资源管理、财务管理,都需要决策。诸如是否要推出一个新品种,是否要进入一个新市场,选择哪一套物流路线,如何对新招聘的员工进行培训,等等。只不过不同层次、不同部门的决策机会或者决策权力有不同,而且人们通常把一些决策不称为“决策”,而称为“决定”。

对有些决策问题,决策者可以根据其经验和判别力做出决断,而许多管理决策问题需要决策者从描述与决策目标有关联的事物和现象状态、规模、变化趋势的数值出发,按照一定的思维方式和思维方法找出方案的评价值,再进行方案的选择。前一种情况是定性决策,后一种情况是定量决策。本书研究的是进行定量决策的科学方法。

数据,就其词义而言,指进行各种统计、计算、科学研究和技术设计等所依据的数值。英文里 data——作为 datum 一词的复数形式,有如下涵义:①论据,作为论据的事实;②资料;③数据。在本书中,“数据”作为集合名词出现,构成数据的单个数字成为“数值”。数据是决策“所依据的数值”,在企业管理决策中,它们表现为各种资料,按存在空间分为企业内部资料和外部资料,按资料的特征分为数据型资料和文字型资料,按其时间属性分为往期资料、当期资料和预测资料。资料的准确性对于决策是至关重要的,从虚假的数据出发,你很难做出正确的决策,“歪打正着”在复杂的决策中能够奏效的可能性几乎不存在。19 世纪美国幽默家亚特姆斯·沃德(Artemus Ward)说过的一句话一针见血地指出了虚假资料的危害:“办事不利非因无知,实因误知。”(It ain't the things we don't know that gets us in trouble. It's the things we know that ain't so.)

模型,是真实事物的人为再现,是它所代表的真实世界中对对应事物的概要复制。它略去了次要枝节,突出了主干,因而浓缩了问题的核心。在管理决策中用到的模型属于图解模型和数学模型。前者以图画形式浓缩决策所投入的要素之间的关系,后者以数学公式刻画决策所投入的要素与决策评价值之间的关系。在一定的意义上,两者又是相通的。前面关于决策的表述那段话已经指明了数据、模型和决策之间的关系。

决策作为一个过程,以数据为投入,以对不同方案的选择为产出,而模型是由数据到决断这一过程中所遵循的思维方式和思维方法。管理者在决策问题面前,有些投入是其能够控制的——通常是管理者组织内部的状况,另有一些投入是管理者所不能控制的——通常是其组织外部的环境,这就使得决策有了确定型与不确定型和随机型之分。

可控制投入、不可控制投入、模型和决策产出的关系如图 1.1 所示。

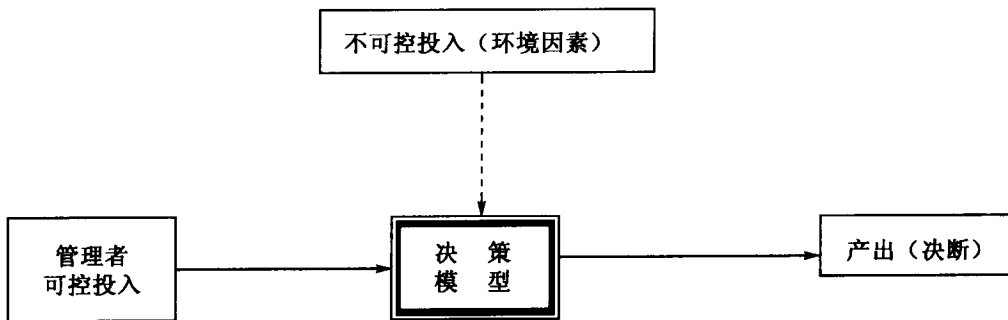


图 1.1 决策模型流程

构建决策模型是问题的核心。各种数学模型具体结构不同,但它们都是含有代表投入因素的变量(或常量)和代表产出结果的变量,并且按特定联系法则构成的数学表达式。建模就是抓住实际决策问题的主要环节,删繁就简,提炼出投入变量,确定他们与结果变量的数理关系,用数学表达式进行表述。一旦模型建成,即可投入已知数据,经运算得出结果,据其对方案做出选择。

第二节 数据的分类

数据是我们所研究的事物、现象的属性的具体描述。这里所谓的“描述”,包括定性的文字型描述和定量描述。如描述一家企业的从业人员的数值、营业收入的数值、资产规模的数值,都能构成数据;其主营业务,可以是化工产品生产、物流、勘探设

计、组合基金运作等,其经济体制可以是国有、集体、私营、股份制、外商投资等,也是数据。就对事物、现象的研究而言,对既存状况的具体描述是数据,对意向的描述和对未来的推测、计划也是数据。如一次市场调查取得的消费者在未来3个月、6个月、12个月、24个月欲购买私家轿车的数量,消费者对某款轿车的性能、款式、价格、销售服务等的评价分数,都属于数据。从不同的角度看问题,数据就有不同的分类方法。社会调查方法学使用四个等级的数据分类法,传统的统计学使用两种类型的数据分类法。

一、数据的四个等级

设有一批旅游者自北京出发,分别到我国东、西、南、北四个边境城市去旅游。这四个城市是普陀、喀什、三亚和漠河。为方便起见,给这四个地方编号,按东、西、南、北分别编号1、2、3、4。当旅游结束时,请每位旅游者都回答由下述四个问题构成的问卷:

问题1. 您去了哪个城市?

