



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 剑桥商务管理证书自学辅导书

# 数量方法

全国高等教育自学考试指导委员会 组编  
倪加勋 贾俊平 编著

4.7

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  
剑桥商务管理证书自学考试辅导书

## 数量方法

全国高等教育自学考试指导委员会 组编  
倪加勋 贾俊平 编著

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合作的课程——商务管理、金融管理专业数量方法课程用教材配套辅导书,是按照课程考试大纲并结合教材编写的。本书各章均包括四部分内容:概述部分说明该部分内容的用途及应掌握的主要内容;基本概念与思考题部分可帮助读者理解主要概念;例题分析部分选择有代表性的重、难点问题,为读者提供基本的解题思路、方法和技巧;练习题部分可供读者选做。本书最后还附有练习题参考答案以及应考指导。

## 前 言

《数量方法自学辅导书》是和高等教育自学考试中英合作商务管理和金融管理两个专业的专业课“数量方法”指定教材《数量方法》(钱小军主编,高等教育出版社,1999年出版)相配套的自学考试用书,它在内容和体系上与《数量方法》教材以及相应的考试大纲一致,是商务管理和金融管理专业自学考试人员必备的学习指导书。

本自学辅导书和《数量方法》教材一样,共分九章。每章包括四部分:第一部分是该章的概述,说明该部分内容的用途、应该掌握的主要内容,以及学习时应当注意的问题;第二部分是基本概念与思考题,围绕该章主要内容提出读者应当掌握的主要概念和思考题,并提供相应的参考答案,力图帮助读者全面完整地理解该章的主要概念;第三部分是例题分析,根据学生在解题过程中的困难,选择有代表性的重点和难点问题做示范分析,力求为读者提供基本的解题思路、解题方法和解题技巧;第四部分是练习题,可供自学考试学员选做,目的是为他们提供更多的实践机会。为了帮助考生更好地参加考试、发挥自己的水平,本书还特别包含了应考指导,希望能对考生有所帮助。本书的最后是各章练习题的参考答案,对于有一定难度的习题还给出了比较详细的解题过程,供自学考试学员和助学人员参考。

本书由中国人民大学统计学系倪加勋(第一章至第六章)和贾俊平(第七章至第九章)编写,全书由清华大学经济管理学院钱小军统一审校。在编写的过程中,我们努力做到,在不失数学的严谨性的同时,以自学考试学员容易理解和接受的方式描述数学概念,力求条理清楚,层次分明,由浅入深,循序渐进。大量的例题和练习题为自学考试人员提供了练习和自我检验的机会。但是,由于

作者水平有限,加上时间仓促,不当和不足之处在所难免,恳请广大读者批评指正。

编者

1999年9月于北京

# 目 录

<b>第一章 数据的整理和描述</b> .....	1
一、概述 .....	1
二、基本概念与思考题 .....	2
三、例题分析 .....	9
四、练习题 .....	15
<b>第二章 随机事件及其概率</b> .....	21
一、概述 .....	21
二、基本概念与思考题 .....	21
三、例题分析 .....	27
四、练习题 .....	30
<b>第三章 随机变量及其分布</b> .....	37
一、概述 .....	37
二、基本概念与思考题 .....	37
三、例题分析 .....	43
四、练习题 .....	45
<b>第四章 抽样方法与抽样分布</b> .....	52
一、概述 .....	52
二、基本概念与思考题 .....	53
三、例题分析 .....	60
四、练习题 .....	64
<b>第五章 参数估计</b> .....	70
一、概述 .....	70
二、基本概念与思考题 .....	70
三、例题分析 .....	76
四、练习题 .....	82
<b>第六章 假设检验</b> .....	87
一、概述 .....	87

二、基本概念与思考题 .....	87
三、例题分析 .....	94
四、练习题 .....	97
<b>第七章 相关与回归分析 .....</b>	<b>105</b>
一、概述 .....	105
二、基本概念与思考题 .....	105
三、例题分析 .....	109
四、练习题 .....	113
<b>第八章 时间数列分析 .....</b>	<b>120</b>
一、概述 .....	120
二、基本概念与思考题 .....	120
三、例题分析 .....	124
四、练习题 .....	129
<b>第九章 指数 .....</b>	<b>137</b>
一、概述 .....	137
二、基本概念与思考题 .....	137
三、例题分析 .....	140
四、练习题 .....	144
<b>练习题参考答案 .....</b>	<b>150</b>
<b>应考指导 .....</b>	<b>187</b>

