

$\int$ TATISTICS

第2版

统计学

向蓉美 王青华 马丹 主编



机械工业出版社
China Machine Press

高等院校精品课程
系列教材

TATISTICS

第2版

统计学

向蓉美 王青华 马丹 主编
苏远琳 黎春 雷敏 夏怡凡 参编



机械工业出版社
China Machine Press

图书在版编目(CIP)数据

统计学 / 向蓉美, 王青华, 马丹主编. —2 版. —北京: 机械工业出版社, 2017.5
(高等院校精品课程系列教材)

ISBN 978-7-111-56719-6

I. 统… II. ①向… ②王… ③马… III. 统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字 (2017) 第 074468 号

本书是关于统计学基本理论和方法的教材, 强调重思想、重应用, 按照先描述统计、后推断统计的思路编写。第 1 章概括性地介绍了统计学的形成及基本概念; 第 2~5 章属于描述统计学的基本理论和方法; 第 6~9 章属于推断统计学的基本理论和方法; 第 9 章相关与回归分析是“计量经济学”课程的起点。

本书可以作为经济类和管理类非统计专业本科生的统计学课程教材, 也可以作为统计学专业本科生的统计学入门教材, 还可以作为广大经济管理人员学习统计学的参考书。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 宋 燕

责任校对: 王 欣

印 刷: 北京诚信伟业印刷有限公司

版 次: 2017 年 5 月第 2 版第 1 次印刷

开 本: 185mm×260mm 1/16

印 张: 16.25

书 号: ISBN 978-7-111-56719-6

定 价: 42.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379210 88361066

投稿热线: (010) 88379007

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzjg@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

作者简介



向蓉美

西南财经大学统计学院教授，博士生导师，四川省教学名师。统计学国家精品课程、统计学国家精品资源共享课程、四川省统计学教学团队负责人。中国投入产出学会常务理事、中国国民经济核算研究会理事、成都市统计学会常务理事。1990~1991年作为访问学者到美国交流，2004年受德国学术交流中心（DAAD）资助赴德国研究访问3个月。先后主持完成3项国家社科基金项目、9项省部级课题，获6项省部级优秀研究成果奖、科技进步奖和优秀教学成果奖，教育部、团中央等第十届挑战杯全国大学生科技作品竞赛优秀指导教师，曾获西南财经大学优秀教师一等奖。主要研究领域为经济统计。出版专著3部，主编和独立编写研究生、本科生教材十余部，其中主编的《统计学导论》被评为“十二五”国家规划教材。



王青华

西南财经大学统计学院副教授，硕士生导师，主要讲授课程有统计学、金融统计分析和宏观经济统计分析等。主编的《统计学导论》被评为“十二五”国家级规划教材，主编及参编教材十余部，发表学术论文十多篇，参加多项国家级和省部级课题研究，曾荣获四川省教学成果二等奖、四川省第十三次哲学社会科学优秀成果二等奖及国家统计局优秀科研成果二等奖。



马丹

女，统计学博士，西南财经大学统计学院经济统计系教授，博士生导师。西南财经大学经济统计系主任，国家级精品课程统计学主讲教师。第十一批四川省科学技术带头人后备人选，四川省统计学会理事、成都市统计学会理事、南方工业统计研究会常务理事。从事经济、金融、统计领域研究工作，在本专业重要学术期刊发表学术论文40余篇，累计逾50万字。参加完成国家社科基金等国家级、省部级课题十余项。主持国家社会科学基金、教育部人文社会科学基金等课题。先后多次获得各级学术嘉奖，获得全国百篇优秀博士论文提名奖、全国统计科研成果奖优秀博士论文二等奖、全国统计科研成果奖优秀教材类一等奖、四川省教学成果奖一等奖等。

第2版前言

统计学作为经济类和管理类专业的必修课，也一直是西南财经大学的大学科基础课，经过多年的建设，西南财经大学的统计学先后被评为四川省精品课程、国家精品课程、国家精品资源共享课程。该课程的教材体系、教学内容、教学资源日趋成熟完善。2015年，党的十八届五中全会公报提出要实施“国家大数据战略”，大数据时代的到来，让“数据即资产”已经成为新的全球发展趋势。在数字浪潮中，统计学无疑将扮演更加重要的角色，统计学教学也面临新环境和新问题，统计学课程教材的适时更新与不断完善势在必行。

