

高等学校经济统计专业系列教材及配套资源

# 统计学

主 编 刘定平

高等教育出版社

高等学校经济统计专业系列教材及配套资源

# 统计学

主 编 刘定平

副主编 周福林

高等教育出版社·北京

## 内容简介

本教材为河南省统计学优秀教学团队和河南省精品在线开放课程的建设成果之一。它融合了描述统计和推断统计、经济统计和数理统计的基本知识和方法,主要内容包括总论、数据收集与整理、统计描述、统计推断、方差分析与回归分析、多元统计分析、时间序列分析、统计指数、统计综合评价、国民经济统计等10章。

本书既可以作为高等院校经济和管理类专业“统计学”课程的教材,也可作为统计类各个专业的“统计学导论”教材。书中配备了拓展知识的二维码,通过扫描就可以获得其资源。同时,还提供了教学课件,供老师们下载使用。

## 图书在版编目(CIP)数据

统计学/刘定平主编.--北京:高等教育出版社,  
2019.3

ISBN 978-7-04-051208-3

I. ①统… II. ①刘… 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(2019)第009199号

策划编辑 吴淑丽  
插图绘制 于博

责任编辑 吴淑丽  
责任校对 王雨

封面设计 于文燕  
责任印制 田甜

版式设计 马云

出版发行 高等教育出版社

社址 北京市西城区德外大街4号

邮政编码 100120

印刷 北京宏伟双华印刷有限公司

开本 787mm×1092mm 1/16

印张 19

字数 470千字

购书热线 010-58581118

咨询电话 400-810-0598

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

网上订购 <http://www.hepmall.com.cn>

<http://www.hepmall.com>

<http://www.hepmall.cn>

版 次 2019年3月第1版

印 次 2019年3月第1次印刷

定 价 41.00元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换

版权所有 侵权必究

物料号 51208-00

## 作者简介

### 刘定平

河南财经政法大学教授,统计学院院长,大数据研究院院长,河南省人文社科重点研究基地“教育统计数据分析与研究中心”主任,中国统计学会常务理事,中国高等教育统计学会常务理事,河南省统计学类专业本科教学指导委员会主任委员,河南省优秀教师、河南省教学名师,2018年入选河南省首批“中原千人计划”专家。主讲“新生成长研讨课”“统计思想与学科前沿”“贝叶斯统计”和“统计案例”等课程,研究方向为经济定量分析与统计应用研究。出版专著1部,主编教材6部,在《数量经济技术经济研究》《中国统计》《中州学刊》等杂志发表论文约30篇,主持或参与完成省和校教改项目约20项,为河南省统计学省级重点学科、统计学省级特色专业、统计学省级专业综合改革试点、统计学省级优秀教学团队、统计学省级优秀基层教学组织、统计学省级精品在线开放课程的主持人或负责人。主持或参与完成国家和省部级科研项目约25项,获得过省级一等奖1项、二等奖5项、三等奖2项。

### 周福林

河南财经政法大学教授,统计学院副院长,2005年毕业于中国人民大学社会与人口学院,人口学博士,河南财经政法大学人口发展研究所所长。主讲课程为《统计学》《社会统计学与SPSS应用》《人口数据分析方法》《调查技术》等,研究方向为人口统计学。曾获中国人口科学优秀成果奖一等奖和二等奖,出版专著《我国留守家庭研究》等3部,主持完成国家社科基金项目《新时期我国家庭结构的变迁研究》,在研国家社科基金项目《基于第三次农业普查的我国农村留守家庭研究》。

这是一个巨变的时代,也是一个甚嚣尘上的时代,更是“比历史上任何时期都更接近中华民族伟大复兴的目标”<sup>①</sup>的时代。寓于这个时代的统计学,也必然受到深刻的影响。互联网、云计算、大数据、快数据、人工智能等色彩斑斓,统计学遇到了前所未有的挑战,但也面临着更多的机遇,嬗变预示着光辉的未来!

从教材出版的繁荣上就能看出端倪。据不完全统计,近三年以来,仅就《统计学》而言,新出版再版的教材就达50部以上。那我们为何还要添砖加瓦呢?

1992年、2000年、2003年和2005年,我们先后主持编写了四本供经济和管理类本科生使用的《统计学》教材,反映效果很好。2013年起改用了其他教材,当时的考虑主要有三点:

一是向其他教材学习。我们自己主编的教材虽然具有我们的特色和传统,但经济统计和描述性统计所占的篇幅明显,而其他教材中“通用”的统计学知识和方法普遍居于中心地位。我们的教学内容是否偏颇或落伍,需要通过教材的更换和社会的需要来检验。

二是与师资队伍的结构变化相适应。由于经济统计博士千金难求,应教学的需要我们引进了较多的数理统计教师,教材的选用也应该与这一现实符合,教材的更换能起到过渡和融合作用。

三是重新编写教材的时机不成熟。不仅上述两个因素影响着我们的决定,更主要的是一些年来,重科研、轻教学弥漫在大陆高校,教材编写无益于职称晋升,很难组织起一个精心编写教材的合理队伍。

五年过去了,在使用其他教材的过程中,我们受益匪浅,但现实迫使我们必须编写一本合乎实际要求的教材,其突出的原因是:

