

福建省“十四五”普通高等教育本科省级规划教材

统计学

主编◎卿 松 王立凤 陶海映



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书作者力求在吸收近些年统计学教材优秀成果,结合本教材编写组 20 多年来长期从事的统计学教学与科研最新成果的基础上,形成融全面性与创新性于一体的统计学教材。本书集统计学基本理论、基本方法、软件应用、学习目的、关键概念、小案例与复习思考题于一体,对各种经济类与管理类专业和不同层次的教学具有广泛的适用性。

图书在版编目(CIP)数据

统计学 / 卿松, 王立凤, 陶海映主编. — 上海:
上海交通大学出版社, 2021.7 (2023.4 重印)

ISBN 978-7-313-25079-7

I. ①统… II. ①卿… ②王… ③陶… III. ①统计学
IV. ①C8

中国版本图书馆 CIP 数据核字 (2021) 第 124035 号

统计学

TONGJI XUE

主 编: 卿 松 王立凤 陶海映

地 址: 上海市番禺路 951 号

出版发行: 上海交通大学出版社

电 话: 6407 1208

邮政编码: 200030

印 制: 北京荣玉印刷有限公司

经 销: 全国新华书店

开 本: 787mm × 1092mm 1/16

印 张: 17

字 数: 385 千字

版 次: 2021 年 7 月第 1 版

印 次: 2023 年 4 月第 2 次印刷

书 号: ISBN 978-7-313-25079-7

定 价: 54.00 元

版权所有 侵权必究

告读者: 如发现本书有印装质量问题请与印刷厂质量科联系

联系电话: 010-6020 6144

P 前言 Preface

进入快速发展的智能经济时代，经济与管理全球一体化，使得社会越来越需要一批融经济与管理于一体的复合型人才。要培养适应现代社会的复合型人才，必须以坚实的理论与方法为基础。统计学正是这样一门融理论与方法于一体，为所有的经济类与管理类专业提供理论与方法的专业基础学科。它可以为经济类与管理类专业知识的学习奠定坚实的基础。为了适应我国经济改革与发展的需要，迫切需要编写一部既为经济类与管理类专业提供理论与方法，又能融实用、简单和创新于一体的统计学教材。近年来也出版了各种统计学或者统计学原理教材，有的教材偏重理论与烦琐的方法，有些教材则偏重实践统计的内容。对于烦琐的统计方法，如果在操作上无法解决，将会极大地限制统计的作用。如果能把烦琐的统计方法用一些工具加以解决，可以大大提高统计的分析功能，充分发挥统计作为一种方法论工具的作用。虽然有些教材也做过这方面的努力，但都不够全面。本书作者力求在吸收近些年统计学教材优秀成果，结合本教材编写组 20 多年来长期从事的统计学教学与科研最新成果的基础上，形成融全面性与创新性于一体的统计学教材。本书集统计学基本理论、基本方法、软件应用、学习目标、关键概念、小案例与复习思考题于一体，对各种经济类与管理类专业和不同层次的教学具有广泛的适用性。

全书共分 9 章。第 1 章为总论，旨在从总体上阐述统计的产生与发展、统计学的性质与特点，以及统计中存在哪些重要的概念和统计学与其他学科之间的关系，为后面的学习内容奠定基础。第 2 章为统计调查与数据整理，主要介绍统计工作中的前两个环节，其中涉及大量的统计调查与数据整理的方法。第 3 章为综合指标分析，充分应用总量指标、相对指标和平均指标来分析现象。每一项指标又会涉及一些具体的方法，它们分别从不同的角度去分析问题。第 4 章为抽样推断，通过抽样推断将统计功能提高到一个新的高度，在大量的科研中，这种分析方法得到了广泛的应用。第 5 章为假设检验，与第 4 章内容紧密相连，这一章主要介绍各种假设检验的方法，既涉及对参数的检验，也包括对非参数的检验。第 6 章为相关分析与回归分析，这两种分析是一种互补关系，相关分析主要分析现象间的相关方向与程度，而回归分析是对现象的相关状态进行分析。第 6 章与第 5 章的关系同样密切，只有通过

假设检验后的回归分析,才更具有实际意义,运用它可以推断或预测经济现象。第7章为时间序列分析,由静态分析转向动态分析,可以分析现象发展变化的趋势。第8章为因素分析,延续了动态分析方法,通过动态的比较,将影响经济现象的因素分离出来,分析其影响的方向和程度,并通过比较,反映出每一因素在经济影响中的地位,为决策提供科学依据。这一方法在很多学科中得到了广泛应用。第9章为统计实践,包括 Excel、EViews 和 SPSS 在统计学中的应用。为了确保将极其烦琐复杂的统计计算简单化,该章详细介绍 Excel、EViews 和 SPSS 统计软件将统计方法在软件中的操作具体化。全书最大的特点就是通过统计软件,将复杂的统计方法计算简单化,既改变同学们认为统计学难学的观念,又充分发挥统计学强大的分析方法的功能,在今后的专业课的学习中,更加体现统计学作为一门工具学科的价值,也为高层次的研究生学习提供研究分析的工具。

全书各章节的内容紧密衔接,逻辑严谨,结构清晰,大大增强了统计学的科学性、严谨性和实用性。本书适用于经济、管理类所有专业的高等院校学生,有助于培养综合型经济与管理人才。

全书由集美大学财经学院卿松教授进行总体设计,并与集美大学工商管理学院王立凤副教授、集美大学财经学院陶海映老师一起进行修改编纂,具体分工如下:卿松老师编写第4、5、6章和第9章的部分内容,王立凤老师编写第1、7、8章和第9章的部分内容,陶海映老师编写第2、3章、附表和第9章的部分内容。全书由卿松老师和杨凯淇老师进行审核、修改和定稿。

在本书出版之际,感谢为本书编写提供资料与宝贵意见的所有同事和朋友。由于编者水平有限,加之时间紧迫,在编写过程中难免出现错漏等不足之处,敬请广大读者批评指正,并提出宝贵的意见和建议,我们将虚心接受,并在后续的修订版中修改和完善。

Contents

第1章

总论 1

1.1 统计的产生与发展 2

1.1.1 统计实践活动的产生与发展 2

1.1.2 统计学的产生与发展 3

1.2 统计学的性质和特点 5

1.2.1 统计的理解 5

1.2.2 统计学的性质和研究对象 7

1.2.3 统计研究的基本方法 9

1.2.4 统计工作的过程 10

1.3 统计学的基本概念 10

1.3.1 总体与总体单位 10

1.3.2 标志与指标 12

1.3.3 差异与变量 13

1.3.4 指标体系 14

1.4 统计学与其他学科的关系 14

1.4.1 统计学与数学的关系 14

1.4.2 统计学与会计学的关系 15

1.4.3 统计学与计量经济学的关系 15

1.4.4 统计学与计算机科学的关系 15

1.4.5 统计学与其他学科的关系 15

第2章

统计调查与数据整理 19

2.1 统计调查 20

2.1.1 统计调查的意义及其分类 20

2.1.2 搜集数据的基本方法 23

2.1.3 统计调查方案设计 24

2.2 数据整理 27

2.2.1 统计整理的意义和内容 27

2.2.2 统计整理的程序和方法 27

2.2.3 统计分组 28

2.3 频数分布 31

2.3.1 频数分布的概念 31

2.3.2 变量数列的编制方法 32

2.3.3 累计频数和累计频率 34

2.3.4 数列的分布类型 35

2.4 数据显示 37

2.4.1 箱线图 37

2.4.2 统计表 38

2.4.3 直方图 40

2.4.4 折线图 41

2.4.5 曲线图 41

2.4.6 饼图 42

2.4.7 环形图 43

2.4.8 线图 44

2.4.9 雷达图 44

第3章

综合指标分析 49

3.1 总量指标 50

3.1.1 总量指标的概念和意义 50

3.1.2 总量指标的种类 51

3.1.3 总量指标的计算和运用	53
3.2 相对指标	54
3.2.1 相对指标的意义	54
3.2.2 相对指标的种类	55
3.2.3 相对指标的表现形式	55
3.2.4 相对指标的计算方法	56
3.2.5 相对指标的运用原则	63
3.3 平均指标	64
3.3.1 平均指标概述	64
3.3.2 位置平均数的计算	65
3.3.3 数值平均数的计算	70
3.4 变异指标	77
3.4.1 变异指标概述	77
3.4.2 标志变异指标的计算	78
3.4.3 分布变异指标	84
3.5 是非标志	86
3.5.1 是非标志的含义	86
3.5.2 是非标志的计算	86

第4章 抽样推断 93

4.1 抽样推断的基本理论	94
4.1.1 随机事件与概率	94
4.1.2 随机变量及其概率分布	95
4.1.3 大数定律和中心极限定律	96
4.2 抽样推断的含义与作用	97
4.2.1 抽样推断的含义及其特点	97
4.2.2 抽样推断的作用	98
4.2.3 抽样推断中的基本概念	99
4.2.4 抽样方法	100
4.3 抽样估计	100
4.3.1 抽样估计的含义与特点	100
4.3.2 抽样估计的优良标准	100
4.3.3 抽样误差	103
4.3.4 抽样估计方法	107

4.4 样本容量的确定	108
4.4.1 估计总体平均数时, 样本容量的确定	108
4.4.2 估计总体成数时, 样本容量的确定	109
4.4.3 确定样本容量时应注意的问题	110
4.5 其他抽样组织方式下的抽样估计	110
4.5.1 等距抽样	111
4.5.2 类型抽样	111
4.5.3 整群抽样	113

第5章 假设检验 117

5.1 假设检验概述	118
5.1.1 假设检验的基本概念	118
5.1.2 双侧检验与单侧检验	119
5.1.3 Z 检验与 t 检验	120
5.1.4 假设检验的两种错误	121
5.2 总体参数检验	122
5.2.1 总体平均数的检验	122
5.2.2 总体成数的检验	123
5.2.3 ρ 值检验	124
5.3 非参数检验	125
5.3.1 非参数检验概述	125
5.3.2 符号检验	126
5.3.3 秩和检验	127

第6章 相关分析与回归分析 131

6.1 相关分析	132
6.1.1 相关关系与函数关系	132
6.1.2 相关关系的种类	132

6.1.3 相关分析	133
6.1.4 相关系数	135
6.2 线性回归分析	140
6.2.1 回归分析	140
6.2.2 一元线性回归分析	141
6.2.3 多元线性回归分析	148
6.3 非线性回归分析	151
6.3.1 非线性回归模型	152
6.3.2 非线性回归模型的估计	153

第 7 章

时间序列分析

7.1 时间序列概述	158
7.1.1 时间序列的概念	158
7.1.2 时间序列的种类	158
7.1.3 时间序列的编制原则	159
7.2 时间序列水平指标	160
7.2.1 发展水平	160
7.2.2 平均发展水平	161
7.2.3 增长量与平均增长量	164
7.3 时间序列速度指标	166
7.3.1 发展速度与增长速度	166
7.3.2 平均发展速度与平均增长 速度	169
7.3.3 时间序列水平分析与速度 分析的结合应用	171
7.4 趋势分析	171
7.4.1 时间序列的构成因素和分析 模型	171
7.4.2 长期趋势的测定	173
7.4.3 季节变动趋势的测定	181
7.4.4 循环变动趋势与不规则变动 的测定	182

第 8 章

因素分析

8.1 指数概述	186
8.1.1 统计指数的概念	186
8.1.2 统计指数的分类	186
8.1.3 统计指数的作用	187
8.2 综合指数的编制	188
8.2.1 综合指数的编制与应用	188
8.2.2 平均指数的编制与应用	192
8.2.3 指数数列	196
8.3 指数体系与因素分析	197
8.3.1 指数体系与连锁替代法	197
8.3.2 两因素分析	198
8.3.3 多因素分析	202

第 9 章

统计实践

9.1 Excel 在统计学中的应用	208
9.1.1 Excel 在数据整理中的应用	208
9.1.2 Excel 在综合指标分析中的 应用	211
9.1.3 Excel 在抽样估计中的应用	214
9.1.4 Excel 在假设检验中的应用	217
9.1.5 Excel 在相关分析与回归分 析中的应用	218
9.1.6 Excel 在时间序列分析中的 应用	220
9.1.7 Excel 在因素分析中的应用	228
9.1.8 利用 Excel 计算平均指数	229
9.1.9 利用 Excel 进行因素分析	230
9.2 EViews 在统计学中的应用	233
9.2.1 EViews 的基本简介	233
9.2.2 EViews 工作特点	233

