

主 编：李红松 邓旭东
副主编：殷志平 刘俊武 李洪斌

统计数据分析 方法与技术



经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

主 编：李红松 邓旭东
副主编：殷志平 刘俊武 李洪斌

统计数据分析 方法与技术



经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

图书在版编目 (CIP) 数据

统计数据分析方法与技术/李红松等编著. —北京: 经济管理出版社, 2014.11
ISBN 978-7-5096-3432-5

I. ①统… II. ①李… III. ①统计数据—统计分析 IV. ①0212.1

中国版本图书馆 CIP 数据核字 (2014) 第 237605 号

组稿编辑: 杨国强
责任编辑: 杨国强 张瑞军
责任印制: 司东翔
责任校对: 赵天宇

出版发行: 经济管理出版社
(北京市海淀区北蜂窝 8 号中雅大厦 11 层 100038)

网 址: www.E-mp.com.cn
电 话: (010) 51915602
印 刷: 三河市延风印装厂
经 销: 新华书店
开 本: 720mm × 1000mm/16
印 张: 21
字 数: 412 千字
版 次: 2014 年 12 月第 1 版 2014 年 12 月第 1 次印刷
书 号: ISBN 978-7-5096-3432-5
定 价: 49.00 元

·版权所有 翻印必究·

凡购本社图书, 如有印装错误, 由本社读者服务部负责调换。

联系地址: 北京阜外月坛北小街 2 号

电话: (010) 68022974 邮编: 100836

前 言

当前,数据挖掘与分析无论是在理论研究领域还是在实际工作中都已得到高度重视,掌握基本的数据分析方法和分析工具已成为管理者应具备的基本技能。为了顺应这一趋势,在高校的教学应注重培养学生掌握方法的应用原理以及方法应用的实现工具并重,这是培养学生知识和技能不可或缺的两个方面。

数据的处理与分析包括数据的获得、数据的整理与展示、数据的分析等环节,其中,数据的分析环节需要运用大量的统计分析方法,而这些方法如果依靠分析工具由计算机自动完成,可大大提高工作效率并确保分析结果的准确性。

本书以统计数据分析的全过程为主线,结合笔者多年的教学经验和体会,介绍了数据的来源与获得、数据的整理与展示、数据处理与分析等内容,其中的分析方法涵盖了统计学中的基本统计方法,主要针对高校的统计学和数据分析类课程教学需要而编写。

本书内容力求简明、通俗易懂。本书注重理论分析与方法应用相结合,在介绍方法应用的过程中,尽量结合 Excel 和 SPSS 软件等广泛使用的数据分析工具,突出工具实现的操作过程,使学习者容易上手,其中,Excel 软件的操作介绍适用于 Office 2003 及以上版本,SPSS 软件的操作介绍适用于 17.0 及以上版本。

全书由李红松、邓旭东担任主编,殷志平、刘俊武、李洪斌担任副主编,各章编写分工如下:第一、二、三、五、七章由李红松、邓旭东撰写;第四章由李红松、刘俊武、殷志平撰写;第六章由李洪斌撰写;第八章由殷志平、刘俊武撰写。本书编写过程中,得到童泽平老师的帮助与支持,并提出了有益的建议,在此表示感谢!

由于笔者水平有限,书中难免存在疏漏、错误和不足之处,希望得到同行和读者的反馈与指正,以便对教材进一步完善。为方便教师教学需要,本书有配套的 PPT 课件可供使用,欢迎使用者索取课件与交流,联系邮箱:WUSTLHS@163.com。

编 者

2014 年 8 月

目 录

第一章 数据分析的基本问题	001
第一节 统计数据类型	001
第二节 数据的来源	005
第三节 数据分析的步骤	008
第四节 常用的数据分析工具	009
第二章 数据的整理与展示	017
第一节 数据的预处理	017
第二节 数据的分组	024
第三节 数据的汇总技术	029
第四节 数据的图形展示	033
第三章 数据的描述性分析	043
第一节 数据的集中趋势分析	043
第二节 数据的离散程度分析	054
第三节 数据分布的偏态与峰度	061
第四章 抽样数据的推断分析	067
第一节 随机抽样方法与抽样分布	067
第二节 总体参数的抽样估计	083
第三节 总体参数的假设检验	094
第四节 方差分析方法	115
第五章 数据的相关与回归分析	137
第一节 相关分析	137
第二节 一元线性回归分析	148