本书是在2013年机械工业出版社出版的《统计学》一书的基础上修订而成的。《统计学》第1版自出版发行以来受到了广大读者的欢迎，先后多次重印，被许多高等院校作为教材或主要参考书。本次修订在保留了原书的基本框架和内容体系的基础上，对各章体例进行了调整，将课后习题单独整理为配套练习册——《统计学学习指导及能力提升训练》，便于教师讲授和学生完成练习。此外，我们适时更新了教材数据和案例，将最新的统计数据引入教材，在教学资料等内容的时效性方面进行了更新和充实。

全书共9章，第1~5章主要涉及描述统计基本概念和方法；第6~9章主要介绍数理统计基本概念和方法。第1章总论论述了统计学的基本概念、特点及历史发展等。第2章主要介绍统计数据的收集、数据整理和展示。第3章介绍数据集中趋势、离中趋势以及分布形态。第4章讲授确定性时间序列分析的基本方法。第5章讲授统计指数理论综合评价。第6章介绍三大抽样分布及抽样分布定理。第7章介绍抽样估计方法和主要的抽样组织方式。第8章讲述假设检验和方差分析的基本理论和方法。第9章讨论了相关与回归分析的基本问题。再版教材由各章原作者进行修订，向蓉美、王青华和马丹三位主编进行统稿。

在本书第2版的修订编写过程中，我们参阅了国内外专家、学者的著作，也参考了同行的相关教材，在此对他们表示崇高的敬意和衷心的感谢！机械出版社的领导及编辑对本书的修订出版付出了辛勤的劳动，也一并表示诚挚的谢意！由于作者水平有限，加上时间仓促，书中错漏和不妥之处在所难免，恳请专家、同行和读者批评指正。

编者

2017年3月于西南财经大学

第 1 版前言

统计的力量是如此强大、无处不在，哪里有数据，哪里就有统计，统计是现代人谋生的手段，因为我们正处于信息化的时代，数字化信息随处可见，所以国家教育部把统计学列为经济类、管理类大学生的专业（核心）基础课。

统计学的内容十分丰富。本书是一本关于统计学基本理论和方法的教材，适用于经济类和管理类非统计专业本科学生，也可以作为统计学专业的统计学入门教材，还可以作为广大经济管理人员学习统计学的参考书。

本书强调重思想、重应用，按照先描述统计、后推断统计的思路编写。只有对样本进行了描述，才能对总体进行推断，描述统计是推断统计的基础。本书第 2~5 章属于描述统计学的基本理论和方法，第 6~9 章属于推断统计学的基本理论和方法。教材最后一章相关与回归分析，是“计量经济学”课程的起点。这样的安排符合知识点的逻辑关系，便于课程之间的合理衔接和教学。

本书把统计知识的学习与计算机软件的运用有机结合，采用 Microsoft Excel 作为数据处理软件。Excel 是最普及的数据处理和分析软件，非常便于统计计算和统计制图。在本书各章中，结合实例介绍其应用。少数 Excel 不能直接处理的计算，我们将简单介绍如何用 SPSS 处理，如单个总体的假设检验、等级相关系数的计算。

为了帮助大家更好地学习、理解和应用所讲知识，每章后面配有本章小结、基本知识梳理以及灵活多样的练习题。

根据学生对象不同或课时安排不同，标记有 * 的内容可以选择讲授。

2007 年西南财经大学的统计学学科被评为国家级重点学科，“统计学”课程被评为国家级精品课程，2012 年“统计学”课程被推荐为国家精品资源共享课。不断提高《统计学》教材的质量，是我们建设重点学科和统计学精品课程最重要的内容之一。本书由西南财经大学统计学院几位长期从事统计学教学的教授和副教授共同编写而成。作者和具体分工如下：向蓉美第 1 章和第 4 章，苏远琳第 2 章，黎春第 3 章，雷敏第 5 章，夏怡凡第 6 章，马丹第 7 章，王青华第 8 章和第 9 章。由向蓉美、王青华和马丹对全书进行总体设计和修改、总撰。谢小燕教授参加了教材总体设计的讨论，提出了很好的建议。