9.2.3 EViews 在相关分析与回归 分析中的应用·····	233	附录 2 t 分布临界值表 ·····	258
9.2.4 EViews 在时间序列中的 应用 ·····	236	附录 3 χ^2 分布临界值表 ·····	259
9.3 SPSS 在统计学中的应用·····	238	附录 4 F 分布临界值表 ·····	260
9.3.1 SPSS 简介 ·····	238	附录 5 二项分布临界值表 ·····	262
9.3.2 SPSS 在统计学中的应用·····	240	附录 6 秩和检验临界值表 ·····	263
附录 1 统计调查与数据整理 ·····	257	参考文献 ·····	264

第1章

CHAPTER 1

总论

学习目标

本章的学习目的在于让学生充分了解统计技能学习的重要性，并通过对统计概念、性质、特点及与其他学科关系的掌握中初步理解什么是统计，在学习过程中应注意哪些关键点。在此基础上，掌握统计学中的几个基本概念，为后续的学习打下基础。

关键概念

统计 (statistics)

标志 (characteristic)

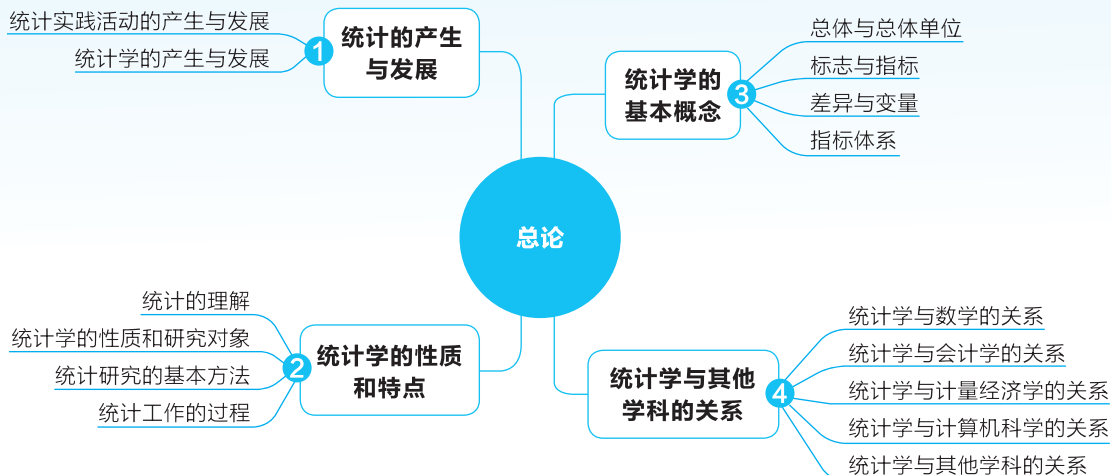
总体 (statistical population)

变量 (variable)

总体单位 (population unit)

指标 (index)

知识导图



1.1 统计的产生与发展

1.1.1 统计实践活动的产生与发展

统计实践活动是顺应人类社会、政治、经济发展以及国家管理的需要而逐步产生并发展起来的，距今已有四五千年的历史。

早期的统计实践孕育于原始社会末期简单的计数活动，人们对猎物的清点与分配可看作最早的统计行为。在我国有所谓结绳记事、结绳计数的方法，郑经的《周易正义》中提到“事大，大结其绳；事小，小结其绳；结之多少，随物众寡”，体现了统计分组、汇总的基本思想。

到了奴隶社会，为了顺利征税、征兵、征劳役，统治阶级有必要详细了解社会基本情况，这些促成了有关人口、财富以及军事方面的统计实践的长足发展。夏朝时资料记载当时全国分九州；到了西周，已建立了基本的统计报告制度，并依报送时间的长短不同分为“日成”“月要”和“岁会”。同处于奴隶社会的古埃及、古希腊、古罗马等国也都有了关于人口、土地等方面的统计实践活动。

进入封建社会后，随着人类社会生产的发展，统计的范围也开始向社会经济生活的各方面扩展，秦时的《商君书》中提到“强国知十三数”，包括反映当时基本国情国力的人口、粮食、农业生产资料以及自然资源等指标，统一中国后，更建立了从中央到地方的“上计”报告制度。长期的中央集权制国家也使得中国在户籍、田亩等领域的统计活动有了很大的发展，不论在统计的方法、制度还是组织上中国都处于当时世界的先进水平。另一方面，自给自足的封建自然经济在阻碍社会生产力发展的同时，也阻碍了统计实践的发展，统计的功能仍局限在国家管理的需要上，其主要工作也多局限在原始调查登记以及简单的汇总分析工作。

统计实践活动的广泛发展始于资本主义社会。在 17—18 世纪的资本主义上升时期，由于工、商、农、贸、交通的发展，统计实践开始扩展到社会经济活动的众多领域，开始出现各种“国情普查”，并形成制度。从 18 世纪末开始，西欧各国开始设立专门的统计机构负责统计工作，收集各方面的统计信息，定期或不定期地举行人口、工业、农业、贸易、交通等专项调查，社会统计、科技统计、环境统计等从经济统计中分离出来，形成了比较完整的统计内容体系，抽样方法的引入更将统计的认识能力推向了新的高峰。由于各国政府的重视与支持，1853 年在比利时的布鲁塞尔召开了第一次国际统计会议，1885 年成立了国际统计学会，1887 年在罗马召开了第一届大会，此后每两年召开一次，1995 年 8 月在北京召开了第 50 届会议。

20 世纪 30 年代，抽样方法的普遍采用，将统计从对现象的客观描述发展到抽样推断、预测分析方面，20 世纪 50 年代后国民经济账户体系和投入产出法的应用、各种新的数学方法的介入更是极大地丰富了统计实践活动。随着计算机等现代科学技术的发展，统计实践活动进入了一个全面发展的新阶段，统计实践已经从早期的国家管理、社会经济管理方面扩展到了人类的一般认知活动，大到一个国家或洲际的管理决策，小到一般的家庭投资决策；从

研究社会经济现象的数量方面,深入到自然技术现象的数量方面,甚至连文科领域的法律、历史、语言等专业也都开始重视统计工作,统计已不仅仅是一个专业领域,还与数学、哲学一样成为众多学科的专业基础。也应该看到,我国现当代统计的发展相对滞后,我国于1906年设立了统计局,直到新中国成立后,统计工作才有了真正繁荣发展的制度环境,但受到苏联理论界的影响,中国的统计工作长期偏向于描述统计,直到20世纪80年代,数理统计才得到大家的认可,在理论与实践发展上开始逐步与世界同步。1984年1月1日颁布施行的《中华人民共和国统计法》经过1996年5月和2009年6月的两次修订,明确指出“统计的基本任务是对经济社会发展情况进行统计调查、统计分析,提供统计资料和统计咨询意见,实行统计监督”,为中国的统计工作提供了有效的法律制度保障。

1.1.2 统计学的产生与发展

与统计实践活动的漫长发展历史相比较,统计理论的发展只有300多年,是人们对统计实践进行的理论概括和经验总结,一般认为起始于17世纪中叶的欧洲。可以大致分为以下三个发展阶段。

1. 古典统计学(17世纪中叶至18世纪末)

古典统计学的代表学派主要有政治算术学派和国势学派。

1) 政治算术学派

政治算术学派的代表人物是英国的威廉·配第(W.Petty)和约翰·格朗特(J.Graunt)。威廉·配第的代表作是《政治算术》(1676年),他针对当时欧洲的时事情况,运用大量的数字资料对英、法、荷三国的国情国力进行了系统的数量对比分析,指出英国的潜在实力,认为英国无须悲观,完全有能力超过法国、荷兰,称霸世界。该书的主要意义在于研究问题的方法,配第创造性地运用了数量、重量和尺度来进行对比分析,表达自己想说的内容,为统计学的创立奠定了方法论的基础。为此,马克思高度评价他是政治经济学之父,在某种程度上也是统计学之父。同一时期的约翰·格朗特发表了《对死亡表的自然观察与政治观察》(1662年)一书,他利用伦敦50多年的人口变动资料进行数量对比分析,编制了第一个死亡率统计表,揭示了人口与社会现象中重要的数量规律,还对伦敦的人口总量进行了较为科学的预测。

政治算术学派可看作统计学的开端,威廉·配第为后来的经济统计开辟了道路,约翰·格朗特则为人口统计奠定了基础,由此发展出农业统计、工业统计、保险统计、卫生统计等各分支学科,形成以社会经济问题为主要研究对象的社会经济统计。

2) 国势学派

国势学派又称“记述学派”,产生于德国,其创始人是德国的海尔曼·康令(H.Conring),主要以文字记述的方式为政治家们提供有关国家管理的人口、财政、军事、土地和法律等方面的知识。其继承者高特弗里德·阿亨瓦尔(G.Achenwall)在《近代欧洲各国国势学纲要》(1749年)一书中运用对比分析的方法研究国家组织、领土、人口、资源财富及国情国力,首次使用了“统计学”一词,认为统计学是关于各国基本制度的学问,其研究对象是一个国家显著事项的整体。严格地说,这一学派并没有把数量对比分析作为

统计学的基本特征，但其后来发展出来的两个分支派，即以丹麦的安彻逊等人为代表的图表派和以德国的布欣为代表的比较派，他们在国势学的研究方法基础上更加重视数字资料的分析研究，吸收了政治算术学派的长处，形成两派的初步综合。到了 19 世纪末，在欧洲各国大学中开设的“国情纪要”或“政治算术”课程逐步被取消，取而代之的是“统计分析科学”课程。

政治算术学派与国势学派同时并存，相互影响，两派之间有许多共同点，诸如都是以社会经济现象作为研究对象，以实际调查资料为根据，以反映基本国情、国力为主要目的。两派的主要区别在于是以定量分析为主还是以定性分析为主，为了统计学的正统之争，双方引发了长达百年之久的争论，直到 1850 年德国统计学家克尼斯（K.G.A.Knies）发表了《独立科学的统计学》一文，主张国势学和统计学的科学分工，从此，统计学具有数量特征的论点被确立下来。

2. 近代统计学（18 世纪末至 19 世纪末）

近代统计学的主要代表学派是数理统计学派与社会统计学派。

1) 数理统计学派

数理统计学派的创始人是比利时学者阿道夫·凯特勒（A.Quetelet），其代表著作有《统计学研究》（1844 年）、《关于概率论的书信》（1846 年）、《社会物理学》（1869 年）等。凯特勒最先将概率论、正态分布等概念应用于人口、犯罪等方面的研究，根据大数定律原理提出了大量观察法，创立了大数法则，完成了概率论与统计学的结合，使统计方法发生重大的飞跃。他认为统计学是可以研究任何科学的一般性研究方法，开创了统计理论与实践的新领域，后经高尔顿、皮尔逊、埃奇沃思、费雪等人的不断丰富和发展，逐渐成为一门完整的方法体系。1867 年韦特斯坦（T.Wittstein）在“数理统计学及其在经济学和保险学中的应用”一文中将其命名为数理统计学，凯特勒被人尊称为“数理统计学之父”“近代统计学之父”，而以随机现象为主要研究对象的数理统计也形成了统计学发展的另一条主线。

2) 社会统计学派

产生于 19 世纪后半叶的社会统计学派的首创者是德国的克尼斯（K.Knies），主要代表人物有梅尔（Mayr）和恩格尔（Engel）。该学派主要认为统计学的研究对象是社会现象，以大量观察法作为主要方法，研究社会现象的内部特点和相互关系，认为统计学是一门实质性研究的社会科学。这一学派融汇了国势学派和政治算术学派的观点，将政府统计与社会调查结合起来，强调全面调查在社会统计中居于重要地位，而抽样调查只在一定范围内具有实际意义和作用。

数理统计学派与社会统计学派也同样并存争论了百年之久，在早期，由于数理统计学派尚未充分发展，社会统计学派在欧洲占有优势地位，对世界各国都有很大的影响。随着数理统计理论的不断产生，其应用领域也不断扩大，数理统计学派开始在国际统计学界占据优势地位，但总的来说，两者已出现了融合趋势。

3. 现代统计学（19 世纪末至今）

一般认为 1908 年戈塞特（W.S.Gosset）创建了小样本分布定理，费雪（R.A.Fisher）在

此基础上提出了方差分析方法、F 统计量、最大似然估计等方法和思想,标志着现代统计学的开端。之后的尼曼、沃尔德等人先后提出了区间估计理论、决策理论、非参数统计方法等,极大地丰富了数理统计理论。另一方面,统计学受计算机科学、系统论、控制论、信息论、混沌理论、协同学理论、突变理论等现代科学技术的影响,新的研究领域及方法也层出不穷,如多元统计分析、时间序列分析、贝叶斯统计、探索性数据分析、数据挖掘、随机模拟、过程控制等,使统计研究向着系统描述、系统推断、系统分析的方向发展,获得的统计资料愈加完善。在应用领域上统计学也不断扩展,其研究对象已不仅仅局限于社会经济现象,还涉及自然技术现象等其他领域,几乎所有的科学研究都离不开统计方法。近年来,随着互联网、大数据、人工智能、云计算等技术的快速发展,统计由单元化向多元化发展,研究范围进一步扩大,调查方法不断改进,其科学性与有效性也得到了大幅提升。