第一,统计学开设的必要性受到非统计院系的质疑。我国的财经类高校开设了概率论与数理统计、统计学和计量经济学三门课程,它们之间的内容严重重复,在其他两门课程饱满开设的情况下,统计学的内容应当回到应该开设的方面上去。由于没有与之适配的教材,无法有效应答兄弟院系的疑问。

第二,社会需求是硬道理。改革开放之前的统计学拘泥于经济统计,而在这之后的统计学教材中,经济统计部分又逐渐被淡化到了点缀的程度。而事实上,统计学的国民教育作用,应当是统计意识的培养为先,能正确领会和解读媒体上的统计信息,这应当是财经类非统计专业教学的首要任务,因此,应当加强统计学中的经济统计内容。

第三,是教学范式改革落地的需要。统计有用,但比较枯燥难学,这是一个普遍的事实。如何能做到有效地学、正确地用,我们曾为此做出过不少努力,并围绕其开展了一些教育和教学的方式与方法研究,形成了一些理念和做法,需要通过教材编写,将其落实到人才培养的过程当中。

第四,是优秀教学团队建设的举措之一。2013年,我们获批为河南省统计学优秀教学团队。

<sup>①</sup> 2012年11月29日习近平在参观《复兴之路》展览时的讲话。

在主导尽量多采用国内外高质量教材的同时,根据需要,尤其是从锻炼团队的角度出发,需要适度鼓励自编教材或合作编写教材,提升团队的层次和影响力。该教材就是其中的第一本教材。

第五,为“以本为本”的一流本科教育做贡献。本科教育具有基础性、完整性和系统性三大特征,统计学是这一特征的基本要素之一,“读和写的能力对社会公众是不可或缺的,总有一天,统计思维也会如此。”<sup>①</sup>随着大数据时代的到来,这正在成为现实。“以本为本”“金课”为先,教材建设是为重中之重。

本教材的特色主要表现如下:

内容现代化。本教材定位于财经类高校的经济与管理专业,较好地解决了与其他课程交叉重叠的问题,重视经济统计,并将其更新到最新的知识点上。加入了逻辑斯蒂回归、多元统计分析等更为现代的统计方法。

方法软件化。本教材将统计方法与 SPSS 软件紧密结合,引导学生的统计思维构建,注重统计方法能解决什么问题,如何解读结果,应当注意哪些事项等方面,而将实现留给了统计软件,所见即所得。

形式信息化。本教材尽量通过统计软件、二维码技术和微课程等手段将纸质教材与数字化资源实现互联,拓宽知识视野,丰富教学手段,提升学习兴趣。

学习在线化。我们建设的统计学是河南省首批精品在线开放课程,我们将根据本教材配套更新“爱课程”网上的知识点视频,这有助于教师将翻转课堂和传统课堂相结合,推动教师主导和学生主体的实现。

作业电子化。本教材后续拟配套统计实验与学习指导书,可以通过二维码获取数据、案例、作业与答案、模拟试题与答案、拓展知识等内容,在立体的环境中,更好地学习和应用统计。

教材中国化。本教材在全面介绍统计学的基本知识与方法的同时,特别强调中国经济统计的理论与实践成果,这在统计描述、时间序列分析、统计指数、统计综合评价、国民经济统计等章节中都有突出的表现。

本教材各章的执笔人分别是:王一惠(第一章)、周福林(第二章)、张志雷(第三章)、金莹(第四章)、蒋青松(第五章)、张坤(第六章第一、二节)、杨文溥(第六章第三、四节)、阚珍珍(第七章)、刘定平(第八章第一至四节)、王超(第八章第五节)、乔晗(第九章)、贺文星(第十章)。初稿出来之后,各章作者相互审阅少则五稿、多则九稿,副教授汪彩玲博士参与了部分章节的修改。主编刘定平教授负责全书的设计、逐字逐句改定和补充了书稿,并进行了总纂,副主编周福林教授负责统筹、通读全稿。

高等教育出版社的编辑吴淑丽博士认真审稿,提出了很好的修改意见和建议,其他编辑的出色工作也让本书锦上添花!本教材在编写的过程中,我们引用和参考了诸多同仁的优秀成果,在此一并表示感谢!

虽然尽心尽力工作,但由于水平所限,书中可能有不当或错误之处,恳请大家匡谬!

刘定平

2019年1月9日

<sup>①</sup> 英国科幻小说大师、社会学家和历史学家韦尔斯(H. G. Wells, 1868-1946)语“Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.”

