

Applied Statistics  
with SPSS

# 应用统计学

## 基于SPSS

21世纪高等教育新理念精品规划教材

申卯兴 编著



21 世纪高等教育新理念精品规划教材

Applied Statistics with SPSS

# 应用统计学

—— 基于 SPSS

申卯兴 编 著

 天津大学出版社  
TIANJIN UNIVERSITY PRESS

## 内 容 摘 要

本书是为了适应新型应用型人才培养的需要而编著的一本适用面宽的应用统计学教材。全书内容共分 11 章,包括导论、数据的收集、数据的图表展示、数据特征的统计描述、统计量与抽样分布、参数估计、假设检验、方差分析、相关与回归分析、时间序列分析、社会经济统计概述。书中基本概念、思想观念、实用方法与实际运用紧密结合,每章都有相应内容的 SPSS 软件操作,章末有思考与练习,书末附有部分思考与练习参考解答供学生参考。教师可以通过联系出版社得到配套的 PPT 课件,以便于教学。

本书可作为应用型本科院校的经济类、管理类、金融类、财经类和各种工程类专业的教材,同时也可以作为应用统计学专业的入门教材,还可以作为广大经济管理人员学习统计学的参考书或在岗人员的培训教材。

### 图书在版编目(CIP)数据

应用统计学:基于 SPSS/申卯兴编著. —天津:天津大学出版社, 2017.7.

21 世纪高等教育新理念精品规划教材

ISBN 978-7-5618-5893-6

I. ①应... II. ①申... III. ①应用统计学—高等学校—教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2017)第 174451 号

出版发行 天津大学出版社

地 址 天津市卫津路 92 号天津大学内(邮编:300072)

电 话 发行部:022-27403647

网 址 publish.tju.edu.cn

印 刷 天津泰宇印务有限公司

经 销 全国各地新华书店

开 本 185mm×260mm

印 张 16.75

字 数 406 千

版 次 2017 年 7 月第 1 版

印 次 2017 年 7 月第 1 次

定 价 40.00 元

凡购本书,如有缺页、倒页、脱页等质量问题,请向我社发行部联系调换

版权所有 侵权必究

# 序

## PREFACE

统计学是随着社会政治经济的发展、为满足国家管理的需要而产生和发展起来的一门方法论学科。统计方法已广泛应用于自然科学和社会科学的众多领域，所以我们常说统计的生命力在于应用。

如今，在我国，统计学是经教育部批准的高等学校经济类、工商管理类专业的核心课程。同时，由于其方法的实用性，众多的工程类专业也将其设置为必修课程。

应用统计学是一门研究客观现象总体数量特征的方法论课程，具有数量性、综合性、应用性和实践性等特征。对现代统计的基本概念、基本理论与基本方法的学习和掌握，将为提供统计信息、应用统计方法、进行统计决策奠定必要的专业基础。

申卯兴教授主持编著的这本应用统计学教材，面向 21 世纪应用型本科人才培养，从通识性的教学需求出发，紧密围绕应用统计学的特征、地位和作用，强调应用、注重实践，既注重应用统计学内容体系的系统性和基本素质的培养，又与软件密切结合，对应用型本科大学生的统计学基础教学具有良好的适用性和实用性。

我相信，本教材的出版必然会对应用统计学方法的通识教学产生积极的影响，帮助老师完美地达到教学目的。

是为序！



2017 年 5 月于中国人民大学统计学院

随着信息化技术的迅猛发展，数据已经成为炙手可热的高频词；在“大数据”概念火热的当今，数据处理和运用已经成为全社会的热门话题。在这种情形下，应用统计学从统计学中脱颖而出，实用的统计概念与方法成为社会各行各业的自然需求。

统计学的研究对象是客观现象的总体数量特征和数量关系，从数据的获取到数据的运用都是统计学的研究范畴，统计学侧重揭示客观事物在数量上的客观规律。因此，从应用的角度讲，统计学不仅是教育部规定的经济类和管理类本科专业的核心课程，也是在很多院校已经普及的各个工程类专业的必修课程之一。