第三节 多元线性回归分析	162
第四节 非线性回归模型	170
第六章 时间序列数据分析与预测	177
第一节 时间序列概述	177
第二节 时间序列的描述分析	180
第三节 长期趋势分析	186
第四节 季节变动分析	192
第五节 循环变动分析	196
第六节 时间序列的预测	198
第七章 统计指数与因素分析法	207
第一节 统计指数的概念与分类	207
第二节 综合指数的编制	209
第三节 平均数指数	213
第四节 指数体系与因素分析	218
第八章 数据分类与降维技术	229
第一节 聚类分析	229
第二节 判别分析	247
第三节 主成分分析	262
第四节 因子分析	275
思考练习题参考答案	295
附 表	303
参考文献	333

第一章 数据分析的基本问题

第一节 统计数据的基本类型

量化分析所依据的数据是对客观现象属性或特征的描述，是对研究对象观测记录或计算处理后得到的结果。每个数据都具有一定的内涵和外延，并由特定的时间、地点及对象概括，比如上海证券交易所股票价格综合指数 2014 年 6 月末收盘点数 2048.3，2013 年我国国内生产总值（GDP）56.9 万亿元，青岛海尔电器股份有限公司的行业类型为家电制造业（隐含了数据时间点为调查时点），等等。

对数据的分析必须首先明确数据的类型，不同类型的数据需要采用不同的处理和分析方法。从不同的角度出发，数据可进行不同的划分，常见的分类视角有四种，分别是计量尺度分类视角、获取方法分类视角、时空分类视角和研究对象范围视角，每种视角下的分类如图 1-1 所示。

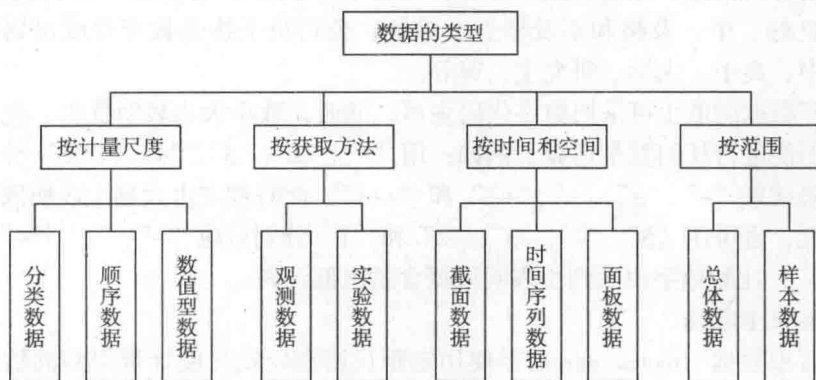


图 1-1 数据的分类

一、分类数据、顺序数据和数值型数据

数据计量的尺度按照由低级到高级、由粗略到精确的标准分为定类尺度、定序尺度、定距尺度和定比尺度。按照计量尺度的不同,数据可分为分类数据(定类尺度)、顺序数据(定序尺度)和数值型数据(定距尺度和定比尺度)。统计中通常又把分类数据和顺序数据称为品质标志数据,将数值型数据称为数量标志数据。

1. 分类数据

定类尺度是按照客观现象的某种属性对其进行分类或分组计量,各类各组之间属于并列、平等并且互相排斥的关系。依据定类尺度计量得到的数据属于分类数据(categorical date),通常用文字表示。例如:将国民经济按产业划分为第一产业、第二产业和第三产业,将学生按性别划分为男生和女生,将交通信号灯的颜色分为红、黄和绿色,等等。

分类数据的值不反映各类的优劣、量的大小或顺序,有时候为了便于数据处理,而采用数字代码表示各个类别,但这种数字并无大小区分,也不能进行任何数学运算。比如:用“1”和“0”分别表示“男生”和“女生”;用“1”、“2”和“3”分别表示“第一产业”、“第二产业”和“第三产业”;等等。

2. 顺序数据

定序尺度是对客观现象按照等级差或顺序差进行分类或分组计量,与定类尺度不同,定序尺度划分的各类各组之间有优劣、量的大小或先后顺序之分。依据定序尺度计量得到的数据属于顺序数据(rank date)。例如:某种疾病治疗方法按疗效划分为治愈、显著有效、好转和无效;学生按某门课程的考试成绩划分为优秀、良好、中、及格和不及格五个等级;公司员工按受教育程度可划分为小学、初中、高中、大学、研究生;等等。

