

高等院校教材

统计学

刘子君 武霞 郭晔 主编



东北大学出版社
Northeastern University Press

统 计 学

主 编 刘子君 武 霞 郭 晔
副主编 包 建 肖 震 王 弢

东北大学出版社

• 沈 阳 •

© 刘子君 武霞 郭晔 2010

图书在版编目 (CIP) 数据

统计学 / 刘子君, 武霞, 郭晔主编. —沈阳: 东北大学出版社, 2010.6
ISBN 978-7-81102-824-9

I. ①统… II. ①刘… ②武… ③郭… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2010) 第 112963 号

出 版 者: 东北大学出版社

地址: 沈阳市和平区文化路 3 号巷 11 号

邮编: 110004

电话: 024—83687331 (市场部) 83680267 (社务室)

传真: 024—83680180 (市场部) 83680265 (社务室)

E-mail: neuph @ neupress.com

http: // www. neupress. com

印 刷 者: 沈阳市北陵印刷厂有限公司

发 行 者: 东北大学出版社

幅面尺寸: 185mm × 260mm

印 张: 12.25

字 数: 314 千字

出版时间: 2010 年 6 月第 1 版

印刷时间: 2010 年 6 月第 1 次印刷

责任编辑: 任彦斌 刘莹

责任校对: 郎坤

封面设计: 唐敏智

责任出版: 杨华宁

ISBN 978-7-81102-824-9

定 价: 35.00 元

前 言

在知识经济时代，数据成为宝贵的经济资源。在国外，银行、保险、证券、科研、教育等行业都广泛采用先进的统计分析技术，从数据中提取有价值的信息和知识，统计方法已经成为理、工、农、医、人文、社科、经管等学科领域科学研究的基本方法。统计知识作为科研的必备工具，成为高等学校各专业的基础性学科，尤其是本科以上层次的经济、管理类专业，更是把统计学作为一门重要的专业基础课，成为这些专业的核心课程之一。

本书以培养应用型人才为目标，参阅了国内外许多优秀的同类教材，既可作为高校经济管理类专业的教材，也可作为其他专业和广大实际工作者的参考用书。本书旨在使读者对统计学这门学科有一个整体、本质的认识，并且反映出统计学中计算机应用的发展，从而将统计学的理论知识和数据分析的实用性在课程中很好地结合起来。

本书共分为十章。第一章统计学总论，介绍了统计学产生发展的基本情况；第二章统计数据的搜集与整理，重点介绍了统计数据的搜集方法和常用的整理方法，即统计分组的方法，同时介绍了统计图和统计表的制作；第三章数据分布特征的描述，主要介绍了数据分布的集中趋势、离散程度、偏态和峰度的测定；第四章概率与概率分布，重点介绍了离散型随机变量的概率分布和连续型随机变量的概率分布，尤其突出了统计中最常用到的正态分布、 t 分布、 F 分布和卡方分布；第五章统计推断，详细介绍区间估计和假设检验的理论与应用；第六章方差分析，主要介绍单因素方差分析和多因素方差分析的理论与应用；第七章相关与回归分析，分别介绍了相关分析和回归分析的基础理论与应用；第八章时间序列分析，重点介绍了时间序列的各种水平分析指标及速度分析指标的计算，以及利用时间序列对现象发展变化趋势进行预测的方法；第九章统计指数，重点介绍了总指数与平均指数的编制与运用，以及统计指数的因素分析；第十章统计决策，重点介绍了各类决策方法及其应用。

本书与经济管理类的基础课程“概率论与数理统计”有机地结合起来，既介绍了通用方法论性质的一般统计理论和方法，又讨论了其在经济管理领域的

应用。主要适合作为经济管理类专业全日制本科学生、大中专院校“统计学”的主讲教材，其内容可使学生具备基本的统计思想，掌握基本的统计方法，培养应用统计方法分析和解决经济管理中实际问题的能力。同时，本教材全程分解了 Excel 和 SPSS 软件，并结合例题，旨在提高学生分析、解决实际问题的能力。本书也可作为统计工作者进行统计实践的参考书。

