

二十一世纪普通高等院校实用规划教材·经济管理系列

统计学原理

TONGJIXUE YUANLI

袁 威 主 编

赠送
电子课件

- 先进性与基础性相统一 •
- 教材建设与教学改革相统一 •
- 综合性与针对性相统一 •



清华大学出版社

二十一世纪普通高等院校实用规划教材 经济管理系列

C8
244

统计学原理

袁 威 主 编

清华大学出版社
北 京

内 容 简 介

本书系统地阐述了统计学的基本概念和基本理论,注重统计思想和统计方法的传达,突出简洁、实用的特点。本书共分 10 章,具体内容包括绪论、统计设计、统计调查、统计整理、综合指标、时间数列、统计指数、抽样推断、相关分析与回归分析、统计综合分析。每章均有导入案例、案例、相关链接等内容,课后有练习题,能够较好地提高专业技能。

本书主要针对应用型本科非统计专业学生编写,适合作高等院校经济类、管理类、金融类、工程类专业本科生统计学课程教材,也可供统计工作人员、企业领导和其他人员学习、培训和工作参考使用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

统计学原理/袁威主编. —北京:清华大学出版社, 2016
(二十一世纪普通高等院校实用规划教材 经济管理系列)
ISBN 978-7-302-43667-6

I. ①统… II. ①袁… III. ①统计学-高等学校-教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2016)第 085007 号

责任编辑:桑任松

封面设计:刘孝琼

责任校对:周剑云

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62791865

印装者:三河市中晟雅豪印务有限公司

经 销:全国新华书店

开 本:185mm×230mm 印 张:23.75 字 数:515 千字

版 次:2016 年 6 月第 1 版 印 次:2016 年 6 月第 1 次印刷

印 数:1~2000

定 价:48.00 元

产品编号:066981-01

前 言

数据遍布于自然界和人类社会的每个角落，而统计是以数据作为研究对象的科学，因此，统计学已经渗透到了各个领域，我们每个人都可能用得到。统计学为我们提供了一套获取数据、分析数据并从中得出结论的原则和方法，掌握统计技术可以帮助我们在信息高速发展和传递的时代获取竞争优势。

本书论述的统计学原理所研究的是统计学的基本理论、基本原则和基本统计方法。这些方法既可以用于对社会现象的数量方面的研究，也可以用于对自然现象的数量方面的研究，它是各种应用统计学的共同基础。

本书具有以下几方面的特点。

(1) 本书是按照统计工作过程各个阶段的顺序编写的。完整的统计工作过程包括统计设计、统计调查、统计整理和统计分析，并借助这些数据资料进行推断和决策。这样可以使本书的整体框架更加严密，对有关内容的理论和方法可以阐述得更具体、更深刻，也更有利于学习者将来在统计实践中开展工作。

(2) 本书注重统计思想和统计方法的阐述。删减了烦琐的数学证明和公式推导，做到通俗易懂、由浅入深。在对统计方法进行阐述后都配有具体应用过程的案例说明，以方便读者的理解，具有较强的实用性和可操作性。

(3) 本书主要针对应用型人才培养目标，注重学生能力的培养。各章都有导入案例、相关链接和案例，在阐述统计方法的同时，培养学生基本的统计思想和技能，有助于培养学生分析问题和解决问题的能力，以满足实际工作需要。

(4) 本书各章后面配有相应的习题。选择题和简答题可以强化学生对基本概念的理解和应用，计算题和案例分析题可以锻炼学生的统计分析能力。本书还有配套的多媒体课件，以满足教学需求。

本书由袁威负责全书整体框架和编写大纲的设计以及全书的总纂和定稿工作。全书共分10章，各章节具体分工如下：袁威负责编写第一、五、七~九章；高彩梅负责编写第二、三、十章；何钦晖负责编写第四、六章。

在本书的编写过程中，参考了许多国内外的书籍和资料，在本书的附录中列出参考文献，若有遗漏，万望见谅。并在此谨向各位作者致以诚挚的谢意！由于作者水平有限，书中难免有不足之处，恳请广大教师和读者不吝批评指正。谢谢大家！