1.2 统计学的性质和特点

1.2.1 统计的理解

1. 统计的含义

“统计”一词最早源于中世纪拉丁语的 *satsus*,意指各种现象的状态和状况,在中国直到明清时代统计的含义也仅指总括、汇总。“统计”作为学科专业名词出现在 18 世纪中叶,阿亨瓦尔把国势学称为 *statistik*,即统计学,认为统计是关于国家显著事项比较和记述的科学,直到 18 世纪末“统计”一词才传入英国,并被赋予新的内容,即用数字表述事实,这已基本接近现代的统计含义。在我国,“统计”一词的专业含义出现在清光绪年间,钮永建等人翻译了横山雅南所著的《统计讲义录》一书,把统计学介绍、传入我国;1907 年彭祖植编写了《统计学》,成为我国最早的一本统计学书籍,“统计”一词也被赋予现在的含义。

在现实生活中,人们往往赋予“统计”以不同的理解,诸如“学统计”“做统计”“据统计”等,概括起来可以将统计的含义归纳为三种内容,即统计工作、统计资料和统计科学。

统计工作(统计实践)是根据科学的方法对客观事物的数量方面进行设计、搜集、整理和分析,并据此得出有关客观事物规律性的工作过程,是对客观事物的数量表现、数量关系以及数量变化等进行描述和分析的计量活动。统计工作的过程也是人们认识客观世界的过程,一般可分为统计设计、统计调查、统计整理和统计分析 4 个阶段,其结果表现为各种数据资料或文字说明,即统计资料。

统计资料是统计工作所获取的反映客观事物规律性的各种数据资料或文字说明,具体表现形式有各种统计表、统计图、统计资料汇编、统计年鉴、统计手册、统计分析报告等。统计资料可分为原始资料(一手资料)和次级资料(二手资料)。原始资料是指直接从各调查单位搜集来的用以反映其特征的数据资料,是原始的、零散的、未经任何加工处理过的信息;次级资料是指由原始资料加工得到的在一定程度上能反映客观事物总体特征的数据资料。相对于原始资料来讲,次级资料更为条理化、综合化、系统化。通常,对统计资料的要求包

括：①准确性。统计资料是用来反映客观事物规律性的，这就要求资料能够如实反映客观事实，即使允许存在误差，也不能超过事先确定好的误差范围。②及时性。统计资料的搜集是为了满足统计目的的需要，要及时搜集、及时加工、及时公布，一旦不能保证，将直接影响到资料的准确性。③全面性。客观事物不是孤立地存在的，彼此之间相互作用、相互联系，这就决定了统计资料不能片面地反映某一方或某一局部的内容或联系，它也将直接影响结论的准确性。

统计科学（统计理论）是统计实践活动的科学总结与理论概括，是研究数据资料搜集、整理、分析的方法论科学。就其理论发展大致可以分为描述统计与推断统计两大部分。描述统计是指对试验或调查过程中所获得的数据资料进行审核、整理、分析，并从中获取反映整个总体量的特征的信息，并以图表等有效方式表达出来。推断统计则是指利用概率论理论，根据所获得的部分单位信息（样本信息），科学推断或检验全部单位（总体）的数量特征。应该说描述统计是统计研究的基础，它为统计推断提供了客观事实根据，但在现实中人们所获得的信息很多属于样本信息，推断统计在实际中的运用越来越广泛，它在当前统计工作研究中的作用也愈加重要。

统计的三种含义具有密切的联系，其中统计工作与统计资料是工作过程与工作成果的关系，统计资料的质量取决于统计工作过程的质量，而统计资料的好坏最终也决定了人们对统计工作的重视程度。统计科学与统计工作是理论与实践的关系，理论来源于实践，反过来又指导着实践工作。可以说，统计工作、统计资料与统计科学三者相互依存、相互联系，共同构成了统计的整体内涵，抛开任何一方去理解统计都是片面的。

2. 统计的职能

《中华人民共和国统计法》在关于统计的任务中明确指出统计应具有以下三项职能。

1) 信息职能

统计工作的最终结果是要获取统计资料，即统计信息。它作为统计的最基本职能，为人们认识世界提供了客观、公正的以数量描述为基本特征的各种信息。威廉·配第在《政治算术》中就曾提到，他所运用的方法不是传统的利用比较级词汇所描述的主观判断，而是利用客观现象本身所表现出来的数据资料说明事实。人们在日常生活中所接触的绝大多数数据信息都是统计信息，是我们认识世界的主要工具之一。

2) 监督职能

统计信息除了可以帮助我们认识客观现象的规律性外，还可以利用统计信息的反馈评判、检验和调控决策方案。人们可以利用统计信息反映客观现象的运行状态，并加以全面、系统的监督和预警，以促使其按照客观规律的要求持续、稳定、协调的发展。

3) 咨询职能

统计咨询职能是信息职能的延续和深化，一方面统计信息可以为本部门制定经营决策提供参考；另一方面也可以成为企业以及各级政府管理部门制定管理决策的依据。此外，统计信息还可以成为人们从事科学研究的资料来源。

1.2.2 统计学的性质和研究对象

1. 统计学的性质

统计学究竟是方法论科学还是实质性科学，这一问题尽管在理论界争论了上百年，但并没有形成一个完全统一的明确说法，归纳起来有三种观点：一是认为统计学是研究客观现象发展规律的，是一门实质性科学；二是认为统计研究随机现象，是一门通用的数理方法学科；三是认为统计学是研究客观现象数量方面的方法论科学。这三种观点站在各自的出发点上来看都有一定的道理，回答这个问题可以让大家明确对待统计学的态度，对全面认识这门科学以及学习、掌握该课程具有重要意义。

从数理统计学的角度来看，尽管在出现伊始就对统计学的影响巨大，不但拓宽了统计的研究领域、丰富了统计的内容，即使在现在也起着不容忽视的作用，但并不能就此说明统计学是一门应用数学。我国在 1992 年确定国家标准学科分类时将统计学与数学、经济学等并列为一级学科，数理统计学为下设的二级学科。从实质性科学来看，主要研究的是某领域现象的本质关系和发展变化规律，诸如经济学、社会学、物理学、生物学等。统计学的任务在于为统计活动提供研究数量、认识规律的方法，也就是说它为我们研究客观现象的规律性提供合适的方法，特别是数量分析的方法。如果说早期的统计学还可以说是实质性研究与方法论研究的紧密结合，那么到了 19 世纪后期，在各行业都陆续创立、使用许多行之有效的统计方法之后，统计也随之形成以统计方法为中心的方法论科学。

2. 统计的研究对象

统计学与统计工作之间理论与实践的关系决定了两者的研究对象是一致的，都是以客观事物的数量方面作为研究对象，包括数量的多少、数量关系、数量特征以及数量界限等。尽管统计是以客观事物的量作为研究对象，但并不阻碍它对事物质的方面的认识。统计研究是从客观事物内涵、外延的定性认识入手，通过定量分析，最终达到对客观事物质的认识，并且是以量的方式来体现事物质的特征，是“质—量—质”的研究过程。

需要注意的是，尽管统计发展到今天所涉及的领域越来越广泛，但并不能就此认为任何一种数量都可以用以统计研究，统计所研究的对象具有自身的特点，主要有以下几个方面。

1) 数量性

统计以客观事物的数量方面作为自己的研究对象，这里的数量方面可归纳为三类：一是反映现象规模的数量多少；二是反映现象之间各种平衡、比例、依存关系的数量关系；还有就是反映现象质的特征和规律。脱离数量性这一特点，统计就不能称为统计，这是统计区别于其他以客观现象作为研究对象的学科的根本点。例如，我们都知道中国是世界上最大的发展中国家，这是大家的定性认识，用统计的语言描述，可以简单通过以下几个数据了解，如 2019 年中国人口总数已达到 14 亿，居世界第一位；国民生产总值保持较高增速，为 6.1%，2019 年达到 99.1 万亿元，稳居世界第二位，人均国内生产总值首次站上 1 万美元的台阶，世界排名第 72 位。按目前联合国的划分标准，中国处于中等偏上收入国家的水平，人均国内生产总值相当于世界平均水平的 90%，仍是一个发展中国家。

2) 总体性 (大量性)

统计研究客观事物的数量方面是为了找出客观事物的规律性, 个别事物因其所处的时间、地点、条件不同有着明显的偶然性和不确定性, 难以说明某一现象总体的本质和规律性。因此, 统计只有对客观事物中的大量个别现象进行观察、分析和研究, 才有可能消除不确定因素的影响, 从而正确揭示出现象总体的本质与规律性。需要提到的是, 客观现象总体由大量的个别单位所构成, 统计在强调研究现象总体的同时, 并不否认个别事物对总体的影响, 应该说统计是从调查个别事物的表现入手, 从个体到总体, 最终找到现象固有的内在变化规律。例如, 要了解城镇居民生活水平的时候, 个别某一户家庭或个人的生活状况并不是我们研究的目的, 我们需要了解的是一个市区、一个城市、一个省, 乃至一个国家的居民的收入、支出、储蓄等状况, 统计所要做的就是消除个别家庭对总体情况的影响, 真实地反映出全体居民的生活水平。

3) 具体性 (客观性)

统计研究对象中的数量是客观存在的、具体的数量, 是在一定的时间、地点、条件下发生的客观上已经存在的事实的量。统计中的每一个量都与现象的质紧密结合在一起, 这与数学中抽象地体现量的数字是不同的。即使在表现研究对象数量方面时有所谓主观项目, 但其反映的也是客观上存在的事实, 是被调查单位主观意愿的客观反映。以前文提到的居民生活水平为例, 统计研究时必须明确是什么时间、地点, 什么范围内的居民的生活水平, 什么样的收入、支出才是所需研究的内容; 必须深入实地调查获取真实信息, 每一数据都必须直接或间接来源于居民的实际情况, 绝不许凭空杜撰或以错误的推断居民的生活状况。

4) 差异性

统计研究现象总体数量方面的前提是总体内各单位之间要存在差异, 而且这种差异不能根据特定原因事先给定, 也不能按照已知条件事先推定。差异应该是由随机因素而引起的个别事物的特殊性或偶然性。正是由于现象总体内各单位存在差异, 才有了统计的必要。前文提到的居民生活水平, 之所以有研究的必要, 主要就是因为每一家庭的生活水平都各不相同, 人们无法通过了解几户家庭的情况来确定一个地区所有居民的真实生活水平, 统计所要做的就是综合所有居民的个别情况来获得全体居民生活水平基本概貌, 它可以代表所有居民的平均生活水准, 是在所有居民生活水平基础上的抽象, 而不是用来准确描述每一居民或家庭的实际情况。

5) 社会性

统计的研究对象是处于一定的社会环境之中的, 是人们有意识、有目的的活动的结果。就其研究的主体——社会经济现象来说, 其本身就是人类有意识的社会活动及其产物, 都与人的利益有关, 体现的也是人与人的关系; 即使把统计的研究对象扩展到自然现象, 也应注意其认识主体——人的社会性, 如果不把自然现象纳入人的研究中, 它就不是统计的研究范畴。统计并不研究自然规律, 只有人基于某种目的运用统计方法来研究自然现象时, 统计活动才会涉及特定的自然现象。也就是说, 统计是一种社会认识活动, 必然要受到社会制度、规范以及社会心理的影响。例如, 在研究职工收入情况时, 如果不区分经营者、管理者、技术人员和普通工人, 势必会造成低估高收入群体收入, 而高估低收入群体收入的情况, 从而引起不必要的矛盾。

1.2.3 统计研究的基本方法

统计有自己的研究对象，在长期的发展过程中也总结并形成了一系列特有的研究方法，这里要介绍的是一些基础性的方法，了解它们不但有利于在统计工作中解决实际问题，同时还有利于初学者从方法体系上了解和认识统计。

1. 大量观察法

大量观察法是指在统计研究过程中，对所研究对象中的全部或足够多的单位进行观察，以反映总体数量特征的方法。大量观察法是统计研究现象总体客观规律性的根本手段。其理论依据是大数定律，认为随机事件在大量重复性试验中的频率是稳定的，随机变量在多次观测中得到的均值也总会稳定在它的期望值附近。这说明大量观察可以促使个体差异的相互抵消，在一定范围内排除个别现象偶然因素的影响，而显现出现象总体的稳定性质。大量观察法广泛应用于各种统计实践，诸如统计报表制度、普查、重点调查和抽样调查等，都是基于大量观察的调查方式。