本书由沈阳航空航天大学经济与管理学院刘子君老师和沈阳航空航天大学北方科技学院武霞老师及沈阳日报报业集团的郭晔同志担任主编。刘子君负责组织第六章至第九章的编写，武霞负责组织第二章至第五章的编写，郭晔负责组织第一章和第十章的编写，本书由包建（沈阳航空航天大学继续教育学院）、肖震（沈阳航空航天大学继续教育学院）、王弢（沈阳市金融学校）等同志担任副主编，负责各章的编写工作。本教材在全体参编人员的共同努力下完成，最后由沈阳航空航天大学刘子君老师定稿。由于编者水平有限、时间仓促，书中难免会有疏漏或错误之处，望各位专家、学者、感兴趣的读者及广大师生批评指正。

编 者
2010 年 4 月

目 录

第一章 统计学总论	1
第一节 统计学的产生与发展.....	1
第二节 统计学的分科.....	3
第三节 统计学与其他学科的关系.....	5
第四节 统计学的基本术语与研究方法.....	5
第二章 统计数据的搜集与整理	9
第一节 统计数据的计量与类型.....	9
第二节 统计数据的搜集	11
第三节 统计数据的整理	15
第四节 统计数据的显示	23
第三章 数据分布特征的描述	32
第一节 数据分布集中趋势的测度	32
第二节 离散程度的描述	40
第三节 分布偏态与峰度的测度	44
第四节 计算工具的使用	46
第四章 概率与概率分布	53
第一节 概率分布	53
第二节 正态分布和正态逼近	58
第三节 大数定律及中心极限定理	60
第五章 统计推断	63
第一节 参数估计的基本理论	63
第二节 总体均值、比例及方差的区间估计	67
第三节 样本容量的确定	73
第四节 假设检验	75
第五节 总体均值的假设检验	78
第六节 总体比率和方差的假设检验	82
第六章 方差分析	88
第一节 单因素方差分析	88

第二节	多因素方差分析	93
第三节	协方差分析	100
第七章	相关与回归分析	103
第一节	相关与回归分析概述	103
第二节	相关系数与相关分析	104
第三节	一元线性回归分析	109
第四节	多元回归与非线性回归	116
第五节	利用软件进行相关与回归分析	120
第八章	时间序列分析	130
第一节	时间序列分析概述	130
第二节	时间序列基本指标	133
第三节	时间序列的测定与预测	138
第四节	利用软件进行时间序列分析	145
第九章	统计指数	150
第一节	统计指数概述	150
第二节	综合指数的编制	152
第三节	平均指数的编制	158
第四节	指数体系与因素分析	162
第十章	统计决策	171
第一节	统计决策概述	171
第二节	风险型决策方法	173
第三节	贝叶斯决策方法	178
第四节	不确定型决策方法	180
附 录		185
参考文献		189

第一章 统计学总论

统计是指根据研究的目的和要求，运用科学的方法，对客观事物或人类实践活动的数据资料进行搜集、整理、分析的过程。统计学则是研究如何对社会总体的数量特征和规律进行描述、推断、认识的一门学科。

第一节 统计学的产生与发展

一、统计与统计学

“统计”一词在各种实践活动和科学研究领域中经常出现，不同的人或在不同的场合，对其解释是有差异的。随着社会经济的发展，“统计”的含义也不断地拓展和延伸。从不同的角度看，对“统计”一词有不同的理解。目前，比较公认的看法是统计有三种含义，即统计活动、统计资料和统计学。

(1) 统计活动

统计活动又称统计实践、统计工作，是指根据统计的目的和要求，运用科学的方法，搜集、整理和分析统计数据，并探索数据的内在数量规律性的活动过程。统计调查、统计整理和统计分析是基本的统计活动，它所提供的统计资料包括原始统计资料、数据整理结果和资料分析结论。进行基本统计活动之前的统计设计和之后的资料保存、利用也属于统计活动范畴。

(2) 统计资料

统计资料或称统计数据，即统计活动过程所获得的各种数字资料和其他资料的总称。表现为各种反映社会经济现象特征的原始记录、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。它是统计活动的成果或产品，反映了事物的数量特征及其变化规律。

(3) 统计学

统计学又称统计理论，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，是经过系统化的知识体系。它以现象总体的数量方面为研究对象，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有着密切的联系。统计活动同统计资料之间是过程同成果之间的关系，统计资料是统计活动的直接成果。就统计活动和统计学之间的关系来说，统计活动属于实践的范畴，统计学属于理论的范畴，统计学是统计活动实践的理论概括和科学总结，它既来源于统计实践，又高于统计实践，并且反过来指导统计实践，统计活动的现代化同统计科学研究的支持是分不开的。