编 者

目 录

第一章 绪论1

第一节 统计学概述2

一、统计的含义2

二、统计的特点3

三、统计的职能5

四、统计的作用6

第二节 统计学的产生与发展8

一、统计的产生8

二、统计学的发展9

三、我国统计发展史12

第三节 统计学的研究对象和研究方法13

一、统计学的研究对象13

二、统计学的研究方法13

第四节 统计工作16

一、统计工作的基本任务16

二、统计工作的过程16

第五节 统计常用的基本概念17

一、统计总体和总体单位18

二、标志和指标19

三、变异、变量和变量值21

四、流量和存量21

课后练习22

实训案例24

第二章 统计设计25

第一节 统计设计概述26

一、统计设计的概念26

二、统计设计的作用26

三、统计设计的分类27

第二节 统计设计的内容29

一、明确统计研究的目的和任务 29

二、确定统计研究的对象和 时空标准 30

三、设计统计指标和统计指标体系 ... 30

四、确定统计分类和分组 30

五、设计统计调查方案 31

六、设计统计资料的整理方案 32

七、设计统计分析的基本方法 33

八、规定各个阶段的工作进度和 时间安排 34

九、明确各部门和各阶段的配合与 协调 35

十、统计力量的组织与安排 36

第三节 统计指标和指标体系设计 36

一、统计指标和统计指标体系 36

二、统计指标与统计指标体系的 设计要求 39

三、统计指标和统计指标体系设计的 内容 40

课后练习 43

实训案例 45

第三章 统计调查 47

第一节 统计数据资料 48

一、统计数据的计量尺度 48

二、统计数据的类型 50

三、统计数据的来源 51

第二节 统计调查概述 52

一、统计调查的概念 52

二、统计调查的要求 53

三、统计调查的种类 54



第三节 统计调查方案设计.....55	二、分配数列的种类.....94
一、确定调查目的.....55	三、分配数列的编制.....96
二、确定调查对象和调查单位.....56	四、分配数列的特征.....100
三、设计调查项目和调查表.....56	第四节 统计表与统计图.....101
四、确定调查时间和调查期限.....57	一、统计表.....101
五、制定调查工作的组织实施计划.....58	二、统计图.....106
第四节 统计调查的组织形式.....59	课后练习.....110
一、普查.....59	实训案例.....113
二、统计报表.....60	
三、重点调查.....62	第五章 综合指标.....116
四、典型调查.....63	第一节 总量指标.....117
五、抽样调查.....63	一、总量指标的概念及特点.....117
第五节 统计调查问卷设计.....66	二、总量指标的作用.....118
一、调查问卷的结构.....66	三、总量指标的计量单位.....119
二、调查问卷的设计原则.....69	四、总量指标的种类.....120
三、调查问卷的内容设计.....70	五、总量指标的计算.....121
四、调查问卷的设计技巧.....75	第二节 相对指标.....122
课后练习.....75	一、相对指标的概念.....122
实训案例.....77	二、相对指标的作用.....122
	三、相对指标的表现形式.....123
	四、相对指标的计算.....125
	五、正确运用相对指标的原则.....134
	六、相对指标的应用条件.....135
	第三节 平均指标.....136
	一、平均指标的概念.....137
	二、平均指标的特点.....137
	三、平均指标的作用.....137
	四、平均指标的计算.....138
	五、正确运用平均指标的原则.....150
	第四节 标志变异指标.....152
	一、标志变异指标的概念.....152
	二、标志变异指标的作用.....152
	三、标志变异指标的种类和计算.....153
	课后练习.....159
	实训案例.....164
第四章 统计整理.....80	
第一节 统计整理概述.....81	
一、统计整理的含义及意义.....81	
二、统计整理的原则.....82	
三、统计整理的步骤.....82	
四、统计整理的组织形式.....84	
五、统计整理的技术.....85	
第二节 统计分组.....87	
一、统计分组的含义.....87	
二、统计分组的作用.....87	
三、统计分组的原则.....89	
四、统计分组的类型.....89	
五、统计分组的方法.....92	
第三节 分配数列.....93	
一、分配数列的含义.....93	

