

统计学原理

主编 周荣辅 刘新建 于 俊

北京工业大学出版社

统 计 学 原 理

主 编 周荣辅 刘新建 于 俊
副主编 张文虎 马菊红 乔 彪

北京工业大学出版社

内 容 提 要

本书是高等学校经管类专业学生使用的教材。本书共分 10 章。内容包括绪论、统计设计与统计调查、统计整理、统计分析指标、时间数列分析、指数分析、抽样调查与统计推断、相关关系分析、多元统计分析及国民经济核算体系等。本书还可用作相关领域的培训教材。

图书在版编目 (CIP) 数据

统计学原理 / 周荣辅, 刘新建, 于俊主编. —北京: 北京工业大学出版社, 2003.9

ISBN 7-5639-1321-1

I. 统… II. ①周…②刘…③于… III. 统计学 - 高等学校 - 教材 IV. C8

中国版本图书馆 CIP 数据核字 (2003) 第 073843 号

统 计 学 原 理

主 编 周荣辅 刘新建 于 俊

※

北京工业大学出版社出版发行

邮编: 100022 电话: (010) 67392308

各地新华书店经销

徐水宏远印刷厂印刷

※

2003 年 9 月第 1 版 2003 年 9 月第 1 次印刷

850mm × 1168mm 32 开本 12.25 印张 307 千字

印数: 0001 ~ 5000 册

ISBN 7-5639-1321-1/G·702

定价: 22.00 元

前 言

统计学是一门非常古老的学科，可以说，不亚于任何一门自然科学。如果说，数学开始于原始人计数的话，那么，同时统计学也就开始了。作为一门独立的专门知识，最晚也从 17 世纪中后期就开始了。到今天，统计学作为一门学科的基本知识内容已经相当成熟。然而，关于统计学的一些基本问题的争论直到今天还在进行，其中争论比较多的一个问题就是描述统计学和推断统计学的名分地位，我国在这方面形成了拉锯战。1949 年以前的大学教材中以数理统计学为主体，1949 年以后，我国主要向前苏联学习，以社会经济统计学为主体。而在最新的发展中，数理统计学派似乎又占了上风，如教育部推荐的面向 21 世纪课程教材系列中，袁卫主编的《统计学》又把数理统计学的相关内容作为主要部分，对描述统计学做了大量删减。

我们编写这本教材主要基于以下理由：(1) 描述统计学是统计学不可或缺的基本组成部分，不应削弱；(2) 经管类大学本科生一般在第二学期就开设统计学原理课程，而作为高等数学组成部分的“概率论与数理统计”或同时开课，或推得更后，所以，在这门课程中不应过多涉及讲授概率论知识；(3) 有些术语在各种书中还有一些乱，需要予以澄清，并且应该照顾到不同学科和课程的统一性，如“集中趋势”、“离中趋势”、变量等；(4) 作为具有现实应用意义的学科，教材中的举例和习题材料应不断更新，与时俱进；(5) 多元统计分析已得到了广泛应用，应该给管理类学生有所介绍。

此外，我们在书中也采取了一些有特色的观点，比如：对统计学的研究对象采用了新的表述；认为，许多统计指标（均值、方差、相关系数等）和概念（如统计分布）不仅适用于随机样本，也适用于一般的总体；对综合评价的有关内容，采用了一些比较新的观点。在编写中我们坚持了系统完整、概念准确、平实易懂、材料更新的原则，努力提高编著质量。当然，由于时间比较紧，也留有一些遗憾，如对于一些典型统计软件使用的介绍未能涉及，有待再版时补充。

本书按照 54 至 60 课时的教学需要编写，在使用中，各学校和专业可以根据各自的学生情况和课程分工选用增删有关内容。比如，课时少或要求低的专业，可以略去第九章；单开国民经济统计学或国民经济核算课的专业可以略去第十章。

本书由周荣辅、刘新建、于俊担任主编；张文虎、马菊红、乔彪担任副主编。参加本书编写的有（以姓氏笔画为序）：于俊、卫爱华、马菊红、邢爱华、乔彪、刘新建、张巍、张文虎、周荣辅、赵玉忠、唐凯、常爱菊。刘新建、周荣辅负责全书的统稿、修改和定稿工作，赵玉忠参与了修改和校对工作。希望我们这本书的出版，能为统计学教材的发展带来一些新气象，真诚期望广大专家、学者和学生不吝批评指正。