正是出于这样的需要，我们经过长时间的探讨和研究，组织了多年从事统计学、应用统计学教学工作的老师，面向应用，从实际需求出发，又不失系统性和创新性，编写了这本适用于应用型本科人才培养的教材。把读者对象定位为应用型本科院校所有需要统计学知识和工具的经管类和工程类专业的本科生。各院校在使用过程中可以根据需要进行适当的取舍。

在本书的编著过程中，我们结合教学实践经验，遵循以下几个原则：（1）不失一般系统性原则，合理取舍，反映统计学原理的基本内容；（2）做到基本概念准确，叙述简洁，反映思想方法和过程；（3）注重方法技术，把数学理论推导通俗化；（4）注重方法应用，紧密结合实际，把技术方法程式化；（5）注重实际运用，全程结合最基础而流行的统计学软件 SPSS 的操作，把应用统计学的运用落到实处。

本书编著的具体分工为：全书内容体系设计和编著策划由申卯兴教授完成，申卯兴编写第 1、10、11 章，令伟锋编写第 2、4 章和 3.1、3.2、3.3 节，何媛编写第 5、6 章和 3.4 节，尚云艳编写第 7、8、9 章。全书由申卯兴教授统稿、修订完成。

在本书的筹划和编写过程中，得到了西京学院、系和教务处等业务机关及其领导的支持和协助。在西京学院院长任芳博士的倡导下，西京学院理学院院长王震教授和经管系主任秦效宏博士等一直关注和敦促，教务处肖建军处长和王登武副处长以及张静教授、杜永红教授、江西省委党校梁玉红教授、上海大学应益容教授等给予了支持和帮助。著名统计学家、中国人民大学何晓群教授为本书作序。同时，笔者参阅和借鉴了国内外不少专家学者和同行的著述和文献。在此，一并致以真诚谢意。

由于编著者水平所限，加之许多主客观的因素，书中难免有疏漏甚或错误之处，敬请读者不吝赐教（反馈意见或索取 PPT 课件请至 [smxyz@sina.com](mailto:smxyz@sina.com)），以臻完善。

编著者 申卯兴

2017 年 5 月于西京园

# 目录

# CONTENTS

## 第1章 导论

- 1.1 统计学/ 1
- 1.2 统计学中的几个基本概念/ 8
- 1.3 统计变量的类型/ 10
- 1.4 统计学的研究类型划分/ 12
- 1.5 统计学软件简介/ 15
- 思考与练习/ 16

## 第2章 数据的收集

- 2.1 数据的几个来源/ 19
- 2.2 抽样调查/ 21
- 2.3 调查方案的制定与调查问卷的设计/ 24
- 思考与练习/ 28

## 第3章 数据的图表展示

- 3.1 数据的预处理/ 29
- 3.2 定性数据的整理与展示/ 30
- 3.3 定量数据的整理与展示/ 33
- 3.4 基于 SPSS 的数据整理与展示/ 40
- 思考与练习/ 51

## 第4章 数据特征的统计描述

- 4.1 数据集中趋势的描述/ 54
- 4.2 数据波动程度的描述/ 57
- 4.3 数据分布形态的描述/ 64
- 4.4 分位数和箱线图/ 67
- 4.5 用软件实现数据特征的计算/ 70
- 思考与练习/ 72

## 第5章 统计量与抽样分布

- 5.1 样本与统计量/ 75
- 5.2 常用的抽样分布/ 78
- 5.3 用软件验证数据的分布/ 86

思考与练习/ 87

## 第 6 章 参数估计

- 6.1 参数估计的基本原理/ 89
- 6.2 一个总体参数的区间估计/ 93
- 6.3 两个总体参数的区间估计/ 97
- 6.4 样本量的确定/ 101
- 6.5 用软件实现参数估计/ 103

思考与练习/ 105

## 第 7 章 假设检验

- 7.1 假设检验的基本问题/ 107
- 7.2 假设检验中的几个重要概念/ 108
- 7.3 一个总体参数的检验/ 112
- 7.4 两个总体参数的检验/ 114
- 7.5 用软件实现假设检验/ 120