第六章 时间数列.....166

第一节 时间数列概述.....167

一、时间数列的概念和意义.....168

二、时间数列的种类.....168

三、时间数列的编制原则.....170

四、时间数列的分析方法.....171

第二节 时间数列水平分析指标.....172

一、发展水平.....172

二、平均发展水平.....173

三、增长量.....179

四、平均增长量.....180

第三节 时间数列速度分析指标.....181

一、发展速度.....181

二、增长速度.....182

三、平均速度.....185

第四节 时间数列变动分析.....189

一、长期趋势分析.....189

二、季节变动分析.....196

三、循环变动分析.....198

四、不规则变动分析.....199

课后练习.....199

实训案例.....203

第七章 统计指数.....206

第一节 统计指数概述.....207

一、统计指数的概念.....207

二、统计指数的性质.....208

三、统计指数的作用.....208

四、统计指数的种类.....209

第二节 综合指数.....210

一、综合指数的编制原理.....211

二、数量指标综合指数的编制.....211

三、质量指标综合指数的编制.....213

四、综合指数的其他编制方法.....215

第三节 平均指标指数.....217

一、加权算术平均数指数.....217

二、加权调和平均数指数.....218

第四节 指数体系与因素分析.....219

一、指数体系的概念与作用.....219

二、指数因素分析法.....221

三、指数因素分析法的应用.....223

第五节 几种常见的指数.....231

一、居民消费价格指数.....231

二、农产品收购价格指数.....237

三、工业生产者价格指数.....238

四、股票价格指数.....240

五、房地产价格指数.....243

课后练习.....244

实训案例.....247

第八章 抽样推断.....250

第一节 抽样推断概述.....251

一、抽样推断的含义.....251

二、抽样推断的特点.....251

三、抽样推断的作用.....252

四、抽样推断中常用的基本概念.....253

五、抽样推断的理论依据.....258

第二节 抽样误差.....260

一、抽样误差的概念.....260

二、影响抽样误差的因素.....261

三、实际抽样误差.....261

四、抽样平均误差.....262

五、抽样极限误差.....268

第三节 抽样估计.....269

一、抽样估计的要求.....269

二、抽样估计的方法.....270

第四节 样本容量确定.....272

一、确定样本容量的意义.....272

二、确定样本容量的因素.....272



三、确定样本容量的方法.....	273
第五节 抽样方案设计.....	276
一、抽样方案设计的基本原则.....	276
二、简单随机抽样.....	276
三、类型抽样.....	278
四、整群抽样.....	281
五、等距抽样.....	283
六、多阶段抽样.....	285
课后练习.....	293
实训案例.....	295

第九章 相关分析与回归分析.....297

第一节 相关分析.....	298
一、变量间关系.....	298
二、相关分析.....	301
三、相关分析方法.....	302
第二节 回归分析.....	311
一、回归分析的概念及特点.....	311
二、回归分析的分类.....	312
三、回归分析的内容.....	312
四、相关分析与回归分析的关系.....	313
第三节 线性回归分析.....	314
一、一元线性回归.....	314
二、多元线性回归.....	318
第四节 非线性回归分析.....	319
一、非线性回归分析的意义.....	320

二、非线性回归方程的测定.....	320
课后练习.....	323
实训案例.....	328

第十章 统计综合分析.....330

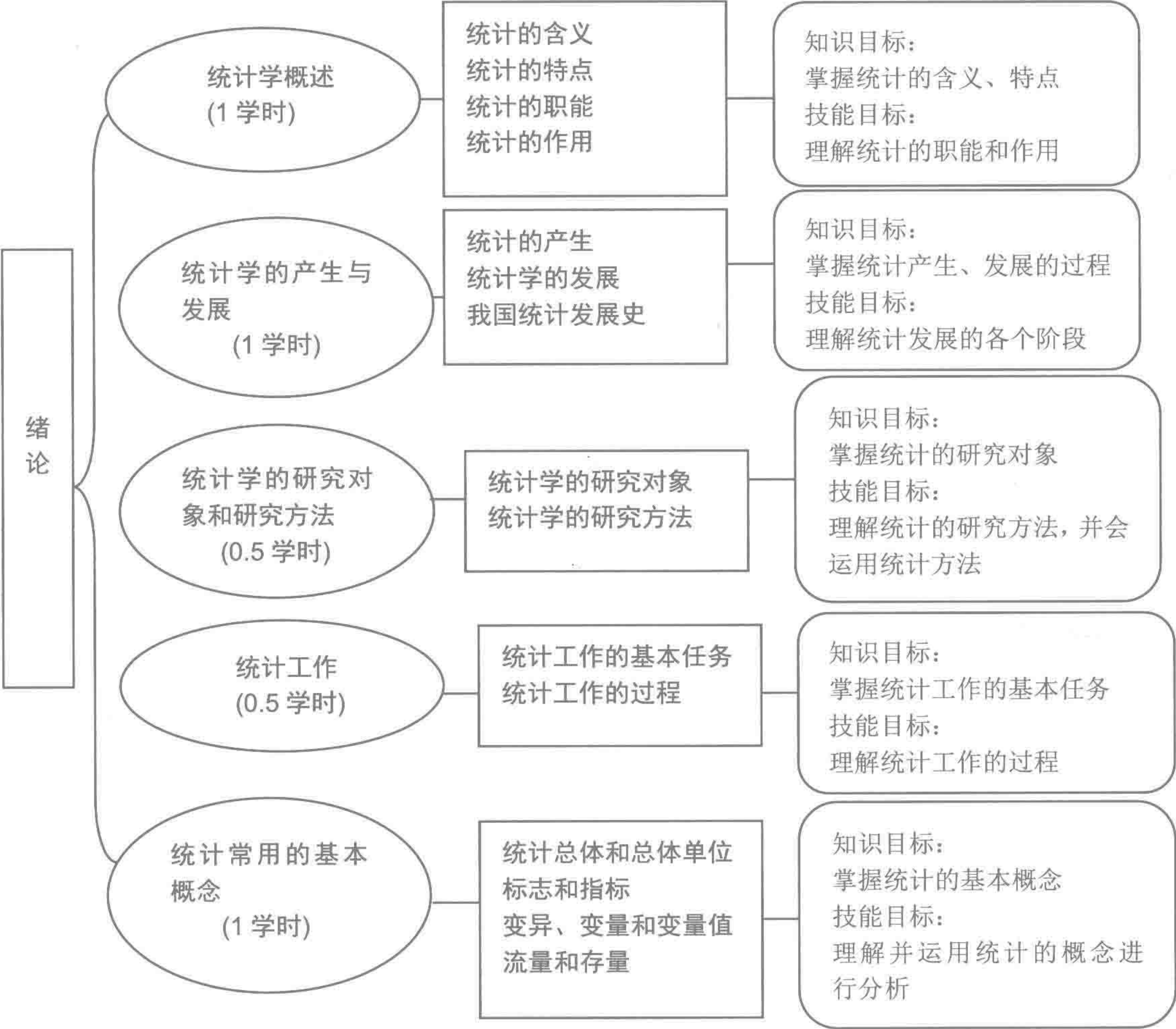
第一节 统计综合分析概述.....	331
一、统计综合分析的概念.....	331
二、统计综合分析的作用.....	331
三、统计综合分析的特点.....	332
四、统计综合分析的程序.....	332
第二节 统计综合分析的方法.....	334
一、综合评价法.....	334
二、统计比较法.....	337
三、平衡分析法.....	339
四、经济景气分析法.....	340
第三节 统计分析报告.....	340
一、统计分析报告概述.....	340
二、统计分析报告的种类.....	341
三、统计分析报告的选题.....	344
四、统计分析报告写作的 基本要求.....	345
课后练习.....	347
实训案例.....	348

