

“十三五” 普通高等院校“十三五”规划教材

工商管理类

管理统计学实验教程

主 编 宋冬梅
副主编 崔琳琳

 南京大学出版社



普通高等院校“十三五”规划教材

工商管理类

管理统计学实验教程

主 编 宋冬梅

副主编 崔琳琳

图书在版编目(CIP)数据

管理统计学实验教程 / 宋冬梅主编. — 南京: 南京大学出版社, 2017. 2

普通高等院校“十三五”规划教材·工商管理类

ISBN 978-7-305-18259-4

I. ①管… II. ①宋… III. ①经济统计学 IV. ①F222

中国版本图书馆 CIP 数据核字(2017)第 013947 号

出版发行 南京大学出版社
社 址 南京市汉口路 22 号 邮 编 210093
出 版 人 金鑫荣

丛 书 名 普通高等院校“十三五”规划教材·工商管理类
书 名 管理统计学实验教程
主 编 宋冬梅
责任编辑 武 坦 尤 佳 编辑热线 025-83597482

照 排 南京南琳图文制作有限公司
印 刷 虎彩印艺股份有限公司
开 本 787×1092 1/16 印张 14.25 字数 356 千
版 次 2017 年 2 月第 1 版 2017 年 2 月第 1 次印刷
ISBN 978-7-305-18259-4
定 价 32.00 元

网址: <http://www.njupco.com>

官方微博: <http://weibo.com/njupco>

微信服务号: njuyuxue

销售咨询热线: (025) 83594756

* 版权所有, 侵权必究

* 凡购买南大版图书, 如有印装质量问题, 请与所购
图书销售部门联系调换

前 言

本书共分为 14 章 35 个实验,融合了 Excel、SPSS 在统计分析中的应用,一个是基础篇,另一个是提高篇。基础篇紧扣 Excel 2007 在统计分析中的应用,相对应的基础知识理论较少,重点放在典型案例和综合实战。提高篇是基于 SPSS 22.0 中文版进行编写的,以统计分析的实际应用为主线,在对主要统计分析方法的基本概念和统计学原理进行简明介绍的基础上,以实例为载体对 SPSS 22.0 中各种分析方法的操作过程进行了清晰说明。

本书主要特点:一是由浅入深、循序渐进;二是图文并茂,直观形象;三是兼顾统计原理与 Excel、SPSS 软件操作应用,对其原理和适用条件作了较详细说明;四是每类统计分析给出具体实例以及软件操作应用,并对输出结果做了详细的分析,章后配备了思考与练习。

本书由宋冬梅担任主编,崔琳琳担任副主编。宋冬梅负责第 6 章至第 14 章的编写;崔琳琳负责第 1 章至第 5 章的编写。宋冬梅负责本书的统稿和最后定稿工作。

统计分析原理和 Excel、SPSS 软件的操作涉及很多基本原理、方法,因此在本书的编写过程中参阅较多国内外已经出版的相关著作(详见参考文献)。在此,谨向这些著作的作者表示衷心的感谢。

盐城工学院、南京大学出版社对本书的立项和出版给予支持和帮助,在此表示衷心的感谢!

由于编者水平有限,书中错误、疏漏之处在所难免,敬请广大读者、同仁批评指正。欢迎读者通过电子邮箱 sdmsds@163.com 与我们联系。

编 者
2017 年 1 月

本书所有数据资源均在此二维码内,请扫码,下载:



目 录

第一部分 基础篇

第 1 章 Excel 统计分析功能概述	1
1.1 Excel 的界面	1
1.2 Excel 中的统计函数及功能	2
1.3 数据分析子菜单的加载及功能	4
【本章小结】	6
第 2 章 统计数据的整理	7
实验一 分布数列的编制	7
实验二 统计图的绘制	11
【本章小结】	17
【思考与练习】	17
第 3 章 统计数据特征的描述	18
实验一 原始数据特征的描述	18
实验二 分布数列特征的描述	21
【本章小结】	22
【思考与练习】	22
第 4 章 相关分析和回归分析	24
实验一 相关分析	24
实验二 回归分析	28
【本章小结】	31
【思考与练习】	32
第 5 章 时间序列分析	33
实验一 时间序列分析指标的计算	33