## 第一章 总论 ..... 1

第一节 什么是统计和统计学 .....	2
第二节 统计学的产生和发展 .....	6
第三节 统计学的分科及其与其他学科的关系 .....	8
第四节 统计学的基本概念 .....	10
第五节 SPSS 使用基础 .....	18
本章小结 .....	24
专业名词 .....	25
思考题 .....	25

## 第二章 数据收集与整理 ..... 27

第一节 统计数据的来源 .....	28
第二节 统计调查方案设计 .....	29
第三节 调查数据的误差、审核和录入 .....	41
第四节 调查数据的整理 .....	45
本章小结 .....	54
专业名词 .....	54
思考题 .....	54

## 第三章 统计描述 ..... 56

第一节 总量指标 .....	57
第二节 相对指标 .....	59
第三节 数据集中趋势——平均指标 .....	62
第四节 数据离散趋势——变异指标 .....	68
第五节 数据分布形态——偏度和峰度 .....	72
第六节 统计图表 .....	74
第七节 SPSS 统计制图 .....	78
本章小结 .....	81
专业名词 .....	82
思考题 .....	83

## 第四章 统计推断 ..... 84

第一节 抽样分布 .....	85
----------------	----

第二节 参数估计 .....	90
第三节 假设检验 .....	102
本章小结 .....	109
专业名词 .....	109
思考题 .....	110

## 第五章 方差分析与回归分析 ..... 111

第一节 单因素方差分析 .....	112
第二节 双因素方差分析 .....	117
第三节 相关分析 .....	125
第四节 回归分析 .....	133
本章小结 .....	146
专业名词 .....	148
思考题 .....	148

## 第六章 多元统计分析 ..... 149

第一节 聚类分析 .....	150
第二节 主成分分析 .....	160
第三节 因子分析 .....	165
第四节 判别分析 .....	172
本章小结 .....	178
专业名词 .....	178
思考题 .....	179

## 第七章 时间序列分析 ..... 180

第一节 时间序列概述 .....	181
第二节 时间序列的水平分析 .....	184
第三节 时间序列的速度分析 .....	190
第四节 长期趋势的测定 .....	193
第五节 季节变动的分析 .....	201
第六节 时间序列的预测 .....	204
本章小结 .....	216
专业名词 .....	217
思考题 .....	217

<b>第八章 统计指数</b> .....	219	思考题 .....	270
第一节 统计指数的概念和种类 .....	220	<b>第十章 国民经济统计</b> .....	271
第二节 总指数的综合形式 .....	222	第一节 国民经济及其分类 .....	272
第三节 总指数的平均形式 .....	229	第二节 国民经济生产条件统计 .....	275
第四节 指数体系和因素分析 .....	233	第三节 国内生产总值 .....	279
第五节 国内外几种重要的经济指数 .....	240	第四节 国际收支 .....	284
本章小结 .....	251	本章小结 .....	286
专业名词 .....	251	专业名词 .....	287
思考题 .....	251	思考题 .....	287
<b>第九章 统计综合评价</b> .....	253	<b>附录</b> .....	288
第一节 统计综合评价概述 .....	254	附表 1 随机数字表(部分) .....	288
第二节 评价指标体系的建立 .....	257	附表 2 标准正态分布表 .....	289
第三节 评价数据的预处理 .....	258	附表 3 $\chi^2$ 分布表 .....	290
第四节 评价指标权重分配方法 .....	262	附表 4 $t$ 分布表 .....	291
第五节 指标综合模型 .....	268	附表 5 $F$ 分布表 .....	292
本章小结 .....	269	<b>参考文献</b> .....	294
专业名词 .....	270		

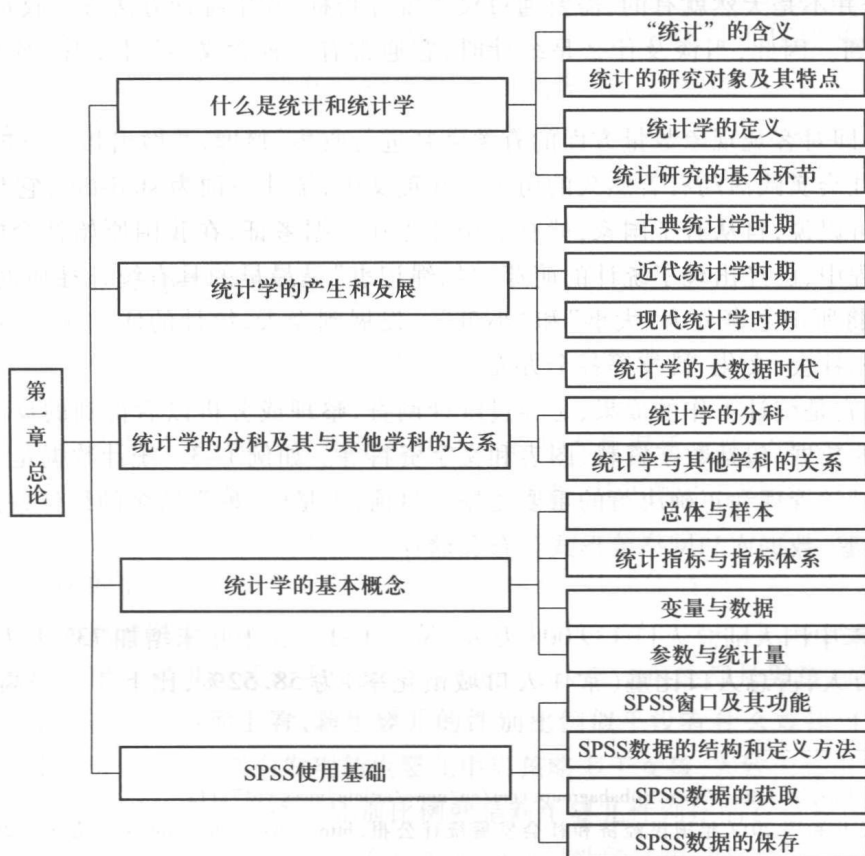


## 章前导言

生活中人们每天都会遇到与统计有关的问题,大众媒体每天也都在不断地出现统计数字。统计已经渗透到经济、社会、自然、科技、生活等的每一个角落,并且正在发挥着越来越重要的作用。那么,究竟什么是统计呢?它是如何产生和发展起来的?作为一门科学的统计学,它与其他学科又有何区别与联系?这些正是本章所要介绍的主要内容。