顺序数据的值也可采用数字代码表示,此时,数字大小表明优劣、先后的意义,但不能进行任何数学运算。例如:用“1”、“2”、“3”、“4”和“5”分别表示尿糖化验结果“-”、“±”、“+”、“++”和“+++”,此时数字由大到小表明尿糖含量由高到低;也可用“5”、“4”、“3”、“2”和“1”分别对应“-”、“±”、“+”、“++”和“+++”,此时数字由大到小表明尿糖含量由低到高。

3. 数值型数据

数值型数据(metric date)是使用定距尺度和定比尺度计量得到的数据。数值型数据用数值表示,依据定距尺度得到总量数据,依据定比尺度得到相对数数据和平均数数据,两者都属于数值型数据。

定距尺度是用数值对现象各类别之间间距进行精确计量测度,通常用于测度

现象的绝对总量、规模和水平，如产量、人口数、销售额、净利润等方面的数据。总量数据根据其数值是否存在间断可进一步分为离散型和连续型数据，数值存在间断、只能用整数表示的数据称为离散型数据，如员工人数、零件个数、企业数等；数值是连续的、可以进行无限分割的数据称为连续型数据，如身高、销售收入、工资、进出口额等。总量数据还可根据反应的时间状态不同分为时点数据和时期数据，时点数据描述现象在某一时刻的状态，不能简单加减，如各月月末库存额相加没有意义；时期数据反映现象在一段时间内累积形成的结果，可以累加，比如第一季度产量与第二季度产量相加表示上半年产量。

定比尺度是在定距尺度的基础上，确定相应的比较基数，然后将两种相关的数据加以对比而形成的比值，使用定比尺度计量得到的数据有相对数数据和平均数数据两种。相对数数据通常用于测度现象的结构、比重、速度、密度、强度等数量关系，如及格率、增长率、资金周转率、利润率等方面的数据；平均数数据反映总体的平均水平，如人均产量、平均年龄等方面。相对数数据和平均数数据不能像总量数据那样直接进行加减，其数学运算时必须结合具体情况，依据特定规则进行。

二、观测数据和实验数据

按照获取数据方法的不同，数据可分为观测数据和实验数据。

1. 观测数据

观测数据 (observational data) 是通过调查或观测而收集到的数据，它是在没有对事物进行人为控制的条件下得到的，社会经济现象方面的数据几乎都是观测数据。例如：民意调查获得的数据；统计报表提交的数据；证券交易数据；等等。

2. 实验数据

实验数据 (experimental data) 是采用科学实验方式，通过控制实验对象而收集到的数据。在实验中，实验环境是受到严格控制的，数据的产生一定是某一约束条件下的结果。例如：新药研制中的临床实验数据；农作物育种培育中种植实验的数据。自然科学研究中通常采取实验的方法获得数据。

三、截面数据、时间序列数据和面板数据

按照被描述对象的数据是否与时间有关，可以将数据分为截面数据、时间序列数据和面板数据。

1. 截面数据

截面数据 (cross-sectional data) 是采用空间维度 (横截面), 描述研究对象不同空间的个体在某一相同时间点表现出的特征和属性。例如: 全国 31 个省市市区 2013 年的国内生产总值和固定资产投资数据; 某高校 2014 年入学新生的身高和体重数据; 等等。

在使用截面数据研究问题时, 有两个方面的问题需要注意: ①由于个体或地域本身的异质性, 不同空间个体会存在较大差异, 这种差异是不可避免的, 属于“无法观测的异质性”, 但引起异质性的因素不能差异太大; ②截面数据要求数据的统计标准及取样时间必须一致。

2. 时间序列数据

时间序列数据 (time series data) 是指对同一对象在不同时间连续观察所取得的数据。时间序列数据采用时间维度描述对象发展变化的轨迹和规律。例如: 1978~2013 年全国国内生产总值和固定资产投资数据; 某企业 2010~2013 年各月的销售收入数据; 等等。

利用时间序列数据时, 一是要求数据的统计标准在各个观察时间上必须一致, 二是对象在不同时间上的空间范围要一致。如对上市公司股票价格变动进行研究时, 在研究期内, 同一上市公司是否存在重大资产重组、是否因利润分配产生股价的除权除息等, 必须综合考虑。