统计活动、统计资料和统计学相互依存、相互联系，共同构成一个整体。这就是通常说的统计。

二、统计学的发展学派

统计实践活动催生了统计学，简要回顾统计学的渊源，有利于把握学科的发展脉络；勾画统计学现在的学科架构，有利于了解本课程所处的学术地位。

统计学始创于 17 世纪中叶至 18 世纪，这是统计学的萌芽时期，也称古典统计学时期。这一时期著名学派有国势学派和政治算术学派。统计学在 18 世纪末至 19 世纪末的一百年中有了巨大的发展，这一时期属于统计学的成长时期，也称近代统计学时期。这一时期主要学派是数理统计学派和社会统计学派。

(1) 国势学派

国势学派也叫记述学派。该学派的代表人物是德国的康令 (H. Conring, 1606—1681) 和阿亨瓦尔 (G. Achenwall, 1719—1772)，他们在大学中开设了新课程——国势学，介绍如何记录国家发展的重要事件。后人把从事这方面研究的德国学者称为国势学派或记述学派。当时记载关于国家组织、人口、军队、领土、居民职业和资源财产等事件主要用的是文字叙述的形式，基本上没有量的描述与分析。

国势学派对统计学的最大贡献是它提出了一个世界公认的名词“统计学”。另外，该学派在研究各国的重要事件时，采用了系统对比的方法来反映各国实力的强弱，统计学分析方法中“对比”的思想来源于此。国势学的缺欠是缺乏数量分析的方法和结论，与现代统计学相去甚远，所以人们将其称为有名无实的统计学。

(2) 政治算术学派

同是研究各国的国情、国力，英国的学者们却采用了数量分析的方法，代表人物是威廉·配第 (William Petty, 1623—1687) 和约翰·格朗特 (J. Graunt, 1620—1674)。配第在其著作《政治算术》中，对当时的英国、荷兰、法国之间的财富进行数量上的计算和比较，开启用数量方法研究社会经济现象之先河。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”。该批学者也因此获得“政治算术学派”之名。格朗特通过对伦敦市五十多年的人口出生和死亡资料的计算，写出了第一本关于人口统计的著作《对死亡率公报的自然观察和政治观察》，统计的含义也因此从记述转变为专指用数量来说明国家的重要事件。

政治算术学派在统计发展的历史上有着重要的地位，它以数量分析为特征，研究客观现象的数量方面，就这点来说，政治算术应该是统计学的起源，威廉·配第被认为是统计学的创始人。由于未使用“统计学”这一科学命名，所以有人评价该学派是“有统计之实，无统计之名”的学派。

国势学派和政治算术学派共存了近二百年，两派长期争论不休，但二者却有一个共同特点——都是以宏观社会经济问题为研究对象，故统称为社会经济统计学派。

(3) 数理统计学派

最初把古典概率论引进统计学的是法国数学家、统计学家拉普拉斯 (P. S. Laplace, 1749—1827)，他阐明了统计学的大数法则，进行了大样本推断的尝试。最终完成统计学和概率论结合的则是比利时统计学家、数学家凯特勒 (Adolphe Quetelet, 1796—1874)，1841 年他出任比利时中央统计委员会会长，1851 年积极筹备国际统计学会组织，并任第一届国际统计会议主席；他在著作《社会物理学》中，利用大数法则，论证了社会生活中的随机偶然现象贯穿着必然的规律性；他运用概率论原理，提出了著名的“平均人”的概念，计算人

类自身各性质标志的平均值,通过“平均人”来探索社会规律,他认为,社会所有的人与“平均人”的差距越小,则社会矛盾就越缓和。凯特勒被统计学界称为“国际统计会议之父”“近代统计学之父”,其贡献就是发现了大量现象的统计规律性和开创性地应用了许多统计方法,为数理统计学的发展奠定了基础。

该学派人才济济,英国的戈赛特(William Sealy Gosset, 1876—1937)用笔名 Student 发表了关于 t 分布的论文,建立了“小样本理论”;英国的费希尔(Ronald Aylmer Fisher)给出了 F 统计量、极大似然估计、方差分析等;内曼(J. Neyman)和皮尔逊(Egon S. Pearson)提出了置信区间估计和假设检验;沃尔德(A. Wald)发明了序贯抽样和统计决策函数。特别要提出的是英国数理统计学家卡尔·皮尔逊(Karl Pearson, 1857—1936),他设计的“直线相关系数”是最常用的相关系数,从其构思过程中,可以体会到“对称美”;他构建的“ χ^2 检验”是最常用的检验方法,从其公式形式 $\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$ 中,可以体会到“简单美”。这再一次验证了“统计学是一门艺术”的说法,也难怪有许多学者把卡尔·皮尔逊尊称为“统计学之父”。