# 第一章 数据的整理和描述

## 一、概述

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,通过数据来研究客观现象,用它来确定有些现象的规模、水平以及完成进度等数量界限;用它来研究事物之间的相互联系;用它来研究事物的发展规律等。因此数据是应用数量方法的基础。俗话说:好媳妇难为无米之炊,这也反映了数据对数量方法这一课程的重要性。但是客观存在的数据需要人们有意识地加以搜集,而且初步搜集的数据往往是杂乱无章的,而杂乱无章的数据不会有很大价值。因此获得数据以后,就要对数据的质量加以甄别,并在此基础上进行整理和描述,使之成为有意义的形式,并初步呈现事物的规律,有时搜集的数据是局部的,要为进一步探索总体规律和统计推断打下基础。

本章的内容是假定数据已经收集,如何对原始的数据进行整理和描述。第一节介绍数据的类型,因为数据可以从不同的角度进行分类,而不同的数据会有一些不同的处理方法;第二节介绍数据的整理与图表显示,这是把零乱的原始数据整理成有意义的形式,并用直观和形象的方法加以显示;第三节是数据集中趋势的度量,集中趋势是刻画一组数据的重要特征之一,它反映数据的中心或重心的位置,使人们能抓住核心,不至于毫无头绪地迷失在数据的海洋中,不得要领;第四节是数据离散趋势的计量,这是刻画数据的另一个特征,与集中趋势相结合,就能较全面地掌握数据的分布状况。本章的重点应掌握数据的分组,特别是构造频数的分布,理解频数分布的主要特征,熟练地计算有关集中趋势和离散趋势的度量指标。



## 二、基本概念与思考题

### (一) 基本概念

1. 分类型数据 又称属性数据,它所描述的是事物的品质特征。从统计的计量水准来说是一种比较原始和低级的计量,称作列名水准。这类数据只能计算各类的频数和比例,不能进行其他数学运算,如人口按性别、民族、宗教信仰等分类,这种分类没有严格的先后顺序。广义的分类数据,也包括顺序的计量水准,如学生的成绩划分为优、良、中、及格和不及格,它们之间有一定的顺序关系,可以比较,但也不能进行其他数学运算。

2. 数量型数据 这类数据是用来说明事物的数量特征。从统计的计量水准来说,包括定距水准(又称间隔水准)和定比水准(又称比率水准)。例如,人的年龄,企业的职工人数,产品产量,国家的国民生产总值(GNP)等用数值的形式表示,这类数据除了计算频数和比例外,还可以进行有关数学运算,如计算平均数、方差等。

3. 截面数据 是指用来描述事物在同一时点(或时期)社会经济各种不同指标的数据。如在同一时期的人口数、国内生产总值、运输量、财政收入等数据,可以观察同一时期各指标之间的相互关系。截面数据还包括同一时期相同指标在不同部门的分布,通常又称横向数据。截面数据可以研究客观现象之间的相互联系。

4. 时间序列数据 将数据按时间的先后顺序排列后形成的数据序列,又称纵向数据。时间序列数据可以反映事物在一定时期范围内的变化情况,研究事物动态变化的规律性并进行预测等。

5. 频数分布 又称次数分布,是按照数据的某种特征进行分组后再计算出各类数据在各组出现的次数加以整理,这种次数也称频数,这种整理后形成的表,称作频数分布表。把频数与全体数据个数之比,称为频率,这样的表就为频率分布表。频数(或频率)分布表可以观察各组数据在全部数据中的状况。

6. 组距 在数量型数列中按单变量分组有时组数过多,不利于观察数据分布特征和规律,需要将数据的大小适当归并,在每组中规定的最大值与最小值之差就称作组距。各组的组距均相等时称作等距数列,不完全相等时称不等距数列。

7. 组界 又称组限,指组距的变量数列的分组中,各组变动范围两端的数值,最小限度的值称作下限,最大限度的值称作上限,上限与下限之差即为组距。

8. 组中值 组距的变量数列中每组上限与下限的平均值,其计算公式为:

$$\text{组中值} = \frac{\text{上限} + \text{下限}}{2}$$

它用来代表该组变量的一般值,可用于计算加权平均数等。

9. 频率分布表 频数分布表的另一种表现形式,它把每组中变量出现的频数转换为相对次数,即得每组次数除以总次数,称为各组的频率,各组的频率相加为 1。

10. 直方图 频数分配表的直观图示形式。它适用于组距数列,图形用一平面的直角坐标,横轴表示变量值,各组的组距大小与横轴上的长度成正比。在一般情况下为了观察变量的规律性,通常采用相同的组距;纵轴表示频数或频率,用高度来表示频数多少,与横轴的各组组距连接垂直直线,这样就形成按频数多少而高低不一的条形图,从而可以直观地显示变量的分布状况,通常可以用条形的面积大小来反映变量在该组的频率高低。

