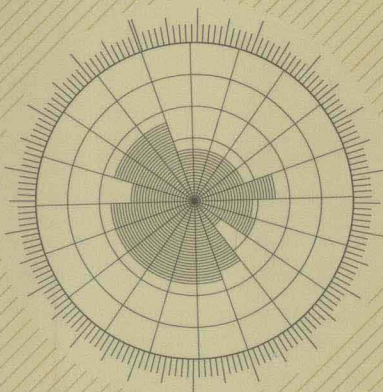


统计学

Statistics

罗良清 魏和清 ○ 主编



中国财政经济出版社

统 计 学

罗良清 主编
魏和清

中国财政经济出版社

图书在版编目 (CIP) 数据

统计学/罗良清, 魏和清主编. —北京: 中国财政经济出版社, 2011. 7
ISBN 978-7-5095-2902-7

I. ①统… II. ①罗…②魏… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2011) 第 095753 号

责任编辑: 樊 闽
封面设计: 汪俊宇

责任校对: 杨瑞琦
版式设计: 兰 波

中国财政经济出版社出版

URL: <http://www.cfeph.cn>

E-mail: cfeph@cfeph.cn

(版权所有 翻印必究)

社址: 北京市海淀区阜成路甲 28 号 邮政编码: 100142

发行处电话: 88190406 财经书店电话: 64033436

涿州市新华印刷有限公司印刷 各地新华书店经销

787 × 1092 毫米 16 开 19 印张 462 000 字

2011 年 7 月第 1 版 2011 年 7 月涿州第 1 次印刷

定价: 35.00 元

ISBN 978-7-5095-2902-7/C · 0019

(图书出现印装问题, 本社负责调换)

本社质量投诉电话: 010-88190744

统计学是一门从事数据采集、整理和分析的方法论科学，其目的是探索数据内在的规律性，以达到对研究总体的科学认识。基于以上认识，本书的编写内容安排如下：

第一章绪论，是对统计学这门学科的高度概述，阐述统计学学科性质、构成及应用情况等，这是我们了解与学习统计的起点。统计研究从采集数据入手，第二章是对统计数据采集、整理方法论的系统介绍。数据本身不是规律，但通过对数据的加工、分析却能显现规律，第三章是常用描述统计数量分析方法介绍。随着时间的推移，所研究事物的数据会不断地波动起伏，如何探索数据在时间上的变动规律，将在第四章时间数列分析中找到满意的答案。在市场经济中，指数是出现频率较高的名词，在新闻媒体中经常听到或看到股票交易价格指数、企业家信心指数和空气质量指数等。第五章介绍了指数的编制原理、方法及应用。要实现对研究总体的认识深化，就必须在描述统计的基础上，通过统计推断的方法探索出总体数量变动的规律性，第六章介绍了抽样和抽样分布原理，它是我们学习推断统计的基础。统计推断从其内容构成来看，包括参数估计与假设检验两部分内容，第七、八章介绍了总体参数估计与假设检验的基本原理与方法。第九章相关与回归分析是对事物之间相互关系、相互依存的一种测度。第十章是统计决策理论的详细介绍。第十一章介绍了统计综合评价与分析的一般理论与方法。第十二章介绍了 EXCEL 软件在常用统计方法中的运用。

在教材写作过程中，我们力求突出以下特色：（1）注重系统性。根据统计教育培养目标，教材内容基本涵盖了我国工商及经济管理类专业统计本科教学所要掌握的知识，克服了国外一些统计教材知识体系松散的不足。（2）强调理论与应用相结合。为了避免国内某些教材理论与实用相互衔接的不足，我们每章以“案例导入”开头，介绍统计方法在现实中的运用，用形象生动的方

式将学生引入统计知识的殿堂。(3) 关注学科发展动向。在教材内容体系设计上体现学科发展的趋势,如通过介绍统计软件在统计中的应用,既提高了学生对统计方法的应用能力,也让学生了解学科的发展趋势。(4) 便利师生使用。本着服务学生、便利教师教学的原则,我们在每章后配置了思考与练习,以便让读者更好地把握学习重点,巩固所学知识。(5) 力求语言形象。在写作风格上我们尽可能少用抽象晦涩的语言,力图简明扼要地阐明统计方法的基本原理和思想,以增强读者学习统计的兴趣。

本教材由江西财经大学金融与统计学院统计学系教师集体合作编写成。它适合统计教学科研人员参考和供经管类各专业从事统计学课程学习的学生使用。

本书写作分工如下:罗良清、魏和清担任主编。各章节编写人员如下:罗良清(第一章)、蔡定萍(第二章)、王拓(第三章)、李海东(第四章)、潘文荣(第五章)、万兆泉(第六章)、李燕辉(第七章)、邓国华(第八章)、魏和清(第九章)、曹俊文(第十章)、詹国华(第十一章)、董成武(第十二章)。