实验二 时间序列的长期趋势分析	36
【本章小结】	39
【思考与练习】	39
第二部分 提高篇	
第 6 章 SPSS 统计分析软件概述	41
1.1 SPSS 的界面	41
1.2 利用 SPSS 进行数据分析的步骤	45
【本章小结】	46
第 7 章 SPSS 数据文件的建立与预处理	47
实验一 建立数据文件	47
实验二 数据文件的编辑	54
实验三 SPSS 数据加工	61
【本章小结】	67
【思考与练习】	68
第 8 章 统计图形的制作与编辑	69
实验一 条形图及其制作	69
实验二 直方图及其制作	74
【本章小结】	78
【思考与练习】	78
第 9 章 描述性统计分析	79
实验一 频数分析	79
实验二 描述性分析	85
实验三 探索性分析	88
实验四 交叉列联表分析	96
【本章小结】	102
【思考与练习】	102
第 10 章 假设检验	103
实验一 单样本 T 检验	103
实验二 两独立样本 T 检验	106
实验三 两配对样本 T 检验	109

【本章小结】	111
【思考与练习】	112
第 11 章 方差分析	113
实验一 单因素方差分析	113
实验二 多因素方差分析	120
实验三 协方差分析	129
【本章小结】	134
【思考与练习】	134
第 12 章 相关分析和回归分析	136
实验一 相关分析	136
实验二 偏相关分析	140
实验三 一元线性回归分析	142
实验四 多元线性回归分析	151
实验五 曲线估计	158
实验六 二元 Logistic 回归分析	162
【本章小结】	170
【思考与练习】	170
第 13 章 聚类分析和判别分析	172
实验一 两步聚类	172
实验二 K 平均值聚类	177
实验三 系统聚类	182
实验四 判别分析	188
【本章小结】	194
【思考与练习】	195
第 14 章 主成分分析和因子分析	196
实验一 主成分分析	196
实验二 因子分析	206
【本章小结】	216
【思考与练习】	216
参考文献	217

第一部分 基础篇

第 1 章 Excel 统计分析功能概述

微软公司的 Excel 是一个快速制表、将数据图表化以及进行数据分析和管理的工具软件包。在当今众多的电子表格软件中,Excel 以其先进的技术、强大的功能、良好的可操作性、简单易用,赢得了用户的广泛认可。Excel 可以管理、组织纷繁复杂的数据,对数据进行分析处理,并能以图表形式给出分析结果。尤其重要的是,Excel 提供了超强的统计分析程序,以分析工具库和统计函数的形式来提供统计分析功能,范围涵盖了最基本的统计分析。Excel 在管理统计学中的应用主要体现在三个方面:一是借助 Excel 中的统计函数实现统计分析指标的计算;二是借助 Excel 加载宏中的分析工具库提供的数据分析功能实现统计数据整理及分析;三是借助 Excel 插入菜单中的图表实现统计图的绘制以及借助单元格的公式编辑计算功能实现统计分析指标的计算等。本章就 Excel 中统计函数的表达式、统计分析功能以及如何加载数据分析、使用数据分析工具进行简要介绍。

1.1 Excel 的界面

现以 Excel 2007 为例,启动 Excel 后可看到它的主界面。Excel 2007 的界面包括:标题栏、快速访问工具栏、菜单选项卡、功能区、工作表等(见图 1-1)。Excel 的统计分析功能主要分布在“插入”、“公式”、“数据”等菜单选项卡中。