2. 统计分组法

统计分组法是根据统计研究目的和研究现象总体的特点，按照一定的标志把所研究现象划分为若干部分或若干组的统计研究方法。分组法是研究现象总体内部差异的主要方法，通过统计分组可以区分现象的类型，研究总体中的结构和比例关系，揭示现象之间的相互依存关系。例如，我们可以通过行业分类来看一个地区的行业分布情况，确定该地的主体经济情况，依其经济规模的大小可以确定与利润、税收等之间的关系。统计分组法存在于统计工作的各个阶段，统计调查离不开统计分组，分组更是统计整理环节的关键所在，许多统计分析方法也都是建立在科学的统计分组之上的。

3. 综合指标法

综合指标法就是利用总量指标、相对指标、平均指标等综合指标对现象总体的数量方面进行描述、分析和概括，从而反映现象本质及规律性的一种分析方法。如前所述，统计研究的是客观事物的数量方面，这些数量特征主要是通过统计指标来进行分析和描述的，综合指标可以从静态和动态上综合反映和分析现象总体的规模、水平、结构、比例、变化程度以及依存关系等方面的数量特征与数量关系，是一种应用简单而且涉及广泛的基础性统计分析方法。

4. 统计模型法

统计模型法是在一定的理论和假定条件下，用数学方程的形式去模拟客观现象数量关系的一种分析方法。统计模型一般包括三个基本要素：变量、基本关系式和模型参数。统计模型法是通过客观现象中存在的数量关系进行较为完整和近似的描述，以达到对客观现象的变化进行数量上的估计与预测的目的。该方法可以看作综合指标法的扩展，是统计研究方法系统化、精确化发展的产物，大大提高了统计分析的认识能力。

1.2.4 统计工作的过程

就一次统计活动而言,一个完整的统计工作过程可分为统计设计、统计调查、统计整理和统计分析四个阶段。

统计设计是根据研究对象的性质及研究目的的要求,对统计工作的各个方面和各个环节所做的通盘考虑与安排。设计阶段涉及了统计工作的各个阶段的工作内容安排和各工作部门之间的协调安排。统计设计的主要内容包括统计指标和指标体系的设计、统计资料搜集方法的设计、统计分类和统计分组的设计、统计表的设计、统计工作各个部门和各个阶段的协调与联系及统计力量的组织与安排。其结果主要表现为各种规定、制度、方案和方法。设计阶段可以说是整个统计工作阶段的前提,不能根据工作的繁复程度来进行取舍。因为设计阶段在其他统计工作之前,我们不能保证设计方案对后续工作的指导完全正确,因此,在实际工作过程中,设计方案往往要根据实际情况随时进行修正,以确保工作的顺利进行。

统计调查是指根据统计方案的要求,采用各种调查组织方式和方法,有组织、有计划地对研究对象的各单位进行调查、登记,准确、及时、全面地搜集统计资料的过程。这个阶段是统计工作从定性认识向定量认识过渡的阶段,搜集资料的质量高低直接决定了后续统计工作的结果,决定了整个统计工作的质量,因此,可以说统计调查是整个统计工作的基础。

统计整理是根据研究目的的要求,对调查阶段所搜集来的原始资料进行审核、分组、汇总,使之条理化、系统化,由反映个别特征的资料信息转化为反映总体或各组数量特征的综合资料的工作过程。统计整理阶段是整个统计工作过程的中间环节,是统计调查阶段的继续,又是统计分析阶段的必要前提,是对客观现象由个体认识过渡到总体认识、由感性认识上升到理性认识的必经阶段,在统计工作中起着承前启后的作用。

统计分析是对统计整理阶段的数据资料,利用科学的统计分析方法进行数量分析,以揭示所研究现象的本质和规律,进而向相关部门或单位提出建议的过程。统计分析是统计工作的最后阶段,在这个阶段所做出的结论是整个统计工作的目的所在,也是统计能否充分发挥信息、咨询和监督职能的关键所在。

1.3 统计学的基本概念

在学习统计的理论和方法时,常用到一些专业概念,这些概念贯穿于整个统计实践过程和统计理论之中,掌握了它们也就掌握了统计学的基础,有必要在此专门加以阐述。

1.3.1 总体与总体单位

总体是统计总体的简称,是客观存在的、具有某种共同性质的许多个别事物所构成的整体。构成统计总体的每一个个别事物就是总体单位。总体是统计研究的具体对象,是客观存在的且可以准确界定的存在。例如,要研究某市民营企业的生产经营状况,那么该市的所有民营企业就构成了一个统计总体,该市每一家民营企业则是一个总体单位。同理,也可以研

究一批灯泡的使用寿命,这批灯泡构成了统计总体,每一个灯泡是一个总体单位。界定某一现象是否为统计总体可以从以下三个方面来看。

1) 同质性

众多个别事物之所以能够作为一个总体存在,其根本原因在于它们之间在某特定研究目的下具有某种相同的性质,这种同质性成为界定个别单位是否为某一统计总体单位的主要依据。在前面的第一个例子中,某市民营企业构成了统计总体,意味着能够成为这个总体中的总体单位必须是该市的企业,且性质上必须是民营的。第二个例子中的灯泡,界定是否为同一总体的前提就是这些灯泡的生产批次,只有同一批次的灯泡才是该例的研究总体。

2) 大量性

前文提到统计是对大量事物的数量表现进行研究以揭示其规律性,总体作为统计的研究对象,也必须存在着足够多的总体单位,也就是说,总体必须由许多个别事物构成,个别或少数单位形不成统计总体。个别事物的现象具有极大的偶然性,只有大量现象的总体才具备稳定性,研究总体就是要在大量观察中找出现象的发展趋势和规律。

3) 差异性

如果说同质性是构成总体的条件,那么差异性(变异)就是统计研究的基础。统计研究的实质就是在明确各单位之间差异的基础上,利用统计方法将总体单位对总体的差异影响抵消掉,从中找出总体的规律性。在前面的例子中,之所以要研究民营企业的生产经营状况,主要在于该市的每一家民营企业无论是经营领域、经营规模,还是效益、产出上都各不相同,无法用某一家企业的情况来代替全体的水平,研究它们就是要去异求同,找出该市民营企业的整体实力水平。

根据总体大量性的状态,可以把总体分为有限总体和无限总体。有限总体的总体单位个数有限,无限总体的总体单位数是无限的。前例都是属于有限总体的情况,这是统计研究中的主体,可适用各种统计调查方法。无限总体在统计中也较常见,例如,要测定某车间的温度或湿度,只要在车间的空间范围内任一地点、任一时间都会得到不同的数据,我们所面临的就是一种无限总体的情况;再比如,要检验新产品在市场上的受欢迎程度,它所面临的消费者也是一种无限总体的情况。对于无限总体,只有采用抽样调查才能够推断其总体的特征。还有一种常见的总体分类,存在统计分组的情况下,按照某种标志将统计总体分组后,除了既有的总体(大总体)外,还会出现许多小总体(类总体、组总体),也就是将总体单位分成的每一类或组,尽管这些组(类)要小于所研究的大总体,但它仍然存在着统计总体的特征,仍是以总体的形态存在的。

在抽样调查中也有一种统计总体的理解,这里的总体又称为全及总体,是指在抽样调查中所要研究的所有事物构成的整体,这个总体的概念与我们在本节中提到的总体概念基本一致。构成总体的个别事物称为总体单位,又叫个体。较为特殊的是,在抽样中还有一部分选中单位,也就是按照随机原则抽取的部分单位构成的集合,这也是一个总体的概念,我们称之为样本总体,简称样本。在一个全及总体中可以抽取许多个样本,抽样推断所要做的就是利用样本总体来推断全及总体的特征。

总体与总体单位之间的关系是整体和个体的关系,两者相互依存、相互联系。总体单位构成了统计总体,要认识统计总体必须先从认识总体单位开始,总体和总体单位的具体形式

也随着统计研究目的的不同而不同，可以是人，可以是物，还可以是组织或时间、空间、行为等。此外，总体和总体单位之间的关系也不是一成不变的，随着研究目的的变动，两者可以相互转化。例如，在前例中研究某市民营企业的生产经营状况，该市的所有民营企业构成了统计总体，该市的某一家民营企业 A 就是总体单位；如果要研究该市民营企业 A 的生产经营状况，那么 A 企业就不再是总体单位而是统计总体了。

1.3.2 标志与指标

1. 标志

标志是说明总体单位所具有的属性或数量特征的名称。标志的具体表现为标志表现。例如，每个人都具有性别、年龄、身高、工种、工龄、收入等特征，这些都是人这个总体单位的标志。又如每家企业都有企业类型、产值、销售收入、职工人数、工资总额等标志。需要注意的是，标志反映的是总体单位所具有的属性或特征，并不包括该属性或特征的具体表现内容，以一个人的收入为例，收入是标志，而收入的具体数额则是标志表现。

标志按照所反映的单位的特征不同可以分为数量标志与品质标志。凡表明总体单位数量特征的是数量标志，而说明总体单位属性特征的为品质标志。如企业的产值、销售收入、职工人数等，人的年龄、工龄、身高、收入等是数量标志；企业的经济类型，人的性别、工种等是品质标志。数量标志的标志表现可用数值来表示，因此，也可将其标志表现称为标志值；品质标志的标志表现一般为文字表述。

按照标志表现是否相同，可以将标志分为不变标志和可变标志。在同一总体内，某一标志的标志表现都相同，即为不变标志。例如，在某市民营企业这个总体中，企业性质这一标志是不变标志，其标志表现都是民营。在一个总体中至少要存在一个不变标志才能把各单位结合成一个总体，不变标志是总体同质性的反映。可变标志指的是在同一总体内，某一标志的标志表现各不相同。例如，前面提到的某市民营企业这个总体中，就销售收入这个标志而言，该市每一家民营企业的销售收入不可能是完全一样的。可变标志体现的是总体的差异性，没有差异也就没有统计的必要了。

2. 统计指标

统计指标是说明总体数量特征的概念和数值。指标不同于标志，可分为两个构成部分：一是指标名称；一是指标数值。例如，2019 年厦门全年地区生产总值为 5 995.04 亿元，按常住人口计算的人均地区生产总值为 142 739 元。这里的地区生产总值与人均地区生产总值是指标名称，而 5 995.04 亿元和 142 739 元就是指标数值，两部分都是指标的构成内容，只有在统计设计阶段，由于尚未正式开始数据资料的搜集工作，设计方案中的指标可以仅指指标名称。从上例中可以看到，单纯指出地区生产总值 5 995.04 亿元，并不能说明什么问题，客观现象中的事物都是具体的，是在一定的时间、地点、条件下发生的，若想用指标反映现实情况，就必须包括更多的信息在内。所以，一个完整的统计指标除了包括指标名称、指标数值外，还应包括指标所属的时间范围、空间范围、计量单位等要素，这五个要素是指标在说明总体数量特征时不可缺少的项目。如上例中，5 995.04 这个数值指的是 2019 年全年厦门市的地区生产总值数据，计量单位是亿元。

指标按所反映的数量特点不同,分为数量指标与质量指标。数量指标反映的是现象总体的总规模、总水平和总工作量,其数值大小一般会随着总体范围的扩大而增加,多用绝对数的形式表示,也称为总量指标。例如,人口总数、收入总额、货物吞吐量、国内生产总值、财政支出等。质量指标反映的是现象总体的相对水平和工作质量,是数量指标的派生指标,一般表现为相对数或平均数。

由于指标是反映研究现象总体的,所以指标也具备了对象总体的一些特点。诸如,指标无论是反映总体质或是量的特征,都毫无意外地以数值来表示;指标反映的也是一定时间、地点、条件下的总体特征,具有具体性的特点;同时,指标也具有综合性的特点,也就是说,指标所反映的总体特征,是由总体单位数量综合而来的,是对总体特征的综合说明。

3. 标志与指标之间的区别与联系

1) 区别

标志是用来说明总体单位特征的,而指标是用来说明总体数量特征的;标志的具体表现可以分为由文字表达的品质标志和由数值表述的数量标志,而指标无论是数量指标还是质量指标都是用数值来表示的。

2) 联系

标志与指标之间的联系来源于总体单位和总体之间的关系。全部的总体单位构成了统计总体,使得许多具有可加性的数量标志经过汇总就可以成为对应总体的指标。同理,总体单位与总体之间存在着转换关系,标志与指标也存在这种转换,即在研究目的不同的情况下,当原来的总体变为总体单位时,相应的指标也就变成数量标志了;反之亦然。

1.3.3 差异与变量

前文提到,统计总体内部必须存在差异性,这是统计研究的必要前提,用来反映总体单位和总体的标志与指标也都体现出了这种差异性。如标志中的可变标志的理解,指标因时间、地点、条件不同所反映出的数量差异。所有这些呈现出来的各种差异,在统计学上称为变异。因此,可以说变异的存在是统计存在的前提,正因为有了差异才有统计的必要。