3. 面板数据

面板数据 (panel data) 是由时间序列数据和截面数据交叉形成的, 是从时间和空间两个维度对研究对象不同空间的个体在不同时间连续观察取得的数据。由于对此类数据进行分析时, 多采用面板模型 (panel model), 故被称为面板数据。例如: 5 家房地产公司在 10 个会计年度投资与利润数据; 全国 31 个省市市区 1978~2013 年的国内生产总值和固定资产投资数据; 等等。

面板数据既可以分析个体之间的差异情况, 又可以描述个体的动态变化特征。比如, 利用全国 31 个省市市区 1978~2013 年的国内生产总值和固定资产投资数据, 既可以分析我国国内生产总值和固定资产投资之间随时间变化的关系, 也可以分析这种关系在各地区是否存在不同。

四、总体数据与样本数据

按照被描述对象涵盖的范围, 数据可分为总体数据和样本数据。所谓总体就是研究对象内全部个体的集合, 每一个体都是一个总体单位; 样本是由总体中的部分个体组成的集合。如果样本是从总体中随机产生的, 称为随机样本, 统计推断要求的样本必须是随机样本。如调查公司欲了解一种饮料在某地区消费者中的

受欢迎程度,通过电话随机采访了120位消费者,这一研究的总体是该地区的所有消费者,每个消费者是总体单位,电话采访的120位消费者构成了样本。

1. 总体数据

如果掌握的是总体中所有个体的数据则称为总体数据(population data),总体数据一般情况下未知,有些总体数据可通过全面调查的方式获得,但有些总体数据即使通过全面调查也不能获得。比如,某大学想要了解2014年入学新生的性格与家庭经济状况存在多大程度的关联,则全部新生是研究总体,通过对每位新生进行问卷调查获得的数据是总体数据。

2. 样本数据

如果掌握的是总体中一个样本的数据则是样本数据(sample data)。随机样本可以较好地反映总体的分布,因而统计中常用样本的数据推断总体的数量特征。如果在上述大学新生性格与家庭经济状况关联程度的研究中,随机抽取100名新生进行问卷调查,100名新生组成了样本,调查获得的数据是样本数据。

第二节 数据的来源

数据的来源可以通过直接方式获得的第一手数据,也可以通过间接方式获得的已经存在的数据即二手数据,前者为数据的直接来源,后者为数据的间接来源。

一、数据的直接来源

如果数据的来源是通过直接组织的调查、观测和科学实验等方式获得,这样的数据属于第一手数据或者原始数据。当所要研究的对象没有现成的数据,或者虽然有现成的数据但数据可靠性存在问题而不能使用,这时需要通过调查、观测或者实验获得研究对象的数据。对于自然现象而言,科学实验和观测是取得数据的主要手段,对于社会经济现象而言,调查是获得数据的重要手段,例如客户满意度调查、电视节目收视率调查、大学生心理健康状况调查等。常见的数据调查方式有普查、统计报表、典型调查、重点调查和抽样调查等。

1. 普查

普查(census)是为了某种特定目的,专门组织的一次性全面调查。普查是基于特定目的、特定对象而进行的,用于收集现象在某一时刻状态下的数据。由于普查涉及面广、调查单位多,需要耗费相当大的人力、物力、财力和时间,通

常间隔较长时间进行一次。对于关系国情国力的基本数据，世界各国通常定期举行普查，以掌握特定社会经济现象的基本全貌。

我国目前已开展的全国普查有人口普查、经济普查、农业普查。自 1990 年开展的第四次人口普查起，之后每 10 年进行一次；自 1996 年底进行的第一次农业普查起，也是每 10 年进行一次；2004 年第一次开展了经济普查，之后每 5 年进行一次。

2. 统计报表

统计报表 (statistical reportforms) 是按国家统一规定的表式、统一的指标项目、统一的报送时间，自下而上逐级定期收集基本统计数据的调查方式。统计报表是我国政府部门收集社会经济统计数据的重要方式。我国大多数统计报表要求调查对象全部单位填报，属于全面调查范畴，所以又称全面统计报表。统计报表具有统一性、全面性、周期性和可靠性等特点。