思考与练习/ 123

## 第 8 章 方差分析

- 8.1 单因素方差分析/ 125
- 8.2 双因素方差分析/ 129
- 8.3 用软件实现方差分析/ 134

思考与练习/ 138

## 第 9 章 相关与回归分析

- 9.1 变量间的相关关系/ 142
- 9.2 相关分析/ 142
- 9.3 一元线性回归/ 146
- 9.4 多元线性回归/ 151
- 9.5 可化为线性回归的非线性回归/ 155
- 9.6 用软件实现回归模型/ 157

思考与练习/ 160

## 第 10 章 时间序列分析

- 10.1 时间序列的几种类型/ 163
- 10.2 时间序列的描述性分析/ 169
- 10.3 平稳时间序列的预测/ 179
- 10.4 时间序列的趋势预测/ 184
- 10.5 复合型时间序列的预测/ 192

10.6 时间序列分析的软件实现/ 200  
思考与练习/ 207

## 第 11 章 社会经济统计概述

11.1 常用社会经济统计指标/ 209

11.2 国内生产总值/ 216

11.3 指数/ 225

思考与练习/ 241

## 附：思考与练习参考解答

## 参考文献



# 第1章

## Chapter

# 01

## 导论

为了使大家从整体上对统计学课程有个基本的认识,本章将结合现代社会经济和管理等领域的应用,对统计学的一些基本概念和名词给予扼要的介绍。



### 本章重点与目标

1. 理解统计学的定义
2. 理解统计学涉及的一些基本概念
3. 熟悉几种不同类型统计变量的含义
4. 了解几种常用的统计软件

## 1.1 统计学

### 1.1.1 统计学的一级学科地位

对中国的普通老百姓来说,“统计”一词并不陌生,GDP(国内生产总值)、CPI(居民消费价格指数)这些统计名词从各种媒体也容易知晓。《大英百科全书》认为,统计学是一门收集数据、分析数据并根据数据进行推断的艺术和科学。1988年出版的《中国大百科全书》经济卷Ⅱ,曾经把统计学定义为一门社会科学。然而,究竟什么是统计学?统计学的学科性质是什么?我们将要学习的统计学理论和方法究竟是社会科学还是自然科学?这些都需要我们去认识和思考。

改革开放对中国的统计学科来说是挑战，更是机遇。严峻的挑战和深刻的矛盾要求统计学界必须坚持实事求是的科学态度，还统计学科的本来面目。经过一大批统计学家的艰苦努力，我国统计学科终于发生了质的变化和飞跃。1992年11月，国家技术监督局正式批准统计学为一级学科，国家标准局颁布的学科分类标准已将统计学列为一级学科；1998年教育部进行的本科专业调整也将统计学归为理学类一级学科；2012年教育部颁布的本科专业目录，除在经济学类中新增经济统计学专业以外，仍在理学大类一级学科的统计学类中设置了统计学和应用统计学两个本科专业，这是适应社会经济和科技发展、适应信息化社会和大数据时代需求的必然选择，更加推动了统计学应用的蓬勃发展。统计学一级学科的地位表明，中国统计学科在与国际接轨的进程中迈出了重要一步。其实，根据国际惯例，根本就不存在社会经济统计学和数理统计学之分。统计学(Statistics)是一门综合性很强的边缘学科，它不是数学，更不是经济学。统计学是研究客观事物数量关系和数量特征的方法论学科。既然是处理数据的方法论学科，那就决定了应用是统计学的本质所在。

统计学到底应该隶属于哪一个学科，中国统计学界至今仍有着激烈的争论，是理学？经济学？还是管理学？无论怎样，我们认为总比把无处不用的统计学单独放在经济学里作为经济学的一个二级子学科要合理得多。其实统计思想和统计方法早已渗透到社会、经济、自然、科技、生活的每一个角落。

需要注意的是，本书介绍的统计学与过去的社会经济统计学相比有着很大的差别。这些差别不仅体现在内容的深度和广度上，更体现在处理问题的思想观点和方法上。在统计学的学习和应用中将充分体现从具体事物到一般概念、再从一般概念回到具体事物的这种辩证观点和较为严格的逻辑推理。统计学给实际应用带来新思维，应用给统计学增添活力。统计学的生命力就在于应用。