(4) 社会统计学派

社会统计学派发源于德国,主要代表人物有克尼斯(K. G. A. Knies)、恩格尔(C. L. E. Engel)和梅尔(G. V. Mayer),他们认为,统计学的研究对象是社会现象,目的在于明确社会现象的内在联系和相互之间的关系,在研究过程中,要采用全面调查,也可以适量地使用抽样调查。恩格尔在 1895 年发表的《比利时工人家庭的生活费》一文中,提出了著名的“恩格尔法则”,从中引申出的“恩格尔系数”作为衡量生活水平的标准,至今仍然被沿用。

第二节 统计学的分科

统计学的内容十分丰富,研究和应用的领域非常广泛。从统计教育的角度,统计学大致有以下两种分类。

一、描述统计和推断统计

描述统计是用图形、表格和概括性的数字对数据进行描述的统计方法。推断统计是根据样本信息对总体进行估计、假设检验、预测或其他推断的统计方法。统计学分为描述统计和推断统计,一方面反映了统计发展的前后两个阶段,另一方面反映了统计方法研究和探索客观事物内在数量规律性的先后两个过程。

统计研究过程的起点是数据,终点是探索到客观事物总体内在的数量规律性。要达到统计研究的目的,搜集到的是总体数据,那么经过描述统计之后,就可以达到探索内在的数量规律性的目的;但如果所获得的数据只是研究总体的一部分数据,那么要探索到总体的数量规律性就必须应用概率论的理论,并根据样本整理出的信息对总体作出科学的推断。显然,描述统计是整个统计学的基础和统计研究工作的第一步。它包括对客观现象的度量,调查方案的设计,科学、及时、快速、经济地搜集与整理数据,用图表显示数据,分析和提取数据中的有用信息,最终能否科学、准确地探索到总体内在数量规律性,与选用何种统计量、选用什么推断方法、如何进行推断有着直接的联系。一个出色的统计工作者的能力和技巧在推

断统计中将得到充分的体现与检验。但如果没有描述统计搜集可靠的数据,并提供有效的样本信息,即使再高明的统计学家和科学的推断方法,也难于得出准确的结论。因而,推断统计对描述统计又有很强的依赖性。

尽管描述统计可以在获得总体数据时,直接探索出总体数量规律性,但这种情况在实际工作中很少见到。自然现象的总体多数是无限的,例如,统计物理研究中要弄清楚分子运动的规律,而分子又是无穷多的,不可能全部观察和试验。社会经济现象的总体虽然多数是有限的,但要考虑获得数据及推断总体的时效性、经济性和准确性,抽样调查往往比普查更有效,因而应用也就更普遍。例如,全国的人口数量和变化、耕地面积、企业个数和经营情况等,虽然可以通过普查得到全部数据,但普查要投入大量的人力、财力和物力,而且要很长的时间才能搜集、整理出所要的数据,不是每年都能做到的,因而我国确定每10年进行一次人口普查和经济普查,其他各年均以抽样调查数据进行推断。此外,大量的管理和研究工作不可能组织普查,例如,城市居民家庭每月的收入支出调查、某种商品的市场调查、某个事件的民意测验等,都只能通过抽样调查的方法,然后对总体数量规律性进行科学的推断。因而,在描述统计中,搜集、整理和分析的多是样本数据。这样,科学地整理样本数据,显示样本数据的特征和规律,提取样本数据中的有用信息等就显得格外重要。

二、理论统计和应用统计

理论统计是指统计学的数学原理。由于现代统计科学用到了几乎所有方面的数学知识,要成为优秀的统计工作者,就必须经过严格的数学训练,特别是从事统计理论和方法研究的人,必须有很好的数学基础。从广义来讲,统计学是应该包括概率论的,因为概率论是统计推断的数学基础,而概率论是数学的一个分支,所以理论统计应该是包括概率论在内的对统计方法数学原理的研究。