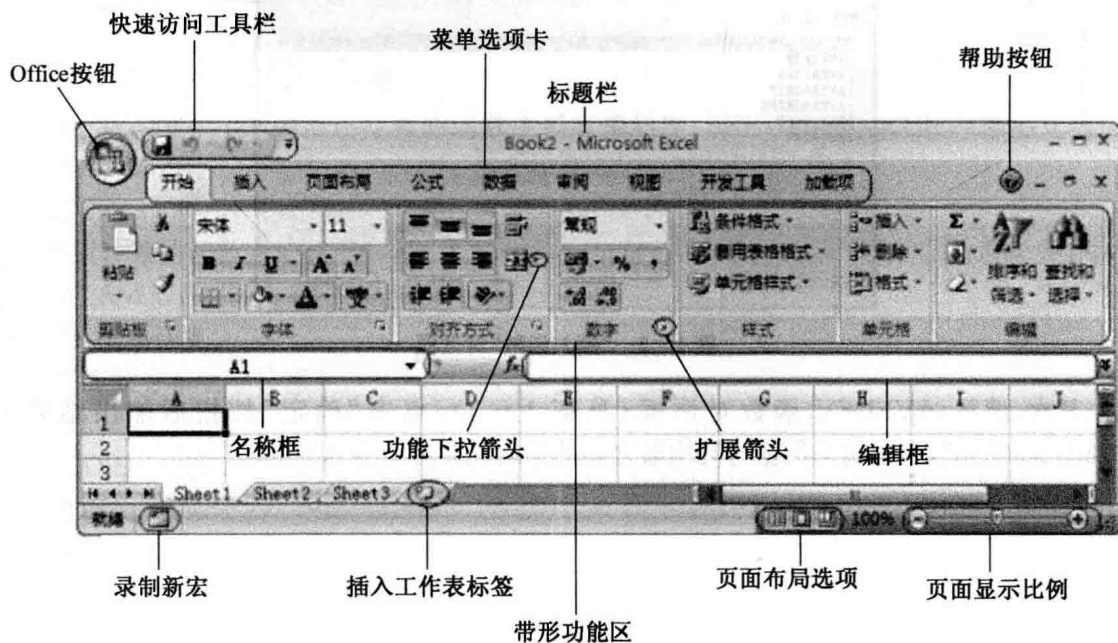


图 1-1 Excel 2007 的界面

1.2 Excel 中的统计函数及功能

1. 统计函数表达式及功能

Excel 中的统计函数共有 80 个(详见前言二维码中学习资料“Excel 统计函数”),现以“公式”菜单下“插入函数”子菜单中的统计函数中的第一个“AVEDEV”为例,说明其表达式、功能及应用如下:

AVEDEV 函数用途:返回一组数据与其平均值的绝对偏差的平均值,该函数可以评测数据的离散度。

语法:AVEDEV(number1,number2,...)。

参数:Number1、number2、... 是用于计算绝对偏差平均值的一组参数,参数的个数可以有 1 到 255 个,可以用单一数组(即对数组区域的引用)代替用逗号分隔的参数。

说明:输入数据所使用的计量单位将会影响函数 AVEDEV 的计算结果。参数必须是数字或者包含数字的名称、数组或引用。逻辑值和直接键入到参数列表中代表数字的文本被计算在内。如果数组或引用参数包含文本、逻辑值或空白单元格,则这些值将被忽略;但包含零值的单元格将计算在内。

平均偏差的公式为:
$$\frac{1}{n} \sum |x - \bar{x}|$$

例如,在工作表 A1~A6 单元格分别输入 4、5、6、5、4、3 六个数据,使用 AVEDEV 函数计算其平均偏差。首先,点击“公式”菜单下“插入函数”子菜单,选择统计函数,选择“AVEDEV”(见图 1-2)。

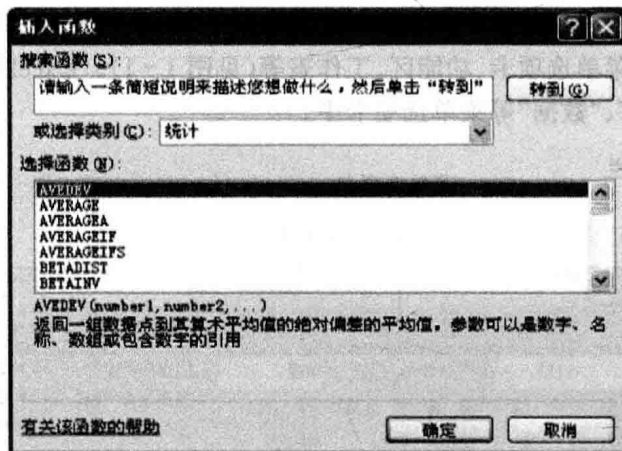


图 1-2 “插入函数”对话框

其次,填写 AVEDEV 函数对话框(见图 1-3),点击“确定”按钮后输出结果为“0.833 333 333”,表示每个数据与其均值之间的平均绝对离差为 0.833 333 333。

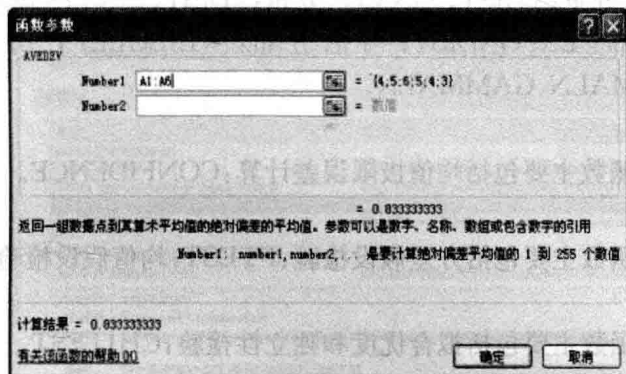


图 1-3 “AVEDEV 函数”对话框

其他统计函数的表达式、用途、语法等参见学习资料:Excel 统计函数。

2. 统计函数的统计分析功能

概括起来,Excel 中的统计函数的统计分析功能主要体现在以下几方面:

(1) 频数分布处理。

用于频数分布处理的函数主要有:COUNTIF、FREQUENCY。

(2) 计算描述统计量。

① 集中趋势计算。

用于集中趋势计算的函数主要有算术平均数: AVERAGE、AVERAGEA; 几何平均数: GEOMEAN; 调和平均数: HARMEAN; 中位数: MEDIAN; 众数: MODE; 四分位数: QUARTILE; K 百分比数值点: PERCENTILE; 内部平均值: TRIMMEAN。