随着社会经济的不断发展和统计理论的不断完善，我们将适时根据其发展变化情况修订本书，也恳请使用本书的读者将建议和意见及时反馈给我们，对此我们表示衷心的感谢。

作者

2013 年 2 月于西南财经大学



目 录

作者简介	
第2版前言	
第1版前言	
第1章 总论	1
引例	1
1.1 为什么要学习统计学	2
1.2 什么是统计	5
1.3 统计学中的基本术语	12
1.4 用 Excel 进行统计分析	19
本章小结	21
思考与练习题	22
第2章 统计数据的收集、整理与显示	23
引例	23
2.1 统计数据的收集	24
2.2 统计数据的整理	28
2.3 统计数据的图表显示	34
本章小结	42
思考与练习题	42
第3章 数据特征的描述	44
引例	44
3.1 总量指标和相对指标	45
3.2 数据集中趋势的描述：平均指标	50
3.3 数据离中趋势的描述：变异指标	59
3.4 数据分布形态的描述：偏度和峰度	63
3.5 利用统计软件计算相关指标和绘制统计图	66
本章小结	70
思考与练习题	71
第4章 时间序列分析	73
引例	73
4.1 时间序列概述	74
4.2 时间序列的水平分析	77
4.3 时间序列的速度分析	81
4.4 时间序列的构成因素分析	88
本章小结	108
思考与练习题	109
第5章 统计指数与综合评价	111
引例	111
5.1 统计指数的概念、作用和分类	111
5.2 总指数的计算	114
5.3 指数体系与因素分析	122

5.4 综合评价	127
本章小结	132
思考与练习题	132

第6章 统计量与抽样分布..... 134

引例	134
6.1 总体分布及其样本特性	135
6.2 总体参数与样本统计量	139
6.3 抽样分布及抽样分布定理	142
本章小结	154
思考与练习题	154

第7章 参数估计 155 |

引例	155
7.1 抽样估计的基本问题	156
7.2 点估计	158
7.3 简单随机抽样的区间估计	166
7.4 其他抽样组织方式的抽样 误差	173
本章小结	177
思考与练习题	178

第8章 假设检验与方差分析 ... 179

引例	179
8.1 假设检验的一般问题	180
8.2 单个总体参数的假设检验 ...	186
* 8.3 两个总体参数的假设检验 ...	190
8.4 多个总体均值的检验: 单因素 方差分析	197
本章小结	203
思考与练习题	203

第9章 相关与回归分析 205 |

引例	205
9.1 相关与回归分析概述	206
9.2 一元线性相关分析	209
9.3 一元线性回归分析	214
* 9.4 多元线性回归分析	226
本章小结	231
思考与练习题	232

附录A 常用统计附表 234 |

部分参考答案 248 |

参考文献 252 |

总 论

引 例

统计量化历史，数据见证辉煌。



中华人民共和国国家统计局

National Bureau of Statistics of the People's Republic of China

走近统计

国家统计局局长致辞
机构职能

统计数据

最新发布
数据查询
数据解读

统计工作

统计动态
通知公告
图片新闻

统计知识

统计百科
统计词典
常见问题解答

统计服务

网上办事
统计咨询
局长信箱

信息公开

公开指南
公开目录
公开年报

当前位置 > 首页 > 专题统计 > 专题分析

专题分析

*金砖国家联合统计手册 (2014)	2014-07-15
*金砖国家联合统计手册 (2013)	2013-10-24
*金砖国家联合统计手册 (2012)	2013-10-24
*金砖国家联合统计手册 (2011)	2013-10-24
*图说13年成就	2013-10-24
*迎接十六大	2013-10-24
*科学发展 成就辉煌	2012-08-15
*“辉煌统计80年”系列报道	2012-08-07
*“十一五”经济社会发展成就系列报告	2013-10-24
*“十五”成就系列报告	2013-10-24
*庆祝新中国成立60周年系列报告	2013-10-24
*纪念改革开放30年	2013-10-24
*工业生产价格指数	2013-10-24
*从十六大到十七大经济社会发展回顾系列报告	2013-10-24
*庆香港回归十周年特别报道	2013-10-24
*其他热点调查分析	2013-10-24