在统计工作者中,从事理论统计研究的人只是很少的一部分,大部分是应用统计方法去解决实际问题的应用统计工作者。统计学是一门数据科学。由于在自然科学、社会科学的所有研究和实际工作中,都要通过数据来分析和解决问题,统计方法的应用就自然而然地扩展到几乎所有的研究领域。

由此可见,现代统计学可以分为两大类:一类是以抽象的数量为研究对象,研究一般的搜集数据、整理数据和分析数据方法的理论统计学;另一类是以各个不同领域的具体数量为研究对象的应用统计学。

理论统计学把研究对象一般化、抽象化,以数学中的概率论为基础,从纯理论的角度,对统计方法加以推导论证,其中心内容是以归纳方法研究随机变量的一般规律。理论统计学的特点是计量不计质,它具有通用方法论的理学性质。

应用统计学是有具体对象的方法论。所谓应用,既包括一般统计方法的应用,更包括各自领域实质性科学理论的应用。应用统计学从所研究的领域或专门问题出发,视研究对象的性质,采用适当的指标体系和统计方法,解决所需研究的问题。应用统计学不仅要进行定量分析,还需要进行定性分析。所以应用统计学通常具有边缘交叉和复合型学科的性质。

理论统计学和应用统计学总是互相促进,共同提高的。理论统计的研究为应用统计提供方法论基础,应用统计学在对统计方法的实际应用中,又常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。

第三节 统计学与其他学科的关系

一、统计学和数学的关系

数学是与统计学关系非常密切的一门学科。数学与统计学都是研究数量规律的，都要利用各种公式进行运算。数学中的概率论为统计学提供了数量分析的理论基础。统计学中的理论统计学以抽象的数量为研究对象，其大部分内容也可以看做数学的分支。

但同时两者之间存在着本质的区别。从研究对象看，数学以最一般的形式研究数量的联系和空间形式，统计学特别是应用统计学，则总是与客观的对象联系在一起的；从研究方法看，数学主要采用逻辑推理和演绎论证的方法，而统计本质上是归纳的方法。统计学家特别是应用统计学家需要深入实际，进行调查或实验去取得数据，研究时不仅要运用统计的方法，而且要掌握某一专门领域的知识。

二、统计学和其他实质性学科的关系

统计方法是适用于所有学科领域的通用数据分析方法，只要有数据的地方，就会用到统计方法。随着人们对定量研究的日益重视，统计方法已经被应用到自然科学和社会科学的众多领域，统计学也已发展成为由若干分支学科组成的学科体系。可以说，几乎所有的研究领域，比如政府部门、学术研究领域、日常生活中、公司或企业的生产经营管理中，都要用到统计方法。

利用统计方法可以简化繁杂的数据，比如，用图表展示数据，建立数据模型。有人认为，统计的全部目的就是让人看懂数据，其实这仅仅是统计的一个方面，统计更重要的功能是对数据进行分析，它提供了一套分析数据的方法和工具。通常是实质性的学科提出了问题，统计学便提出相应的方法，两者之间联系密切，但同时两者之间也有着明显的区别。实质性学科研究该领域现象的本质关系，并对有关规律作出合理的解释和论证，这一过程中虽然统计学提供了分析的方法和工具，却并不直接对规律产生的原因和机理作出进一步的分析。例如，利用统计方法对吸烟和不吸烟者患肺癌的数据进行分析，可以得出吸烟是导致肺癌的重要原因的结论，而吸烟为什么会患癌症，仅仅依靠统计学是无法说明的，而必须由医学作出解释。

第四节 统计学的基本术语与研究方法

一、统计学的基本术语

在统计学的学习过程中，要用到一些专门的术语，熟悉这些基本术语是统计学学习的基本前提。

1. 总体、总体单位与样本

(1) 统计总体

统计总体简称总体，就是统计所要研究的事物或现象的全体，即由客观存在的，具有某种共同特征的许多个别事物构成的整体。例如，如果要研究某高等学校全体学生的英语成绩

时,各院系的所有学生构成统计总体,各院系每个学生均在校注册这一共同特征成为构成这个总体的前提条件。

(2) 总体单位

总体单位是构成统计总体的个别事物。在上例中,该校的每一名学生便是总体单位。对于不同的研究对象,总体单位既可能是人、物,也可能是企业、机构,甚至可能是时间、地域。

(3) 样本

样本是指从统计总体中抽取出来作为代表这一总体的、由部分个体组成的集合体。抽取样本的目的是用来推断总体,这就必然要求样本能够代表总体。样本代表总体的程度越高,由样本计算的指标与总体指标的误差就越小。因此总希望样本具有较高的代表性。遵循随机原则的抽样,能够排除主观因素的影响,保证取样的客观性;采用非随机原则的抽样,有时会更快捷、更经济,只是抽出的样本无法计算误差。