11. 条形图和柱形图 一种用来对各项信息进行比较的图示方式。在平面上用相同宽度但不同长度的条形来表示数值的大小,其条形可以是横的,也可以是竖的,当条形竖立时,也称柱形图。条形图可以用来比较不同国家、地区、行业以及公司单位之间同一指标的差异,也可以比较同一单位在不同时间指标值的差异。如果把两个或两个以上指标的条形图合成一组条形图,称复式条形图,如果把一个条件的全部长度分割成几个小段,每小段长度代表总体的一个组成部分,就称作结构条形图。

12. 饼形图 又称圆形结构图,一般用来描述和显示总体中各种类型占全体的比例。通常以圆面积表示研究对象的总量,把圆形分成若干个扇形部分,每个扇形部分代表一种组成部分,该组成部分的大小与扇形的面积大小成正比,从而表示总量的构成状况,形象地显示总量的结构。

13. 折线图 有两种折线图,一是在研究动态趋势时,以横坐标表示时间,纵坐标表示现象的数值,将所形成的点逐点相连,就形成动态折线图,可以反映动态的变化趋势;另一是在直方图的基础上,将顶端中间的点,其临近两点用直线加以连接,就形成频数分配的折线图,把折线两端延伸与横轴相连,则折线图下所覆盖的面积,相当于直方图的面积,在频率分布时其面积表示 1,可以反映变量的分布状况。

14. 曲线图 是折线图的修匀,折线图在各点连接时会产生突变,而客观事物的发展往往是逐渐变化的,通过修匀后的曲线图则弥补了这一不足,反映了逐渐变化的过程。反映时间数列的曲线图又称动态曲线图或历史曲线图,反映变量数列的曲线图又称频率曲线图。

15. 散点图 又称散布图,通常用来描述两个变量之间的关系,当一个单元具有两个标志值时,在坐标轴上分别用横坐标和纵坐标表示,在它们取值的交叉点上作点,这些点所形成的图形,就称散点图。它可以观察两个指标之间是否存在关系,如果有关系又是何种形状的关系,在相关与回归分析中具有重要作用。

16. 茎叶图 形象地把每个数据分为茎和叶两部分,把数字的主干部分加以归类作为茎,然后在分类时把其余部分作为叶,列在相应的茎上,其优点是可以把统计的分组和频数分配的划记工作一次完成,既保持了直方图的直观形象,又保留了原有数据的原始信息,从中可以得到平均数,中位数和众数等特征值。

17. 平均数 又称均值,其中最常用的是算术平均数,是指一组数据之和除以数据的个数,其计算公式为:

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

平均数代表了一组数据的一般水平,因为它是把高低数据相互抵消的结果,它也是数据集的重心位置,正好是一个平衡点,它反映了数据的位置或集中趋势。因为在多数情况下,大部分数据分布在它的左右。

18. 中位数 将一组数据按照由小到大依次排序后处于中间位置上的变量值,也就是说中位数将整个数据一分为二,正好有一半的数据比中位数小,另一半的数据比中位数大,如果数据集为偶数,则应是中间两个数值的平均数,如果是频数分配表,中位数往往位于某一组距之内,需要用插入的方法计算。中位数具有稳健性的优点,即它不受个别极端值的影响,因此当数据中含有极端值时,用中位数来描述集中趋势比平均数更为恰当。

19. 众数 是指一组数据中出现次数最多的那个变量值,众数的优点在于反映了数据中最常见的数值,它不仅适用于数量型数据,也适用于分类型数据。如在商务统计中可用来反映最普遍、最流行的款式、尺寸、规格等特征,有助于作生产计划决策。但其缺点是有些数据集可能没有众数(如每个数据只出现1次),也可能有几个众数(有几个数据出现的次数相同)。

20. 方差 是一组数据的每一个观察值与其平均数离差平方的平均数。设有  $n$  个数据  $x_i (i = 1, \dots, n)$  其平均数为  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , 则方差定义为:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