作为长期从事统计理论研究的群体,我们愿在充分借鉴国内外同行相关研究成果的基础之上,以多年在统计园地的潜心研究与积累,为构建新的统计教材的内容体系尽绵薄之力。但由于统计学是一门发展中的学科,对其认识和研究也是一个渐进、不断深化的过程,加之自身水平的限制,书中不足恳请读者批评指正。

编者

2011年4月

第一章 绪 论	(1)
§ 1.1 统计学的性质及其种类	(1)
§ 1.2 统计的应用领域	(4)
§ 1.3 统计学中的几个基本概念	(5)
§ 1.4 统计研究方法	(8)
思考与练习	(10)
第二章 数据的采集、整理和显示	(11)
§ 2.1 数据的采集	(12)
§ 2.2 数据的整理	(17)
§ 2.3 频数分布	(20)
§ 2.4 数据的显示	(25)
思考与练习	(31)
第三章 统计数据的描述	(32)
§ 3.1 集中趋势的测度	(33)
§ 3.2 离散趋势的测度	(44)
§ 3.3 偏度和峰度的测度	(49)
§ 3.4 相对位置的测度及异常值的检测	(52)
思考与练习	(55)
第四章 时间序列分析	(57)
§ 4.1 时间序列概述	(58)
§ 4.2 时间序列的传统分析	(60)
§ 4.3 时间序列的构成因素分析	(68)
§ 4.4 时间序列的趋势变动分析	(69)

2 统计 学

§ 4.5 时间序列的季节变动、循环变动和不规则变动分析	(77)
思考与练习	(81)
第五章 相对指标和指数	(83)
§ 5.1 相对指标概述	(84)
§ 5.2 指数的概念与种类	(88)
§ 5.3 指数的编制及计算	(90)
§ 5.4 指数的分析与运用	(99)
§ 5.5 常用的几种经济指数	(107)
§ 5.6 综合评价指数	(117)
思考与练习	(120)
第六章 抽样和抽样分布	(123)
§ 6.1 抽样方法概述	(124)
§ 6.2 概率分布	(131)
§ 6.3 常见的概率分布	(132)
§ 6.4 大数法则与中心极限定理	(139)
思考与练习	(144)
第七章 总体参数估计	(146)
§ 7.1 参数估计的基本问题	(147)
§ 7.2 单个总体均值和比率的区间估计	(149)
§ 7.3 样本容量的确定	(156)
§ 7.4 两个总体均值和比率差异的区间估计	(160)
§ 7.5 其他抽样组织形式的总体参数区间估计	(165)
思考与练习	(171)
第八章 假设检验	(174)
§ 8.1 假设检验概述	(175)
§ 8.2 单一总体	(179)
§ 8.3 两个总体	(183)
思考与练习	(187)
第九章 相关与回归分析	(188)
§ 9.1 相关分析与回归分析概述	(189)
§ 9.2 一元线性相关和一元线性回归	(194)
§ 9.3 多元线性回归和相关分析	(207)
§ 9.4 非线性回归分析	(212)

思考与练习	(219)
第十章 统计决策	(222)
§ 10.1 统计决策概述	(223)
§ 10.2 确定型统计决策方法	(225)
§ 10.3 完全不确定型决策	(225)
§ 10.4 风险型统计决策	(229)
§ 10.5 贝叶斯决策	(233)
思考与练习	(234)
第十一章 统计综合分析与评价	(236)
§ 11.1 综合评价概述	(237)
§ 11.2 综合评价指标的选择及预处理方法	(238)
§ 11.3 综合评价指标权重的确定方法	(246)
§ 11.4 常用的综合评价方法	(253)
§ 11.5 综合国力评价	(257)
思考与练习	(258)
第十二章 EXCEL 在统计中的应用	(261)
§ 12.1 统计图的生成	(261)
§ 12.2 描述性变量的生成	(264)
§ 12.3 区间估计及假设检验	(268)
§ 12.4 相关与回归分析的应用	(279)
§ 12.5 时间序列指标分析	(281)
思考与练习	(285)
附录 常用统计表	(287)
主要参考文献	(295)