编者

2003 年 7 月

目 录

前 言	(1)
第一章 绪论	(1)
第一节 统计学的学科性质	(1)
第二节 统计学的产生与发展	(3)
第三节 统计学的基本范畴	(8)
第四节 统计工作与统计组织	(10)
第二章 统计设计与统计调查	(15)
第一节 统计设计的概念、种类及其内容	(15)
第二节 统计指标体系设计	(23)
第三节 统计调查的概念与种类	(31)
第四节 统计调查方案设计	(34)
第五节 统计调查方法概述	(35)
第三章 统计整理	(48)
第一节 统计整理的一般问题	(48)
第二节 统计分组	(51)
第三节 统计汇总	(57)
第四节 统计分布	(61)
第五节 统计表	(73)

第四章 统计分析指标	(83)
第一节 统计指标的分类	(84)
第二节 总量指标和相对指标	(85)
第三节 平均指标	(106)
第四节 变异指标	(125)
第五章 时间数列分析	(140)
第一节 时间数列的概念和种类	(140)
第二节 时间数列的水平指标	(145)
第三节 时间数列的速度指标	(152)
第四节 长期趋势分析	(160)
第五节 季节变动分析	(170)
第六章 指数分析	(184)
第一节 统计指数的概念与种类	(184)
第二节 综合指数	(188)
第三节 平均数指数	(194)
第四节 指数体系及因素分析	(199)
第五节 指数数列	(213)
第七章 抽样调查与统计推断	(224)
第一节 统计推断的特点与相关概念	(224)
第二节 推断分布的概率论基本知识	(227)
第三节 抽样分布与抽样误差	(238)
第四节 参数估计初步	(246)
第五节 假设检验初步	(256)
第六节 抽样调查的组织形式	(264)

第八章 相关关系分析	(283)
第一节 相关关系的概念及种类	(283)
第二节 相关性分析	(286)
第三节 一元回归分析	(290)
第四节 多元线性回归初步	(297)
第五节 自相关分析	(302)
第九章 多元统计分析初步	(308)
第一节 多元统计分析的基本概念与数据预处理	(309)
第二节 聚类分析	(312)
第三节 主成分分析	(322)
第四节 因子分析	(336)
第五节 综合评价	(341)
第十章 国民经济核算体系	(358)
第一节 国民经济核算概述	(358)
第二节 国民经济核算体系的内容和方法	(361)
第三节 国民经济核算的主要分类	(363)
第四节 国民经济统计的常用指标	(372)

第一章 绪 论

第一节 统计学的学科性质

统计学是关于客观事实数据采集和处理分析的一门方法论学科。在我们的日常社会生活中，到处都有统计的身影。描述说明我们国家的经济、社会、国防等方面的发展状况离不开统计；了解企业的经营状况、采取经营决策离不开统计；进行科学研究也离不开统计；在现代日益丰富多彩的社会生活中，要执掌好一个家庭，做好收入开支计划也离不开统计。统计对现代社会中的每一个人都十分重要。

在日常生活、企业与公共管理及科学研究的交流语言中，“统计”一词在不同的场合有不同的含义，通常讲有三种：统计工作、统计资料和统计学。

统计资料是关于客观事实的系统的数字记录。也许你经常不断地在你的笔记本上记一些数字，但如果这些数字没有按一定的规则排列，也就是没有系统化，那么，它们就还不是统计资料。比如，你某日在某页上写着：买白菜 3 斤，1.5 元；跑步 20 分钟，逛街 2 小时，体重 65 公斤。这些数字只有当你为了某种目的，把它们分门别类地组织起来，才成为有意义的统计资料。

统计工作是采集获取统计资料以及对统计资料进行统计整理和分析的实践活动。统计工作作为一种社会实践活动，存在于所有的组织管理活动和大量的科学研究之中。统计工作的最基础活

动是采集数据，然后是整理数据，最后是分析数据。统计工作的规律和方法就是统计学的基本内容。

统计学是关于统计工作的规律和方法的知识体系。作为一门学科，统计学的学科性质和研究对象在其发展史中一直存在着争论。作为一门入门教程，这里不准备讨论这些争论，只是把我们的观点叙述出来，其他某些观点在脚注中给出，有兴趣者可以去找原文读。