变量是指可变的数量标志。与标志一样,变量也可以分为两个部分:一是变量的名称;一是变量的表现,即变量值。如某企业的销售收入为2 000万元,这里的销售收入是变量,而2 000万元是变量值。品质标志并不能称为变量,只有在经过量化处理后,才可称为变量。此外,还需要注意的是,人们在研究过程中为了描述方便,往往把有变异情况的指标也统称为变量,这更多体现的是实际的操作需要,而非变量的理论界定。

变量按取值是否连续,可分为离散变量和连续变量。离散变量在测度时只能整数断开,其数值可一一列举,多以计数的方式获得,例如人口数、企业数、桥梁数、牲畜数、机器台数等。而连续变量在测度时可以做无限分割,相邻两数之间还可能呈现出无限个不同位数的小数数值,一般通过度量工具测得,例如,人的身高、体重、年龄,产品的产值、销售额等。

变量按其所受因素影响的不同,还可分为确定性变量和随机变量。确定性变量是指受确定性因素影响的变量,如中奖人数、企业的收入、成本、利润等,多为描述统计研究的内

容。这里的确定性因素指的是明确的、可解释的、人为的或受人控制的因素。随机变量则是受随机因素影响的变量，如中奖号码、股票成交量等，一般在推断统计中出现。随机因素是那些不确定的、偶然的、非人为控制的、不可解释的因素。

1.3.4 指标体系

任何一个统计指标反映的都只能是统计总体某一方面的数量特征，而统计所研究的客观现象总体是一个复杂的总体，要想全面系统地认识总体，描述事物发展的全过程，就要设立统计指标体系。统计指标体系指的是具有某种内在联系的一系列指标所构成的整体。例如，要了解企业的财务状况，企业的资产、负债、成本费用、收入、利润、所有者权益等都要在指标体系中有所反映，才能确保对企业财务状况的全面掌握。

由于每个现象总体的内在联系具有不同的特点，统计指标体系一般可分为两种类型：一是存在确定的数量关系的以数学式联系的指标体系，如“产量 $\times$ 价格 = 产值”，在这个指标体系中有产量、价格和产值三个指标；二是存在某种共性的框架式联系的指标体系，如财政部制定的企业经济效益评价指标体系包括了销售利润率、总资产报酬率、资本收益率、资本保值增值率、资产负债率、流动比率（或速动比率）、应收账款周转率、存货周转率、社会贡献率、社会积累率等 10 项指标。

设计统计指标体系是一项科学性很强的复杂工作，在构建体系时，要贯彻少而精的原则，注重指标之间的层次性以及彼此之间的联系，使其统一在统计研究目的的需要上。指标体系在设计上还要注意它的可行性，设计出来的指标应该能够满足研究的需要，在此基础上，还要兼顾指标体系的通用性和稳定性，使指标体系尽可能长时间地服务于统计工作。

1.4 统计学与其他学科的关系

统计作为一门独立的社会科学，在成长过程中借鉴了众多学科的知识理论，同时，也为其他学科提供了大量的数量研究方法，统计学并不孤立地存在于人们所创造的知识体系中，它与其他学科有着千丝万缕的联系。这里概括地介绍统计与几门学科之间的关系，以加强读者对统计学的认识。

1.4.1 统计学与数学的关系

统计学与数学都是研究数量的关系和规律的科学，都要用大量的数据进行分析和论证。不同的是，数学研究的主要是抽象的数字、抽象的数量规律，主要运用逻辑推理、论证的方法，抛开具体的对象，研究最基础性的数量关系。统计学在成长过程中借鉴了大量的数学方法，如微积分、线性代数、概率论、数量统计、模糊数学等都为统计提供了有力的分析工具，在今后统计学也仍会从数学中吸取养分，以丰富自身的方法论体系。需要注意的是，统计研究的量是客观存在的、与所研究的客观对象联系在一起的数量，这些数据从客观实际中搜集而来，最终还是要返回实际中去反映或指导实践工作。统计学尽管也经常使用数学公式

推导的方式，但其主要采用的是归纳与演绎相结合的方式，重在探讨数量规律所对应的客观现象的实际情况，而非单纯地研究数量关系。

1.4.2 统计学与会计学的关系

统计学与会计学在某种程度上说都存在数据的搜集、处理和分析过程，但两门学科的目标、职能各不相同，这决定了两门科学的根本区别。前文说过，统计学是研究如何进行数据调查、整理和分析以反映客观现象规律的科学，它的研究对象是客观事物的数量方面，其数据多数直接或间接来源于统计调查。会计学的目标则是对会计主体的活动进行核算以提供会计主体活动的信息，它的研究对象是会计主体的资本运动的数量方面，利用的也是各种原始凭证，有着自己的一套数据登记方法。会计的这些规定要求会计核算必须真实地记录、如实地反映会计主体的活动，一般不能采用推算或估计的方法。应该说会计所研究的范畴要比统计小很多，但这并不影响两者在信息资料上的相互利用，统计更是会计信息的主要使用者。两者在研究方法上也互相借鉴，财务分析中就大量地使用了统计的分析方法，而在国民经济核算体系中也运用了会计的复式记账原理和账户体系。

1.4.3 统计学与计量经济学的关系

计量经济学是数学、经济学和统计学的综合，它利用计量的方法论述经济关系，侧重于对经济现象的验证、评价和预测。计量经济学主要使用的是模型分析，尤其是回归模型的分析，通过参数估计与假设检验来拟合模型，并对模型的有效性进行判断，目的是研究多个变量之间的相关关系或依存关系。相较于计量经济学来说，统计学的研究工作内容以及所涉猎的领域要更加复杂化、多元化，在研究方法上也比计量经济学丰富得多。

1.4.4 统计学与计算机科学的关系

在计算机没有发明出来之前，统计对大量数据的处理是依靠人工完成的，速度慢、效率低下，准确性也受到一定的影响。随着计算机技术的发明和普遍应用，统计工作也开始采用计算机技术，并研发了各种各样的软件包，加快了统计信息的处理与发布，也使得统计信息的存储更加方便，保存的时间也更加长久。目前国际上较为流行的统计分析软件有 SAS (statistical analysis systems)、SPSS (statistical package for the social sciences) 系统，Excel 也能处理一些基础性的统计问题，掌握其中的一种或几种软件对于学习统计是必要的。

1.4.5 统计学与其他学科的关系

统计学作为一门方法论科学，在各个学科领域的应用形成了各种不同的专业统计学科，如经济统计学、社会统计学、人口统计学、科技统计学、环境与生态统计学等。这些专业统计学主要研究的是各专业领域现象总体的数量方面，相应地也形成有别于其他领域的专业统计术语，这些专业术语的理解是建立在相应领域的基本理论之上的，如经济统计学的理论基

础是经济学、人口统计学的基础是人口学等，脱离了这些理论基础，也就无法掌握这些专业统计学。可以说，专业统计学是统计方法论与各专业领域理论结合的产物。

本章小结

统计实践活动产生于原始社会末期，在资本主义社会得到广泛发展，统计理论也产生于统计实践活动丰富的资本主义发展初期，主要包括三个时期：古典统计时期、近代统计时期和现代统计时期。统计一词包含统计工作、统计资料和统计科学。统计工作包括设计、调查、整理、分析四个环节。统计是研究客观现象总体数量方面的方法论科学，统计所研究的对象要具有数量性、总体性、具体性、差异性和社会性的特点。基本的统计研究方法有大量观察法、统计分组法、综合指标法和统计模型法。统计具有信息、监督与咨询的职能。

统计总体是客观存在的、具有某种共同性质的许多个别事物所构成的整体。构成统计总体的每一个个别事物就是总体单位。总体具有大量性、同质性和差异性的特点。总体与总体单位之间的关系是整体和个体的关系，两者相互依存、相互联系。标志是说明总体单位所具有的属性或数量特征的名称。标志的具体表现为标志表现。统计指标是说明总体数量特征的概念和数值。指标不同于标志，具有六个构成要素，分别是指标名称、指标数值、指标所属的时间范围、空间范围、计算方法和计算单位。变量是指可变的数量标志。变量按取值是否连续，可分为离散变量和连续变量。

小案例

2019 年我国国民经济和社会发展统计公报（节选）

2019 年，面对国内外风险挑战明显上升的复杂局面，在以习近平同志为核心的党中央坚强领导下，各地区各部门以习近平新时代中国特色社会主义思想为指导，全面贯彻党的十九大和十九届二中、三中、四中全会精神，按照党中央、国务院决策部署，坚持稳中求进工作总基调，坚持新发展理念和推动高质量发展，坚持以供给侧结构性改革为主线，着力深化改革扩大开放，持续打好三大攻坚战，统筹稳增长、促改革、调结构、惠民生、防风险、保稳定，扎实做好稳就业、稳金融、稳外贸、稳外资、稳投资、稳预期工作，经济运行总体平稳，发展水平迈上新台阶，发展质量稳步提升，人民生活福祉持续增进，各项社会事业繁荣发展，生态环境质量总体改善，“十三五”规划主要指标进度符合预期，全面建成小康社会取得新的重大进展。

一、综合

初步核算，全年国内生产总值 990 865 亿元，比上年增长 6.1%。其中，第一产业增加值 70 467 亿元，增长 3.1%；第二产业增加值 386 165 亿元，增长 5.7%；第三产业增加值 534 233 亿元，增长 6.9%。第一产业增加值占国内生产总值比重为 7.1%，第二产业增加值比重为 39.0%，第三产业增加值比重为 53.9%。全年最终消费支出对国内生产总值增长的贡献率为 57.8%，资本形成总额的贡献率为 31.2%，货物和服务净出口的贡献率为 11.0%。人均国内生产总值 70 892 元，比上年增长 5.7%。国民总收入 988 458 亿

元,比上年增长6.2%。全国万元国内生产总值能耗比上年下降2.6%。全员劳动生产率为115 009元/人,比上年提高6.2%。

年末全国大陆总人口140 005万人,比上年末增加467万人,其中城镇常住人口84 843万人,占总人口比重(常住人口城镇化率)为60.60%,比上年末提高1.02个百分点。户籍人口城镇化率为44.38%,比上年末提高1.01个百分点。全年出生人口1 465万人,出生率为10.48‰;死亡人口998万人,死亡率为7.14‰;自然增长率为3.34‰。全国人户分离的人口2.80亿人,其中流动人口2.36亿人。

年末全国就业人员77 471万人,其中城镇就业人员44 247万人,占全国就业人员比重为57.1%,比上年末上升1.1个百分点。全年城镇新增就业1 352万人,比上年少增9万人。年末全国城镇调查失业率为5.2%,城镇登记失业率为3.6%。全国农民工总量29 077万人,比上年增长0.8%。其中,外出农民工17 425万人,增长0.9%;本地农民工11 652万人,增长0.7%。

全年居民消费价格比上年上涨2.9%。工业生产者出厂价格下降0.3%。工业生产者购进价格下降0.7%。固定资产投资价格上涨2.6%。农产品生产者价格上涨14.5%。12月份,70个大中城市新建商品住宅销售价格同比上涨的城市个数为68个,下降的为2个。

年末国家外汇储备31 079亿美元,比上年末增加352亿美元。全年人民币平均汇率为1美元兑6.8985元人民币,比上年贬值4.1%。

供给侧结构性改革继续深化。全年全国工业产能利用率为76.6%,比上年提高0.1个百分点。其中,黑色金属冶炼和压延加工业产能利用率为80.0%,提高2.0个百分点;煤炭开采和洗选业产能利用率为70.6%,与上年持平。年末商品房待售面积49 821万平方米,比上年末减少2 593万平方米。其中,商品住宅待售面积22 473万平方米,减少2 618万平方米。年末规模以上工业企业资产负债率为56.6%,比上年末下降0.2个百分点。全年教育、生态保护和环境治理业固定资产投资(不含农户)分别比上年增长17.7%和37.2%。“放管服”改革持续深化,微观主体活力不断增强。全年新登记市场主体2 377万户,日均新登记企业2万户,年末市场主体总数达1.2亿户。全年减税降费超过2.3万亿元。

新动能保持较快发展。全年规模以上工业中,战略性新兴产业增加值比上年增长8.4%。高技术制造业增加值增长8.8%,占规模以上工业增加值的比重为14.4%。装备制造业增加值增长6.7%,占规模以上工业增加值的比重为32.5%。全年规模以上服务业中,战略性新兴产业企业营业收入比上年增长12.7%。全年高技术产业投资比上年增长17.3%,工业技术改造投资增长9.8%。全年服务机器人产量346万套,比上年增长38.9%。全年网上零售额106 324亿元,按可比口径计算,比上年增长16.5%。