绪论

【案例导入】

在《中国问题报告》国土篇中，有关抗日战争的一组文字如下：“在德国选择了忏悔时，日本人却选择了靖国神社，选择了军备扩张。由于教育等人为误导及官方的错误态度，日本青少年对二战缺乏正确的了解和认识。据《读卖新闻》一次舆论调查显示，在青少年中，将侵华战争看做是侵略战争的人数不到 1/10，对日本的战败，其态度是‘此仇必报’、‘下次一定要打一场胜利的战争’。1994 年 8 月，《读卖新闻》访问包括首相在内的内阁大臣，提问他们对‘大东亚战争’的认识，结果 19 人中，只有 5 人承认有过侵略，而这 5 人中没有一人对侵略中国表示悔罪之意。”

以上所展现的均为统计数据。由此我们清楚地认识到国耻教育的必要性，并了解了日本人对侵华战争错误认识之根本及其顽固性。读完以上内容，人们自然会提出一些问题，如：

- (1) 以上数据是怎样获得的，数据的常用形式有哪些，怎样对数据进行统计处理？
- (2) 什么是统计，统计与数据之间是什么关系？
- (3) 常用的统计方法有哪些？

回答以上问题，必须建立在对统计学基本问题的了解之上。本章主要讲述统计的产生与发展、统计学的学科性质、应用领域及统计学的基本概念及方法等问题。

§ 1.1

统计学的性质及其种类

1.1.1 统计学的性质

统计学是以现象的数量特征为研究对象，利用自身特有的方法，发现现象的规律的一

方法论科学。

现代统计学具有以下几个特点：

1. 数量性。统计学的核心内容就是数据的搜集、整理、分析的原理和方法。所以，统计的语言就是数字，没有数字，不从数量方面来研究问题，就谈不上统计。只要有数据的地方，就会有统计方法，包括政府部门、学术研究领域、日常生活、公司或企业的生产经营管理，甚至法律，都要使用统计方法。统计研究客观事物的数量，如果数字资料准确，统计方法运用得当，同样可以达到正确认识和反映本质的目的。统计研究客观事物的数量，主要包括数量状态、数量关系和数量变化规律。统计学的数量性质能够把它与那些非数量性质的学科，如哲学、经济学、政治学、历史学等区分开来。

2. 总体性。数量有个体数量与总体数量之别，统计学主要研究后者，它对大量同类现象的数量方面进行综合研究。单个数字不能称为统计，也不可能指望从它身上发现什么有价值的东西，只有对大量的现象或对某一现象进行多次重复的观察，才有可能找到统计关系和统计规律。统计虽然研究总体数量，但必须从个体数量的调查入手，遵循由个体数量到总体数量的认识逻辑。

3. 变异性。统计学主要研究不确定性现象。不确定性，是指由于受到偶然性的、随机因素的作用，使得客观事物的实际数量表现存在一定程度的“不可确知性”。正是这种不确定性，决定了统计的变异性。例如，各种股票的价格和成交量每天不同，这才需要对其进行统计，编制股票指数等指标。如果说总体各单位的变异表现出个别现象的特殊性和偶然性，而对现象总体的数量研究，则是通过量观察，从各单位的变异中归纳概括出它们的共同特征，显示出现象的普遍性和必然性。在现代统计中，处理不确定问题是统计研究的主要课题和任务。

4. 归纳推断性。统计方法具有归纳推断的特点。统计主要是通过两条途径来认识总体的：一是对总体中的个体进行逐个的调查，取得全面调查的数据；二是从总体中按照一定的方法抽取部分事物组成调查样本，通过对样本的调查来获得对总体的推断。对于第一种方法，运用算术方法和统计描述手段可以达到；而后者相对复杂，需要运用概率论知识和统计方法。在实践中，全面调查和非全面调查都会用到，但由于全面调查受到诸多因素的制约，从时效性、经济性、实用性和可行性考虑，大多是利用样本资料来对总体进行推断。在这里要明确，我们所说的统计的归纳推断性与逻辑意义上的归纳推断是不同的概念，统计推断只是根据观察到的样本情况，对总体的可能情况作出判断，而逻辑上的归纳推断是从假设、命题出发，按严格的逻辑推理程序进行推断。

5. 广泛性。统计学研究的数量来自社会各个领域的数量方面。它的这一特点，使它与研究某一特定领域的其他社会科学（如政治学、经济学、法学）区别开来。它研究的领域包括整个社会，既研究生产关系，也研究生产力与生产关系以及生产力之间的关系；既研究经济基础，也研究上层建筑以及经济基础和上层建筑之间的关系。除此之外，统计还研究生产、流通、分配、消费等社会再生产的全过程以及社会、政治、经济、军事、法律、文化、教育等领域的数量方面。

