



高等教育自学考试中英合作商务管理专业与
金融管理专业 剑桥商务管理证书指定教材

数量方法

(数量方法课程用书)

全国高等教育自学考试指导委员会 组编

钱小军 主编



高等教育出版社

高等教育自学考试
中英合作商务管理专业与金融管理专业
剑桥商务管理证书指定教材

数量方法

(数量方法课程用书)

全国高等教育自学考试指导委员会 组编
钱小军 主编

高等教育出版社

内 容 提 要

本书是教育部考试中心与英国剑桥大学考试委员会共同合作的课程——商务管理、金融管理专业数量方法课程专业用教材,是按照课程考试大纲、参考国外相关教材并结合我国实际编写的。通过本书的学习,可使读者具有收集、概括和呈现数据的意识,能够理解基本的概率论原理,理解和运用基本的统计概念,理解关于统计推断的有关概念并对数据进行统计分析,识别两个变量之间是否存在线性关系并作出估计和预测,理解简单的时间数列模型并进行预测,学会运用指数综合分析和评价经济现象,以更好地适应商务管理和金融管理工作的需要。

前 言

为了适应中国市场经济的运行和发展的需要,给工商企业界人士及希望在商务管理和金融管理方向发展的人员提供具有国际标准的专业教育和资格认证,中国教育部考试中心和英国剑桥大学考试委员会在中国自学考试领域合作、共同设计“商务管理”和“金融管理”专业的部分课程。

合作课程的设计工作首先在专科层次进行,无论是否具有工作经验,只要愿意提高商务管理、金融管理方面的知识和技能,均可以参加到这个学习中来。合作课程采用模块式组合,灵活机动的学分制,为学员根据需求和兴趣自由选考课程提供了可能。

两个专业共有 19 门课程,公共基础课 4 门;大学语文、政治经济学、哲学、剑桥商务英语;共同专业基础课 7 门;企业组织与环境、商务交流、数量方法、经济学、会计学、信息技术管理、财务管理;专业课 4 门,商务管理专业的专业课为:商法、市场营销、人力资源管理 and 国际贸易实务;金融管理专业的专业课为:金融法、管理会计、财务报表分析和金融概论。在所有的 19 门课程当中,大学语文、政治经济学和哲学与国内自学考试中的课程相同,其余课程由双方专家共同设计。为了使课程具有国际标准,考试大纲经过反复商讨,分头编写,共同审订;教材采用合适的英文教材的中译本或者重新量体编写;为了配合学习,教学辅导材料则采用国外流行的学习包的形式。

合作课程得到了中英双方的承认。对于每模块(课程)学习合格的学员,教育部考试中心和剑桥大学考试委员会联合签发合作模块(课程)合格证书。对于取得指定的 8 门合作模块(课程)合格证书的学员,剑桥大学考试委员会将颁发“商务管理证书”。如果学员学完了 15 门课程取得规定的学分,还可以获得高等教育自学

考试的大专毕业证书。这两个专业的毕业生如愿自费留学英国，可以得到剑桥大学考试委员会的推荐。

为了设计这两个专业、翻译编写这一套教材和辅导材料，中英两国的一些专家多次研讨，精心筹备，倾注了大量的精力和心血。他们认真而又紧张的工作，圆满完成了任务。在此谨向他们致以衷心的感谢。

愿我们这项工作能够为愿意进入商界涉足管理的人们提供一个很好的起点。

教育部考试中心

1998年4月

编者的话

《数量方法》一书经过一年多的准备、写作和修改,终于完成并与读者见面了。这本书是高等教育自学考试中英合作商务管理专业和金融管理专业的专业课之一——数量方法课程——的教材。

数量方法课程是教育部考试中心和英国剑桥考试委员会合作设计的、中国高等教育自学考试商务管理、金融管理两个专业的必考课程之一,是为了培养考生具有对数据进行初等处理,正确理解和有效运用基本的统计方法,以及能够运用统计和数学方法进行经济和商业决策等能力而设置的一门专业基础课。

我们处在信息时代,信息时代充满数据,商务管理和金融管理工作更是每天需要处理大量的数据。为了能够对数据进行科学的分类、筛选和量化处理,以便从杂乱无章的数据中得出有助于决策的结论,并指导实践,必须掌握数量方法这门课程所讲述的知识。因此,数量方法是一门实用性很强的科学,它强调了理解和解释商业环境中会普遍遇到的统计问题所必需的基本知识。学习这门课程的目的不是使考生成为专业统计人员,它所讲述的内容不能代替常识和经验,但它能帮助考生评价和分析自己或他人的观点。