本章的学习要求是:①理解统计的含义及其研究对象;②知晓统计研究的基本环节;③了解统计学的发展简史,特别是一些有代表性的学派;④认识统计学的学科性质及其与其他学科的关系;⑤掌握统计学的基本概念,如:总体、样本、标志、指标、指标体系、变量、数据、参数及统计量等;⑥熟知统计软件 SPSS 的基础操作。

## 本章知识结构图



## 第一节 什么是统计和统计学

### 一、“统计”的含义

在当今信息时代,人们经常可以从各种媒体上听到或看到各种各样的数据。

#### 例 1-1:

2017年11月11日,阿里巴巴集团控股有限公司旗下中国及国际零售平台通过支付宝结算的商品成交额(以下简称GMV)达到人民币1682亿元(约253亿美元),较2016年增加39%;其中,通过支付宝结算的移动端GMV占比,由2016年的82%上升至2017年的90%。全日支付总数达到14.8亿笔,全日物流订单达8.12亿笔,交易覆盖全球225个国家和地区。<sup>①</sup>

#### 例 1-2:

2017年中国国内生产总值(GDP)达827122亿元,比上年增长6.9%。其中,第一产业增加值65468亿元,增长3.9%;第二产业增加值334623亿元,增长6.1%;第三产业增加值427032亿元,增长8.0%。第一产业增加值占国内生产总值的比重为7.9%,第二产业增加值比重为40.5%,第三产业增加值比重为51.6%。全年人均国内生产总值59660元,比上年增长6.3%。全年国民总收入825016亿元,比上年增长7.0%。<sup>②</sup>

这些数据是如何得到的?为什么要得到这些数据?它们又意味着什么?这就是统计问题。然而,这些数据并不是天然就有的,需要通过设置统计指标,利用统计方法进行收集、整理或适当的分析才能得到。因此,当谈及什么是统计时,它通常有三种含义:统计工作、统计数据和统计理论。

统计工作,即对客观现象数量方面的有关资料进行收集、整理,并做出相应分析的活动过程。统计作为一种社会实践活动有着悠久的历史。在英文中,统计一词为statistics,它与“国家”有着相同的词根。可以说,自从有了国家,就有了统计工作。据考证,在我国原始社会的末期及奴隶社会形成的过程中,已经出现了统计的萌芽,“结绳记事”是最早的具有统计性质的调查活动,它以“结”的大小将所记之事分为“大事”和“小事”。发展到今天,统计的活动领域越来越广,它已涉及政治、经济、社会、文化、科技等各个方面。

统计数据,它是统计工作的成果,是经过统计调查、整理或分析以后得到的反映客观现象特征的各种形式的数字资料,包括数字资料、图表和文字资料等。如例1-3。统计数据是一个单位乃至一个国家实施科学管理和正确决策的重要依据。目前,大量的、种类繁多的统计数据多以各种统计公报、统计年鉴、数据库和网络等形式公布和储存。

#### 例 1-3<sup>③</sup>:

2017年年末中国大陆总人口139008万人(见表1-1),比上年末增加737万人,其中城镇常住人口81347万人,占总人口比重(常住人口城镇化率)为58.52%,比上年末提高1.17个百分点。

<sup>①</sup> 阿里巴巴集团网站, <http://www.alibabagroup.com/cn/news/article?news=p171112>.

<sup>②③</sup> 中华人民共和国2017年国民经济和社会发展统计公报, [http://www.stats.gov.cn/tjsj/zxfb/201802/t20180228\\_1585631.html](http://www.stats.gov.cn/tjsj/zxfb/201802/t20180228_1585631.html).

点。户籍人口城镇化率为42.35%，比上年末提高1.15个百分点。全年出生人口1723万人，出生率为12.43‰；死亡人口986万人，死亡率为7.11‰；自然增长率为5.32‰。

表 1-1 2017 年年末中国大陆人口数及其构成

指标	年末数/万人	比重/%
全国总人口	139 008	100.00
其中:城镇	81 347	58.52
乡村	57 661	41.48
其中:男性	71 137	51.20
女性	67 871	48.80
其中:0-15岁	24 719	17.80
16-59岁	90 199	64.90
60周岁及以上	24 090	17.30
其中:65周岁及以上	15 831	11.40

统计理论,是对统计工作进行总结和理论概括的系统学说,它阐述统计工作的基本理论和基本方法,统计理论的系统化和科学化结晶成为统计学。

统计工作、统计数据和统计理论三者之间存在着密切的联系。统计工作与统计数据是工作过程与工作成果的关系。工作过程的好坏直接影响到工作成果质量的高低。统计数据来源于统计工作过程,同时,一定积累的统计数据,又是下一步开展统计工作的基础。统计理论与统计工作是理论与实践的关系,随着统计工作的不断深入,一方面可以获得更加丰富的统计数据资料,另一方面又不断地丰富统计理论,促使统计理论更加完善。反过来统计理论指导统计工作,使统计工作更加科学、高效。因此,统计一词的三种含义是相互联系的,不能将它们分割开来。