 **中华人民共和国国家统计局**
National Bureau of Statistics of the People's Republic of China

近 国家统计 统计 最新发布 统计 统计动态 统计 统计百科 统计 网上办事 信息 公开指南
代 计 局 机构职能 数据 数据查询 工作 通知公告 知 统计词典 服务 统计咨询 公 公开目录
统计法规 数据解读 图片新闻 常见问题解答 失信企业公示 公开年报

当前位置 > 首页 > 统计数据 > 最新发布

新常态新战略新发展——“十二五”时期我国经济社会发展成就斐然

来源：国家统计局 发布时间：2015-10-13 08:51 [打印本页](#) [关闭窗口](#)

新常态 新战略 新发展^[1]

——“十二五”时期我国经济社会发展成就斐然

“十二五”时期的五年，是改革开放伟大历史进程中具有鲜明里程碑意义的五年。这五年，世情国情发生深刻变化，我国经济发展步入新常态。从国际看，世界经济处在危机后的深度调整期，呈现低增长、不平衡、多风险的特征，地缘政治等非经济因素影响加剧，我国发展面临的外部环境更趋复杂。从国内看，“三期叠加”的阵痛持续加深，多重困难和挑战相互交织，改革转型任务繁重。面对复杂多变的国际环境和艰巨繁重的改革发展任务，党中央、国务院总揽全局，审时度势，紧紧围绕“四个全面”战略布局，主动适应引领经济发展新常态，扎实推动“大众创业、万众创新”，坚持稳中求进工作总基调，不断创新宏观调控政策、思路、方式；坚持实施创新驱动发展战略，大力推进结构调整和转型升级；坚持改革开放不动摇，着力开拓发展新空间、激发新动力，实现了经济平稳较快发展和社会和谐稳定，为全面建成小康社会奠定了坚实基础。

统计是静止的历史，历史是流动的统计。每逢大事，在国家统计局的网站上都会出现专题文章，因为发展、成就等需要统计数据说明。这些数据是怎么得到的？统计数据的含义是什么？如何利用这些数据？本书各章会为你一一解读。

1.1 为什么要学习统计学

1.1.1 感悟统计

统计是社会认识最有力的武器之一。

——列宁

某些人不喜欢统计这个名词，但我却发现其中充满了乐趣……它们处理各种复杂现象的能力是非凡的，它们是追求科学的人从荆棘丛生的困难阻挡中开辟道路的最好工具。

——英国著名生物学家、统计学家高尔顿

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计划的合理基础的贡献，涉及重大科学突破……我很荣幸地向你们转达瑞典皇家科学院的祝贺，并且请你——丁伯根教授，从国王陛下手中接受1969年度阿尔弗雷德·诺

贝尔经济学奖金。

——爱立克·伦德伯教授在第一届诺贝尔经济学奖颁奖大会上的讲话^①

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；在理性的基础上，所有的判断都是统计学。

——C. R. 劳《统计与真理》^②

好的数据确实胜过轶闻。比起轶闻和光大声嚷嚷预测未来，数据要客观得多。统计和其他的公开论述比起来，它根据事实且较科学又较理性。对于争议性的议题，统计研究应该比其他大部分证据受到更多的重视。

——戴维 S. 穆尔^③

学者不能离开统计而研究，政治家不能离开统计而施政，事业家不能离开统计而执业。

——我国著名经济学家、人口学家马寅初

众所周知，《红楼梦》一书共 120 回，自从胡适做《红楼梦考证》以来，一般都认为前 80 回为曹雪芹所写，后 40 回为高鹗所续；然而长期以来这种看法一直都饱受争议。从 1985 年开始，复旦大学的李贤平教授带领他的学生从统计角度做了考证。一般认为，同一情节大家描述得都差不多，但由于个人写作特点和习惯的不同，所用的虚词是不一样的。他们创造性的想法是将 120 回看成是 120 个样本，然后确定与情节无关的 47 个虚词（之、其、或、亦……呀、吗、咧、罢……可、便、就等）出现的次数（频率），作为《红楼梦》各个回标志，利用统计方法果然能将 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《红楼梦》不是出自同一人的手笔；之后又进一步分析前 80 回是否为曹雪芹所写，又找了曹雪芹的其他著作，做了类似计算，结果证实了用词手法完全相同，断定前 80 回为曹雪芹一人手笔，而后 40 回不是高鹗一个人所写。这个论证在红学界轰动很大，使红学界大为赞叹。

——红楼梦作者考证^④

1.1.2 统计学是一门应用范围很广的科学

统计学是一个研究领域非常丰富、应用范围非常广泛的科学。我们正处于信息经济

① 王宏昌，林少宫．诺贝尔经济学奖金获得者讲演集（1969～1995）[M]．北京：中国社会科学出版社，1997.