② 离散程度计算。

用于离散程度计算分析的函数主要有平均差: AVEDEV; 样本标准差: STDEVA、STDEV; 总体的标准偏差: STDEVP、STDEVPA; 样本方差: VAR、VARA; 总体方差: VARP、VARPA; 样本偏差平方和: DEVSQ。

③ 数据分布形态的描述计算。

用于数据分布形态的描述计算的函数主要有偏斜度: SKEW; 峰度: KURT; 标准化值 z : STANDARDIZE。

④ 数值计算。

数值计算的函数主要包括计数: COUNT、COUNTA; 极值: MAX、MAXA、MIN、MINA、LARGE、SMALL; 排序: RANK、PERCENTRANK。

(3) 概率计算。

① 离散分布概率计算。

用于离散分布概率计算的函数主要包括排列: PERMUT; 概率之和: PROB; 二项分布: BINOMDIST、CRITBINOM、NEGBINOMDIS; 超几何分布: HYPGEOMDIST; 泊松分布: POISSON。

② 连续变量概率计算。

用于连续变量概率计算的函数主要包括正态分布: NORMDIST、NORMINV; 标准正态分布: NORMSDIST、NORMSINV; 对数正态分布: LOGINV、LOGNORMDIST; 卡方分布:

CHIDIST、CHIINV; t 分布: TDIST、TINV; F 分布: FDIST、FINV; β 概率分布: BETADIST、BETAINV; 指数分布: EXPONDIST; 韦伯分布: WEIBULL; Γ 分布: GAMMADIST、GAMMAINV、GAMMALN、GAMMALN。

(4) 参数估计。

用于参数估计的函数主要包括均值极限误差计算: CONFIDENCE。

(5) 假设检验。

用于假设检验的函数主要包括方差假设检验: FTEST; 均值假设检验: TTEST、ZTEST。

(6) 卡方检验。

用于卡方检验的函数主要包括拟合优度和独立性检验: CHITEST。

(7) 相关、回归分析。

用于相关、回归分析的函数主要包括相关分析: COVAR、CORREL、PEARSON、FISHER、FISHERINV; 线性回归分析: FORECAST、RSQ、LINEST、INTERCEPT、SLOPE、STEYX、TREND; 曲线回归分析: LOGEST、GROWTH。

1.3 数据分析子菜单的加载及功能

1. 数据分析子菜单的加载

Excel 中如果没有数据分析选项, 可以进行加载。

(1) Excel 2003 版本的, 可以通过工具菜单中的加载宏加载, 可在工具菜单下选择加载宏, 在弹出的对话框中选择分析工具库, 便可出现数据分析选项, 具体步骤为: 点击“工具”菜单, 选择“加载宏”, 选中“分析工具库”(打钩)。如果 Excel 2003 组件齐全, 就会直接加载, 加载后工具菜单下多一个数据分析模块; 如果 Excel 2003 组件不全, 则需进行光盘安装。

(2) Excel 2007 版本的, 可以通过自定义快速访问工具栏加载。点击 Excel 2007 界面最上一行自定义快速访问工具栏(右边的下拉箭头)(见图 1-4)。点击后选择“其他命令”, 出现 Excel 选项对话框, 点击左边选项中的“加载项”后, 点击右边的“分析工具库”, 再点击右下方“确定”按钮, 可完成了“数据分析”子菜单的加载(见图 1-5)。