首先，我们认为，统计学是一门方法论技术科学，因而它的研究对象是统计工作的方法技术或更确切地说，是统计分析（结果）质量与统计方法的矛盾体^①。毛泽东在《矛盾论》中曾指出：“科学研究的区分，就是根据科学对象所具有的特殊的矛盾性。因此，对于某一现象的领域所特有的某一种矛盾的研究，就构成某一门科学的对象。”^② 所以，我们要确认一门学科的研究对象，就是要找出该学科所研究的事物的特殊矛盾。对于一般方法论学科讲，其研究对象就是对应的实践活动的效果和采用的方法的矛盾，围绕这一核心矛盾就展开了一系列研究内容。如，统计工作规律的研究是为了更好地设计和应用统计方法；概率理论的应用，是为了使在有限样本下对统计总体的特性推断的可靠性有更确切的把握，并提供随机现象规律的量的表示；统计指标设计方法的研究是为了更好地、更科学地表现所研究事物的本质特征，等等。关于一般统计学原理应用于各种领域的管理和科学研究的方法论内容，就形成了各专门领域的统计学，如社会经济统计学、生物统计学、教育统计学等。

在统计学与实体性科学的关系上，各专门领域统计学具有双

① 有人认为：“统计学是研究现象总体数量方面的方法论的科学”，参见下页注解中郑德如主编的《统计学》，第7页。

② 《毛泽东选集》第1卷，第309页，北京：人民出版社1991年版。

重身份，它们既是各专门领域的学科组成部分，如教育统计学属于教育科学、生物统计学属于生物科学，同时也是统计科学这一大学科的组成部分。一般统计学原理本身属于统计科学这一领域的基础学科。

其次，我们要注意区分统计学的研究对象和统计（工作）研究对象。统计学的研究对象如上所述，是指一门科学的研究对象，而统计研究对象则是指在一次实体性研究（包括管理研究和科学研究）活动中的实体对象，这一对象我们称之为总体（详见本章第三节）。如果我们正在进行中国儿童的健康水平研究，那么，这里的研究对象就是全体中国儿童这个总体；如果研究的课题是“中国儿童健康水平的统计方法研究”，那么这里的研究对象就是统计学的研究对象，只是更具体地属于卫生统计学的研究。

第二节 统计学的产生与发展

一、古代统计实践活动的发展^①

统计学作为一个学科的发展是与统计实践活动的发展紧密联系的。

统计活动是人类进行社会生产和管理活动的重要环节，有了人类的社会存在，就有了统计活动。所以，自从人类结绳记事开始就有了统计。由于收取赋税的需要，所以较早开展的国家统计活动是人口统计和土地统计。中国从公元前 1 000 多年前的夏朝开始就有人口、土地等方面的记载，如当时人口约 1 355 万，土

^① 本节有关内容可以参见郑德如主编：《统计学》（修订本）第 1~7 页，上海：立信会计出版社 1998 年版；于涛主编：《社会经济统计学原理》第 18~26 页，武汉：武汉大学出版社 1992 年版。

地 2 438 万顷。在古希腊、罗马等国也有人口、财产、世袭领地等的统计。到了封建社会时期, 这些国家统计制度进一步完善、加强和扩充。近代统计制度的形成是与资本主义社会的产生同步的。由于社会经济的快速发展, 对各种社会资源数据和市场数据的需要迅速增加, 所以, 人口、工业、农业、商业、海关、外贸、物价等方面的统计先后形成规模, 进而统计组织作为政府的一个职能部门建立起来。

二、统计学的产生和发展

虽然统计活动的出现非常久远, 零散的统计知识也有产生, 但是, 系统的统计理论即统计学的产生则仅有三百余年的历史。这个历史被划分为三个时期: 古典统计学、近代统计学和现代统计学。

1. 古典统计学 (17 世纪中末叶至 18 世纪中末叶)

这个时期主要诞生了两个统计学流派: 记述学派和政治算术学派。

记述学派又称国势学派, 主要以文字形式定性记录一个国家各方面的国情国势基本特征, 数字的作用尚不突出。该学派的主要代表人物是德国的康令 (H. Conling, 1606—1681 年) 和阿亨瓦尔 (G. Achenwall, 1719—1772 年)。

康令从 1660 年 11 月 20 日起, 在不伦瑞克城西尔姆斯特 (Hermsteda) 大学以“国势学”为题, 讲授政治活动家应具备的知识, 主要是记述和比较各国间的领土、人口、财政、军事、政治和法律等方面的情况, 从研究目的、研究对象和研究方法方面初步形成了社会经济统计的体系。其主要后继者之一施梅泽尔 (Schmetzer, 1679—1747 年) 曾于 1723—1731 年在耶拿大学、1741—1746 年在哈利大学讲授《政治统计学讲义》, 但施梅泽尔在统计学理论上没有多少贡献。