② C. R. 劳是当代国际最著名的统计学家之一。C R 劳．统计与真理 [M]．北京：科学出版社，2004.

③ 戴维 S 穆尔．统计学的世界（第 5 版）[M]．郑惟厚，译．北京：中信出版社，2003.

④ <http://www.qintai.net/forum.php?mod=viewthread&tid=6254>.

时代,世界充满了数据,数字化信息随处可见。

无论我们是否学过统计学、懂得统计学,我们生活中的每一天都会遇到大量统计问题,新闻和大众媒体每天都在表现统计数字。例如,每年的《政府工作报告》总是要列举大量的水平、比例、结构、速度等数据,说明国民经济的发展状况;统计部门每月、每年都要公布居民消费价格指数(CPI),反映一定时期内城乡居民所购买的生活消费品和服务项目价格变动趋势和程度。确实,用“好”“比较好”“很好”“非常好”“比任何时候都好”等这样的比较级或最高级的词语来反映国家的经济形势就太苍白、太空洞了,而2016年我国现价国内生产总值达到744 127亿元,按可比价比2015年增长6.7%^①,比1978年增长27倍多,对世界经济增长的贡献率为33.2%;2016年全年全国居民人均可支配收入23 821元,比上年名义增长8.4%,扣除价格因素实际增长6.3%。全国居民人均可支配收入中位数20 883元,比上年名义增长8.3%等。这样的统计数据实实在在、具体地反映了中国改革开放以来和2016年取得的成就。

宇宙间万事万物,林林总总,各种事物可依特定的性质予以归类,形成各种群体——现象总体。无论是属于自然的、实验的,还是社会的、经济的,凡是可以数据表现的总体,都可以作为统计的研究对象。在很多学科中,统计学都是必需的基础知识,统计方法和统计思想渗透到社会、经济、自然、科技、生活的每一个角落,甚至一篇文章是否有质量,能否在高级别的杂志上发表,很大程度上要看其统计数据和统计方法应用是否正确。统计学不仅在社会经济领域得到发展,而且一些过去与数量毫无联系的学科,如政治学、法学、历史学、艺术学、考古学等,都在对应用统计方法技术进行研究和实践。

有人说统计学是21世纪最有前途的学科之一。因为从20世纪后半半个世纪起,人文社会科学的发展与统计学的关系越来越紧密,统计学的发展已经渗透到人文社会科学的许多领域,并由此产生许多新的学科,如人口统计学、历史统计学、教育统计学、心理统计学、社会统计学等。从本质上讲,信息经济所依赖的不只是信息处理手段的先进性,更重要的是信息收集、整理的准确性,而准确的信息收集、整理离不开统计学学科的发展;在一些发达国家,统计学是大学里最受重视的学科,统计学发展得如何是衡量某一大学学术水平的标志。在这些国家,统计学是强势学科。

英国作家、历史学家韦尔斯(H. G. Wells)曾经说:“统计思维总有一天会像读与写一样,成为一个有效率公民的必备能力”,“像今天有能力的公民能读会写一样,将来有一天要求有能力的公民必须会计算,而且能够利用平均值、最大值和最小值。可以

① 国家统计局“中华人民共和国2016年国民经济和社会发展统计公报”以下有关国民经济2016年统计数据均来自此公报。

预期,这样的时代已经不远了”^①。

统计分析涉及大量数据,所以统计人员、利用统计数据的人员、研究人员总是利用计算机软件进行数据处理和计算。统计分析中能够使用的软件很多,本书使用最为普遍的是 Excel 软件,少数不能用 Excel 软件处理的分析,我们介绍 SPSS 软件。

1.1.3 本书的内容结构

本书按照先描述统计学、后推断统计学的思路编写。只有对总体进行了描述,才能认识总体,只有对样本进行了描述,才能对总体进行推断,描述统计是推断统计的基础。第2~5章属于描述统计学的基本理论和方法,第6~9章属于推断统计学的基本理论和方法。本书第9章相关与回归分析是“计量经济学”课程的起点。这样的安排符合知识点的逻辑关系,便于课程之间的合理衔接和教学。