1.1.2 统计学的种类

从横向看，可以发现各种统计学都具有以上特点，因而形成了一个学科“家族”。从

纵向看,统计学的方法应用于各种实质性的学科,同它们相结合,产生了一系列的专门领域的统计学(见图1-1)。

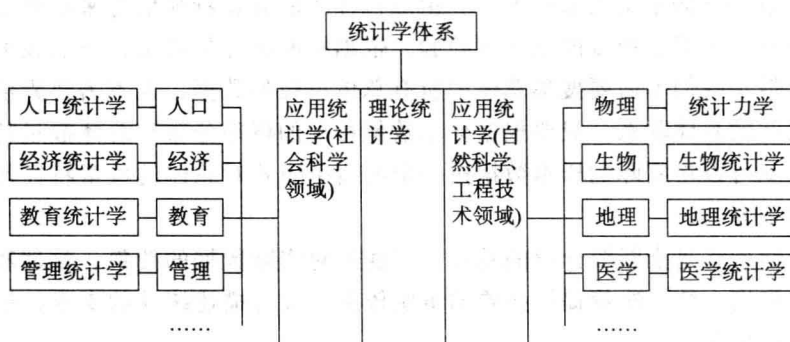


图 1-1 统计学的整体结构

由图1-1我们可以看到,统计学分为两大类:一类是研究抽象的数量的,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;另外一类是以各个不同领域的具有实际意义的数量为研究对象的应用统计学。

理论统计学又被称为数理统计学,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。其中心内容是以归纳方法研究随机变量的一般规律,例如统计分布理论、统计估计与假设检验理论、相关与回归分析、方差分析、时间序列分析、随机过程理论等等。这些方法适用于自然现象和社会现象。

应用统计学则与各不同领域的实质性学科有着非常密切的联系。它从所研究领域专门问题出发,视研究对象的性质,采用适当的统计指标体系和统计方法,以解决所研究的问题,是有具体对象的方法论。所谓应用,不仅包括统计方法的应用,还包括各自领域实质性科学领域的应用。应用统计学不仅要定量上来进行分析,还要定性上来进行分析。它总是先从对现象的分析中获得需要考察的指标,建立指标体系,然后采集数据,对数据进行分析,并结合对现象的定性分析,最后得到符合客观现实的结论,并以此作为决策的依据。所以,应用统计学需要实质性学科的理论作为指导,它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互相促进,共同提高。理论统计学的研究为应用统计学提供了方法论基础,提高了统计分析认识问题、解决问题的能力;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,促进其发展。

就统计方法的特征和功能来看,又可以把统计分为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想象的数据,并通过图表等形式对所搜集到的数据进行加工处理和显示,进行通过综合、概括与分析,得出反映客观想象的规律性数量特征。描述统计学是一切统计活动所运用的基本方法。由上可知,统计活动是要认识事物的总体数量特征和之间的数量关系,但如果仅仅依靠描述统计,无法到达这个目的。例如,我们要了解全国的物价水平,而市场上销售的货物数不胜数,我们只能取得局部的样本单位的数据,总体的物价水平是确定,但商品的价格水平却会因为所选取商品的不同而发生

变化,因此从挑选的样本商品来推断总体的物价水平,就需要依靠建立推断统计学来解决问题。

推断统计学是在概率论的基础上研究由随机样本的数量特征信息来推断总体的数量特征,并作出具有一定可靠程度的估计或检验。依据推断统计的理论,虽然我们不知道总体的数量特征,但是我们不需要搜集总体中所有总体单位的数据,也不需要弄清样本每一单位与总体的之间的具体联系,只要根据样本数量特征的概率分布与总体数量特征之间存在的客观联系,就可以将实际的样本数据按一定的概率模式对总体的数量特征作出符合一定精度的估计和检验。

描述统计学始终是推断统计学的基础,它通过对实际数据的搜集、整理和分析,为统计推断、统计咨询、统计决策提供必要的事实依据。没有描述统计的支撑,最先进的推断方法也将是空中楼阁。

§ 1.2

统计的应用领域

1.2.1 统计的应用领域

统计学的性质决定了其应用领域的广泛性,在各个学术领域都要使用统计。生物学、经济学和心理学三个学科对统计的使用尤其多,以至于它们已经发展了自己的一套统计方法,即生物统计、计量经济和心理测试。还有一大批的历史学家、考古学家、地理学家、语言学家利用统计知识得出各种结论,如由于中世纪大鼠疫而导致的死亡人数。可以看出,几乎所有的经验学术领域都以各种方式依赖于统计。我们仅就统计学在社会科学中的应用作粗略的介绍。