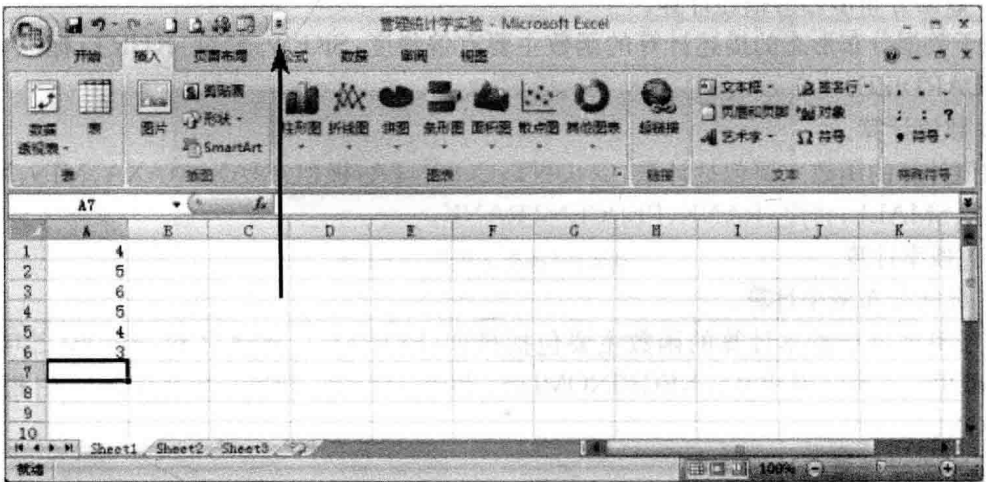


图 1-4 自定义快速访问工具栏



图 1-5 “Excel 选项”对话框

2. “数据分析”工具的统计分析功能

Excel 软件提供了 15 个数据分析工具,称为“分析工具库”。在进行统计分析时只需为每一个分析工具提供必要的数据和参数,分析工具就会使用适宜的统计函数,在输出表格中显示相应的结果。因此,使用分析工具比使用统计函数更加简便,可节省操作步骤。其中,有些分析工具在生成输出表格时还可以同时生成图表。15 个数据分析工具的统计分析功能主要体现在以下几个方面:

(1) 统计绘图、制表。

利用 Excel 分析工具库中的“直方图”分析工具,可以进行频数分布处理和绘制直方图。

(2) 计算描述统计量。

利用 Excel 分析工具库中的“描述统计”分析工具,可以计算常用的集中趋势描述、离散程度描述、数据分布特征描述及其他基本统计量。

① 集中趋势描述的指标主要有:平均值、中位数、众数。

② 离散程度描述的指标主要有:极差(全距)、标准误差(相对于平均值)、标准偏差、方差。

③ 数据分布特征描述的指标主要有:峰值、偏斜度。

④ 数值统计的指标主要有:最小值、最大值、总和、总个数。

利用“排位与百分比排位”分析工具,可以产生一个数据列表,在其中罗列给定数据各个数值的大小次序排位和相应的百分比排位,用来分析数据中各数值间的相互位置关系。

(3) 参数估计。

利用“描述统计”分析工具,可以计算正态分布下方差未知的样本均值极限误差,从而实现单一总体均值的区间估计。

(4) 假设检验。

利用 F-检验分析工具、t-检验分析工具、z-检验分析工具,可以进行总体均值、方差的假设检验。

① 两个总体均值检验:包括利用“z-检验:双样本平均差检验”分析工具,可以在两总体方差已知时,进行两总体均值的假设检验;利用“t-检验:双样本等方差假设”分析工具,可以在

两正态总体方差未知但相等时,进行两总体均值的假设检验;利用“t-检验:双样本异方差假设”分析工具,可以在两正态总体方差未知且不相等时,进行两总体均值的假设检验;利用“t-检验:成对双样本均值分析”分析工具,可以进行均值的成对检验。

② 两个总体方差检验:利用“F-检验:双样本方差分析”分析工具,可进行两个总体的方差检验。

(5) 方差分析。

利用方差分析工具,可进行单因素和双因素的方差分析。

① 单因素方差分析:利用“单因素方差分析”分析工具,可以对两个以上总体均值的显著性差异进行检验。

② 双因素方差分析:利用“无重复双因素分析”分析工具,可以对两个因素各自对实验结果影响的显著性进行检验;利用“可重复双因素分析”分析工具,可以对两个因素各自对实验结果及两因素交互作用对实验结果影响的显著性进行检验。

(6) 相关分析、回归分析。

利用“相关系数”分析工具和“协方差”分析工具,可以对两个及两个以上变量间的相关关系进行分析计算。利用“回归分析”分析工具,可以建立简单线性回归和多元线性回归模型,并可对模型的有效性进行检验分析。

(7) 时间序列分析。

利用“指数平滑”分析工具,可对时间序列基于前期预测值导出相应的新预测值,进行趋势分析。利用“移动平均”分析工具,可对时间序列数据进行移动平均处理,进行数据的趋势分析。

(8) 抽样。

利用“随机数发生器”分析工具,可以按照用户选定的分布类型,在工作表的特定区域中生成一系列独立随机数。利用“抽样分析”分析工具,可以以输入区域为总体构造总体的一个样本。当总体太大而不能进行处理或绘制时,可以选用具有代表性的样本。如果确认输入区域中的数据是周期性的,还可以对一个周期中特定时间段中的数值进行采样。例如,如果输入区域包含季度销售量数据,以4为周期进行。

(9) 数据变换。

利用“傅里叶分析”分析工具,可以对数据进行快速傅里叶变换和逆变换,变换后的数据用于相关系数检验和分析。

本章小结

1. 了解 Excel 的界面,“公式”菜单中“插入函数”子菜单,理解 Excel 统计函数的含义。
2. 理解 Excel 数据分析子菜单的加载及主要分析工具的统计分析功能。