目前我国统计报表由国家统计报表、业务部门统计报表和地方统计报表组成，报表内容涉及国民经济各行业、各部门，其中国家统计报表是统计报表体系的基本部分。按照报送周期不同统计报表又分为月度、季度和年度统计报表。

3. 典型调查

典型调查是根据调查目的和要求，在对总体进行全面分析的基础上，从全部单位中选取少数有代表性的单位进行深入调查的一种非全面调查方式。比如研究资源在促进县域经济发展中的作用时，可以从全国经济百强县中挑选具有资源优势的县作为典型县进行典型调查，搜集数据。

典型调查具有如下特点：第一，典型调查是选取少数有代表性的单位进行调查，由于调查单位少，因此省时省力；第二，典型调查适用于对现象进行深入细致的调查，既可以搜集反映研究对象属性和特征的数字资料，分析现象量的属性，也可以搜集不能用数字表示的文字资料，以深入分析数量关系形成的原因，并提出解决问题的思路 and 办法。

典型调查的关键是典型单位的选取。典型单位的选取与研究目的有关，如果研究目的是推广成功的经验，则典型单位从先进单位中产生；如果是为了总结教训，则典型单位从后进单位中产生；如果是为了近似了解总体的一般情况，则适合划类取典。

4. 重点调查

重点调查是从研究对象中选取部分重点单位进行调查以获得数据的一种非全面调查方式。重点单位是指对于所要研究的属性特征在总体中具有重要地位的单位。比如，要想了解我国房地产业当前的生产经营状况，重点单位是销售额、开工竣工面积在全国排名靠前的房地产企业，它们在全国近十万家房地产企业中只占百分之一，但房地产销售额可能占到全国的一半以上，选取排名前 100 的房地

产企业进行调查能够反映整个行业的基本情况。

重点调查省时、省力，当研究任务是为了了解研究对象的基本情况，并且总体存在重点单位时才适合采用重点调查收集数据。

5. 抽样调查

抽样调查 (sampling survey) 是按照随机性原则，从研究对象中抽选一部分单位（或者个体）进行调查，并据以对研究对象作出估计和推断的调查方法。在抽样调查中，研究对象是由许多单位组成的整体，称为总体或母体，抽出的部分单位组成的整体称为样本，样本包含的单位数称为样本容量。显然，抽样调查属于非全面调查。

抽样调查是用样本数据特征推断总体特征，必然产生代表性误差，即使严格遵守随机性原则，代表性误差也不可能完全消除，但在满足一定可信度的情况下，代表性误差可以通过样本容量加以控制。

抽样调查被用于某些不能进行全面调查的事物，或者理论可行但实际上不能进行全面调查的事物，比如存在破坏性实验的灯泡使用寿命调查、海洋渔业资源调查等。此外，许多研究虽然可以进行全面调查，但限于时间、成本等因素，运用抽样调查可能更好。

典型调查和重点调查虽然也属于利用样本数据反映总体特征的非全面调查，但其样本的产生属于非概率抽样，这与抽样调查中按随机性原则产生的样本完全不同。为了对两者加以区别，通常所指的抽样调查为随机抽样。

二、数据的间接来源

数据的间接来源是指使用别人调查的数据或者对原始数据进行加工整理后的数据，也称为二手数据。所有间接得到的数据都由原始数据过渡而来，间接来源的途径包括：公开的出版物，如年鉴、期刊、著作、报纸和报告等；官方统计部门、政府、组织、学校、科研机构等通过媒体公开的数据，如政府统计部门定期通过网站公布的各种统计数据。

使用间接来源的数据时应注意以下问题：一是要了解间接数据中变量的含义、计算口径、计算方法，以防止误用、错用他人的数据；二是引用间接数据时要注明数据的来源或出处。

第三节 数据分析的步骤

在已经获得研究分析所需要的数据后，接下来便进入数据的整理与分析阶段。数据分析的目的是从描述研究对象的数据信息中，发现其内在特征和规律。完整的数据分析过程包括数据的整理、数据的展示、分析方法选取、分析结果的评价等环节。

一、数据的整理

数据整理是将获得的数据进行加工汇总，以使之条理化的过程，数据整理的内容包括审核与订正、分组、汇总等。

1. 数据的审核与订正

数据的审核就是检查数据是否准确、完整，是保证数据质量的关键。准确性检查一般依靠逻辑常识判断数据是否存在错误，比如当数据中存在异常值时、当数据有不符合逻辑之处时，都要进行鉴别核实；对于存在数量关系的数据，还应核实是否存在计算上的错误。一旦发现数据存在错误，应对数据进行订正。