施梅泽尔的学生阿亨瓦尔对国势学派的理论发展做出了重要贡献。阿亨瓦尔当时在哥丁根大学开设了“国家学”课程，1749年发表《欧洲各国国势学概论》。在其著作的序言中，阿亨瓦尔提出了“统计学（Statistik）”一词，并搜集了大量实际资料，分门别类地记述有关国情国力的系统知识。

政治算术学派的创始人是英国经济学家威廉·配第（W. Petty, 1623—1687年）和约翰·格朗特（J. Graunt, 1620—1674年）。配第的代表性著作是《政治算术》（1690年出版），该书运用大量的实际统计数据，对英国、法国、荷兰三国的国情国力作了系统的数量对比分析，阐明了英国的地位 and 发展的方向道路。格朗特的代表性著作是《对死亡率公报的自然观察和政治观察》，他收集整理了1603年以来公布的死亡表，加以观察比较，论证了出生、死亡、死因、男女性比率等人口动态存在一定的规律，并用新颖的方法编制出死亡率表，以死亡率表为依据进行了人口统计的推算。

与国势学相比较，政治算术更接近于现代的国民经济统计或统计核算，格朗特的“死亡表”接近于现代推断统计的萌芽。恩格斯对配第在其著作中阐明的统计学研究目的、研究任务、研究对象和研究方法给予了高度评价。在批判杜林对《政治算术》的态度时，恩格斯说：“杜林先生……对配第创造《政治算术》，即一般所说的统计，也采取同样的态度。他对于配第所用方法的奇特，只是恶意地耸耸肩膀，如果我们想到，现在的统计同配第向它概要地提出的目的还相距很远，那么，在两百年以后这种自以为高明的骄傲就只是表现无法粉饰的愚蠢。”

不过，无论是国势学派还是政治算术学派，他们的著作和教学内容还属于实质性社会科学。

2. 近代统计学时期（18世纪末至19世纪末）

近代统计学包括数理统计学和社会统计学两大分支。这两个

分支既是统计学原理的两个相对独立的内容体系，也是关于统计学的学科性质相互争论的两个学派。近代统计学也可以称为现代统计学的奠基时期，因为其内容体系和发展线索形成了现代统计学的基础。

(1) 数理统计学派

18世纪中末叶以后，概率论的发展和数学方法在社会经济统计方面的运用，促进了概率论和政治算术的结合，形成了数理统计学派。

比利时的生物学家、数学家和统计学家阿道夫·凯特勒（A. Quetelet, 1796—1874年）被认为是数理统计学的创始人和奠基人，其主要著作有《比利时王国统计研究》（1829年）、《犯罪倾向》（1831年）、《论人类及其能力之发展或社会物理学论》（1835年）、《关于概率论书简》（1846年）、《社会制度》（1848年）、《社会物理学》（1869年）、《人体测定学》（1871年）。他把法国的古典概率引入社会经济统计，用这种方法既研究自然，又研究社会，对生物学和经济学的发展起过一定的推动作用。在凯特勒工作的基础上，后经英国的高尔登（Goldon, 1822—1911年）和皮尔逊（Pearson, 1857—1936年）等人的发展，形成了现代数理统计学中的许多重要内容，并形成了统计学研究的数理统计学派。凯特勒认为统计学是一门独立的方法论科学。

(2) 社会统计学派

19世纪后半叶，在德国兴起了社会统计学派，由克尼斯（K.G.A. Klies, 1821—1898年）首创，主要代表人物为恩格尔（C. L. E. Engel, 1821—1896年）和梅尔（G.V. Mayr, 1841—1925年）。他们认为，统计学所研究的是社会总体而不是个别的社会现象。由于社会现象的复杂性和总体性，必须通过大量观察和分析，方能反映社会现象的规律。强调研究对象的总体性和采用大量观察法构成了社会统计学派研究方式的两大特点，这也是现代统计

研究的两大特点。但这一派仍把统计学看做是一门实质性科学。

3. 现代统计学时期（自 20 世纪初始）

近代时期以来，统计学一直沿着两个方向发展。

首先，数理统计学在随机抽样的基础上建立起推断统计的理论和方法。英国统计学家 W.S.Gosset (1876—1936 年)^① 被认为是推断统计学的奠基人，他提出了用于小样本理论的 t 分布。该分布的现在形式是由 R.A.Fisher (1890—1962 年) 发展出来的。Gosset 和 Fisher 曾相互通信讨论问题。在 Fisher、J.Neyman (波兰统计学家) (1894—1981 年) 及 E.S.Pearson (K.Pearson 之子) 等人的推动下，建立起了统计假设检验理论。美国统计学家 A.Wald (1902—1950 年) 将统计学中的估计和假设理论予以归纳，提出把统计推断理论称为“统计决策理论”。R.A.Fisher 还是实验设计统计学的开创者，开篇著作是 *The Arrangements of Field Experiments* (*Journal of the Ministry of Agriculture, England*, 1926, 33: 503~513)。1935 年，Fisher 出版了 *The Design of Experiments* (Edinburgh: Oliver and Boyd)，以后经美国的 G.W.Snedecor 及其学生 Certrude Cox 和 W.G.Cochran (1909—1980 年) 进一步发展了该学科。