第 2 章 统计数据的整理

本章主要介绍如何利用 Excel 的“公式”、“插入”、“数据”等菜单中的子菜单“插入函数”、“图表”、“数据分析”等实现统计数据的整理,根据研究需要编制分布数列,同时运用统计图表展示统计分布数列。

实验一 分布数列的编制

【实验目的】

学会对统计数据的整理,并掌握统计分布数列的编制方法。

【实验内容】

【例 2-1】 表 2-1 是某电脑公司某年在 108 个商场的销售量数据。

表 2-1 某电脑公司某年在 108 个商场的销售量数据

(单位:台)

224	186	206	220	218	212	196	191	271
169	200	233	270	212	227	235	238	172
257	202	208	214	221	251	211	192	224
179	253	235	281	222	227	235	182	193
214	202	209	184	223	228	192	193	180
268	204	198	215	223	210	236	194	274
262	216	210	173	214	229	236	196	184
215	205	279	215	224	208	209	235	199
258	280	210	226	284	233	238	197	185
216	206	228	206	224	227	240	271	209
210	252	248	217	246	234	241	198	187
235	206	211	218	226	234	242	198	244

要求:用 Excel 编制分布数列(组距取 20)。参见数据文件:Excel 2-1. xls。

【实验步骤与实验结果】

1. 使用统计函数——COUNTIF 进行汇总

(1) 输入数据(见图 2-1)。

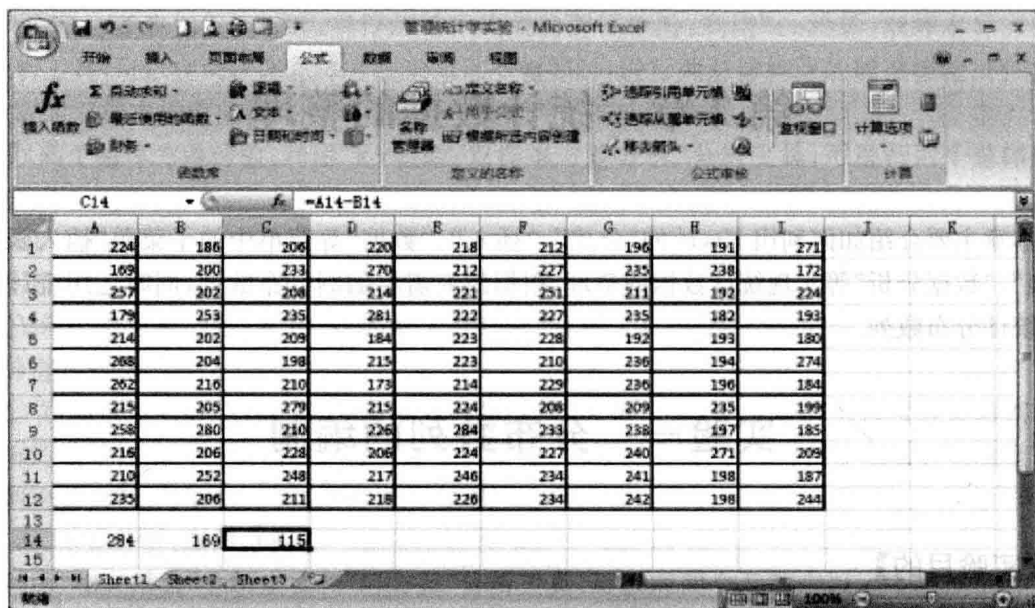


图 2-1 输入原始数据

(2) 确定最大值与最小值,利用函数 MAX(见图 2-2)和函数 MIN 函数(见图 2-3)确定,最大值为 284,最小值为 169,全距为 115(=284-196)。