## 1.1.2 统计与统计学

### 1.1.2.1 统计学的概念

“统计”一词经常出现在各种社会实践活动和科学研究领域中。然而，在不同场合，不同的人对它的理解常常有差异。其实，我们常说的统计一词具有如下三种含义。

(1) 统计活动：统计活动又称统计工作，是指收集、整理和分析统计数据，并探索数据的内在数量规律的活动过程。

(2) 统计资料：统计资料也称统计数据，是在统计活动过程中获得的各种数字资料和其他资料的总称，表现为反映社会经济现象数量特征的原始记录、统计台账、统计表、统计图、统计分析报告、政府统计公报、统计年鉴等各种数字和文字资料。

(3) 统计学：统计学是指阐述统计工作基本理论和基本方法的科学，是对统计工作实践的理论概括和经验总结。它以客观现象总体的数量为研究对象，阐明统计设计、统计调查、统计整理和统计分析的理论与方法，是一门方法论科学。

统计活动、统计资料和统计学之间相互依存、相互联系，共同构成了一个完整的整体——统计。统计工作与统计资料之间的关系是过程与成果之间的关系，统计资料是统计工作的直接成果。就统计工作和统计学的关系来说，统计工作属于实践的范畴，统计学属于理论的范畴，统计学是统计工作实践的理论概括和科学总结，它来源于统计实践，又高于统计实践，反过来又指导统计实践，统计工作的现代化与统计科学研究的支持是密不可分的。

基于上述对“统计”概念的理解，我们认为统计学是一门研究数据的方法论学科，通过搜索、收集、处理、分析等手段，从数据中得出结论，从而达到推断所考察对象的本质，甚至预测对象未来的目的。统计学的研究对象是客观现象的总体数量特征和数量关系，从数据的获取到数据的运用都是统计学的研究范畴。统计学侧重揭示客观事物数量上的客观规律，因此必然会用到大量的数学及其他学科方面的专业知识。

统计学的基本特点：总体性、数量性、客观性、随机性、广泛性。

(1) 总体性：统计的研究对象是客观总体现象的数量方面。如人口统计是要反映和研究一个国家或一个地区全部人口的综合数量特征，而不是要了解和研究某个人的特征，但是它是从调查每个人开始的。人口统计是这样，其他统计活动也是这样。

(2) 数量性：统计的研究对象是客观现象的数量方面，包括数量的多少、数量之间的关系、质量互变的数量界限。

(3) 客观性：统计研究的现象是自然或社会的客观事物在数量方面的自然表象，是对客观事物由表及里的客观认识。

(4) 随机性：统计研究的是同类现象总体的数量特征，它的前提是总体各单位的特征表现存在差异，而且这种差异并不是由于某种固定的原因事先给定的。

(5) 广泛性：统计研究的观念、思想、方法已经广泛地被各行各业的人员所认识和接受，统计方法的使用范围几乎覆盖了社会科学和自然科学的各个领域。

在历史上，从方法的功能来看，统计学可以分为描述统计学(Descriptive Statistics)和推断统计学(Inferential Statistics)；从方法研究的角度来看，统计学可分为理论统计学(Theoretical Statistics)(即数理统计学，Mathematical Statistics)和应用统计学(Applied Statistics)。应用统计学探讨如何运用统计方法解决实际问题。将理论统计学的原理应用于各个学科领域，就形成了应用统计学。应用统计学着重阐明这些方法的统计思想和具体应用，而不是统计方法数学原理的推导和证明。更加具体深入些，还可以形成形形色色的应用统计学分支。例如，统计方法在生物学中的应用形成了生物统计学，在医学中的应用形成了医疗卫生统计学，在农业试验、育种等方面的应用形成了农业统计学，在经济领域的应用形成了经济统计学及其若干分支，在管理领域的应用形成了管理统计学，在社会学研究和社會管理中的应用形成了社会统计学，在人口学中的应用形成了人口统计学，等等。应用统计学除了包括各领域通用的方法，如参数估计、假设检验、方差分析等，还包括某领域特有的方法，如经济统计学中的

指数法、现代管理决策法等。

### 1.1.2.2 统计学简史

统计学是一门古老而又现代甚或时髦的科学，一般认为其学理研究始于古希腊的亚里士多德时代，迄今已有 2 300 多年的历史。它起源于社会经济问题的研究，在 2 000 多年的发展过程中，统计学至少经历了“城邦政情”“政治算术”和“统计分析科学”三个发展阶段。不同阶段的各学派之间相互争论、相互渗透，使统计理论最后发展成为统一的现代统计学。