课后习题参考答案.....350

参考文献.....370

第一章 绪 论

【教学导图】



【导入案例】

在日常生活中，常常要对生活中的不确定性现象做出决策。对此，不懂统计的人可能无所适从，而懂得统计的人却能将眼前的不确定性现象进行量化，进而使其精确化，依此



做出较为明智的选择。有位教授朋友曾讲过这样一件事情：一天早上，他打车出门准备去某地讲课，不巧途中堵车，幸亏他对当地的环境比较熟悉，脑海中马上就出现了多条可供选择的路线，对比之下，最终选择了一条路线相对较短、堵车概率相对较低的道路，如约赶到了讲课现场，避免了上百名学生等待的局面。这位朋友面对不确定性现象所做的思考，进而做出的决策，其实就是统计的应用。再看看我们身边经常发生的事情——居家购物，尤其是大件物品的购买，人们一般会坚持货比三家，比价格、比质量、比服务，然后确定是否购买。这个比较的过程，也就是询问、了解、收集数据，然后据此分析评价作出判断的过程，其实质也就是统计的基本活动过程。

第一节 统计学概述

一、统计的含义

统计语源最早出现于中世纪拉丁语的 Status，意思是各种现象的状态和状况。统计作为一种社会实践活动已有悠久的历史，最初统计只是为统治者管理国家的需要而收集资料，掌握国家人力、物力和财力的信息，作为管理国家的依据。统计学是一门收集、整理和分析统计数据的方法科学，其目的是探索数据的内在数量规律性。

统计是随着社会生产的发展、适应国家与社会管理的需要而产生和发展的，其涉及的范围也逐步扩展到社会的各个领域。统计具有不同的含义，即统计工作、统计资料和统计学。

1. 统计工作

统计工作即统计实践活动，是指对社会经济现象数量方面进行搜集、整理和分析工作的总称，它是一种社会调查研究活动。例如，我国进行人口普查时要经过方案设计、入户登记、数据汇总、分析总结和资料公布等一系列过程，这些都是统计工作。又如，企业的企管部每月要编制生产报表，这个过程也是统计工作。

2. 统计资料

统计资料即统计信息，是指通过统计工作过程所搜集、整理、编制的各种统计数据的总称。统计资料是进行国民经济宏观调控的决策依据，是社会公众了解国情、国力和社会经济发展状况的信息主体。

【相关链接 1-1】

2010 年 1 月 1 日起实施的新《统计法》第三章第二十三条明确规定：“县级以上人民政府统计机构按照国家有关规定，定期公布统计资料。国家统计数据以国家统计局公布的数据为准。”