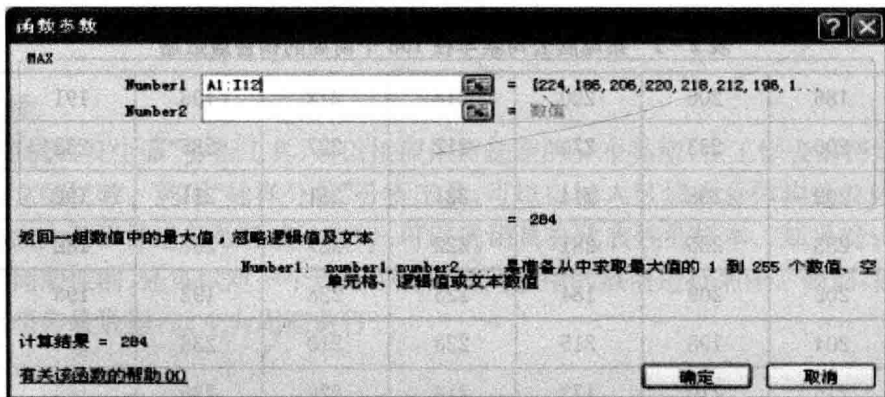


图 2-2 “MAX 函数”对话框

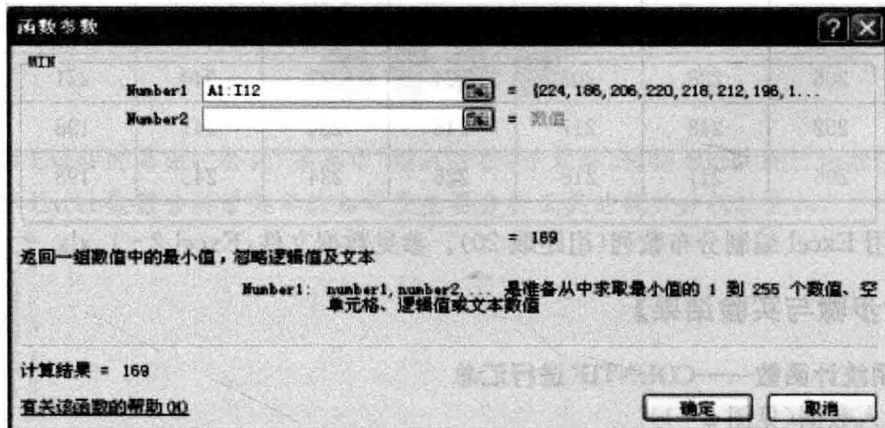


图 2-3 “MIN 函数”对话框

(3) 确定组数与组距。用户确定,本例组数可定为7个组,组距可定为20。

(4) 确定分组状况。上下组限由用户确定,根据本例数据,可分为以下7个组:160~180,180~200,200~220,220~240,240~260,260~280,280~300。

(5) 汇总每组频数(利用统计函数 COUNTIF 汇总)。

将光标移动到输出汇总结果的单元格,点击“插入函数”子菜单,选择类别“统计”,在选择函数中找到“COUNTIF”,填写“COUNTIF”对话框(见图2-4),以统计第一组单位数为例,“Range”中输入原始数据区域,本例为A1:J12,“Criteria”为汇总时的组上限条件区域,第一组为160~180之间的,输入“<180”,点击“确定”按钮,汇总出第一组频数,为4。

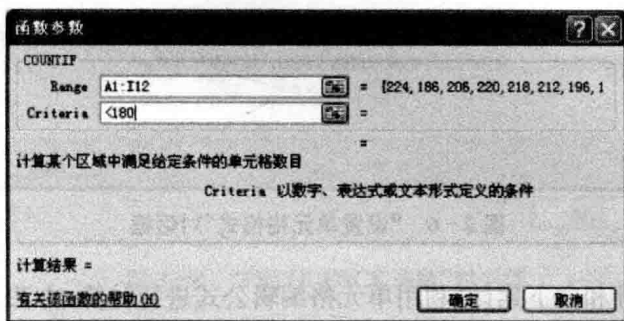


图 2-4 “COUNTIF 函数”对话框

依次汇总其他组的频数。需要注意,采用 COUNTIF 汇总结果,组限大的组包含组限小的组的频数,因此组限大的组的频数应为:汇总结果减去组限小的组的频数。

(6) 计算各组频率(利用单元格编辑公式进行计算,见图2-5)。

14	分组	各组单位数	频率(%)
15	160-180	4	=B15/\$B\$22*100
16	180-200	20	
17	200-220	33	
18	220-240	30	
19	240-260	11	
20	260-280	7	
21	280-300	3	
22	合计	108	

图 2-5 频率计算公式

比如第一组频率的计算公式为:

$$=B15/ \$ B \$ 22 * 100$$

此公式编辑好后,点击“回车”,得到此组频率。由于一般情况数值保留两位小数,可将光标移至此单元格后点击鼠标右键,点击“设置单元格格式”,设置数值保留位数(见图2-6)。这样第一组的频率公式编辑计算完成,其他组可采用光标拖拽方式复制此公式,具体操作:将光标移至编辑好公式的单元格,当光标右下方空心“十”字形变成实心“十”字形时,进行拖拽,复制此公式计算各组频率。