其次，社会经济统计学在 19 世纪中叶以后也成熟和发展起来。这方面的现代工作包括社会调查、经济计量学和投入产出核算以及更完备的国民经济核算体系。在另一方面，社会经济统计学并不排斥数理统计学，相反，许多人认为，数理统计学方法（样本推断方法）应是社会经济统计学的重要组成部分。所以在现代的许多统计学原理书中都把推断统计学作为重要的组成部分。

^① W.S.Gosset 1876 年生于 Conterbury。在他读完 Winchester 和牛津大学 New 学院的教育以后，于 1899 年进入 Dublin 的一个公司做酿酒师，并进行检验数据解释和大麦实验安排，后来他被送往伦敦大学学院于 1906—1907 年跟随 K.Pearson 研究统计。

一般把统计学原理分为描述统计学和推断统计学两个相对独立的部分。描述统计学的内容主要来源于社会经济统计学,有些术语和方法也借鉴了数理统计学;推断统计学是数理统计学的应用形式,数理统计学本身成为应用数学的一个分支学科。但是关于统计学原理的内容结构问题仍在争论,如最近出版的袁卫主编的《统计学》给予了推断统计学以主体地位。

第三节 统计学的基本范畴

一个学科的基本范畴是其特殊研究对象的反映,这些基本范畴贯穿了其领域中任何一种著作的始终,是它的行话。本节给出目前比较公认的统计学的几个基本范畴。

一、统计总体与总体单位

由具有某种共同特征的事物组成的集合,作为统计研究的工作对象,我们称之为统计总体,简称总体。组成总体集合的那些元素通常被称为总体单位,也称为个体。我们要研究当前某地区一线工人的收入生活现状,这个地区所有的一线工人就是其统计总体,其中的每个工人就是总体单位。要研究某地区小学生的生活学习时间分配状况,该地区所有的小学生就是其统计总体,其中的每个学生就是总体单位。要研究全国农村的耕地利用状况,全国农村的所有耕地就是其统计总体,其中的每一单位耕地就是总体单位。要研究某省 1970 年到 2000 年的 GDP 增长情况,该省 1970 年到 2000 年的 30 年 GDP 数据就是其统计总体,其中每一年的 GDP 数就是总体单位。

总体是统计研究要认识的对象,但认识总体要从调查总体单位的基本数据入手。用个比喻来说,分子是保持物质化学性质的最小单位,是化学研究的对象,而原子是组成物质的更小单元。

总体就是统计研究中的分子，总体单位就是原子。不是任何与统计数字有关的研究都是统计研究。

在统计研究实践中，一个事物的统计角色是相对的，有时是总体，有时是总体单位。一所学校的学生作为一个集合，当研究该校的学生状况时是总体，而当研究一个地区或全国同类学校在校内状况方面的校级水平时，每个学校的学生集合就构成一个总体单位。

二、标志与指标

总体的属性特征是统计研究的具体内容。在统计研究中，总体的属性特征值是由总体单位的属性特征值计算得出的。在统计学中，把总体单位的属性特征统称为标志，作为标志的函数的相应总体的属性特征统称为统计指标，简称指标。

指标在统计学中是一个专门的概念，与我们日常语言或其他学科中的指标意义不同。我们在学统计以前也常说指标，这些指标的含义包括：①某一个任务目标，如生产指标；②任意一个事物的属性，如智商指标、病理指标、检验指标、产品质量指标等。标志通常可分为品质标志和数量标志两类。品质标志值域是总体单位的分类名称集合，具有不同品质标志值的总体单位之间显示在某一方面具有较大的性质差异，如性别、产品等级、学历等。数量标志的值域是实数或其一个子集，总体单位在这种标志上取值表现出某种程度的连续性，如年龄、产量、利润、工时等。

三、样本与统计量

在统计研究中，许多时候我们难以获得完整总体的资料，只有部分总体单位的数据，这时就产生了我们是否可以根据部分推断全体的问题，这就是推断统计学的研究任务。

广义而言，我们可以把总体单位全体的任意一个子集称为该