区域协调发展扎实推进。分区域看,全年东部地区生产总值511 161亿元,比上年增长6.2%;中部地区生产总值218 738亿元,增长7.3%;西部地区生产总值205 185亿元,增长6.7%;东北地区生产总值50 249亿元,增长4.5%。全年京津冀地区生产总值84 580亿元,比上年增长6.1%;长江经济带地区生产总值457 805亿元,增长6.9%;长江三角洲地区生产总值237 253亿元,增长6.4%。

脱贫攻坚成效显著。按照每人每年 2 300 元（2010 年不变价）的农村贫困标准计算，年末农村贫困人口 551 万人，比上年末减少 1 109 万人；贫困发生率 0.6%，比上年下降 1.1 个百分点。全年贫困地区农村居民人均可支配收入 11 567 元，比上年增长 11.5%，扣除价格因素，实际增长 8.0%。

（资料来源：国家统计局官网）

复习思考

1. 什么是统计？统计的研究对象是什么？举例说明统计的研究对象有哪些特点。
2. 从统计学的发展历史来看，统计学的性质有哪些变化？为什么说统计学是一门方法论科学，统计学的这种性质要求我们在学习统计时应注意哪些问题？
3. 阅读本章案例导读中的材料，了解统计的基本作用，并说明统计是如何反映客观事物的特征的。通过阅读材料了解 2019 年我国社会经济发展情况。
4. 怎样理解统计总体的大量性、同质性和差异性？
5. 举例说明为什么总体和总体单位会随着研究目的的改变而改变。
6. 从统计学的发展史来看，试说明描述统计和推断统计的演化，并试着解释什么是描述统计，什么是推断统计，两者在研究现象总体时有何不同？
7. 阅读下列材料：“2014 年厦门全市共有各类卫生机构 1 255 个，其中医院 43 家、社区卫生服务中心 25 个、卫生院 13 个、门诊部 127 个、妇幼保健机构 7 个、疾病预防控制中心 7 个、专科防治机构 1 个；专业卫生技术人员 2.31 万，其中，中山医院拥有床位 2500 张，卫生人员 2 600 余名。全市人口平均期望寿命为 80.02 岁。”请解释统计总体、总体单位、标志、指标以及变量的概念，并从上述材料中指出这些总体、总体单位、标志、指标和变量是哪些。
8. 标志、指标和变量的分类有哪些？试举例说明。
9. 什么是统计指标体系？试根据自己所熟悉的社会经济问题或周围的日常生活问题设计一个统计指标体系，了解指标体系在设计时应注意的问题。
10. 结合自己所学的专业，探讨统计与该专业是否存在联系，以及如何在自己的专业中发挥统计工具的作用。



第一章更多习题可
扫描二维码进行测试

第2章

CHAPTER 2

统计调查与数据整理

学习目标

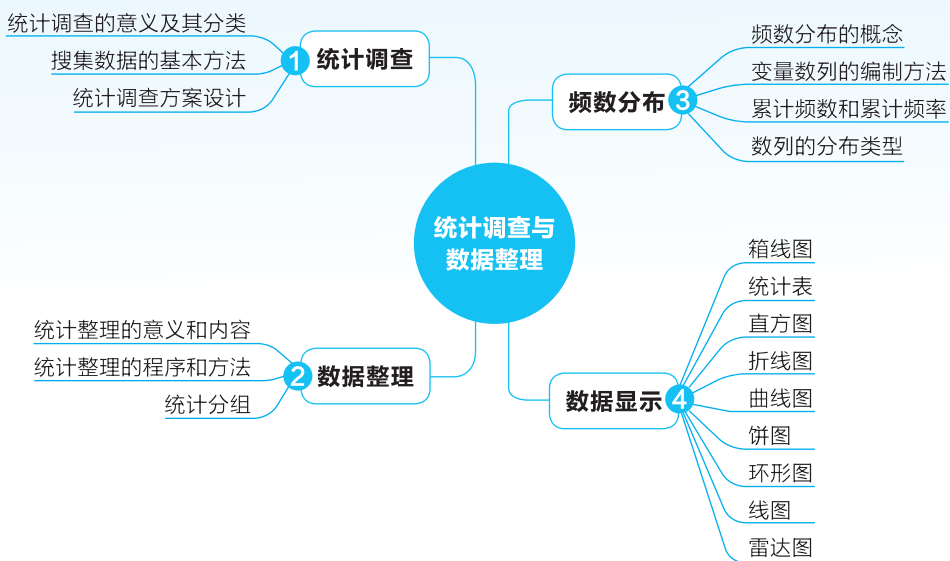
通过本章的学习,使读者认识统计调查与数据整理的基本理论,掌握统计调查与数据整理的基本方法及其在实际研究中的应用,了解数据显示的方法。

关键概念

统计调查 (statistical investigation) 统计分组 (statistical grouping)

频数分布 (frequency distributing)

知识导图



2.1 统计调查

2.1.1 统计调查的意义及其分类

数据是统计的根本，它从何而来呢？一般而言，可以通过三种途径获得数据，①对现象进行直接的调查和记录；②对现象进行特殊的实验和观察；③引用现有的数据。相应地，统计数据的收集方式也分为统计调查、实验研究和利用现有数据三种类型。大量的经济管理活动中获取统计数据的主要途径是统计调查，因此，统计调查是获得社会经济管理等方面数据的主要来源。

1. 统计调查的意义

统计调查是按照统计设计所确定的指标和指标体系，采用科学的方法，有组织、有系统地搜集原始资料的过程。其任务是取得准确、及时、全面、系统的原始资料。

统计调查在统计工作中具有重要意义。首先，从统计工作的全过程来看，统计调查是认识事物的起点。统计调查处于基础阶段，统计资料的整理、计算汇总与分析研究都必须在调查搜集资料的基础上进行。因此，调查工作的好坏，取得的资料是否完整与正确，将直接影响以后各个阶段工作的好坏，影响整个统计工作任务的完成。其次，从统计调查的主要特征来看，统计调查能够搜集到大量数字资料信息。

统计调查必须达到准确性、及时性和完整性三个基本要求。准确性是统计调查最基本的要求，也是统计工作的生命。因此，在调查中必须坚持统计制度和纪律，如实反映客观实际，防止各种错漏发生。及时性要求按规定的时间及时完成各项调查任务，使决策者能及时得到信息，因为过时的资料犹如“雨后送伞”。完整性要求统计调查按调查项目所列内容进行全面系统的资料搜集。

2. 统计调查的分类

统计调查有五种基本方式，即统计报表制度、普查、重点调查、典型调查和抽样调查。它们主要解决了要调查哪些单位的问题。

1) 统计报表制度

统计报表制度是根据有关法规的规定，自上而下地统一布置、自下而上地逐级按照统一要求提供基本统计数据的调查方式。它由政府统计部门向列入调查范围的全部统计调查单位发放统计报表，并由这些单位以原始记录为依据，按照统一的指标定期填好后，按照统一的报送时间和程序向发放报表的部门报送。

统计报表制度是国家搜集国民经济和社会发展情况基本资料的手段，为国民经济和社会发展规划的制定与执行服务。其主要优点有：①设计精心周密、高度统一、规范。统计报表制度的指标体系、表格形式、报送程序和报送时间均由国家统一规定；②回收率高，内容相对稳定，便于资料的积累和对比；③层层上报、逐级汇总，可以满足各级部门的管理需要；④能够满足不同时期的分析需求。日报、旬报可用于报告重要项目的进度，月报、季报和半年报可用于经常性计划和合同执行情况的检查，年报则可用于年度总结和分析。

2) 普查

普查是指为了详尽了解某时某地某种事物的全面情况而专门组织的一次性全面调查。普查比其他方式的调查所获得的资料要更为详细全面,但是普查涉及面广,调查单位多,耗费的人力、物力和财力也较多,经历的时间也较长。因此,普查一般主要用来调查重要项目。例如,一个国家或地区的人力资源、财力资源和物质资源的数量、分布及其利用情况等。

实际工作中,普查通常采用两种组织方式进行:一种是组织专门的普查机构进行普查,另一种是不设立专门的普查机构。前者通过成立专门的普查机构,并配备大量的专业普查员对调查单位进行直接登记并开展普查工作,例如我国的人口普查等。后者是直接利用调查单位的原始记录和核算资料,结合盘点清查,由调查单位自行填报调查表格而开展普查工作,例如我国的物资库存普查等。

与前者相比,后者组织方式简便,适用于内容单一涉及范围较小的情况,尤其适用于出于某种紧迫性而需要快速进行的普查。但是,它们都是一次性或周期性的,并且数据标准化程度高,有着较高的准确性,可以为其他调查提供基本依据。同时也都要规定统一的标准时点和普查期限。

3) 重点调查

重点调查是指在全部单位中选择一部分重点单位进行调查,以取得统计数据的一种非全面调查方法。所谓重点单位是指在总体中标志值占绝对比重的那一部分单位,它可以是一些企业、一些行业、一些城市或一些地区。确定重点单位的关键在于确定所研究现象的标志总量及其比重,如为了掌握“三废”排放情况,就可选择冶金、电力、化工、石油、轻工和纺织等重点行业的工业进行调查,再如为了掌握铁路物流运输和客运情况,可以选择大秦铁路、铁龙物流等企业进行调查,因为这些行业、这些企业在各自领域占了较大比重。

由于重点单位在全体调查对象中只占较小部分,调查的标志量在总体中却占较大的比重,因而对这部分重点单位进行调查所取得的统计数据能够反映社会经济现象发展变化的基本趋势。与抽样调查不同的是,重点调查取得的数据只能反映总体的基本发展趋势,不能用以推断总体,因而也只是一种补充性的调查方法。目前主要是在一些企业集团的调查中运用。重点调查的优点是花费力量较小,能及时提供必要的资料,便于各级管理部门掌握基本情况,采取措施。如我国1979年大中型企业环境保护基本情况调查和1985年全国工业污染源调查就是重点调查。

4) 典型调查

典型调查是根据调查目的和要求,在对调查对象进行初步分析的基础上,有意识地选取少数具有代表性的典型单位进行深入细致的调查研究,借以认识同类事物的发展变化规律及其本质的一种非全面调查。使用典型调查时须注意所选的对象要具有代表性,能够集中、有力地体现问题和情况的主要方面,同时注意点与面的结合,切忌“一刀切”,并将定性分析与定量分析相结合,以提高分析的科学性和准确性。

典型调查要求调查者搜集大量的第一手资料,搞清所调查的典型中各方面的情况,并做系统细致的解剖,从中得出用以指导工作的结论和办法。它适用于调查样本较大、总体同质性较高,且调查者对总体情况比较了解,能准确地选择有代表性对象的情况。实施典型调查的主要步骤是:①“解剖麻雀”,根据研究目的,通过多种途径了解研究对象的总体情

况；②“划类选典”，从总体中初选出备选单位，加以比较，慎重选出有较大代表性的典型；③“抓两头”，进行调查搜集资料并分析研究资料，得出结论。

典型调查与重点调查的主要区别表现在两个方面：首先是选择对象的标准不同。典型单位是在对总体情况分析的基础上有意识地抽选出来的，重点单位是根据其标志值是否占全部总体单位标志总量的绝大比重这一标准来确定的，这一标准是客观存在的，是不容挑选的。其次是调查目的不同。典型调查需要说明的是总体的发展变化规律，而重点调查则是为了了解总体的一般情况和发展水平。

5) 抽样调查

抽样调查是指按照随机原则从调查总体中抽取一部分单位作为样本进行调查，根据调查所取得的有关数字特征推断总体相应的数字特征的一种调查方式。抽样调查是市场调查中最常用的调查方式，它可按抽样方式分为简单随机抽样、分层抽样、整群抽样、等距抽样等方式，关于抽样调查的基本理论，本书第4章将会予以详细介绍。

上述五种基本调查方式按调查范围分为全面调查和非全面调查；按调查登记时间的连续性分为经常性调查和一次性调查；按组织形式分为统计报表制度和专门调查。它们之间的关系如图2-1所示。