1. 普陀 2. 喀什 3. 三亚 4. 漠河

问题2. 与以往旅游活动总的感觉相比,您认为此次旅游

☐ 很满意 ☐ 较为满意 ☐ 无所谓 ☐ 不太满意 ☐ 很不满意

问题3. 您在此地时最热一天的最高气温是____℃。

问题4. 您此行的全部支出是_____元。

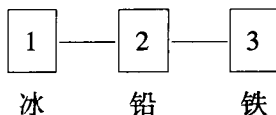
按照本问卷的设计格式,四个问题得到的答案都是数码。但是四组答案数码所反映的与去各地旅游的有关特征的含义都不相同。第一个问题的答案,或1,或2,或3,或4,纯粹是人为的城市代码,这些数码之间不存在任何数量关系。对于第二个问题,有1,2,3,4,5五种答案,它们不再仅仅是代表符号,而是反映旅游者满意程度的符号,数值较大的数码代表较大幅度的满意。第三个和第四个问题的答案都是通常意义上的数值,但两者之间仍有区别。我们可以说4 000元支出是2 000元支出的2倍,但是不能说40度是20度的2倍,最多只能说从0到20的间隔与从20到40的间隔相等。因为如果用数轴表示支出,0表示没有支出,这是自然的起点;而对于温度数轴,0并不表示没有温度,它只是为计量方便人为确定的原点。摄氏温度和华氏温度各有其不同的原点。上述四个问题的答案,向我们展示了数据的四个层次:定名数据、定秩数据、定距数据和定比数据。

定名数据(Nominal Data),是仅对事物、现象所属类别的具体描述。它们不表示顺序,没有多少大小之分。在人口调查时,可以用0代表男性,用1代表女性,也可以反过来用1代表男性,用0代表女性。前面曾经用1、2、3、4分别代表普陀、喀什、三亚、漠河,现在改为用1、2、3、4分别代表普陀、漠河、喀什、三亚,也没有什么关系,因为它们的顺序同我们所研究的问题没有联系。

定秩数据 (Ordinal Data), 是对事物、现象所具有的属性顺序的具体描述。它们描述属性的强弱, 测定结果按档次排列, 但是各档次之间的间隔是不确定的。对个体进行测定得一序列尺度数码, 不仅表明该个体与具有相同数码的个体属性程度相同, 还表明它与具有与之不同数码的其他个体属性程度的不同。本例中对普陀等四城市旅游的满意程度即是一组典型的定秩数据。

定距数据 (Interval Data) 的描述功能高于定秩数据, 它们不仅能表明事物、现象的强弱程度, 而且以确定的间距表明这种强弱程度。温度是典型的间距尺度。

如果仅说冰、铅和铁三者相比较, 冰最易熔化, 铅次之, 铁最难熔化, 就产生了一组定秩数据:



但是, 如果测定三种固体的熔化温度点, 则不仅能排出融化的难易顺序, 还可以确切标明三个熔点的间隔。这就是定距数据。三种固体的熔点在两条不同的数轴——摄氏温度计和华氏温度计上的读数如图 1.2 所示。

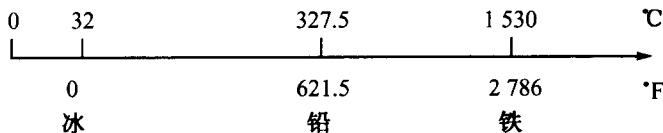


图 1.2 冰、铅、铁熔点读数

定距数据的数值间有确定的间距, 但是没有自然确定的原点, 换言之, 它的零点是人指定的。因此, 计算定距数据之间的比例没有意义。在上例中, 不能用 327.5 去除 1530 从而得出“铁的熔点是铅的熔点的 4.67 倍”的结论, 同样也不能用 621.5 去除 2786, 从而得出“铁的熔点是铅的熔点的 4.48 倍”的结论, 因为摄氏温度和华氏温度的零度都是人为规定的。

经济和工商管理中常见的定距数据的例子有本书后面将用到的效用 (Utility) 和统计指数。

定比数据 (Ratio Data) 是最高级的数据, 它不仅有确定的间距, 而且有自然确定的零点。用人数表示家庭规模, 用公里表示旅游路线长短, 用货币金额表示销售量, 等等, 都属于比例尺度。在这里, “有”与“无”的概念是明明白白的, 因此“零”值是公认的。

四种数据的描述功能是依次增大的, 相应地, 它们的可计算性也是依次增大的, 具体内容将在第二章中展开。由于定距数据使用的机会较少, 而且在许多场合可以同定比数