设置本课程的具体目的是:使自学应考者具有收集、概括和呈现数据的意识,能够理解基本的概率论原理,理解和运用基本的统计概念,理解关于统计推断的有关概念并对数据进行统计分析,识别两个变量之间是否存在线性关系并作出估计和预测,理解简单的时间数列模型并进行预测,学会运用指数综合分析和评价经济现象,为更好地适应商务管理和金融管理工作的需要作好必要的准备。

学习本课程必须具备高中毕业生所应当具备的数学知识,但不要求考生掌握高等数学知识(虽然懂得高等数学有助于掌握和理解基本的概率统计原理)。

本书的第一章是关于数据的整理和描述。通过本章的学习,考生应当理解和掌握如何对数据进行整理、分组、制表和画图,能够适当地选择和解释数据的各种综合指标,以便能够突出地显示数据的本质和统计含义,从而更有效地交流数据和使用数据。

第二章是关于随机事件及其概率的介绍,它是后续章节的基础。通过本章的学习,考生应当理解和掌握随机试验、随机事件、样本空间、事件的概率、条件概率以及事件的独立性等概率论的一些基本概念,学会运用事件的关系与运算、概率的性质以及概率的计算法则,计算一些简单事件的概率,为后续章节的学习打下基础。

第三章讨论随机变量及其分布。通过本章的学习,考生应当理解随机变量的概念,了解概率函数、分布函数和概率密度函数等概念及其性质,掌握常见的随机变量及其分布,特别是二项分布和正态分布,理解数学期望和方差的概念。初步了解二元随机变量以及协方差和相关系数概念。能利用数学期望和方差,以及决策树知识进行简单的分析决策。

第四章是关于抽样方法和抽样分布的讨论。通过本章的学习,考生应当了解抽样方法的基本原理,掌握几种简单的抽样方法,能正确使用随机数设计随机抽样,能够正确处理抽样中常见的问题,了解样本容量与估计量精度和费用之间的关系;理解统计量的概念,掌握样本均值和样本方差的计算方法,理解样本均值和样本方差在实践中的应用。

第五章讲述参数估计。对总体的特征,即总体的参数作出估计,并且给出这种估计的可靠程度是统计推断的两大任务之一。通过本章的学习,考生应当了解和掌握对总体未知参数进行点估计和区间估计的方法,学会解释它们的实际意义,并能够应用所学的知识,解决实际问题。

第六章讨论假设检验。统计假设检验是统计推断的另一任务。通过本章的学习,考生应当了解统计假设检验是统计推断的另一任务,掌握假设检验的有关概念和一般步骤,理解假设检验的

基本思想,学会针对不同情形选择适当的统计量对统计假设进行检验。

第七章是关于相关与回归的分析。通过本章的学习,考生应当了解回归分析是数理统计中研究变量之间相关关系的一种常用方法,掌握回归分析的有关概念和回归直线或非直线方程的估计方法,以及对相关关系的显著性进行检验,并利用回归直线方程或非线性方程的线性化进行估计和预测的方法。

第八章是有关时间数列的分析。通过本章的学习,考生应当了解时间数列概念以及时间数列分析的目的,了解时间序列的四大构成因素,学会用各种分析方法测定时间数列的长期趋势、季节变动和循环波动,描述和分析时间序列的行为,并对各观察值作出对比分析。

第九章是关于指数的讨论。通过本章的学习,考生应当理解学习和掌握指数的必要性,了解指数的基本概念及其分类,掌握指数的编制原则和方法,并运用指数综合分析和评价经济事物变化的方向和大小。

本书注重对考生应用能力的培养,力求浅显易懂,便于自学。本书包含各种应用实例,并详细解答。每章后都附有一定数量的习题供考生练习,书末给出了答案。

清华大学经济管理学院钱小军副教授(第一至第三章)、中国人民大学倪加勋教授(第四至第六章)和贾俊平副教授(第七至第九章)参与了本教材的编写,钱小军任主编并对全书的内容进行了统校审定。在编写过程中,我们努力从实用的角度出发,尽量使内容浅显易懂,但是疏漏错缪、繁琐难懂之处在所难免,恳请广大读者指正。我们恳切地希望广大考生和助学人员能将自己学习和使用本教材的意见提供给我们,以使我们在今后的工作中做得更好。