## 二、统计的研究对象及其特点

统计的研究对象是客观现象总体的数量方面,包括:数量表现、数量关系和数量变动规律等。

### 例 1-4:

1955年10月,联合国在其出版的《用于总体估计的基本数据质量鉴定方法》(手册Ⅱ)中认为:出生性别比偏向于男性。一般来说,每出生100名女婴,其男婴出生数位于102~107之间。此分析明确认定了出生性别比的通常值域为102~107之间。从此之后,出生性别比下限不低于102,上限不超过107的值域一直被国际社会公认为正常值,其他值域则被视为异常。

就单独的一个家庭来观察,每个家庭新生婴儿的性别有一定的随机性,可能是男孩,也可能是女孩。如果不对生育人口进行任何限制,有些家庭的几个孩子可能都是男孩,而有些家庭的几个孩子则可能都是女孩。从表面上看,新生婴儿的性别比例似乎没有什么规律可循。但如果对大量的家庭新生婴儿进行观察,就会发现新生婴儿中男孩略多于女孩,大致为每出生100个女孩,相应地就有102~107个男孩出生。这一性别比例就是新生婴儿性别比的数量规律。通过对大量看起来随机的现象进行统计观察和分析,就有可能得到规律性的认识,即随机之中有规律性,而这

种规律就是统计规律。因此,统计通常被看作是在随机性中寻找规律性,在偶然性中发现必然性。

统计的研究对象具有以下特点:

### 1. 总体性

所谓总体性,是指统计要通过对现象总体中的构成元素,即个体,进行大量观察和综合分析,从整体上去认识和把握客观现象的数量特征与规律。也就是说,统计研究的目的,不是认识个体现象的数量特征,而是具有综合意义的总体现象的数量特征。例如,在对新生儿性别的研究中,目的并不在于了解某一家庭新生儿的性别,而是通过对大量家庭新生儿性别的调查,来反映某一国家某一时期的出生性别比。某一家庭的新生儿性别具有随机性和特殊性,而一个国家某一时期的新生儿性别特征则具有规律性或一般性。

### 2. 数量性

客观现象都存在质和量两个方面,是质和量的统一。统计是从量的方面对客观现象进行研究的,其本质属于定量认识。当然,对客观现象数量方面的研究并不是孤立进行的,而是建立在对客观现象定性认识的基础之上的,统计虽然属于定量认识,但开展统计工作和统计研究的前提条件,是人们必须对研究对象具有一定程度的定性认识。统计的目的在于通过对现象量的方面的观察和量变规律的研究,逐步把握现象的质。统计对客观现象总体数量方面的研究包括数量规模、数量变化规律、现象间的数量关系以及质与量互变的数量界限等。

### 3. 变异性

统计研究客观现象总体的数量特征,其前提是构成总体的各个体的特征表现存在着差异,而且这些差异并不是事先可以预知的。例如,人口统计中,每个自然人的性别、出生时间、出生地、寿命等都有差别,因此才需要研究人口总体的性别比例、出生率、地区分布、期望寿命等特征。如果各自然人不存在这些差异或者这些差异是可以按已知条件完全推知的,也就不需要进行统计了。统计研究就是要从每个个体之间的差异中概括出共同的、普遍的数量特征,并对差异情况做出必要的反映。例如,一个地区的居民间一般都存在着收入差异,所以,统计就用平均收入来反映该地区居民的收入水平,用基尼系数等指标来反映居民收入的差异程度。

## 三、统计学的定义

统计学是一门收集、整理、分析数据,并从中得出数量结论的方法论科学。其目的是探索数据的内在数量规律性,以达到对客观现象的科学认识。

如前所述,统计学来源于统计工作,是对统计工作的理论概括和高度总结,反过来又用于指导统计工作。统计学是一门应用性很强的学科,如今,它几乎与所有的学科领域都有着或多或少的联系,可以说,凡是有数据的地方,就有统计学的用武之地。

一些有代表性的  
统计学定义  
[可通过扫描  
封底的二维码  
获取]

## 四、统计研究的基本环节

统计研究的基本环节包括以下四个阶段:统计设计、数据收集、数据整理、数据分析与解释。

### 1. 统计设计

统计设计指的是根据统计研究问题的性质、研究的目的和任务,在相关学科的理论指导下,

科学地确定统计研究的对象,制定相应的统计指标及指标体系,给出统一的定义、标准,提出数据收集、整理、分析和发布的方案,规定研究进度安排和质量要求,拟定研究工作的资源配置和组织实施方式等。由于统计是一项需要高度集中统一的工作,没有预先的科学的设计,就很难达到预期的目的,因此,在一项大规模的统计活动开始前,必须进行统计设计,而且要求设计者不仅要有扎实的统计理论基础,还要具有所研究领域的相关知识。