1.2 什么是统计

1.2.1 统计的含义

“统计”是社会经济生活中经常使用的名词。在英语中单数 statistic 译为中文是统计学,复数 statistics 译为中文是统计,是指描述事物属性的实际数据,即统计数据。“统计”一般具有三种含义:统计工作、统计资料和统计科学。

1. 统计工作

统计工作即统计实践活动,是人们为了说明所研究对象的某种数量特征和数量规律性,而对该现象的数据进行收集、整理与分析的活动。例如,为了获得粮食产量而进行的抽样调查活动,为了获得全国人口的数量和构成等而进行的人口普查活动等。

统计实践活动有很久远的历史,早在公元前 2000 多年,统治者为了征兵、征税、管理奴隶的需要,就有了人口、土地、财产的统计活动。《通典》记载了我们历史上最早的统计数据:“禹平水土,为九州”,“九州之地凡 24 388 024 顷,人口 13 553 923 人”。到中世纪,西欧各国都有了人口、军队、领地、财产等的统计活动。

2. 统计资料

统计资料即统计数据,是通过统计工作所获得的能够说明现象总体某种特征的数据,是统计实践活动的成果。例如,2016 年我国现价国内生产总值为 744 127 亿元,按

^① C. R 劳. 统计与真理 [M]. 北京: 科学出版社, 2004.

可比价计算比 2015 年增长 6.7% 等, 这些数据就是广大统计部门和统计工作者辛勤劳动得到的能够说明我国经济发展水平的统计资料。统计数据最集中、最系统地反映在各种统计年鉴中, 如《中国统计年鉴》《国际统计年鉴》《中国经济年鉴》《中国金融年鉴》《中国物价及城镇居民家庭收支调查统计年鉴》等。《中国统计年鉴》涵盖了国民经济所有主要数据, 包括国民经济核算、各行业生产、就业、人民生活、对内对外贸易、社会活动、环境等方面的数据。随着互联网技术的发展和普及, 网络提供的数据成为统计数据主要来源已是趋势, 如中经网统计数据库、中宏产业数据库、万方商务信息数据库、EPS 全球统计数据库、CSMAR 期货股票分析高频数据库等。一些国际组织也有自己的统计数据库, 如联合国、世界银行、国际货币基金组织等。统计数据还可以从经济分析报告、专著、期刊里获得, 如《中国经济数据分析》《经济形势分析与预测 (经济蓝皮书)》, 一些期刊, 如《中国统计》里刊登的最新月度数据等也是统计数据的来源渠道。

3. 统计科学

统计科学即统计学, 是一门研究收集数据、表现数据、分析数据、解释数据, 从而认识现象数量规律、帮助人们更有效地进行决策的方法论科学。统计学源于统计实践活动, 是对统计实践活动的理论概括和总结, 又用于指导统计实践活动。

收集数据需要对客观现象做周密细致的调查; 表现数据需要对调查得到的数据加以整理, 使之成为反映现象总体的条理化、系统化的数据; 分析数据需要用科学的方法从数据中得出反映现象本质数量规律性的结果; 解释数据需要用有关知识对数量规律性做出说明。所有这些就构成了统计学的研究内容。

统计学是一门为定量分析提供方法的方法论学科, 因此它是一门应用性很强的学科, 它几乎与所有的学科领域都有着或多或少的联系, 凡是有数据的地方, 就有统计学的用武之地。统计学为其他学科提供研究数量规律性的方法, 但是统计学绝不是万能的, 各个学科数量规律的解释还需要由各学科的理论来完成。比如, 大量观察发现: 我国 2015 年人口的性别比, 从 0~4 岁组的 116.23, 到 70~74 岁组的 96.36, 再到 95 岁以上组的 39.36^①, 随着年龄的增大呈下降趋势。形成这样的比例和趋势的原因, 不是统计学能够解释的, 要用遗传学、医学甚至社会学的理论来解释。又如大量观察表明, 吸烟者患肺癌的比例大于不吸烟者患肺癌的比例, 吸烟是否会导致患肺癌? 为什么会导致患肺癌? 这些是医学研究的问题。