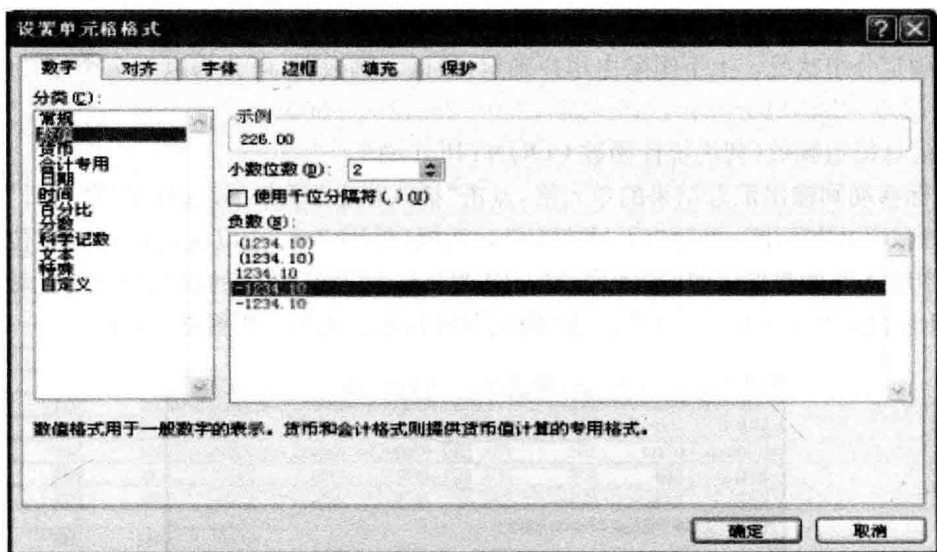


图 2-6 “设置单元格格式”对话框

(7) 进行向上累计和向下累计(利用单元格编辑公式进行计算,方法与频率计算类似),编制统计表描述统计数据整理结果(见表 2-2)。

2. 使用统计函数——FREQUENCY 进行汇总

函数“FREQUENCY”可以汇总在给定的组距范围内的频数。组数、组距、分组状况(上下组限)同上述步骤。“FREQUENCY”可以一次汇总所有组的频数,其步骤如下:

(1) 在单元格中输入每组上限。由于本例中采用了重叠组限,根据“上限不在内”的统计规则,每组上限比设定的稍小一点,此例中可分别取“179、199、219、239、259、279、300”(用单元格 G16~G22 表示),频数的输出区域可以设定为 B16~B22(见图 2-7)。

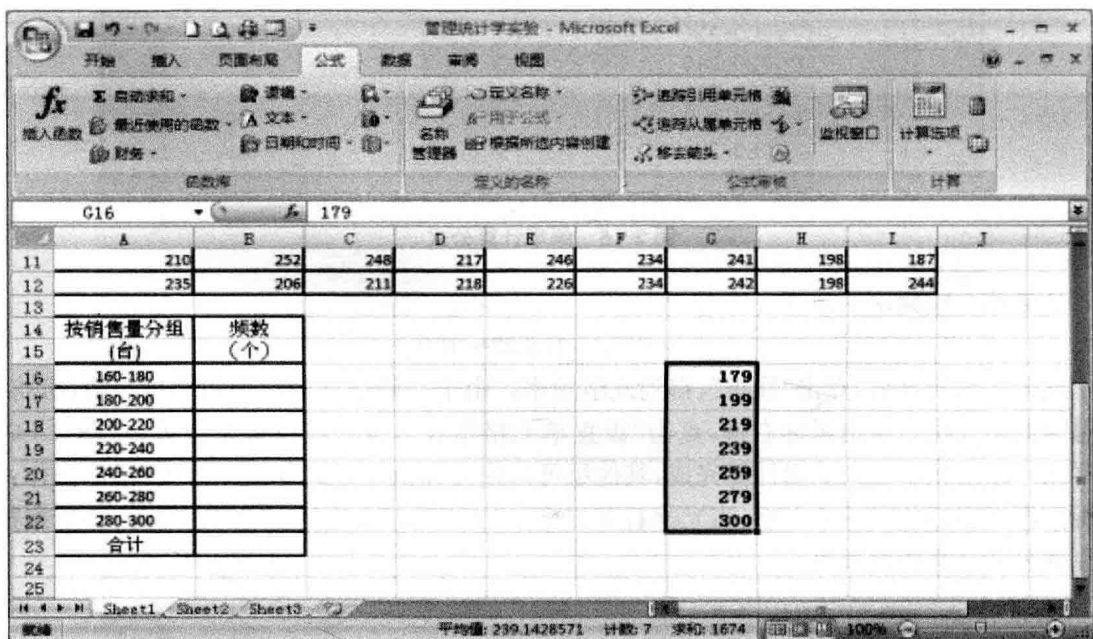


图 2-7 各组组限及频数输出区域图