## 2. 数据收集

数据收集是在明确统计研究的对象和目的的基础上,根据统计设计的内容、指标和指标体系的要求,有计划、有目的、有组织地搜集资料的工作过程。数据收集是否真实、准确、及时、完整,将直接影响到统计分析的质量。数据收集,可以是通过自己的调查或实验活动,直接获得原始数据,即一手数据;也可以是找到并使用别人通过调查或实验的方式搜集到的数据,即二手数据。对于大多数自然科学和工程技术研究来说,有可能通过有控制的科学实验去取得数据;而对于社会经济现象来说,一般无法进行重复实验,要取得有关数据就必须到社会经济总体中去选取足够的个体进行调查以取得数据。

## 3. 数据整理

数据整理是统计研究的第三个环节,指根据统计研究的目的,将通过统计调查或实验所得到的一手数据,以及通过各种途径获得的二手数据,进行科学的分类、汇总和计算,以在一定程度上说明总体的数量特征。数据整理的结果一般要通过统计指标、统计表、统计图等来表达。

## 4. 数据分析与解释

数据分析与解释是统计研究过程的最后一个环节,是在数据整理的基础上,根据研究的目的和任务,运用科学的统计方法,对数据进行计算、推断与分析,以揭示客观现象总体的状况、特点、规律性和存在的问题等。在这个环节,既可能会用到统计描述的方法,也可能会用到统计推断的方法。对于数据分析的结果,还应与现实问题紧密结合起来,进行数据的阐释,即说明为什么会得出这些数据结果,这些结果的含义是什么,从中能得出哪些具有规律性的结论,需要进一步探讨的问题又有哪些,等等。

统计研究的基本环节可以通过图 1-1 来表示。

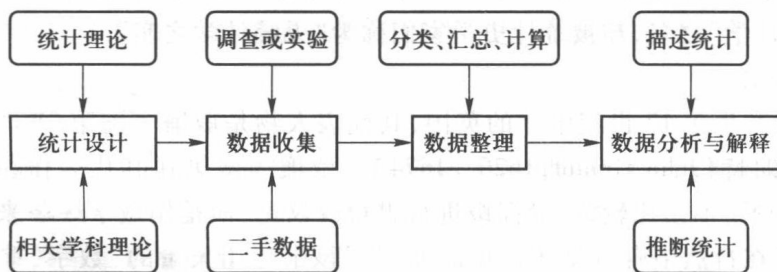


图 1-1 统计研究的基本环节

值得注意的是,统计研究的基本环节虽然可以划分为统计设计、数据收集、数据整理、数据分析与解释等四个部分,但这在很大程度上只是理论上的划分,是相对而言的,在实际统计研究中,这些环节往往会交织在一起,很难明确分开,更不必要硬性分开。

## 第二节 统计学的产生和发展

人类的统计活动可以追溯到相当遥远的古代。在原始社会,人类最初的一般计数活动蕴藏着统计的萌芽。到奴隶社会,统治阶级为了对内统治和对外战争,开始了人口、土地和财产等的统计。例如:在中国,据记载,公元前 21 世纪的夏朝分为九州<sup>①</sup>,人口 1 355 万人,土地 2 438 万顷;而同一时期的古希腊、古罗马等国家,也有关于国情国力的记载。经过漫长的历史长河,到封建社会末期和资本主义社会初期,随着社会生产力的迅速发展和社会分工的愈益发达,统计工作除服务于政府管理需要外,还逐步扩大到社会、经济、科技、环境等各个领域。此时,统计工作客观上需要从理论上对其进行概括和总结,于是统计学应运而生。作为距今有三百多年历史的统计学,按统计方法特征的演变顺序,一般可将其发展划分为四个阶段:古典统计学时期、近代统计学时期、现代统计学时期和统计学的大数据时代。

### 一、古典统计学时期

17 世纪中叶,欧洲各国相继进入资本主义工场手工业的经济迅速发展阶段,为了适应各国经济发展的不同需要,在欧洲各国不约而同地从不同领域开始了统计学的奠基工作,因此,从 17 世纪中叶到 18 世纪中叶这一时期被认为是统计学的初创时期,即古典统计学时期,当时主要有两大学派,即国势学派和政治算术学派。

#### 1. 国势学派

国势学派的创始人是德国人康令(H. Coning, 1606—1681),他第一个在德国赫尔莫斯达德大学讲授《欧洲最近国势学》课程,主要用对比的方法讲授国家比较方面的知识,于是“国势学”由此产生。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,阿亨瓦尔(G. Achenwall, 1719—1772)发扬了康令的学术思想,并最早使用“统计学(德文 Statistik)”这一术语。

国势学派认为统计学是关于国家显著事项的学问,主要通过对国家组织、人口、军队、领土、居民职业和资源财产等事项的记述对国情国力进行研究。但由于国势学派主要使用文字记述的方法进行研究,在进行国势比较分析中偏重事物性质的解释而不注重数量分析,所以尽管这一学派首先提出了“统计学”之名,却被统计史学家们称为“无统计学之实”。