编者

1999年8月

· 5 ·

目 录

第一章 数据的整理和描述	1
第一节 数据的类型	1
第二节 数据的整理与图表显示	3
第三节 数据集中趋势的度量	15
第四节 数据离散趋势的度量	24
第二章 随机事件及其概率	33
第一节 随机试验与随机事件	33
第二节 事件间的关系与运算	37
第三节 事件的概率与古典概型	42
第四节 条件概率与事件的独立性	50
第三章 随机变量及其分布	62
第一节 随机变量	62
第二节 离散型随机变量	63
第三节 连续型随机变量	74
第四节 二元随机变量	85
第五节 决策准则与决策树	91
第四章 抽样方法与抽样分布	102
第一节 抽样的作用与抽样方法	102
第二节 抽样中经常遇到的几个问题	111
第三节 抽样中的三种分布及中心极限定理	118
第四节 一些常用的抽样分布	124
第五节 几个重要的小样本抽样分布	129
第五章 参数估计	134
第一节 参数估计的一般问题	134
第二节 总体均值的区间估计	140
第三节 总体比例的区间估计	147
第四节 两个均值或两个比例之差的区间估计	151

第五节 样本容量的确定	155
第六章 假设检验	164
第一节 假设检验的基本概念	164
第二节 参数的假设检验	170
第三节 非参数假设检验	183
第七章 相关与回归分析	202
第一节 简单线性相关	202
第二节 一元线性回归	208
第三节 多元线性回归和非线性回归	218
第八章 时间数列分析	227
第一节 时间数列的对比分析	227
第二节 长期趋势分析及预测	239
第三节 季节变动分析	257
第四节 循环波动分析	264
第九章 指数	269
第一节 指数的含义与分类	269
第二节 加权指数	272
第三节 指数体系	281
第四节 几种常用的重要指数	287
习题参考答案	296
附表 1 标准正态分布表	307
附表 2 泊松分布表	309
附表 3 t 分布表	311
附表 4 χ^2 分布表	313
附表 5 F 分布表	317
附表 6 曼·惠特尼检验 U 的临界值表	329
附表 7 威尔科克森带符号的秩检验 T 的临界值表	331
附表 8 斯皮尔曼等级相关系数检验 t 的临界值表	332
附表 9 平均增长速度累计法查对表	333

第一章 数据的整理和描述

从事商务管理和金融管理的人员,不可避免地经常接触大量的统计数据。杂乱无章的数据体现不出它的价值,对实践也没有直接的指导意义。因此,学习、理解和掌握如何对数据进行整理、分组、制表和制图,以及如何选择适当的综合指标,以便能够突出地显示数据的本质和统计含义,从而适合商务管理和金融管理工作的需要,是十分必要的。本章主要介绍统计数据的整理和描述方法。

第一节 数据的类型

根据描述事物所采用的不同度量尺度,数据可分为分类型数据和数量型数据。

分类型数据描述的是事物的品质特征。例如,人的性别、民族、职业等等,都是分类型数据。虽然有时我们给不同类别赋予不同的数值,但分类型数据的本质表现是文字形式。如性别表现为男、女,民族表现为汉、回、蒙、藏等,产品分为合格品和废品等等。分类型数据的分类必须是清楚唯一的,如果产品分为合格品和废品两种,那么一个产品不能同时既是合格品又是废品。

数量型数据说明的是事物的数量特征。例如,人的年龄、企业的职工人数、产品的产量和寿命、企业的营业额、股票的价格、产品的市场占有率、国民总产值、国家的人口等等,都是数量型数据。数量型数据用数值形式表示。如某人的年龄为25岁,企业的职工人数为250人,企业某产品的市场占有率为31.7%等等。

数据按照被描述的对象与时间的关系分为截面数据、时间序列数据与平行数据。

截面数据描述的是事物在某一时刻的变化情况,即所谓横向数据。例如表 1.1 列出了 1996 年我国各地区国内生产总值数据,这是一组截面数据。

表 1.1 1996 年我国各地区国内生产总值 (单位:亿元)

地区	国内生产总值	地区	国内生产总值	地区	国内生产总值
北京	1 615.73	浙江	4 146.06	海南	389.53
天津	1 102.40	安徽	2 339.25	四川	4 215.00
河北	3 452.97	福建	2 606.92	贵州	719.83
山西	1 305.50	江西	1 517.26	云南	1 491.62
内蒙古	984.78	山东	5 960.42	西藏	64.76
辽宁	3 157.69	河南	3 683.41	陕西	1 175.38
吉林	1 337.16	湖北	2 970.20	甘肃	714.18
黑龙江	2 402.58	湖南	2 647.16	青海	1 83.57
上海	2 902.20	广东	6 519.14	宁夏	193.62
江苏	6 004.21	广西	1 869.62	新疆	912.15