构成样本的个体数目称为样本容量。通常用小写英文字母 n 表示,相对于 N 而言, n 一般只是一个很小的数。比如想了解流水线上产品的合格率,会随机地抽选 100 个产品进行检验,这 100 个产品就构成了一个样本,样本容量是 100。会根据对这 100 个产品检验的结果,计算出样本对应的合格率,并用它来代表总体的合格率。

在实际工作中,统计总体称为全及总体,由于样本是从总体抽取出来并代表总体的,所以全及总体又称为母体,样本则称为子体。

2. 统计标志、统计指标与指标体系

(1) 标志

标志是指个体具有的属性或特征,它是说明个体的属性或特征的具体名称,但对应于某个标志,各个个体的具体表现(即标志值)是可以不同的。正是由于各个个体就某些标志而言具有相同的标志值,才构成了统计总体;也正是由于各个总体单位就某些标志而言具有不相同的标志值,所以才有了统计的必要。

标志按其性质划分可分为品质标志和数量标志。品质标志表明个体属性方面的特征,品质标志的标志表现只能用文字说明,不能用数量表示,即得到分类数据和顺序数据;数量标志表明个体数量特征,其标志表现只能用数字表现,即得到数值型数据。可变的数量标志称为变量,其值称为变量值。

标志按照其变异情况分为不变标志和可变标志。不变标志是指某一标志对所有个体而言,具体表现都是相同的,可变标志又称变异标志,是指某一标志在各个个体之间的具体表现不尽相同。

有时会按照某一个品质标志,将总体划分为具有某一特征的个体的集合和不具有某一特征的个体的集合两类。如在全部产品中,分为合格品与不合格品;在全部人口中,分为男性和女性。这种用“是”“否”或“有”“无”来表示的标志,叫做是非标志,也叫交替标志或 0、1 标志。

(2) 统计指标

统计指标是反映统计总体数量特征的概念和数值。统计指标由两项基本要素构成,如 2008 年我国国内生产总值为 20.93 万亿元。即指标的概念(名称),指标概念为国内生产总值和指标的取值为 20.93 万亿元。

统计指标可以分为数量指标和质量指标。反映现象总规模、总水平的统计指标称为数量

指标，如人口总数、企业总数、商品进出口总额等，这些指标反映现象或过程的总规模 and 水平，所以也称为总量指标，用绝对数来表示；凡是反映现象相对水平和工作质量的统计指标称为质量指标，如职工平均工资、人口密度、工人出勤率等；质量指标是总量指标的派生指标，用相对数或平均数来表示，以反映现象之间的内在联系和对比关系。

(3) 指标体系

统计指标体系是由一系列相互联系的统计指标所组成的有机整体，用以反映所研究现象各方面相互依存、相互制约的关系，如工业企业统计指标体系。

(4) 标志和指标的关系

标志反映总体单位的属性和特征，而指标则反映总体的数量特征。标志和指标的关系是个别和整体的关系。需要通过对各单位标志的具体表现进行汇总和计算，才能得到相应的指标。

总体和总体单位的概念会随着研究目的不同而变化，因此指标与标志的概念也是相对而言的。例如，所要研究的是全国工业企业的情况，则各企业的职工人数、固定资产、工业增加值等都是总体单位（即各个企业）的标志；如果研究目的变成研究某一企业的职工状况，那么该企业变成一个总体，企业职工人数变成了统计指标，每个职工的文化程度、技术等级、性别等就成为标志。

3. 总体参数和样本统计量

(1) 总体参数

总体参数简称参数，是想要了解的总体的某个特征值。通常所关心的参数有总体平均数、标准差、总体比例等。在统计中，总体参数通常用希腊字母表示，如总体平均数用 μ 表示，总体标准差用 σ 表示，总体比例用 π 表示。总体参数是一个未知的常数，是希望知道的，如不知道整个国家的收入差异，不知道流水线上的产品合格率，但它们都是想要得到的数据，为得到这些数据可以采用全面调查，即对这个国家的每个人进行调查、对流水线上的每一个产品进行检验，但如果某个产品质量检验是破坏性的，那么需要采用抽样调查，根据从抽取的样本中所获得的资料来推断总体参数。