国家统计局发布的关于 2014 年粮食产量的公告：根据对全国 31 个省(区、市)农业生产经营户的抽样调查和农业生产经营单位的全面统计，全国粮食播种面积为 112 738.3 千公顷(169 107.4 万亩)，比 2013 年增加 782.7 千公顷(1174.1 万亩)，增长 0.7%。全国粮食单位面积产量 5385 公斤/公顷(359 公斤/亩)，比 2013 年增加 8.4 公斤/公顷(0.6 公斤/亩)，提高 0.2%。全国粮食总产量 60 709.9 万吨(12 142 亿斤)，比 2013 年增加 516 万吨(103.2 亿斤)，增长 0.9%。其中谷物总产量 55 726.9 万吨(11 145.4 亿斤)，比 2013 年增加 457.7 万吨(91.5 亿斤)，增长 0.8%。

(资料来源：<http://www.chinabgao.com/stat/stats/39284.html>)

3. 统计学

统计学是系统论述统计理论和方法的科学，是长期统计工作实践的经验总结和理论概括。统计学由一套收集和处理统计数据的方法组成，这些方法来源于对统计数据的研究，其目的是探索数据内在的数量规律，以达到对客观事物的科学认识。若统计数据不用统计方法去分析，则仅仅是一堆杂乱无章的数据而已，无法得出任何有益的结论。

统计工作、统计资料、统计学三者之间存在着密切的关系：统计工作的成果是统计资料，统计资料和统计学的基础是统计工作，统计学既是统计工作经验的理论概括，又是指导统计工作的原理、原则和方法。如图 1-1 所示。

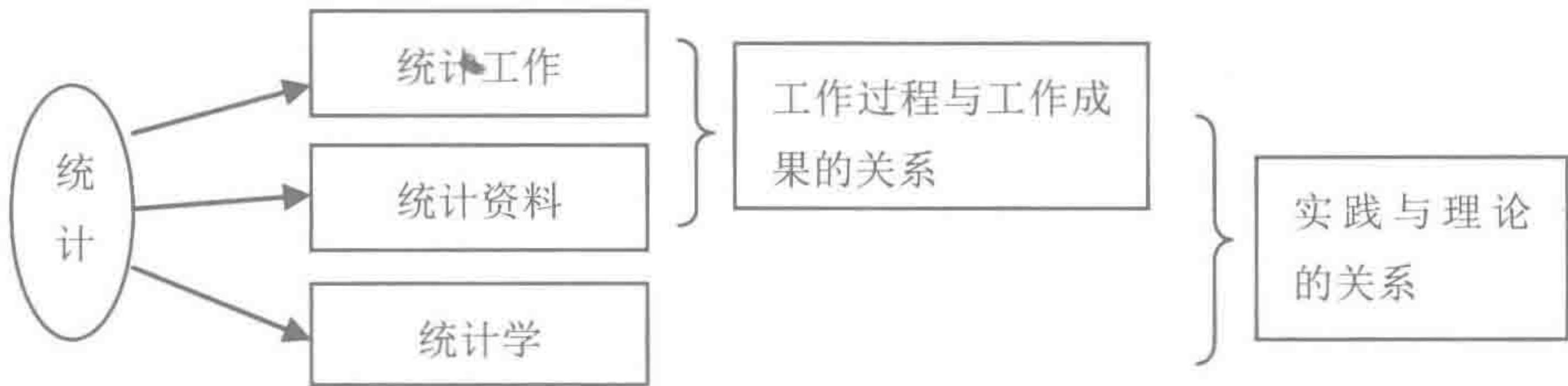


图 1-1 统计含义及两重关系

二、统计的特点

统计是对社会经济现象的一种调查分析活动，它具有以下 6 个特点。

1. 数量性

统计是用大量的数字资料说明事物的规模、水平、结构、比例关系、差别程度、普遍程度、发展速度、平均规模 and 水平、平均发展速度等。数量性是统计的基本特性，但应该注意，统计不是单纯地研究经济现象的数量方面，而是在质与量的密切联系中研究经济现象的数量方面。社会经济现象有质量和数量两个基本方面，两者是辩证统一的。要了解事物的存在和发展并掌握其变化规律，就必须分析、研究事物的量的方面及其发展规律性在具体时间、地点、条件下的数量表现。认识事物的数量表现，是深入把握事物质量的前提和基础。因此，事物的数量是认识客观现实的重要方面，通过分析研究统计数据资料，研究和掌握统计规律性，就可以达到统计分析研究的目的。具体地说，是用规模、水平、速度、结构和比例关系等去描述和分析社会经济现象的数量表现、数量关系和数量变化，揭示事物的本质，反映事物发展的规律，推测事物的发展前景。例如，一个国家的人口数量、结构和分布，国家财富的数量、构成和利用情况，国民经济的规模、发展速度、人们生活水平等数字，都是反映该国国情、国力的基本方面。通过一系列统计指标对这些基本情况有所了解，才可能对该国有一个客观的认识。