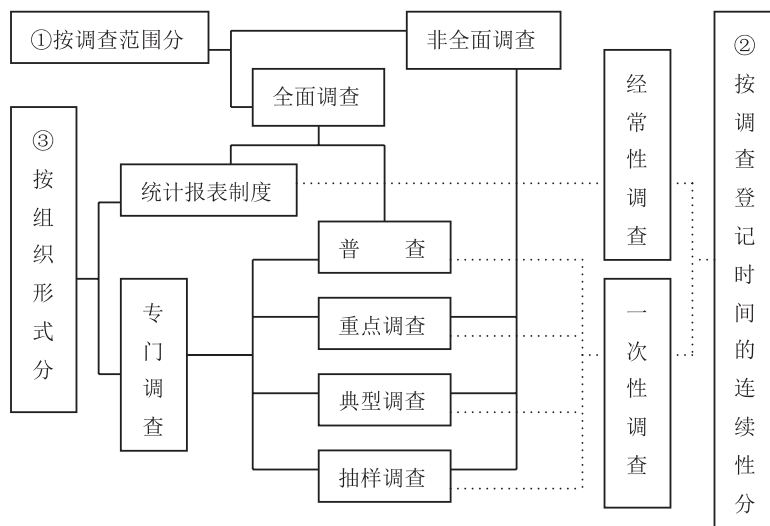


图 2-1 统计调查方式示意图

全面调查是指对被研究现象总体的所有个体进行调查的一种调查方式，非全面调查则是指对被研究现象总体的一部分个体进行调查的一种调查方式。例如要了解厦门市居民家庭的居住水平，如果对厦门市居民家庭一一调查则是全面调查，如果只对厦门市集美区和同安区的家庭进行调查则是非全面调查。

经常性调查是指随着研究对象的发展变化，连续不断地进行调查登记。一次性调查是指间隔一段时期进行的调查，它可以定期进行，也可以不定期进行。

专门调查是为研究某些专门问题、由进行调查的单位专门组织的调查，我国的人口、房屋等普查即为我国某一时期的专门调查。

2.1.2 搜集数据的基本方法

统计数据搜集的基本方法主要有直接观察法、采访法、报告法、通信法、专家调查法和实验设计等方法。

1. 直接观察法

直接观察法是指由调查人员到现场对调查对象直接进行观察和计量,以取得原始资料的一种调查方法。在观察过程中,调查人员所处的地位是被动的,也就是说调查人员对所观察的事件或行为不加以控制或干涉。例如,在进行商场调查时,调研人员并不访问任何人,只是观察现场的基本情况,然后记录备案。一般调研的内容有某段时间的客流量、顾客在各柜台的停留时间、各组的销售状况、顾客的基本特征及售货员的服务态度等。

在特定条件下,当被调查者难以配合调查研究的实施时,应采取直接观察法。如被调查者面对某些敏感性、私密性的问题而不好回答或没有足够的时间来回答提问时,就应通过直接观察法直接观察他们的行为表现,记录他们的行为方式或行为时间,以此来达到研究的目的。在直接观察法下,观察结果容易受观察者个人的态度、观念和周边具体环境的影响,即使观察者有着较高的专业水平。因此,观察者除了凭借自己的注意力和记忆力之外,还要借助先进的记录工具(如录音、录像、照相等器材)完整地记录下被调查者复杂的行为活动。

2. 采访法

采访法是指由调查人员携带调查表向被调查者逐项询问,将答案填入表内的一种调查方法。采访法根据填写的方式分为口头询问法和自填法两种,口头询问法是由调查人员对被调查者逐一采访并记录;自填法是由调查人员把调查表交给被调查者并说明填写方法和要求,被调查者填写好调查表后由调查人员审核收回的调查方法。

在一对一的采访中,调查人员与被调查者可以直接交谈,双方可以进行当面提问和答询。在集体采访中,调查人员与被调查者可以相互讨论,相互补充。因此采访法具有所获取的数据信息相对准确和丰富的优点,但是采访法同时也具有耗费过大的缺点,在调查经费紧张时不宜采用这种方法。

3. 报告法

报告法是指由报告单位根据原始记录和核算资料,按照统计调查机关颁发的统一表格和要求,按一定的报送程序向统计调查机关提供统计资料的方法。

我国现在各企业、各机关向上级部门填报统计报表就是报告法,这种调查是各地区、各部门、各单位按照有关法规的规定,必须对国家履行的一种义务。报告法的特点是有统一项目、统一表格、统一要求和统一上报程序,其资料来源于原始记录,可以同时进行大量的调查。在报告系统健全、原始记录和核算工作完整的情况下,报告法可以保证所提供资料的准确性和及时性。

4. 通信法

通信法是指调查人员向被调查者邮寄或传真调查资料或调查表格,被调查者按要求做出回答并将结果邮寄或传真给调查人员的方法。

通信法的特点是以自愿提供资料为前提,没有强制性,虽然具有节省人力、物力、财力

和时间的优点，但是通信法可能回收率比较低，同时所回收的数据资料可能具有较低的可信度，影响调查质量。因此，通信法一般在对被调查者个人信息比较了解时采用。

5. 专家调查法

专家调查法依据系统的程序，采用匿名发表意见的方式，即专家之间不得互相讨论、不发生横向联系，而只能与调查人员发生关系，通过多轮次调查专家对问卷所提问题的看法，经过反复征询、归纳、修改，最后汇总成专家基本一致的看法，作为调查的结果。

专家调查法具有广泛的代表性，调查结果也较为可靠。如某代理商采用专家调查法调查某饮品的销售量，该代理商首先选择若干专卖店经理、营养师、顾客、销售代表和海外公司经理组成专家小组。然后将该饮品和一些相应的背景材料发给各位专家，要求大家给出该饮品的最低销售量、最可能销售量和最高销售量，并说明自己作出判断的主要理由。

代理商将专家们的意见收集起来后，进行归纳整理再返回给各位专家，然后要求专家们参考他人的意见对自己的判断重新考虑。专家们完成第一次判断并得到第一次判断的汇总结果以后，除销售代表 Peter 外，其他专家在第二次判断中都做了不同程度的修正。如是重复进行前面的步骤，在第三次判断中，大多数专家又一次修改了自己的看法。第四次判断时，所有专家都不再修改自己的意见。因此，专家意见收集过程在第四次以后停止。最终调查结果最低销售量 100 万罐，最高销售量 175 万罐，最可能销售量 150 万罐。

6. 实验设计

实验设计是一种通用的科学合理地安排实验和分析实验数据的方法。将一组随机抽取的实验对象随机分配到两种或多种处理组，观察比较不同处理的效应，这种研究称为实验研究。

实验设计是为实验研究做的周密计划，它可以用于收集测试某一新产品、新工艺、新方法或新方案使用效果的数据资料。实验设计能有效地控制误差，节省人、财、物力，提高效率，但须遵循四个基本原则：①对照原则，即实验要设立对照，使得除实验因素外，对照组与实验组其余因素保持一致；②重复原则，即研究对象要有一定的数量，或者说样本含量应足够。根据每个具体研究，可有不同的方法进行样本含量估计；③随机原则，即应保证每个实验对象都有同等机会进入实验或接受某种处理；④均衡原则，即各处理组非实验因素的条件基本一致，以消除其影响。

随着现代信息技术的发展，计算机、网络、光电技术、卫星遥感、地理信息系统等高新技术已经或正在被广泛地引入统计调查领域中，统计调查人员应根据调查目的与调查对象的具体特点，选择合适、恰当的调查方法。

2.1.3 统计调查方案设计

统计调查是一项理论性与实践性非常强的系统性工作，大型调查通常需要多人协同工作，工作过程中必须统一认识、统一内容、统一方法、统一步调。因此，事先制订一个科学周密的统计调查方案是顺利完成调查任务的必要条件。统计调查方案是指关于统计调查的工作计划，它是调查工作的指导性文件。

1. 统计调查方案的基本内容

一项完整的调查方案应该确定进行统计调查的目的、明确调查对象和调查单位、确定调查项目和选择调查方式方法,并规定调查地点、时间及调查的具体措施等。换言之,完整的统计调查方案应该包括以下六个方面的基本内容。

1) 统计调查的目的

统计调查方案应该首先明确统计调查的目的,同时调查目的要符合客观实际,因为它是统计调查行动的指南。调查目的不明确,就无法确定调查对象、调查范围、调查内容和调查方法,其结果必然使调查混乱,甚至导致调查中断,从而贻误工作。因此,明确统计调查的目的是统计调查方案应首先解决的问题。

2) 调查对象、调查单位和报告单位

明确调查对象和调查单位是统计调查方案的又一基本内容。调查对象是需要调查的现象的总体。调查单位是调查对象中的总体单位,它们是集合与元素的关系。报告单位是负责上报调查资料的单位。

调查单位和报告单位有时是同一事物,有时又不是同一事物。例如在班级同学本周末是否愿意去厦门南普陀寺游玩的调查中,全班同学这个整体便是调查对象,班上的每一位同学个人就是调查单位,同时也是报告单位,在这里调查单位和报告单位是相同的。再如,在我班同学拥有的电脑情况调查中,全班所有同学的电脑这个整体便是调查对象,班上每一位同学的电脑这个个体就是调查单位,每一位同学本人是报告单位。在这里,调查单位和报告单位是不同的。

3) 调查项目和调查表

调查项目是调查中所要登记的调查单位的特征,即调查中规定的调查单位的标志,它由一系列品质标志和数量标志构成。例如在我班同学电脑拥有情况调查中,电脑的品牌、种类配置等就是调查项目。

调查表是将所有调查项目按一定顺序排列所形成的表格,它可以分为单一表和一览表两种形式。只能登记一个调查单位的内容的调查表是单一表,可以同时填写多个调查单位的内容的调查表是一览表。单一表便于整理分类,适用于调查项目较多的调查;一览表简洁明了,适用于调查项目不多的调查。统计调查中采用哪一种调查表由调查目的、调查任务确定。

4) 调查地点、调查时间和调查期限

调查地点是指调查对象所在的地点,它是统计资料所属的空间范围,明确调查地点可以防止统计资料发生错漏。调查时间是指调查资料所属的时间,即指调查资料所反映的现象客观存在的时间。调查期限是指调查工作的起讫期限,包括搜集资料和报送资料的整个工作所需要的时间。

例如,为了分析海西经济区的经济发展状况,在2021年3月1日—3月10日对海西经济区2010—2020年的GDP等经济发展状况指标进行了调查。那么在这项调查中,海西经济区是调查地点,2010—2020年是调查时间,2021年3月1日—3月10日是调查期限。

5) 统计调查的方式和方法

统计调查的方式和方法有很多种,各种方法各有其优缺点,分别适用于不同的条件。某次统计调查采用何种调查方式和调查方法,应当根据统计调查的目的并结合各种调查方式和

调查方法的优缺点综合考虑。

在一次统计调查中可以单独运用某种调查方式和调查方法，也可以综合用多种调查方式和调查方法，但是所运用的调查方式和调查方法应该都在统计调查方案中加以规定。

6) 调查组织实施计划

调查组织实施计划是指确保实施调查的具体工作计划。在调查组织实施计划中，应该明确规定本次调查的组织领导机构，具体地规定宣传教育方式、人员培训内容、文件印刷、经费的来源与开支办法等。此外，调查组织实施计划中还应该明确调查资料的报送程序与报送方式、调查结果的公布时间和公布渠道等内容。

2. 统计调查问卷的设计

调查问卷是调查人员向被调查者发出的调查提纲或调查表，设计科学的调查问卷是顺利取得规范的数据资料的基本保证。调查问卷按回答问题的形式可以分为“开放式”和“封闭式”问卷，但一般都有开头、甄别、主体和背景等部分。

开头包括问候语、填表说明和问卷编号等，一般用来说明调查的目的、内容、填表须知和问卷识别等。开头部分力求文字简洁、准确，语气谦虚、诚恳。例如，某卫视新媒体组织的第一次休闲网游有奖问卷调查的开头部分如下。

您好！首先感谢您参与此次有奖问卷调查活动！

2010年初我卫视新媒体即将推出一款融合了时尚、交友、竞技等众多热门元素的休闲网游《网球宝贝》。为了能够进一步了解市场，让游戏更贴近玩家的实际需求，为玩家提供更为优质的服务，特通过此次有奖问卷调查活动来收集广大玩家《网球宝贝》的反馈信息。一等奖一名，奖品为Turbospeed的加速软件；二等奖一名，奖品为蓝牙鼠标；三等奖五名，奖品为PinStripe 4GB的U盘。活动截至2010年2月15日，请您留下有效地址和联系方式，以便联系获奖者和邮寄奖品。

甄别部分主要是通过一些问题排除不符合条件的被调查者，其目的是确保被调查者能够作为调查项目的有效代表，使之符合调查研究的需要。主体部分包括了要调查的全部问题和问题的答案，是问卷的核心。背景部分包括被调查者的性别、民族、职业、收入等主要个人特征或者是单位名称、地址、负责人、主管部门等单位主要情况。

乔治·盖洛普曾言，没有什么比对问题的选择、措辞还要困难和重要。设计一份良好的问卷应该注意以下几个问题。

1) 提问内容要具体、尽可能的短

问题要确切，避免专业术语和模棱两可，力求明确。例如，“您对您家庭目前的住房状况是否满意”就过于笼统，无法达到预期效果，如果改为“您对您家庭目前的住房面积是否满意”和“您对您家庭目前的住房设计布局是否满意”两个问题就显得具体，能够达到调查的要求。