(2) 样本统计量

样本统计量简称统计量，是根据样本数据计算出来的一个量。通常所关心的样本统计量有样本平均数、样本标准差、样本比例等。样本统计量通常用英文字母来表示，比如，样本平均数用 $\bar{x}$ （读做 x -bar）表示，样本标准差用 s 表示，样本比例用 p 表示。由于样本是从总体中抽取的，所以统计量总是可以计算获得的。抽样的目的就是要用样本统计量来计算总体参数。

二、统计学研究方法

在统计理论的研究与实践过程中，主要应用到的统计学研究方法有以下五大类。

(1) 大量观察法

大量观察法是对所研究现象总体的全部或足够多的单位进行调查并加以综合研究的方法。

(2) 试验设计法

试验设计法是通过设计实验取得所要研究数据的方法。试验设计要遵循的原则：重复性原则，即在相同条件下重复多次试验；随机化原则，即在实验中对实验对象的分配和试验次

序是随机安排的；区组化原则；尽量扩大组内标志值差异，缩小组间差异。

(3) 统计描述法

统计描述法是用综合指标、统计表、统计图等形式描述研究总体现象的数量特征的方法。包括统计分组法和综合指标法。

(4) 统计推断法

统计推断法是指在一定的置信标准要求下，由样本信息推断总体数量特征的归纳推理方法。包括参数估计、假设检验、方差分析等。

(5) 统计模型法

统计模型法是根据统计资料，运用统计方法，对研究现象的结构或过程建立一种统计表达式，进行有关分析的方法。

本章要点

本章主要介绍了有关统计学的初始思想、发展历程、学科体系、统计研究的特点和方法等。通过本章的学习，要求学生在了解统计学的发展历史的基础上，理解统计的三种含义——统计活动、统计资料、统计学；知晓各统计学派的代表人物和主要思想；熟悉统计学的学科体系，掌握统计学的基本术语，它们包括总体、总体单位与样本，统计标志、统计指标与指标体系，总体参数和样本统计量，要求能够理解并掌握它们之间的关系与区别。

思考题

- 1.1 “统计”一词有哪几种含义？它们之间的关系如何？
- 1.2 简述统计总体与指标、个体与标志之间的关系和区别。
- 1.3 简述统计学的研究方法。
- 1.4 统计学有哪些分科？它们之间的关系是什么？
- 1.5 统计总体与总体单位的区别与联系是什么？
- 1.6 请举例说明指标与标志之间的区别与联系。
- 1.7 统计学与数学的区别与联系是什么？
- 1.8 统计学发展中有哪些主要学派？其学术观点是什么？例举主要代表人物及其对统计学的贡献。
- 1.9 请举出统计应用的几个例子。
- 1.10 请举出应用统计的几个领域。

第二章 统计数据的搜集与整理

第一节 统计数据的计量与类型

统计数据的搜集和整理是统计研究的初级阶段，这个阶段的基本任务是搜集所需要的统计资料并进行加工整理，为下一步的定量分析提供必要的前提和条件。

一、统计数据的计量

统计数据是对客观现象进行计量的结果。在搜集数据之前，需要对现象进行计量和测度。根据对研究对象计量的不同精确程度，人们将计量尺度由低到高、由粗略到精确分为四个层次，即定类尺度、定序尺度、定距尺度和定比尺度。

(1) 定类尺度

定类尺度也称分类尺度、列名尺度等，是这样一种品质标志：按照它，可以对研究客体进行平行的分类或分组，使同类同质、异类异质。例如，按照性别，将人分为男、女两类；按照经济性质，将企业分为国有、集体、私营、混合制企业等。定类尺度是最粗略、计量层次最低的计量尺度，利用它只可测度事物之间的类别差，而不能了解各类之间的其他差别。定类尺度计量的结果表现为某种类别，但为了便于统计处理，如为了便于计算和识别，也可用不同数字或编码表示不同类别。比如用 1 表示男，0 表示女等。但这些数字只是不同类别的代码，决不意味着它区分了大小，更不能进行任何数学运算。定类尺度能对事物做最基本的测度，是其他计量尺度的基础。