2. 总体性

总体性又称大量性或综合性。统计研究的着眼点是大量社会经济现象总体，而不是少量或个别现象，它是通过对个别事物的大量观察，占有丰富材料并加以分析综合，来反映现象总体的数量特征，揭示现象的本质和规律性。例如，城镇居民家庭消费状况的调查，虽然是对每个具体的家庭户进行调查，但并不是研究个别居民家庭的基本生活状况，而是要反映一个城市、一个部门、一个国家的居民收入水平、收入分配、消费水平、消费结构等情况。客观事物的个别现象通常有其特殊性、偶然性，而总体现象则具有相对普遍性、稳定性，是有规律可循的，统计研究现行总体的数量特征，有助于对现象规律性的认识。

3. 具体性

统计研究的是自然、社会经济领域中具体现象的数量方面，而不是抽象的数量关系，是具有明确的现实含义的，这一特点是统计学与数学的重要区别。数学是研究事物的抽象空间和抽象数量的科学，而统计学研究的数量是客观存在的、具体实在的数量表现，是一个有具体时间、具体地点、具体条件限定的数量。例如，2013 年中国 GDP 总量是 94 946 亿美元。但要注意，统计在研究数量关系时，也要遵守数学表明的客观现象量变的规律，并在许多方面运用数学方法。

4. 客观性

统计数量是客观事物的反映,表示客观现象在具体时间、空间条件作用下,实际已经达到的水平和程度。它独立存在于外部世界,不是人的意志所能转移的。统计数据的客观性是统计质量的基础,统计数据应客观、真实。

5. 变异性

变异性又称差异性,是指总体各单位由于随机因素引起的某一标志表现的差异。构成统计研究对象的总体各单位,除了在某一方面必须是同质的以外,在其他方面又要有差异,而且这些差异不是由某种特定的原因事先给定的。例如,一个企业中的职工,其工龄、文化程度、工资水平等都存在着差异,这才有必要研究其平均水平及其各个层次的结构等状况。如果各个总体单位之间不存在差异,也就没有必要进行统计。如果各个总体单位之间的差异是按照已知条件事先可以推定的,也就不需要统计了。比如,昼夜时间长短因季节变化而有所不同,这个可以按照已知条件事先推定,无须统计。

6. 社会性

统计学研究社会经济现象,这一点与自然技术统计学有所区别。自然技术统计学研究自然技术现象(如天文、物理、生物、水文等现象),自然现象的变化发展有其固有的规律,在其变化进程中,通常表现为随机现象,即可能出现或可能不出现的现象。而统计学通过研究大量社会经济现象总体的数量方面来认识人类社会活动的条件、过程和结果,反映物质资料的占有关系、分配关系、交换关系以及其他社会关系。

三、统计的职能

统计是国家实行科学决策和科学管理的一项重要基础工作,是党、政府和人们认识国情国力、制定国家政策的重要依据。因此,统计部门不仅要提供信息,而且要提供咨询与建议,所以统计应具有以下几项职能。

1. 信息职能

信息职能是指统计具有一整套科学统一的统计指标体系和统计调查方法,能够灵敏地、系统地为决策和管理采集、处理、传递、存储和提供大量综合反映客观事物总体数量特征的社会经济信息。统计部门是提供和管理统计信息的主要职能部门,提供全面、及时、可行的社会经济统计信息是统计部门最基本、最普遍的职能。



2. 咨询职能

咨询职能是指利用已经掌握的统计信息资源,运用科学的分析方法和先进的技术手段,深入开展综合分析和专题研究,为科学决策和管理提供各种可供选择的咨询建议和对策方案。统计咨询水平的高低是衡量统计工作水平的重要标志之一,统计信息咨询除了为党政领导加强宏观调控提供具有量化特点的咨询意见和对策建议外,还必须面向许多部门、研究机构、社会团体、基层企业、个人和国外多类用户,大力发展统计信息咨询业,建立公共统计信息咨询服务机构,培育统计信息市场。

3. 监督职能

监督职能是指通过统计调查和统计分析,及时、准确地从总体上反映经济、社会和科技运行状态,并对其实行全面、系统的定量检查、监测和预警,揭示社会经济运行中的偏差,使社会经济决策及其执行过程按客观规律的要求持续、稳定、协调地发展。