尽管统计学不能解决各门学科的所有问题, 但是, 各门学科离不开统计学, 统计学的理论和方法在各门学科的研究中会发挥越来越重要的作用。

① 《中国统计年鉴 2016 年》。

作为一门科学,统计学初创于17世纪中叶至18世纪初,当时主要的学派有政治算术学派、国势学派和社会统计学派。

政治算术学派的创始人是英国人威廉·配第。配第首先提出了用数量科学地研究社会经济现象的方法——政治算术。他的名著《政治算术》(1676年)就是指数字和统计学方法,在序言中配第明确指出:“我进行这项工作所使用的方法,在目前还不是常见的。因为和只使用比较级或最高级的词语以及单纯做思维的论证相反,我却采用了这样的方法(作为我很久以来就想建立的政治算术的一个范例),即用数字、重量和尺度的词汇来表达我自己想说的内容,只进行诉诸人们感官的论证和考察在性质上有可见的根据的原因。”他还说,“用数字、重量和尺度(它们构成我下面立论的基础)来表示的展望和论旨,都是真实的,即使不真实,也不会有明显的错误……因为,能够证明为确实的东西,也就是确实的。”他在研究社会经济现象的规律时,还应用推算法、分组法,编制了原始数据的图表,计算了一系列的总量指标、相对指标和平均指标,是最早估算国民收入的人。但是配第始终没有用“统计学”三个字,所以政治算术学派是有统计学之实,而无统计学之名。

国势学派的创始人是德国人赫尔曼·康令。1660年以后,康令开始定期地、系统地用对比的方法讲授国家比较方面的知识,他不仅讲述事实,而且试图探讨事实的因果关系。他把这个课程叫作“欧洲最近国势学”,于是“国势学”由此产生。1660年他首次在印刷品上使用 Statistik,这个单词的前半部 state 源于国家政策,是指政府部门记录人口出生和死亡信息的工作。当时,康令的学说在学术界产生了很大的影响,德国大学中的许多教授都称赞并追随他的学说。其中,戈特弗里德·阿亨瓦尔发扬了康令的学术思想,把这门课程定名为“统计学”,开始有了“统计学”这个名称。但是国势学派只是对各国情况做一般性的比较记述,如“某国人口众多”“土地辽阔”之类,而没有进行数量研究和描述。所以国势学派是有统计学之名,而无统计学之实。

近代统计学产生于18世纪末至19世纪末,当时主要的学派是数理统计学派和社会统计学派。数理统计学派的代表人物是比利时著名的统计学家、数学家、物理学家、天文学家和人类学家兰贝特·阿道夫·雅克·凯特勒。他融会贯通各家各派的统计思想,博采众长,把德国的国势学、英国的政治算术,以及意大利、法国的古典概率论加以协调、统一、改造和融合成为具有近代意义的统计学,把统计学推向了新的发展阶段。他将统计方法用于研究人类,促进了人口统计学的发展;他提出“平均人”理论,用平均值作为实际值的一种代表值;他用大量统计数据对犯罪问题的研究,被人称为“道德统计”;他创建“国际统计会议”组织,促进了国际统计交流与合作。可以说凯特勒是古典统计学的终结者,近代统计学的先驱者,在统计学发展史上起着承先启后、继往开来

的作用。

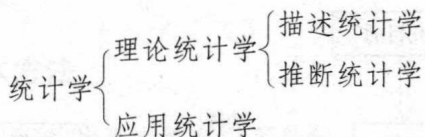
19 世纪下半叶,德国统计学界在英国政治算术学派的影响下,努力使统计学成为一门用统计数量表达社会经济现象及其规律的社会科学,从而社会统计学派逐渐形成。该学派的先驱者 K. G. A. 克尼斯把统计学的性质定义为“具有政治算术内容的社会科学”。他在《作为独立科学的统计学》(1850)一书中,提出了“国家论”与“统计学”科学分工的主张。认为国家论是用文字记述的国势学的科学命名,统计学则是用数值研究社会经济规律的政治算术的科学命名。社会统计学派的创始人 G. V. 迈尔在《社会生活中的规律性》(1877)一书中提出,统计学是通过对数据的大量观察,对人类社会生活的状态及其产生的规律做有系统的说明与研究。他明确指出统计学的研究对象是社会经济现象的规律,认为脱离规律性的研究就不能获得科学的认识,统计学不仅要确定事实数量的记述,还必须研究其规律性。另一位代表人物德国统计学家 E. 恩格尔提出了统计调查、整理和分析三阶段的统计方法。他通过对英国、法国、德国和比利时等国的工人家庭调查,撰写了《比利时工人家庭的生活费》(1895)一书,提出著名的恩格尔定律,即一个家庭(或个人)的收入越低,其食品支出在收入中所占比例就越高;反之,其比例就越低。恩格尔系数等于 $(\text{食品支出总额}/\text{收入}) \times 100\%$ 。