统计学的英文单词 *statistics* 最早源于现代拉丁文 *statisticum collegium*（国会）以及意大利文 *statista*（国民或政治家）。德文 *statistik*，最早由 *Gottfried Achenwall* 于 1749 年使用，代表对国家的资料进行分析的学问，也就是“研究国家的科学”。19 世纪，统计学在广泛的数据以及资料中探究其意义，并且由 *John Sinclair* 引入英语世界。“城邦政情”阶段始于古希腊的亚里士多德撰写的《城邦政情》（或《城邦纪要》），他撰写了一百五十余种纪要，其内容包括各城邦的历史、行政、科学、艺术、人口、资源和财富等社会和经济情况的比较、分析，具有社会科学的特点。“城邦政情”式的统计研究延续了一两千年，直至 17 世纪中叶才逐渐被“政治算术”这个名词所替代，并且很快演化为“统计学”。

#### 1. 古典统计学阶段

从 17 世纪中叶到 18 世纪中叶是古典统计学阶段，当时主要有政治算术学派和国势学派。政治算术学派产生于 17 世纪中叶的英国，主要代表人物是威廉·配第。英国古典政治经济学的创始人威廉·配第的代表作《政治算术》一书，是经济学和统计学史上的重要著作。书中用“数字、重量、尺度”等定量的分析工具，对英国和当时主要发达国家的经济实力进行了比较分析。该书的出版标志着统计学的产生，因此威廉·配第被推举为统计学的创始人。

国势学派产生于 18 世纪的德国，国势学派又称记述学派，创始人是 17 世纪德国的海尔曼·康令。国势学派收集大量实际资料，分门别类、系统地记述了有关国情国力的重要事项，如人口、领土、政治、军事、经济、宗教、地理、风俗、货币等。国势学派使用了“统计学”这个名称。但由于其几乎不用数字而只用文字对国情国力进行系统的描述，所以人们也把它叫作记述学派，并认为国势学派有统计学之名而无统计学之实。

#### 2. 近代统计学阶段

近代统计学阶段是 18 世纪末到 19 世纪末。该阶段的主要贡献是建立和完善了统计学的理论体系，并逐步形成了以随机现象的推断统计为主要内容的数理统计学和以传统的政治经济现象描述为主要内容的社会统计学。

数理统计学派产生于 19 世纪中叶, 创始人是比利时的阿道夫·凯特勒, 他在统计学发展中的最大贡献是把概率论引入了统计学, 从而使统计学产生了质的飞跃。凯特勒的研究成果在自然科学、经济学、生物学等科学中得到不断的应用, 逐渐形成一门独立的学科。因此, 阿道夫·凯特勒被称为数理统计学的奠基人、“近代统计学之父”。

社会统计学派产生于 19 世纪后半叶, 创始人是德国的克尼斯。克尼斯在《独立科学的统计学》中, 提出了把“国势论”作为“国势学”的科学命名, 把“统计学”作为“政治算术”的科学命名, 从而结束了国势学派与政治算术学派长达 200 年之久的争论。

### 3. 现代统计学阶段

从 19 世纪末开始, 统计学进入现代统计学阶段。在这个阶段, 数理统计学与社会统计学逐步融合成为统一的现代统计学。“数理统计”是统计学在第三个发展阶段所形成的所有收集和分析数据的新方法的一个综合性名词。概率论是数理统计方法的理论基础, 它不属于统计学的范畴, 而属于数学的范畴。

我国的统计学起初是从国外传入的, 1949 年以前一直存在数理统计学派和社会经济统计学派。中华人民共和国成立初期, 人们认为只有社会经济统计学才是唯一的统计学, 从而在根本上否定了数理统计学是统计学的组成部分, 这严重妨碍了整个统计学的发展。改革开放以后, 人们被禁锢的思想终于获得解放, 经过长期、广泛的认识和探讨, 我国统计学科建设取得了重大突破和质的飞跃。1996 年 10 月, 中国统计学会、中国数理统计学会、中国现场统计学会联合举办全国统计科学研讨会, 这次会议形成了中国各统计学科和各统计学派之间相互借鉴、相互融合、共同发展的思想, 确立了统计学科体系的基本框架, 肯定了统计学是包括社会经济统计学和数理统计学在内的一般方法论性质的科学, 这为今后我国统计学的发展奠定了坚实的基础。