(2) 定序尺度

定序尺度也称序数尺度、顺位尺度等，是这样一种品质标志：利用它，不仅能将事物分成不同的类别，还可确定这些类别的等级差别或序列差别。例如“产品等级”可将产品分为一等品、二等品、三等品、次品等。定序尺度对事物的计量要比定类尺度精确一些，但它至多测度了类别之间的顺序，而未测量出类别之间的准确差值。因此，定序尺度的计量结果只能比较大小，不能进行加、减、乘、除等数学运算。

(3) 定距尺度

定距尺度也称间隔尺度、等距尺度、区间尺度等，是能测度事物类别或次序之间间距的数量标志，更具体地说，定距尺度是可将事物区分为不同类别，对这些类别进行排序，并较准确地度量类别之间数量差距的一种计量尺度。该尺度通常使用自然或物理单位作为度量单位，如收入用人民币“元”度量；长度用“米”度量等。定距尺度的计量结果表现为数值。定距尺度的数值可做加、减法运算，例如，一个地区的温度 20°C 与另一个地区的 25°C 相差 5°C ，等等；但不能做乘、除法运算，而且定距尺度没有绝对的零点。

(4) 定比尺度

定比尺度也称比率尺度，定比尺度的计量结果也表示为数值，定比尺度这种数量标志不仅能测度各类别的大小和多少，而且有一个绝对零点作为起点。这个绝对零点是它与定距尺

度的明显差别，就是说，定距尺度中没有绝对零点，即使其计量值为“0”，这个“0”也是有客观内容的数值，即“0”水平，而不表示“没有”或“不存在”。例如，一个地区的温度为 0°C ，这表示一种温度的水平，并不是说没有温度。而定比尺度中绝对零点的“0”表示“没有”或“不存在”。例如，一个人的收入为“0”，表示这个人没有收入等。现实中，大多数场合人们使用的都是定比尺度。

定比尺度与上述三种计量尺度相比，还有一个特性，就是可以计算数值之间的比值。例如，一个人的月工资收入为600元，另一个人的为300元，可以得出前一个人的收入是后一个的2倍。但定距尺度由于不存在绝对零点，就只能比较数值差，而不能计算比值。比如，可以说 30°C 与 15°C 之差为 15°C ，而不能说 30°C 比 15°C 热1倍。可见，定比尺度可以做加、减、乘、除法运算。尽管如此，定比尺度和定距尺度还属于同一层次，在实际应用中，通常对两者可不作区分，而统称为定量尺度。

上述四种计量尺度对事物的计量层次是由低级到高级、由粗略到精确，逐步递进的。高层次的计量尺度可以计量低层次计量尺度能够计量的事物，但不能反过来。显然，可以很容易地将高层次计量尺度的计量结果转化为低层次计量尺度的计量结果，将考试成绩的百分制转化为五等级分制就是一例。

二、数据的类型

(1) 按照计量尺度划分

按照所采用的计量尺度不同，可以将统计数据分为定类数据、定序数据、定距数据和定比数据。统计数据是采用某些计量尺度对事物进行计量的结果，但采用不同的计量尺度，会得到不同类型的统计数据。就上述四种计量尺度计量的结果来看，可以大体上将统计数据分为定性的数据和定量的数据两种类型。

定性数据也称品质数据，说明事物的品质特征表现的具体类别，不能用数值表示；因为这类数据由定类尺度和定序尺度计量形成，所以又可细分为分类数据和顺序数据。

定量数据也称数量数据或数值型数据，是说明现象数量特征表现的，能够甚至必须用数值来表现；因为这类数据由定距尺度和定比尺度计量形成，所以又可细分为区间数据和比例数据。

对不同类型的数据，可以采用不同的统计方法来处理和分析，比如，对定性数据一般只采用分组法计算，分析各组的频数或频率；而对定量数据则可以采用更多的统计方法去处理，计算、分析更多的统计指标或统计量。

(2) 按照数据的搜集方法划分

按照数据的搜集方法划分，可以将统计数据分为观测数据和实验数据。观测数据是通过调查或观测而搜集到的数据，这类数据是在没有对事物人为控制的条件下而得到的，社会经济现象的统计数据几乎都是观测数据。试验数据是在实验中控制实验对象而搜集到的数据。如医药研究试验数据、动植物杂交品种试验数据等。自然科学领域的大多数数据都是试验数据。

(3) 按照数据的时间关系划分

按照被描述对象与时间的关系，可以将统计数据分为截面数据和时间数据。截面数据是指同一时间不同空间上的数据。时间数据是指同一空间不同时间上的数据。