数据的完整性审核是检查应调查的对象中是否存在遗漏的单位，要调查的数据项目是否填写齐全等，如果有遗漏，则应想办法补充完整。

2. 数据的分组

所谓分组就是按照某种标准将数据分成几个不同的部分。对于分类数据和顺序数据，每一类实际上就是一个组，一般不需要做进一步的划分；对于数值型数据，有时候则需要对数据进行分组。关于数据的分组，在第二章中将有详细介绍。

3. 数据的汇总

数据的汇总是指采用某种标准分组后，将原始数据按组别分别加总归类的过程。比如：收集到了班上每位同学的英语考试分数，按不同的分数段分别计算人数，属于最简单的汇总；对于人口普查取得的原始数据进行汇总则相当复杂，需要采用专门的组织形式和技术方法。

二、数据的展示

已经整理好的数据需要按某种方式将其展示出来，选择合适的数据展示方式，不仅可以使数据的特征规律较好地得到体现，还可以达到形象、美观的效

果。数据的展示方式包括表格和图形。

表格是最基本的数据展示方式，既可以展示数据本身，也可以对数据进行整理汇总，还可以对数据进行分析并展示分析的结果。

相比表格而言，图形展示数据时，更简明、生动、形象，可以使数据的特征规律一目了然，但在数据的容量方面，图形不如表格。利用图形不仅可以展示数据，还可以对数据进行简单的分析。关于数据的图形展示与分析将在第二章中详细介绍。

三、数据分析方法的选取

对数据进行量化分析，必须选取恰当的分析方法，以揭示研究对象的本质特征和规律。数据分析既有数据特征的简单描述分析，也有复杂的推断量化分析，具体采用什么方法必须结合研究目的、数据的类型、数据的范围等因素综合确定。

描述分析是对研究对象的规模、水平、内部结构或比例、发展变化的速度、数据的集中趋势、离散性、分布特征等进行的一般性分析，具体可以运用图、表或计算综合指标等方法实现。

推断分析是运用特定的数量方法，就某一假设、总体未知的特征、现象变化的规律、现象之间的关系等方面进行分析，以验证假设或得出结论。具体方法有总体参数估计分析法、假设检验法、方差分析法、相关分析法、回归分析法、时间序列趋势分析法、聚类分析法、判别分析法、主成分分析法、因子分析法等。这些具体方法的应用将在本书的相关章节中进行系统介绍。

四、数据分析结果的评价

运用数据分析方法得到结果后，需要用语言文字对结果进行分析评价。评价的内容包括对结果包含的含义、特征和规律的解释，对存在的问题（或者成功的经验）及原因的分析，提出解决问题的措施或建议等。

第四节 常用的数据分析工具

运用量化分析工具，不仅可以快速、准确地对数据进行计算、汇总和处理，对于一些复杂的数据分析方法，往往需要运用专门的工具才能实现。因而量化分析工具是数据分析的有效助手。目前，运用最为广泛的初级量化分析工具软件是

Office 系列办公软件中的组件 Excel，此外，还有一些专门的分析工具软件如 SPSS、SAS、EViews、MATLAB 等。本书后续各章的实例分析中，将重点介绍 Excel 和 SPSS 的数据处理与分析功能的具体实现过程。

一、Excel 与数据分析简介

1. Excel 的特点及主要功能

Excel 是 Office 系列办公软件组件之一，自 Microsoft 公司 1985 年推出最初版本以来，经过不断升级，功能日益增强。它具有界面直观、简单易操作、数据存储和处理功能强大、数据格式兼容许多专门分析工具等特点，被广泛应用于社会经济管理的各个领域。

Excel 的主要功能包括数据存储管理、数据组织与运算、数据分析与预测、图表制作等，本书后续各章在介绍数据分析各种方法时，将会涉及这些功能的运用。

2. Excel 的工作表

Excel 文件也称工作簿，一个工作簿文件由许多单独的工作表构成，早期 Excel 版本的工作表最多 255 个，自 1997 版开始突破了这一限制，理论上可以有任意多个，但具体数量取决于计算机的物理内存。