方差是反映一组数据离散程度的重要指标。当与平均数一起应用时,可以说明平均数代表该数据集的代表性,方差愈小,平均数的代表性愈强。

21. 标准差 方差的平方根,即

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

也是反映数据离散程度的指标,由于方差是变量与平均数离差平方的平均数,因而方差的量纲与原来数据的量纲不一致,标准差将其开平方根,就恢复了原来数据的量纲。

22. 极差 又称全距,指一组数据中最大值与最小值之差。显然它也是度量一组数据的离散或集中程度的,极差愈小表示数据的集中程度愈高,极差愈大表示数据愈分散。它的优点是计算简便,但缺点是易受极端值的影响而不够稳定,而且也没有充分利用所有数据的信息。

23. 变异系数 又称离散系数,是指一组数据的标准差与平均数之比。用  $v$  表示,则

$$v = \frac{\sigma}{\bar{x}} \times 100\%$$

它反映数据的相对离散程度,便于在不同均值的两组数据中以及对于具有不同属性的两组数据中比较离散程度。

24. 四分位点 将一组数据由小到大顺序排列,用  $Q_1$ 、 $Q_2$  和  $Q_3$  三个点将整个数据的个数进行四等分,它们分别位于 25%、50% 和 75% 的位置,这三个点就称为四分位点,这三个点的数值称为四分位数。由上述定义可知,该组数据中有 25% 小于  $Q_1$ ,  $Q_1$  为第 1 四分位数;  $Q_2$  就等于中位数,也可称为第 2 四分位数;有 25% 的数据大于  $Q_3$ ,  $Q_3$  称为第 3 四分位数。

25. 四分位差 基于四分位点计算的数据值之差,又分为四分位极差和四分位半距。四分位极差是指第 3 个四分位数  $Q_3$  与第 1 个四分位数之差,即  $Q_3 - Q_1$ ,它表明除去两端各 25% 的数据后的极差;四分位半距是将四分位极差除以 2。两者都是度量数据离散程度的指标。与全距相比,它排除了少数极端数值的影响。

## (二) 思考题

1. 将原始数据编制成频数分布的作用是什么?

原始数据往往是杂乱无章的,杂乱无章的数据就没有多大价值。通过频数分布的整理可以显示出数据的规律性,如看出最大

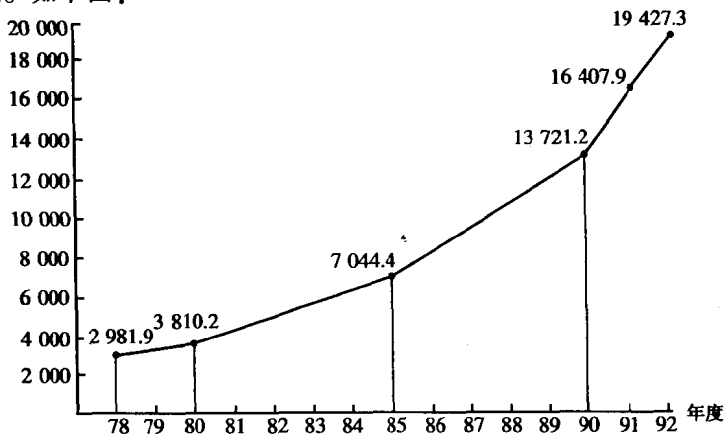
值是多少、最小值是多少,它是在多大的范围内变动,也有利于平均数等的计算。通过直方图的编制还可以看出数据是否对称,它的离散程度等。

2. 一组数据用原始数据计算平均数与编制频数分布后计算平均数,其结果是否一致?

通常的情况下不会一致,因为利用频数分布计算平均数时是以组中值来代表该组的全部数据,它是以该组数据在组内是均匀分布为前提的,这种假设会与实际情况有所出入而不一致,但一般不会相差很大,因为在有些组内组中值可能偏高,而在有些组内则可能偏低,会相互抵消。但计算平均数会带来很大方便,因此损失一些微小的精度往往是值得的。

3. 下面是“北京一日”的统计资料中摘录的国内生产总值:1978年2 981.9万元,1980年3 810.2万元,1985年7 044.4万元,1990年13 721.2万元,1991年16 407.9万元,1992年19 427.3万元。若用统计图加以显示,应选用什么图,要注意什么问题?