## 1.1.3 应用是统计学的生命

统计学在中国被列为一级学科, 这只是统计与国际接轨的第一步。由于统计界长期受前苏联模式的影响, 与中国其他人文社会科学一样, 统计学面对现代市场经济也越来越不适应。

中国仍处在新旧经济体制的变革时期, 面对市场经济, 任何学科都需要调整自己的位置, 重新审视和制定自己的发展目标。市场经济的理论和实践对统计信息的需求急剧增加, 对统计学理论和方法提出了更高的要求, 前苏联模式的统计学面对市场经济的需求已显得无能为力。现代市场经济理论与时代的潮流使中国人文社会科学、财经管理类学者的弱点暴露无遗。中国几千年的优秀文化赋予中国学者思辨性的思维和深厚的底蕴, 但也留下了忽视形式逻辑的弊端。先秦诸子以来, 我国学者的思辨性思维是较为发达的, 但这些思辨性思维常常缺乏经验事实作为基础, 空泛的议论较多, 实质性的内容较少, 逻辑推理的思维方式和定量分析淡化, 更不追求严密的公理化体系, 以至于我国人文社会科学、财经管

理类学者面对现代社会经济的发展有力不从心之感。

在 21 世纪,如何摆脱困境、走向未来,置身于古老东方文化氛围之中的中国学者需要冷静反思。时代要求我们必须抛开偏见,克服由于文化背景的差异所产生的观念屏障,正确理解并批判地吸收建立在发达商品经济基础上的外来文化,加强数学方法、统计学方法的学习,提高定性分析与定量分析相结合的能力。这样,在新世纪的实践进程中,中国人才才能大步赶上世界发达国家。现在部分高校在人文社会科学、财经管理类专业开设统计方法与技术的课程,就是向前迈进的一种有力举措。

哪里有数据,哪里就有统计。统计学是总结经验的学科。统计方法可以帮助我们获得利益,统计方法可以保护我们已获得的利益。统计数据中隐含着非常丰富的宝贵信息,如果用简单的描述统计方法,就只能对统计数据进行粗加工,难以看到数据中隐含的规律性。要从杂乱无章的统计数据中发现事物质的规律性就像雾里看花,只有靠现代统计分析方法才能真切地看到其统计规律性。

人文社会科学、财经管理类学者要跟上时代的步伐,就必须在研究事物质的规律性的同时,研究其量的规律性。任何现象都是质和量的统一,质是量的基础,而一定的量是质存在的必要条件。定性分析是定量分析的基础,定量分析则使定性分析更加准确可靠,使人们对质的规律性的认识更加深入和全面。在马克思主义理论的指导下,在人文社会科学、财经管理类研究中融入量化分析,是学科本身进一步完善的表现。

当今中国不仅需要经济理论家,更需要高级经济管理人才。在世界经济贸易中,随着 ISO 9000 系列标准的形成和发展,质量体系认证在世界各国兴起,20 世纪末一股 ISO 9000 热潮席卷全球,质量体系认证日益成为国际贸易中衡量供方质量保证能力和水平的标志。对统计技术在质量管理中的作用,最早提出全面质量管理理论的美国通用电气公司质量管理部部长费根堡姆博士曾说:“在全面质量管理活动中,无论何时、何地都会用到统计方法,但是统计方法只是全面质量管理的一个内容,它不等于全面质量管理。”“统计方法所表达的观点,对全面质量管理的整个领域都有深刻影响。”ISO 9000 系列标准的基本点是要求企业在生产过程中建立有效的质量保证体系,并对质量体系中相互关联、相互作用的若干要素进行有效的控制。在过程质量控制中,科学、有效的方法之一就是统计方法。因此,在 ISO 9000 系列标准的各个模式中以及质量管理和质量体系要素指南中,都要求在市场分析、产品设计、工序控制、性能评定、数据分析等方面广泛使用统计技术,其范围包括实验设计、方差分析、显著性检验、累计和控制图、抽样检验等技术。今天世界管理界被普遍认可的六西格玛管理更是统计学思想和方法的典型应用。