统计信息职能是统计最基本的职能,是保证统计咨询和统计监督职能有效发挥的基础。统计咨询职能是统计信息职能的深化。统计监督职能是在统计信息、咨询职能基础上的进一步拓展,并促进统计信息和咨询职能的深化。统计的 3 种职能相辅相成、相互作用,构成了一个有机整体。只有进一步优化统计信息职能,提高统计咨询水平,强化统计监督职能,充分、有效地发挥统计信息、咨询、监督职能的整体功能,才能充分体现和发挥统计在社会经济管理中的重要地位和作用。

四、统计的作用

统计的应用领域非常广泛,无论是从宏观管理上还是微观决策上,甚至在日常生活中,都起着重要的作用。具体主要包括以下几个方面。

1. 统计是认识自然和社会的必要方法和途径

列宁曾经说过:“统计是认识社会的最有力武器。”人们要认识社会、认识自然就离不开各种各样的数据,并且要对这些数据进行分析研究,这就需要运用一整套的统计理论和方法指导。所以,统计是认识自然和社会的手段,统计学则为这种手段提供了理论和方法。

2. 在宏观经济管理中,统计是国家宏观调控和管理的重要工具

政府在进行宏观经济管理中要实现四项宏观经济管理目标,即保持国民经济持续健康发展、实现劳动力的充分就业、保持物价稳定及保持国际收支均衡。要实现这四大宏观经济管理目标,就需要政府利用统计方法,科学合理地收集数据,对国民经济和社会发展状况进行跟踪监测和预警,对各种社会经济问题进行定量分析,从而为制定和调整宏观

经济政策提供决策依据。

3. 在微观企业管理中，统计是企业管理与决策的依据

企业管理者不仅要掌握人财物、产供销的情况，而且要及时了解市场经济信息，依据市场需求安排企业的生产经营活动，在激烈的市场竞争中要做到知己知彼，才能立于不败之地。这就要借助统计调查，搜集各方面的统计资料，进行统计分析，为企业的生产经营决策提供可靠的科学依据。

4. 统计为科学研究提供有力手段

无论是在社会科学领域，还是在自然科学领域，进行某个课题研究时，为使观点与结论具有事实依据和说服力，必须根据调查研究或实验取得的统计数据来说明问题，这就需要运用科学的统计原理与方法来进行。例如：在医学界，人们利用统计方法来研究疾病的原因或影响因素，判断药物或治疗方法的有效性；在考古学界，人们凭借统计方法来推断特定发掘物的历史年代；在心理学界，人们用统计方法分析特定刺激的心理效应；在经济学界，人们用统计方法研究国民经济运行状况和各种决策方案的优劣；在生物学界，人们用统计方法来研究基因定律。几乎所用的科学研究领域都离不开统计学。历史上许多有关领域的著名专家，往往也是著名的统计学家。

【案例 1-1】

《红楼梦》版权问题

《红楼梦》一书共有 120 回，一般认为前 80 回为曹雪芹所写，后 40 回为高鹗所续，长期以来学术界对这个问题一直存有争议。

1980 年 6 月，美国威斯康星大学陈炳藻教授在首届国际《红楼梦》研讨会上宣读了一篇《从词汇上的统计论〈红楼梦〉的作者问题》的文章，引起了国际红学界的关注和兴趣。1986 年，陈炳藻教授公开发表了《电脑在文学上的应用：〈红楼梦〉与〈儿女英雄传〉两书作者用词的比较》一文，之后又出版了《电脑红学：论〈红楼梦〉作者》的专著。陈炳藻教授通过利用电脑对《红楼梦》前 80 回和后 40 回的用词进行了测定，并从数理统计学的观点出发，探讨《红楼梦》前后用词的相关程度。他将《红楼梦》120 回按照顺序编成 3 组，每组 40 回，并将《儿女英雄传》作为第 4 组进行比较研究，从每组中任取 8 万字，分别挑出名词、动词、形容词、副词、虚词这 5 种词，借助统计上的相关分析计算各组相关程度。结果用词的正相关程度是 32.14%。由此推断，前 80 回和后 40 回的作者均为曹雪芹一人的结论。

然而，复旦大学数学系李贤平教授却给出了相反的回答。1987 年，李贤平教授带领



学生将 120 回看成 120 个样本，然后确定与情节无关的 47 个虚词(之、其、或、亦、呀、吗、可、在、但、儿等)作为变量，统计出每一回里变量出现的次数作为数据，用多元统计中的聚类分析法进行合并，果然将 120 回分成两类，即前 80 回为一类，后 40 回为一类，很形象地证实了此书不是出自同一人的手笔。之后他们又进一步分析了前 80 回是否为曹雪芹所写这一问题，根据曹雪芹的另一本著作做了类似计算，结果证实了用词手法完全相同，断定为曹雪芹一人所写。类似的论证还推翻了后 40 回为高鹗一人所写的传统认识。