(资料来源:《中国统计年鉴 1997》)

时间序列数据描述的是事物在一定的时间范围内的变化情况,即所谓纵向数据,例如表 1.2 中的北京市总人口从 1982 年到 1990 年的数据。

表 1.2 1982~1990 年北京市总人口 (单位:万人)

年份	1982	1983	1984	1985	1986	1987	1988	1989	1990
总人口	917.83	933.20	945.20	957.90	971.23	987.97	1 001.20	1 021.11	1 032.21

(资料来源:《80 年代中国人口变动分析 中国人口续篇》)

平行数据是截面数据与时间序列数据的组合,例如表 1.3。对于表 1.3 中的每一个省份(每一行),都有 1993 年到 1996 年该省的国内生产总值(时间序列数据);而对于表中的每一年(每一

列),都有东北三省在该年的国内生产总值(截面数据),它表示的是一个平行数据。

表 1.3 东北三省 1993~1996 年国内生产总值

地区	1993 年	1994 年	1995 年	1996 年
辽宁	2 010.82	2 461.78	2 793.37	3 157.69
吉林	717.95	936.78	1 129.20	1 337.16
黑龙江	1 203.22	1 618.63	2 014.53	2 402.58

(资料来源:《中国统计年鉴 1997》)

在统计中,我们把对事物现象特征的描述称为变量。如果它是分类型数据,称为分类型变量;如果它是数量型数据,则称为数量型变量。很多情况下,我们所研究的变量都是数量型变量,大多数的统计分析方法也都是对于数量型变量的分析,因此有时把数量型变量简称为变量。

第二节 数据的整理与图表显示

一、数据的分组与频率直方图

统计分组是数据整理的一项初步工作,它是根据实际需要,将数据按照数据的某种特征或标准分成不同的组别。对数据的分组需要按照数据的特征来进行。因此,我们将用不同的方法对分类型数据和数量型数据进行分组。

按照数据的某种特征对数据进行分组后,再计算出所有类别或数据在各组中出现的次数或频数,就形成了频数分布表。我们称全部数据在各组内的分配状况为数据的频数分布,分配在各组内的数据个数为频数,频数与全体数据个数之比称为频率。

分类型数据按类计算出各类的频数或频率,就形成了频数或频率分布表。我们用下面的例子具体说明如何制作分类型数据的频数和频率分布表。

【例 1.1】某车间有 25 名工人,加工某种零件,下表记录的是他们某月的平均日产量。车间规定平均日产量超过 120 件的工人当月奖金为一等,平均日产量在 110~120 件之间的工人奖金为二等,平均日产量低于 110 件的工人没有奖金。试统计该车间的 25 名工人中,获得一等奖、二等奖以及没获奖的人数。

平均日产量	平均日产量	平均日产量	平均日产量	平均日产量
115	103 * *	121 *	107 * *	111
109 * *	108 * *	116	119	116
113	112	125 *	113	123 *
124 *	117	105 * *	128 *	126 *
99 * *	120	92 * *	114	106 * *

这个数据集不是分类型数据集,但是其数据可以按照奖金等级标准分类。可以看出该车间该月共有 6 名工人的平均日产量高于 120 件(带 * 号的数据),有 11 名工人的平均日产量在 110~120 件之间,平均日产量低于 110 件的工人有 8 名(带 * * 号的数据);分别占全体工人的 24%,44% 和 32%。将结果归纳成频数和频率分布表(同时计算出累积频率),我们得到:

按奖金等级分组	频数(人)	频率(%)	累积频率(%)
>120 件	6	24.0	24.0
110~120 件	11	44.0	68.0
<110 件	8	32.0	100.0
合计	25	100	

从上面的频数和频率分布表,我们就能很容易地看出,获得不同奖金等级的人数,及其占全体工人的百分比。同时由累积频率,我们也能方便地得知有 68% 的工人获了奖。

分类型数据按类分组时,一定要注意既不能重数也不能漏数。

这就要求所有类别必须有明确的界定。比如上例中,日产量正好为 110 件和 120 件的工人应分在获得二等奖金的组内。

对于数量型数据,我们只简单介绍两种方法——单变量值分组法和组距分组法。

单变量值分组法就是把每一个变量值作为一个组。我们用下面的具体实例对这一方法进行说明。

【例 1.2】某单位有职工 20 人,下面是六月份该单位职工请假天数的记录:

0,0,1,0,2,1,0,0,0,1,2,0,5,1,1,0,0,0,0,10,0

观察这个记录,我们不难发现,所有不同的请假天数一共只有 5 个,即 0、1、2、5 和 10。因此,采用单变量值分组方法分组的话,应该分成 5 个组,具体结果如下:

请假天数	频数(人)	频率(%)	累积频率(%)
0	11	55	55
1	5	25	80
2	2	10	90
5	1	5	95
10	1	5	100
合计	20	100	

由以上频率分布表可以容易地看出,该单位六月份有 55% 的全勤职工,有 90% 的职工请假天数不超过 2 天,等等。

在数据较多且比较分散的情况下,单变量值分组法由于组数过多,不便于观察数据的分布特征和规律。因此,单变量值分组方法适用于数据较少或分布比较集中的情形。

对于变量值较多的情况,可以采用组距分组法。我们以表 1.4 中的数据为例具体讲述组距分组法的具体做法。

表 1.4 我国各地区 1995 年死亡率 (单位:‰)

地区	死亡率	地区	死亡率	地区	死亡率
北京	5.12	浙江	6.75	海南	5.61
天津	6.23	安徽	6.41	四川	7.21
河北	6.32	福建	5.90	贵州	7.60
山西	6.12	江西	7.28	云南	8.03
内蒙古	6.70	山东	6.47	西藏	8.80
辽宁	6.15	河南	6.28	陕西	6.57
吉林	6.09	湖北	6.91	甘肃	6.49
黑龙江	5.33	湖南	7.15	青海	6.89
上海	7.05	广东	5.70	宁夏	5.49
江苏	6.56	广西	6.53	新疆	6.45

(资料来源:《中国人口统计年鉴 1996》)

用组距分组法对上述数据进行分组的具体步骤:

第一步:将全部数据按上升顺序排列,找出最大值 max 和最小值 min。表 1.4 中的最大值和最小值分别为

$$\max = 8.80 \quad \min = 5.12$$

第二步:确定组数,计算组距。一个具体的数据集应当分成几组比较合适呢? 分组过多或分组过少都达不到简化综合数据集的目的,因此在具体分组之前,必须选择一个适当的分组个数。确定组数的一般原则为

数据个数 n	分组数
50 以下	5~6
50~100	6~10
100~250	7~12
250 以上	10~20

表 1.4 中共有 $n = 30$ 个数据,可以取分组数 $m = 6$ 。组距 c 的定义为 $c = \frac{b-a}{m}$,其中 a 是一个不大于 \min 的数, b 是不小于 \max 的数,即 $a \leq \min, b \geq \max$ 。应当选择 a 和 b ,使得组距 c 不为无限小数。以表 1.4 为例,我们可以取 $a = 4.9, b = 9.1$,则组距 $c = \frac{9.1-4.9}{6} = 0.7$ 。

第三步:计算每组的上、下限(分组界限)、组中值,以及数据落入各组的频数 v_i (个数)和频率 f_i (频数/ n),形成频率分布表。特别需要注意的是各组的边界定义一定要明确,不要遗漏也不要重复。

第一组的下限为 a ,上限为 $a + c$;第二组的下限为 $a + c$,上限为 $a + 2c$;一般地,对于任意 $1 \leq i \leq m$,第 i 组的下限为 $a + (i-1)c$,上限为 $a + ic$ 。第 m 组的下限为 $a + (m-1)c$,上限为 $a + mc = b$ 。每组的组中值为(上限+下限)/2。

表 1.4 数据的频率分布表如表 1.5 所示。

表 1.5 我国各地区 1995 年死亡率频率分布表

组号	分组界限	频数 v_i	频率 f_i (%)	组中值
1	[4.9,5.6)	3	10.00	5.25
2	[5.6,6.3)	8	26.67	5.95
3	[6.3,7.0)	12	40.00	6.65
4	[7.0,7.7)	5	16.67	7.35
5	[7.7,8.4)	1	3.33	8.05
6	[8.4,9.1]	1	3.33	8.75

根据频率分布表,可以画频率直方图。具体做法如下:在横坐标 x 轴上,以每个分组区间为底边,画出高为 f_i/c 的矩形。表 1.5 的频率直方图如图 1.1 所示。