#### 2. 政治算术学派

政治算术学派产生于 17 世纪中叶的英国,其代表人物是威廉·配第(William Petty, 1623—1687)和约翰·格朗特(John Graunt, 1620—1674)。威廉·佩第在其代表作《政治算术》(1676 年)一书中写道:本书“不用比较级、最高级进行思辨或议论,而是用数字……来表达自己的问题……借以考察在自然中有可见的根据的原因。”该书运用大量的“数字、重量、尺度”等定量分析工具对英国、法国和荷兰三国的财富和力量进行了比较分析,以批驳当时英国国内的悲观论调。约翰·格朗特在其代表作《对死亡表的自然观察和政治观察》(1662 年)中,通过大量观察

<sup>①</sup> 据《尚书·禹贡》记载,大禹治水后按照山川河流的走向,把全国划分为九州:冀州、兖州、青州、徐州、扬州、荆州、豫州、梁州、雍州。

的方法,发现了新生儿的性别比例和不同原因的死亡人数比例等一系列人口数量规律,并运用各种方法对统计数据进行间接的推算和相互印证。

政治算术学派关于运用大量观察、分类分析和对比分析等方法对社会经济现象进行研究的主张,为统计学的发展开辟了广阔的前景。然而,这一派学者一直没有使用“统计学”这一术语,而是用“政治算术”来表明其研究的特色,因而被称为“政治算术”学派,虽无“统计学”之名而实为统计学之正统起源。

## 二、近代统计学时期

从18世纪末到19世纪末,统计学的理论体系得以不断发展和完善,逐渐形成了以随机现象的推断统计为主要内容的数理统计学派和以社会经济现象描述为主要内容的社会统计学派。

### 1. 数理统计学派

数理统计学派产生于19世纪中叶,代表人物是比利时著名的统计学家阿道夫·凯特勒(L. A. J. Quetelet, 1796—1874)。凯特勒将概率论正式引入到统计思维中来,融会贯通德国的国势学、英国的政治算术等各家各派的统计思想,博采众长,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。他主张用研究自然科学的方法来研究社会现象,正式把概率论引入统计学,并最先用大数定律论证了社会生活中随机现象的规律性,还提出了误差理论和“平均人”思想。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前启后、继往开来的作用。

### 2. 社会统计学派

自凯特勒之后,统计学的发展开始变得丰富而复杂起来。19世纪后半叶,致力于自然领域研究的英美数理统计学派开始发展的时候,在德国兴起了与之不同的社会统计学派,其首创人物为克尼斯(K. G. A. Knies, 1821—1898)。他认为统计学是一门社会科学,是研究社会现象变动原因和规律性的实质性科学,其显著特点是强调对总体进行大量观察和分析,通过研究其内在联系来揭示社会现象的规律。各国专家学者在社会经济统计指标的设定与计算、指数的编制、资料的收集与整理、统计调查的组织和实施、经济社会发展评价与预测等方面取得了一系列重要成果,如:德国统计学家恩格尔(C. L. E. Engel, 1821—1896)提出的恩格尔系数,美国经济学家库兹涅茨(Simon Kuznets, 1901—1985)和英国经济学家斯通(Richard Stone, 1913—1991)等开创的国民收入和国内生产总值核算方法等,都是这一学派19世纪至20世纪的代表性学术成果。

## 三、现代统计学时期

现代统计学时期一般认为从19世纪末开始,成熟于20世纪,这一时期的显著特点是数理统计学同自然科学、工程技术科学紧密结合,且被广泛应用于各个领域,从而得到迅速发展,各种新的统计理论与方法(特别是推断统计)大量涌现。例如:英国生物学家、统计学家高尔顿(Francis Galton, 1822—1911)首次提出并阐述了“回归”的概念;英国统计学家卡尔·皮尔逊(K. Pearson, 1857—1936)导出了一般化的次数曲线体系,提出了 $\chi^2$ 检验和标准差的概念,发展了相关理论等;英国数学家戈塞特(W. S. Gosset, 1876—1937)导出了 $t$ 分布,开创了小样本理论的先河;英国统计学家费希尔(R. A. Fisher, 1890—1962)论证了方差分析的原理和方法,并将其应用于试验设计,还阐明了极大似然估计方法等;波兰统计学家内曼(J. Neyman, 1894—1981)和英国统计学家

小皮尔逊(E. S. Pearson, 1895—1980)共同完善了现代统计学的核心内容,即区间估计和假设检验的理论;美籍罗马尼亚统计学家瓦尔德(A. Wald, 1902—1950)发展了统计决策理论,建立了序贯分析,等等。20世纪五六十年代以来,稳健统计、时间序列、抽样理论、统计诊断、探索性分析、贝叶斯统计等,都取得了重要的进展。

#### 四、统计学的大数据时代

20世纪以来,随着计算机技术及网络的发展,社会各行各业逐步走上了信息化的道路并积累了海量的数据。数据的爆炸式增长正引领一场新的时代变革,大数据时代已经来临,它正迅速地改变着人们的行为与思维。大数据不再是基于人工设计、借助传统方法而获得的有限、固定、不连续、不可扩充的结构型数据,而是基于现代信息技术与工具,可以自动记录、存储和连续扩充的、大大超出传统统计记录与存储能力的一切类型的数据。<sup>①</sup>大数据具有规模大、种类多、速度快、价值密度低等特点,在数据增长、分布和处理等方面具有更多复杂的性质。