现代统计学一般认为从 19 世纪末 20 世纪初开始,其标志是推断统计学的问世。英国生物学家、统计学家卡尔·皮尔森于 1894~1895 年提出包括正态分布、矩形分布、J 形分布、U 形分布等 13 种曲线及其方程式。他的这一研究成果,打破了以往次数分布曲线的“唯正态”观念,推进了次数分布曲线理论的发展和应用,为大样本理论奠定了基础;他提出了著名的统计量卡方(χ^2)和卡方检验;他还提出了标准差概念及其符号 σ , 发展了相关回归理论等。英国统计学家威廉·西利·戈塞特于 1908 年在《生物统计学》上以 Student 笔名发表了《平均数的概率误差》,创立了 t 分布,开创了小样本理论的先河,解决了他多年使用小样本中许多悬而未决的问题。

统计学的发展史表明,随着社会的发展与实践的需要,统计学越来越多地依赖和吸收数学方法,使统计方法不断丰富和完善,不断发展和演变,成为研究社会经济现象、自然技术现象数量方面的方法论科学。目前,统计学越来越多地向其他学科领域渗透,形成各种以统计学为基础的边缘学科,随着统计学应用日益广泛和深入,特别是借助电子计算机,统计学所发挥的功效必将日益增强。

1.2.2 统计学的类型

由于统计学是一门应用非常广泛的科学,所以其内容体系也非常丰富。统计学可以做如下的分类:



理论统计学是把统计研究对象一般化、抽象化后，形成的可以应用于各种统计活动的一般统计理论和方法。本书的内容属于理论统计学。

统计学是一门收集和分析数据的科学，在社会科学和自然科学的领域中，都需要通过数据分析来解决实际问题，因而，统计方法的应用几乎扩展到了所有的科学研究领域。

应用统计学是统计的一般理论和方法应用到各个领域形成的科学，如国民经济统计学、货币金融统计学、管理统计学、人口统计学、心理统计学、医学统计学、生物统计学等。

1. 描述统计学

描述统计学是关于如何对现象的数据特征进行观测、整理、计量、表述的理论和方方法论科学，研究如何取得反映客观现象的数据，并通过图表形式对所收集的数据进行加工处理和显示，进而通过综合概括与分析得出反映客观现象的规律性数量特征。其特点是用从一个总体或样本中收集到的数据，来对这个总体或样本进行描述或得出有关这个总体或样本的结论。例如，全班有 50 名同学或从全校 15 000 名同学中抽取 400 名同学，把这 50 名同学或 400 名同学的成绩用图、表或特征值（如平均分数、分数的标准差和及格率等）表示出来，从而得出针对该班 50 名同学或 400 名同学学习情况的结论。

2. 推断统计学

推断统计学也称为统计推断，是关于如何抽取样本并利用样本数据推断总体有关数据的理论和方方法论科学。其特点是用从总体中随机抽取的样本数据，得出关于这个总体的结论。例如，从全校 15 000 名学生中，随机抽取 400 名学生进行学习情况调查。用这 400 名学生平均分数、分数的标准差和及格率等推算出全校 15 000 名学生的平均分数、分数的标准差和及格率等，从而得出全校学生学习情况的结论。“你可以将推断统计学看作在样本信息基础上对总体水平的‘最优猜测’”^①。

3. 描述统计与推断统计的关系

描述统计是基础，只有依靠描述统计收集、整理和显示可靠的统计数据并提供有效的样本信息，推断统计才能进行，其关系如图 1-1 所示。

① 道格拉斯 A 林德，等．商务与经济统计方法：全球数据集（英文版·原书第 13 版）[M]．冯燕奇，叶光，聂巧平，译．北京：机械工业出版社，2009.