统计图可以任意创作,但要求简明、生动、准确。上述动态发展趋势可以选用条形图或动态折线图。但应注意由于年度之间的间隔有长短,在作折线图时应注意适当距离才能准确显示各年速度。如下图:



#### 4. 简述平均数的作用。

在大量的数据观察中,平均数抵消了一些偶然变动的影响,从而呈现了数据的一般水平,在频数分布中平均数体现了大量数据的集中趋势,其他数据围绕平均数上下波动,在统计推断中也离不开平均数。平均数还有许多良好的数学性质,如各个观察值与均值离差之和为零,与均值离差平方之和为最小等。但平均数在反映数据的一般水平时也有局限性,就是它易受少数极端值的影响。

#### 5. 简述中位数的作用。

中位数也是一种集中趋势的计量,但它是一个位置的中间值。当数据中有极端的数字时,平均数易受极端值的影响,而中位数则不受其影响,因此在统计中称为有比较稳健的性质。但反过来说它不如平均数灵敏,它没有充分利用数据的信息。

#### 6. 简述众数的作用。

一个表明位置的集中趋势,它表明数据出现次数最多的那个值,它有其特定的作用,如电视机有各种不同的型号,需要了解哪种型号是销售最多的,这就是众数的型号,它不能由平均数或中位数来替代。另外它不仅可用于数量型变量,也可用于属性变量。

#### 7. 简述在一组数据中,平均数、中位数和众数之间的关系。

当一组数据呈中间大两头小的对称分布(包括正态分布)时,平均数、中位数和众数是一致的。但如果呈现偏态时这三者就不一致了。当分布为右偏(右边拖一个长尾巴)时,则平均数受极端值的影响而比较大,中位数在中间而众数最小。当分布为左偏时,则平均数最小,众数最大,而中位数仍在中间。当数据不存在明显的规律性时,也有例外的情况。

#### 8. 简述标准差与变异系数的关系。

标准差与变异系数都是测定一组数据的离散程度,有相同的作用,但标准差是测定绝对的离差大小,具有量纲,它的计量单位与测量数据的单位是一致的,变异系数则是测量相对的变异程度,因而没有量纲。它们在不同场合各有其不同的作用。

#### 9. 简述极差与四分位极差的关系。

极差与四分位极差都是计量一组数据离散程度的,极差比较直观且容易计算,在质量管理中测量产品变动情况时经常用到极差。但极差是最大值减去最小值,容易受极端值的影响。而四分位极差是两端各去掉 1/4 的单位后,计算其差值,因此它消除了两端极值的影响,但在计算时要先算出  $Q_1$  和  $Q_3$ ,相对比较麻烦一些。

### 三、例题分析

例题 1(填空题):

下面是一个班 40 个学生经济学的考试成绩:

88	79	90	88	71
83	65	98	86	33
71	49	100	77	93
60	91	80	63	67
99	66	42	75	62
69	68	75	96	83
63	87	84	58	90
21	87	68	89	80

要求整理成频数分布表和计算该组数据的有关特征值填入下面的空格:

**经济学考试成绩频数分布表**

组 距	人 数	人数的百分比
20—30 分以下		
30—40 分以下		
40—50 分以下		
50—60 分以下		
60—70 分以下		



续表

组 距	人 数	人数的百分比
70—80 分以下		
80—90 分以下		
90—100 分以下		
100 分		

从该频数分布表计算平均数  $\bar{x} =$  \_\_\_\_\_, 方差 = \_\_\_\_\_, 标准差 = \_\_\_\_\_, 变异系数 = \_\_\_\_\_, 从原始数据可知中位数 = \_\_\_\_\_, 众数 = \_\_\_\_\_,  $Q_1 =$  \_\_\_\_\_,  $Q_3 =$  \_\_\_\_\_, 四分位极差 = \_\_\_\_\_。

参考答案:

组 距	人 数	人数的百分比
20—30 分以下	1	2.5%
30—40 分以下	1	2.5%
40—50 分以下	2	5.0%
50—60 分以下	1	2.5%
60—70 分以下	10	25.0%
70—80 分以下	6	15.0%
80—90 分以下	11	27.5%
90—100 分以下	7	17.5%
100 分	1	2.5%

平均数 = 75, 方差 = 295.61, 标准差 = 17.193,

变异系数 = 0.229, 中位数 = 78, 众数 = 无明显众数(共有 8 个数值出现两次),  $Q_1 = 65.5$ ,  $Q_3 = 88.5$ ,  $Q_3 - Q_1 = 23$

分析:通过茎叶图来把原始数据整理成频数分布表是一种比较方便的方法,可以将十位数作茎,个位数作叶。