现代经济学一个很重要的标志就是模型技术的应用,而这里的模型技术更多是指统计模型技术。诺贝尔经济学奖获得者萨缪尔森曾说,第二次世界大战以后的经济学是计量经济学的时代。计量经济学的模型技术就是以统计模型为基础的。自 1969 年诺贝尔经济学奖设立以来,已有 80 多位学者获奖,其中 2/3 以上的获奖者是统计学家、计量经济学家、数学家。从大多数获奖者的经历和著作看,他们对统计方法的运用都很娴熟。瑞典皇家科学院爱立克·伦德伯教授在诺贝尔经济学奖的授奖仪式上精辟地指出:“你们都是把经济

学发展为数学和定量科学的先行者，你们借助于发展成熟的理论和统计分析来创造经济政策和经济规划的合理基础的贡献，涉及重大科学突破。”这足以说明统计量化分析研究已成为现代经济科学研究的重要手段。

在 21 世纪的进程中，中国的人文社会科学、财经管理类学者肩负着时代的重托。社会和谐发展，社会主义新农村建设，金融、投资风险，人口与社会保障，经济持续增长，环境污染治理等问题，都迫切地等待着我们去深入研究。这些问题的研究都将借助统计方法与技术的应用而获得成功。

### 1.1.4 统计学名人名言

1925 年，英国著名小说家赫伯特·乔治·威尔斯（Herbert George Wells）说：“有一天，统计思想会和读写能力一样，成为一个有效公民的必需能力。”威尔斯尤以科幻小说创作闻名于世，毕业于英国皇家理学院，任教于伦敦大学。一个现代人如果没有读写能力将是一个不合格的公民，可见统计学课程对于我们是多么重要。



Herbert George Wells

“在终极的分析中，一切知识都是历史；在抽象的意义下，一切科学都是数学；在理性的世界里，所有的判断都是统计学。”

——C. R. Rao

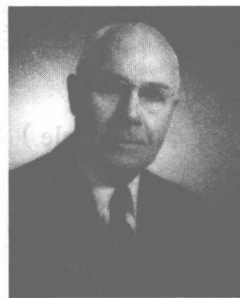
C. R. Rao 是美国科学院院士、英国皇家统计学会会员、国际上最伟大的统计学家之一，他于 1920 年 9 月 10 日出生于印度的一个贵族家庭，1940 年获印度安德拉大学数学学士学位，1943 年在印度统计研究所取得统计学硕士学位，随后赴英国剑桥大学师从现代统计学的奠基人费歇尔（R. A. Fisher）教授，并于 1948 年获得剑桥大学博士学位。



C. R. Rao

C. R. Rao 认为，统计学基本上是寄生的，靠研究其他领域内的工作而生存。这不是对统计学的轻视，而是因为对很多寄主来说，如果没有寄生虫就会死。对有的动物来说，如果没有寄生虫就不能消化它们的食物。因此，人类奋斗的很多领域，如果没有统计学，虽然不会死亡，但一定会变得很弱。

萨维奇（L. J. Savage）是美国坝工专家，早年就读于威斯康星大学工程系，1903 年获博士学位，1924 年任垦务局工程和研究中心设计总工程师。他一生参与了 60 余座大坝的建设。萨维奇是长江三峡工程的倡导者之一，1944 年及 1946 年应中



L. J. Savage

国政府邀请两次来华工作，编写了《扬子江三峡计划初步报告》。

萨维奇说：“统计学没有任何固定的对象，是一门独特的学问！统计学研究的是来自各领域的的数据，因解决其他领域内的问题而存在并兴旺发达。”

品读完几位统计学名人的名言，对于什么是统计学，统计学在人们日常生活和现代社会、经济中的地位，我们大概已心中有数了。

## 1.2 统计学中的几个基本概念

可以想象作为一级学科的统计学内容非常多，名词概念自然不少，这里先来介绍几个最常用的概念。

### 1.2.1 总体与个体

总体 (population) 是包含研究对象全体的集合，组成总体的每个元素称为个体。如要研究某专业高等数学的考试情况，则该专业全体学生构成总体，而每一位学生是一个个体。