2) 一个问题只包含一项内容

如果包含的内容过多，就会令人无从答起，也会给统计处理带来困难，影响调查效果。例如，“你喜欢唱歌跳舞吗”这个问题就包含了两个方面的内容。正确的问题设计应该是将其分解为“你喜欢唱歌吗”与“你喜欢跳舞吗”两个问题。

3) 避免诱导性问题

诱导性问题通常会使被调查者不敢表达真实的想法,引出与事实相反的结论。例如,在“艾森豪威尔将军说,陆军部和海军部应当合并为统一的作战部,你同意吗”这一问题中,同意的比例为49%,而将其修改为“陆军部和海军部应当合并为统一的作战部,你同意吗”后,结果同意的比例为29%。因为,问话中去掉了“艾森豪威尔将军说”这一暗示语。因此,问卷设计应坚持客观的态度,避免诱导性。

4) 要注意设计问题的开放式回答

有些问题的答案不能一一列举完毕,为了统计分析的需要,应设有开放式回答。例如,“以下哪个网球休闲游戏的广告语最能引起你的注意?①我向网,我追球;②手指激情;③一网情深,球起心动;④其他”。

5) 问题应按照逻辑类型和难易程度排序

引导性的问题宜放在前,敏感性的问题宜放在后,以便引起被访者的兴趣和使其免处于守势地位,并有循序渐进的感觉。

2.2 数据整理

2.2.1 统计整理的意义和内容

统计整理是根据统计研究的目的和要求,对调查得到的原始记录进行分组和汇总,使其条理化、系统化的工作过程。统计整理在统计工作中具有十分重要的意义。统计调查所取得的原始资料是反映各个总体单位的资料,它们仅说明各个总体单位的具体情况,是不系统的、分散的,甚至还可能带有一定的片面性。统计需要的是反映总体特征的统计指标,因此需要进行统计整理。

统计整理介于统计调查和统计分析之间,是统计工作的第二阶段。统计整理在统计工作中起到承上启下的作用,既是统计调查的继续,又是统计分析的基础和前提,它实现了从个别单位的标志值向说明总体数量特征的指标值过渡,是人们对社会经济现象从感性认识上升到理性认识的过渡阶段,是进一步进行统计分析的必要前提。可见,统计整理绝不是一个单纯的技术问题,而是统计工作中一个极其重要的理论问题。其主要内容通常包括三个方面,即设计整理方案、对统计资料进行汇总并计算各项指标和通过统计表描述整理的结果。

2.2.2 统计整理的程序和方法

统计整理的主要内容决定了统计整理的全过程,统计整理的全过程体现了统计整理的主要内容。

1. 统计整理的程序

统计整理的全过程大体上可以分为以下五个步骤。

1) 设计统计整理方案

统计整理方案与调查方案应紧密衔接,其指标体系要与调查项目一致,或者是其中的一部分,绝不能矛盾、脱节或超越调查项目的范围。整理方案是否科学,对于统计整理乃至统计分析的质量都是至关重要的。

2) 审核、修订调查资料

在汇总前,要对调查得来的原始资料进行审核,审核它们是否准确、及时、完整,发现问题,加以纠正。统计资料的审核也包括对整理后次级资料的审核。

3) 进行科学的统计分组

用一定的组织形式和方法,对原始资料进行科学的分组,是统计整理的前提和基础。在统计整理中,根据统计研究的目的和要求,抓住最基本的、最能说明问题本质特征的统计指标进行统计分组,并根据分析的需要确定分组的具体形式。

4) 统计汇总

对分组后的资料,进行汇总和必要的计算,使得反映总体单位特征的资料转化为反映总体数量特征的资料。

5) 编制统计表

统计表是统计资料整理的结果,也是表达统计资料的重要形式之一。根据研究的目的可编制出各种统计表。

2. 统计整理的方法

统计整理的方法有统计分组、统计汇总和编制统计表。统计分组是根据任务的要求,对调查所得的原始资料,确定要进行哪些分组或分类,在分组的基础上确定应该汇总得到哪些统计指标。

统计汇总是继统计分组后的一个重要步骤,它是指将多单位的各种标志值相加进行汇总。统计汇总技术主要有手工汇总和电子计算机汇总两种。手工汇总常用的方法主要有划记法、过录法、折叠法和卡片法。电子计算机汇总大致需经过编制程序、编码、数据录入、数据编辑和计算与制表等步骤。

2.2.3 统计分组

统计分组是指根据统计研究的目的与要求以及研究现象的内在特点,将统计总体按照某一个或某几个标志划分为若干个性质不同又有联系的部分。统计分组时必须同时满足穷尽和互斥的原则。

所谓穷尽是指每一个总体单位都有组可归。例如将一群小学生按性别分为男生和女生两组,没有一名小学生既不归属于男生组又不归属于女生组的情况,这就是穷尽原则。所谓互斥是指每一个总体单位只能归属于某一个分组。例如,上例中,不能存在一名学生既归属于男生组又归属于女生组的情况,这就是互斥原则。在连续式组距数列中,用“下限记入原则”来保证互斥原则。

1. 统计分组的分类

分组标志是指作为分组依据的标志。进行统计分组的关键在于选择分组标志，而选择分组标志的关键在于所选标志要服从研究任务的需要，能够反映总体的本质特征。统计分组的分类方法有很多，其中之一就是按分组标志进行分类。

1) 按照分组标志的多少分类

将统计分组按照分组标志的多少划分，可以分为简单分组和复合分组。其中，只有一个分组标志的是简单分组，如表 2-1 所示。复合分组是先按某个分组标志分组，再在此基础上按另外一个分组标志将每一个分组进行进一步细分的统计分组，如表 2-2 所示。表 2-2 为在已按身高分组的基础上再将每一组按性别分组的复合分组。复合分组的分组标志可以是两个或两个以上。

表 2-1 某车间零件加工数量统计表

零件数 / 件	工人数 / 人
20~29	21
30~39	53
40~49	49
50~59	27
合 计	150

表 2-2 初三（1）班学生身高统计表

身高 /cm	学生人数		
	男	女	小计
140 以下	2	1	3
140~160	15	8	23
160~180	9	13	22
180 以上	1	1	2
合 计	27	23	50

2) 按照分组标志的类型分类

将统计分组按照分组标志的类型划分，可以分为品质分组和数量分组。如果分组标志是品质标志，那么统计分组就是品质分组；如果分组标志是数量标志，那么统计分组就是数量分组。表 2-1 中的分组标志是零件加工数量，因为零件加工数量是数量标志，所以表 2-1 是数量分组；表 2-3 中的分组标志是服务态度，因为服务态度是品质标志，所以表 2-3 是品质分组。

表 2-3 某银行窗口服务态度统计

服务态度	服务人次
很满意	11
满意	27
不满意	9
很不满意	3
合计	50

3) 按照分组的任务和作用分类

将统计分组按其任务和作用的不同，可分为类型分组、结构分组和分析分组。进行这些分组的目的是，分别是划分社会经济类型、研究同类总体的结构和分析被研究现象总体诸标志之间的联系与依存关系。

类型分组和结构分组的界限比较难区分，也没有截然的界限，通常两者结合使用。一

般认为，现象总体按主要的品质标志分组，多属于类型分组，如银行服务窗口按服务态度分组，如表 2-3 所示；按数量标志分组多是结构分组，进行结构分组的现象总体相对来说同类型性较强，如车间工人按零件加工数量分组，如表 2-1 所示。

分析分组是为研究现象总体诸标志依存关系的分组，如表 2-4 所示。分析分组具有易与类型分组、结构分组相区别的明显特征。分析分组的分组标志称为原因标志，与原因标志对应的标志称为结果标志。原因标志既可以是数量标志，也可以是品质标志，但结果标志一定是数量标志，而且要求计算其相对数或平均数。在表 2-4 中，企业类型是原因标志，员工总数、工资总额和平均工资等都是结果标志。

表 2-4 某地区企业情况统计表

企业类型	员工总数 / 万人 (1)	比重 / % (2)	工资总额 / 亿元 (3)	比重 / % (4)	平均工资 / (元 / 人) (5) = (3) / (1)
国有企业	10 765.9	73.4	7 211.0	76.68	6 698
集体企业	2 817.0	19.20	1 253.4	13.33	4 449
股份制企业	460.1	3.14	350.9	3.73	7 627
外资企业	565.2	3.85	546.8	5.81	9 674
其他企业	60.1	0.41	42.2	0.45	7 022
合计	14 668.3	100	9 404.3	100	6 411

2. 统计分组的方法

一般而言，统计分组的方法主要有单项式分组和组距式分组两种，它们各自有不同的适用场景。

1) 单项式分组

单项式分组是指将分组标志的每一个取值作为一组的分组方法，表 2-3 和表 2-4 都是单项式分组。

单项式分组操作简便迅捷，分组标志有几个取值就分几组，不需要考虑分多少组等问题。但是，单项式分组通常聚合度不够，不能很好地体现组内的共性，难以较好地显现总体的分布规律，因此仅适用于分组标志的取值较少的场合。一般而言，品质分组主要采用这种分组方法，对于分组标志为取值不多的离散变量时，根据需要也可以采用单项式分组，但更多的是采用组距式分组。

2) 组距式分组

组距式分组是指将分组标志的若干个取值作为一组的分组方法。所以，在组距式分组中，分组标志在每组的取值表现为一个区间，表 2-1 和表 2-2 都是组距式分组。

在组距式分组中，每组分组标志的取值区间的长度称为该组组距。按每组的组距是否相等可以把组距式分组分为等距组和异距组。如果相等称为等距组，否则称为异距组。采用等距组或异距组，取决于研究对象的性质和特点。标志值的变动比较均匀的采用等距组，标志值的变动幅度较大或很不均匀的采用异距组。例如收入、产品产销量、产值等宜用异距组。

分组标志在每一组的端点标志值称为组限，其中较小的称为下限，较大的称为上限。如果统计分组每一组的上限等于后一组的下限，则该统计分组为连续式分组，如表 2-2 所示；

否则称为间断式分组，如表 2-1 所示。上限和下限都有的组称为闭口组，如表 2-1 所示；缺少下限或者上限的组称为开口组，如表 2-2 所示。

间断式分组一般主要用于离散型变量分组，而在实际工作中，离散型变量也主要采用连续式分组，很少采用间断式分组。至于开口组和闭口组则根据需要确定，两者都有应用。需要注意的是，采用连续式分组时，如果总体单位的取值恰好等于组限，则应遵循下限记入的原则。

3) 组距式分组的基础计算

在组距式分组中，把每组分组的标志的取值区间的中间值称为该组的组中值。组距式分组的基础计算包括每组组距和组中值的计算，由于组距式分组是统计分析中常用的分组方法，因此，组距式分组的基础计算也就称为统计分析指标的计算基础。

通常用 x 、 U 、 L 和 d 分别表示组距式分组的组中值、上限、下限和组距，式 (2-1) 适用于间断式分组的基础计算；式 (2-2) 适用于连续式分组的基础计算；式 (2-3) 适用于开口组的基础计算。

$$d = U - L + 1, x = \frac{U + L}{2} \quad (2-1)$$

$$d = U - L, x = \frac{U + L}{2} \quad (2-2)$$

$$d = \text{相邻组组距}, x = L + \frac{d}{2} = U - \frac{d}{2} \quad (2-3)$$

2.3 频数分布

统计分组只是将总体按照分组标志划分为若干个不同的小组，并没有讲究这些若干个不同的小组之间的排列顺序，因此统计分组还不能很好地体现总体分布的规律。为了更好地研究总体的数量属性，可以引入频数分布。

2.3.1 频数分布的概念

频数分布是指将总体进行统计分组后，按分组标志的取值顺序将各组进行排列，形成总体单位在各组间的排列，又称次数分布，简称分布或数列。分布包括两个基本要素，即分组标志和各组的总体单位数，如图 2-2 所示。

在分布中分组标志又称为变量，用 x 表示。各组的总体单位数是用绝对数表示的，称为频数或次数，用 f 表示；如果以相对数表示的称为频率，它们都可以作为权数。

× × × × × × 统计表		
分组标志 (x)	频数 (f)	频率 ($f/\Sigma f$)
...	...	...
...	×	...
...	...	...
...	×	...
合计	Σf	1

图 2-2 展示了频数分布和频率分布的关系。表格下方有一个示意图，显示了一个包含“合计”字样的框，该框下方有一个指向“频数分布”的箭头，另一个指向“频率分布”的箭头。

图 2-2 分布的基本要素