打开或新建一个工作簿，界面为当前活动工作表（见图 1-2），包括标题栏、菜单栏、工具栏、地址栏、编辑栏及表格，工作表格由行列标示的单元格组成，

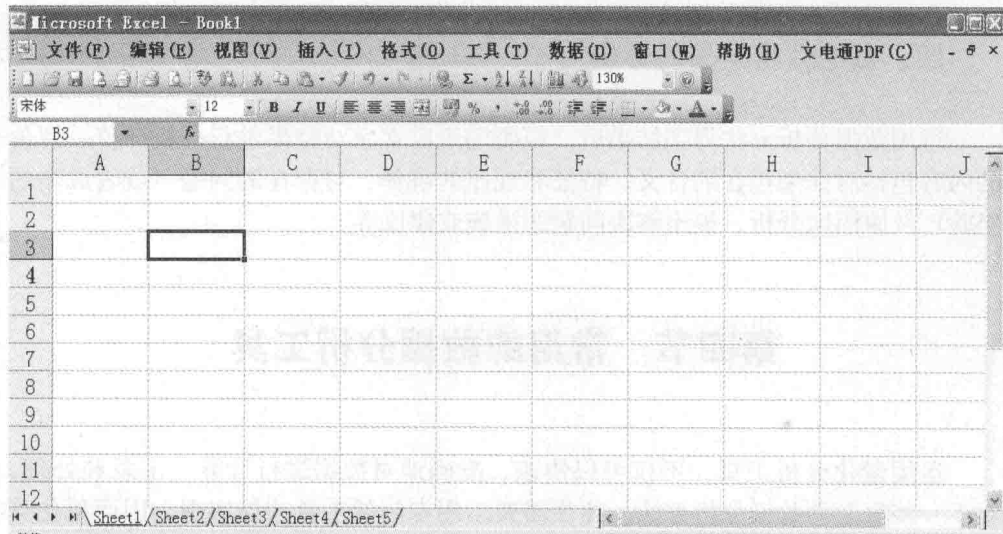


图 1-2 Excel 工作表的界面

在单元格内可以进行数据的输入和编辑操作。针对工作表,可以通过全方位的格式设置,以达到美化效果。

3. Excel 的函数

Excel 提供了许多函数,涉及数学与三角函数、统计、文本、日期与时间、信息、逻辑、财务、工程、查找与引用及数据库十大类共 300 多种,利用这些函数并结合编辑公式,可以方便地对数据进行各种计算、组织、处理。下面是数据处理与分析时经常用到的 Excel 函数。

- (1) AVEDEV——平均差计算函数。
- (2) AVERAGE——算术平均数计算函数。
- (3) CHIDIST——根据给定的 χ^2 分布的区间点和自由度,返回右侧收尾概率。
- (4) CHIINV——根据给定的右侧概率和自由度,返回 χ^2 分布的区间点。
- (5) CONFIDENCE——给定显著性水平和总体标准差,返回总体均值的置信区间。
- (6) CORREL——返回两个变量之间的 PERSON 相关系数。
- (7) COUNT——计算包含数字的单元格以及参数列表中数字的个数。
- (8) COUNTIF——计算区域内符合指定条件的非空单元格个数。
- (9) COVAR——返回两变量之间的协方差。
- (10) DEVSQ——返回各数据点与均值之间的离差平方和。
- (11) FDIST——给定区间点和分子分母自由度,返回 F 分布的右侧概率。
- (12) FINV——给定右侧概率和分子分母自由度,返回 F 分布的区间点。
- (13) FORECAST——返回一元线性回归拟合线的一个拟合值。
- (14) FREQUENCY——数组函数,以垂直数组的形式返回一组数据的频率分布。
- (15) FTEST——返回 F 检验的结果,即两组数据方差无显著差异时的单侧概率。
- (16) GEOMEAN——返回一组数的几何平均数。
- (17) HARMEAN——返回一组数的调和平均数。
- (18) INTERCEPT——返回一元线性回归拟合方程的截距。
- (19) KURT——返回一组数据的峰值。
- (20) LINEST——数组函数,返回线性回归方程的参数。
- (21) MEDIAN——返回一组数的中位数。
- (22) MODE——返回一组数的众数。
- (23) NORMDIST——返回正态分布的累计左侧概率或概率密度函数值。
- (24) NORMINV——根据给定的左侧概率和正态分布特征值,返回对应的区间点。