(资料来源：吕光明.《红楼梦》作者的统计论证[J].数据，2008，4)

第二节 统计学的产生与发展

统计作为一种社会实践活动，是随着人类社会经济的发展、治国和管理的需要而产生和发展起来的，至今已有四五千年的历史。而统计学或统计理论则是在长期统计实践活动基础上形成和发展起来的，距今只有 300 多年的历史。

一、统计的产生

人类的统计实践是随着计数活动而产生的，早在原始公社时期，就有结绳记事、结绳计量的方法，即“事大，大结其绳；事小，小结其绳；结之多少，随物众寡”，这可以说是统计的萌芽。当时居住在“两河流域”一带的人们也在石头上和泥板上刻出一个个像楔子形的符号，如“丫”和“<”，分别表示 1 和 10。但那时的计数活动，还主要表现在人们对仅有的剩余劳动成果或其视线所及的劳动对象加以清点与度量。而统计实践的真正萌芽是在古代奴隶社会。当时的统治阶级为了治理国家的需要，常常进行征税、征兵、服劳役等统治活动，因此有了了解社会基本情况的需要。所以，一般认为有了国家就有了统计实践活动。

我国是世界上最早开始统计活动的国家之一，早在公元前 21 世纪的夏朝，就有了人口与土地数字的记载，并依山川土质、人口物产及贡赋多寡，将全国分为九州。以后的周王朝，在统计方面更为完善，不仅制定了乡的定期报表制度，在统计方法上还应用了专门调查、统计图示及账册。进入封建社会后，随着人类社会的发展，统计的范围逐渐由人口、土地发展到社会经济生活的各个方面。但由于自给自足的自然经济占主导地位，生产力水平低下，经济落后，长期的封建生产关系阻碍了社会生产力的发展，相应地也阻碍了统计实践的发展。统计实践的广泛发展始于资本主义社会。17 世纪以来，资本主义国家由于工

业、商业、农业、贸易、交通的发展,统计实践从国家管理领域扩展到社会经济活动的许多领域。从18世纪起,各资本主义国家先后设立专业的统计机构,收集各方面统计资料,定期或不定期举行人口、工业、农业、贸易、交通等项调查,出版统计刊物,建立国际统计组织,召开国际统计会议。

二、统计学的发展

统计学起源于研究社会经济问题,今天仍然是研究社会经济问题的基本方法。从统计学的产生和发展的过程来看,大致可以划分为3个阶段。

(一)统计学的萌芽期

统计学初创于17世纪中叶至18世纪,当时主要有政治算术学派和国势学派。

1. 政治算术学派

政治算术学派起源于17世纪的英国,其主要代表人物是威廉·配第(W. Petty, 1623—1687)和约翰·格朗特(J. Graunt, 1620—1674)。17世纪的英国学者威廉·配第在1671—1676年间写成《政治算术》一书,当时正值第三次英、荷战争,国内经济困难,国外面临着荷、法两国的威胁。配第为了让人们知道和确信“英国的事业和各种问题并非出于可悲的状态”,在这本书中用数字比较分析了英国、荷兰、法国三国的经济实力和造成这种实力差异的原因,并从贸易、税制、分工、资本和利用闲散劳动力等多方面提出了英国的强盛之道。这种用数字来表达,用数字、重量和尺度来计量,并配以朴素的图标,正是现代统计学广为采用的方法和内容。正是在这个意义上,马克思称配第是“政治经济学之父,在某种程度上也可以说是统计学的创始人。”

配第的朋友约翰·格朗特通过对伦敦市50多年的人口出生率和死亡率进行分类计算,编制了世界上第一张“死亡率”统计表,写出了第一本关于人口统计的著作——《对死亡表的自然观察和政治观察》(1662年)。但遗憾的是,该学派的学者都还没有使用“统计学”这个名称。从此,统计的含义从记述转变为专指在“量”的方面来说明国家的重要事项。这就为统计学作为一种数量方面认识事物的科学方法,开辟了广阔的发展前景。

2. 国势学派

国势学派产生于18世纪的德国,主要代表人物是康令(H. Conring, 1606—1681)和阿亨瓦尔(G. Achenwall, 1719—1772),康令以叙述国家显著事项和国家政策关系为内容,在大学开设了“国势学”课程,很受当时学者的欢迎。主要继承人阿亨瓦尔的主要著作是《近代欧洲各国国势学概论》。他继续开设“国势学”课,并于1749年首次使用统计学代替国