在大数据时代,统计学的研究对象并没有变,但数据的来源、体量、类型、速度与量化的方式改变了。大数据时代的统计思维具有两个崭新的特点:一是,随着数据获取、存储与计算能力的提升,可以很容易获得“总体”中的全部数据,且可以在总体上直接进行计算,不再需要进行“抽样操作”。二是,在海量、动态、异构数据环境中,人们更加关注的是数据计算的“效率”而不再盲目追求其“准确度”,重视对“相关性”的分析而不是等到发现“真正的因果关系”之后才解决问题。大数据时代给统计学的发展带来了挑战和机会,将统计、计算机与应用对象相互结合,是这个时代对人们认识事物和发现规律提出的新要求。

### 第三节 统计学的分科及其与其他学科的关系

#### 一、统计学的分科

随着统计方法和理论不断地丰富和完善,以及统计学被广泛地应用到自然科学和社会科学的各个领域,统计学已经发展成为由较多分支科学组成的学科体系。现代统计学,从研究目的分工角度看,可分为理论统计学和应用统计学;从统计方法的特点角度看,可分为描述统计学和推断统计学。

##### 1. 理论统计学和应用统计学

理论统计学是把研究对象一般化、抽象化,以概率论为基础,从纯理论的角度,对统计方法加以推导论证,其中心内容是统计推断问题,实质是以归纳方法研究随机变量的一般规律。理论统计学包含的内容很多,例如统计分布理论、统计估计和假设检验理论、相关与回归分析、方差分析、时间序列分析等。理论统计学是统计方法的理论基础,没有理论统计学的发展,统计学也不可能发展成为像今天这样一个完善的科学知识体系。

应用统计学研究的是如何应用统计理论和方法去解决实际问题。它从所研究的领域或专门

<sup>①</sup> 李金昌. 大数据与统计新思维[J]. 统计研究, 2014(01).



问题出发,根据研究对象的性质和特点采用适当的指标体系和统计方法以解决所需研究的问题,因此,应用统计学需要有相关专业实质性科学的理论做指导,通常具有边缘交叉和复合型学科的性质。目前,统计理论和方法的应用几乎扩展到了所有的自然科学和社会科学的研究领域,其在经济领域中的应用形成了经济统计学,在人口学中的应用形成了人口统计学,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,等等。

理论统计学和应用统计学的发展相辅相成。理论统计学的研究为应用统计学解决实际问题提供了方法论基础,理论统计学的不断发展和完善大大提高了统计分析解决问题的能力;而应用统计学在对统计理论和方法的实际应用中,又经常会对理论统计学提出新的问题,开拓理论统计学的研究领域,促进理论统计学的发展。

## 2. 描述统计学和推断统计学

描述统计学研究如何取得反映客观现象特征的数据,对所收集到的数据进行加工整理并以图表形式展示,进而通过综合、概括与分析等方法得出客观现象的数量规律性。描述统计学的内容包括:统计数据的收集方法、数据的加工整理方法、数据的展示方法、数据分布特征的概括与分析方法等。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,从而得出有关这个总体或样本的结论。例如:调查全班 60 名同学或从全校一年级 5 000 名同学中抽取 60 人进行观测,收集某学期英语考试的成绩,并以图表的形式展示出来,同时计算平均分数、分数的标准差、及格率等指标,从而得出关于该班学生或抽取的 60 名同学英语学习状况的结论。

推断统计学研究如何根据样本数据去推断总体数量特征,它是在对样本数据进行描述的基础上,根据样本与总体之间存在的客观联系,对总体的未知数量特征做出具有一定可靠性保证的推断。例如,从全校一年级 5 000 名同学中抽取 60 人,以这 60 名同学某学期英语考试成绩的平均分数以及分数的标准差,来推断全校一年级 5 000 名同学该次英语考试的平均成绩和分数的标准差。

描述统计学和推断统计学的划分,一方面反映了统计方法发展的前后两个阶段,另一方面也反映了应用统计方法探索客观现象数量规律性的不同过程。由于在对现实问题的研究中,所获得的数据主要是样本数据,因此,推断统计在现代统计学中的地位和作用十分重要,已成为统计学的主体内容。当然,这并不是说描述统计不重要,如果没有描述统计收集可靠的统计数据并提供有效的样本信息,即使再科学的统计推断方法也难以得出切实合理的结论。从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学发展成熟的重要标志。

## 二、统计学与其他学科的关系

### 1. 统计学与数学的关系

统计学与数学关系密切,二者都是研究数量规律的,都要利用各种公式进行运算。现代统计学中运用了大量的数学理论与方法,例如,数学中的概率论研究随机现象的数量关系和变化规律,为现代统计学提供了数量分析的理论基础。但统计学与数学之间也存在着本质的区别。从研究对象上来看,统计学研究的是具体的、客观事物的数量规律,而数学研究的是抽象的数量规律;统计学研究的是有具体实物或计量单位的数据,而数学研究的则是没有量纲或单位的抽象的数。从思维方式上看,统计学与数学也是不同的,统计学是将演绎与归纳相结合,以归纳为主要的思维方式;而数学所采用的一般是纯粹的演绎思维。由此可见,统计学的方法都是在应用的推