总体根据其所包含的单位数目是否可数可以分为有限总体和无限总体。有限总体是指总体的范围能够明确确定，而且元素的数目是有限可数的。例如，由若干个企业构成的总体就是有限总体，一批待检验的电子元件也是有限总体。无限总体是指总体所包括的元素是无限的、不可数的。例如，在科学实验中，每个实验数据可以看作总体的一个元素，而实验可以无限地进行下去，因此由实验数据构成的总体就是一个无限总体。

实际上，在统计问题中，我们关心的总体是事物的某一个特征，并非事物本身。如上所述，对于一批待检验的锂电池，我们更关注每个锂电池的寿命，而不是锂电池本身，所以也可以把这批锂电池的寿命集合作为总体，这个总体是由一些实数构成的集合。通常情况下，统计上的总体是一组观测数据，而非一些物品的集合。

### 1.2.2 样本与样本量

样本 (sample) 是从总体中抽取的一部分元素的集合，构成样本的元素的数目称为样本量 (sample size)。抽取样本的目的是根据样本提供的信息推断总体的特征。例如，从一批锂电池中随机抽取 100 块，这 100 块就构成了一个样本，然后根据这 100 块锂电池的平均使用寿命推断该批锂电池的平均使用寿命。



### 1.2.3 参数

参数 (parameter) 是用来描述总体特征的概括性数字度量, 它是人们想要了解的总体的某种特征值。人们所关心的参数通常有总体平均数、总体标准差、总体比例等, 总体参数通常用希腊字母表示。例如, 总体平均数用  $\mu$  (mu) 表示, 总体标准差用  $\sigma$  (sigma) 表示, 总体比例用  $\pi$  (pi) 表示, 等等。

由于总体数据通常是不知道的, 所以参数往往是一个未知的常数。例如, 我们不知道某大公司所有员工的平均薪水, 不知道一个城市老年人口的比例, 不知道一批产品的合格率, 等等。因此, 才需要进行抽样, 并根据样本去估计总体参数。

### 1.2.4 统计量

统计量 (statistic) 是用来描述样本特征的概括性度量。它是根据样本数据计算加工出来的一个量, 由于抽样是随机的, 因此统计量是样本的一个函数, 它不依赖于任何的未知参数。人们所关心的统计量主要有样本平均数、样本标准差、样本比例等, 样本统计量通常用英文字母表示。例如, 样本平均数用  $\bar{x}$  (读作 x-bar) 表示, 样本标准差用  $s$  表示, 样本比例用  $p$  表示, 等等。除样本均值、样本比例、样本方差这类统计量外, 还有一些是为满足统计分析的需要而构造出来的统计量, 比如用于统计检验的  $z$  统计量、 $t$  统计量、 $\chi^2$  统计量、 $F$  统计量, 等等, 它们的含义将在后面相关的章节中作介绍。另外, 有关总体、样本、参数、统计量的概念可以用图 1-1 来直观描述。

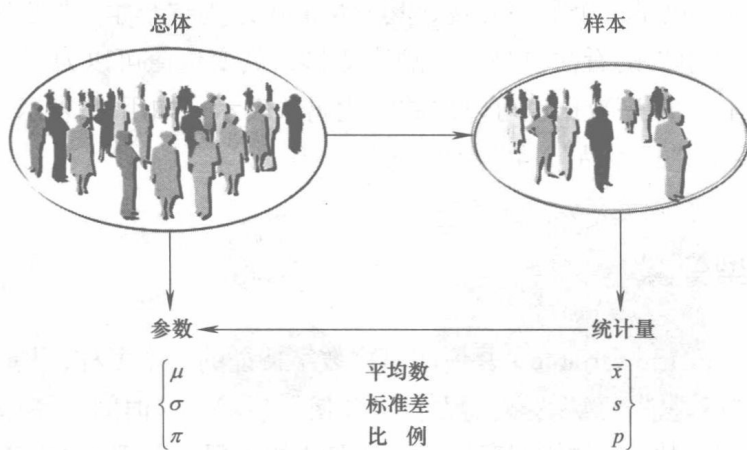


图 1-1 总体和样本、参数和统计量