经济统计学:经济管理是统计方法得到较早利用和发展的领域。经济统计的传统内容主要包括国民收入核算、价格指数编制、时间序列的古典分析、经济前景预测等。此外,经济统计学重要的研究课题还有经济结构、经济增长、经济效益、通货膨胀、政策效应等。

管理统计学:就是运用统计分析方法分析和解决企业经营管理活动中遇到的各种需要作出决策的问题。管理统计学研究的问题有市场调查、商情预测、产品试验设计、成本预算、库存管理、排队模拟、工序控制、抽样调查等。

社会统计学:社会学家在研究社会问题时大多要作实地调查,统计学方法在确定样本规模和制订合适的调查方案时有很大的用途,而且统计分析和推断方法对搜集来的资料进行正确使用也具有一定的指导意义。社会分层和流动性、收入与财富分配、贫困化问题、地区差异比较、国家现代化、人文发展水平、社会犯罪现象、劳动与休闲、职业质量等,都是社会统计学的重要研究课题。

人口统计学:原始的统计活动就是对部落、氏族的人口进行计数和清点,所以人口统

计学是最古老的一门统计学科。人口统计的研究课题主要有调查系统、人口寿命表的编制、人口统计核算体系、人口增长与控制、人口与资源、环境、生活质量和社会发展的关系等。

教育统计学：教育是提高居民素质和让社会成员获取一定谋生技能的手段，在不同时期或不同的地区，人们关心教育的侧重点可能差别很大。教育统计学的研究课题有社会受教育的水平、教育资源的分配和教育效益、教育的内容和方法的合理性、继续教育与职业培养、影响学生成绩的因素等。

这里只是挂一漏万地介绍了运用统计的部分领域，希望在后面的统计方法的学习过程中，能够尽量把学到的方法应用到实践中去，进一步体会统计科学的意义。

1.2.2 统计的误用

政治家 Benjamin Disraeli 曾经说过：“有三类谎言：谎言、糟透的谎言和统计。”因为在现实生活中，人们总是有意无意地滥用统计，比如，错误地对统计进行定义、错误的图表展示、选取不合理的样本、数据的遗漏或者逻辑的错误。这些误用有些是常识性的，有些是技术性的，有些则是故意的。统计成了有些人歪曲事实的工具。有些人为了达到自己的目的，选取不恰当的样本来得到貌似合理的数据。例如，某某药物的有效率是百分之百，看来似乎此药有奇效，可是如果它选取的样本就是一个呢，这样的百分之百就没有任何说服力。这个例子看上去是用事实说话，实际只是统计陷阱。有些人利用统计作数据分析只是为了寻找支持他们心目中已有的某种“结论”性的东西，他们希望看到一种符合他们需要的数据分析结果，这种做法是歪曲数据分析实质的。数据分析的真正目的是要从已知的数据中找出规律，从中得到启发，在分析之前是没有结论的，只有分析结束后才得到结论。

此外，在统计的运用方面还一个问题就是简单问题复杂化。一些管理决策者为了证明自己管理的科学性，从而为一个简单的问题建立一个大家看不懂模型，编一大堆的程序，得出一大堆难懂的数字和符号，用晦涩的语言陈述结论，得出一些似是而非的管理建议或决策。这样的做法很不可取，不仅脱离管理问题，对实际的决策也未必起到作用。管理决策离不开统计，但是把问题复杂化的做法同样不可取。所以，应该把统计应该恰当地应用到它可以充分发挥作用的地方。不能把统计神秘化，更不能歪曲统计，利用统计来掩盖事实。

§ 1.3

统计学中的几个基本概念

1.3.1 总体与样本

总体是统计学中最基本的概念，是统计学研究的对象。在商务与经济统计中，总体是

指客观存在的,在某种相同性质基础上许多个体组成的集合。如研究江西财经大学学生的学习和生活情况时,江西财经大学所有的学生即为一个统计总体。又如在医疗卫生统计中,总体为“一组性质相同的数据集合”,若研究某种治疗高血压药物的疗效时,此时的总体为所有(已患高血压者和将患高血压者)患者的高血压数据。

总体又可以分为有限总体和无限总体。如果总体包含的单位数是有限的,则被称为有限总体,如某班的人数、某地区的银行数。而当其包含的单位数是无限的时候,则称之为无限总体,如要检验新设计的服装在市场上是不是受消费者的青睐。对于有限总体,我们既可以采取全面调查的方式,也可以采取抽样调查的方式。无限总体只能抽取其中的一部分样本单位进行非全面的调查,据此推断总体的性质。

样本是从总体中获得的一个群或组,样本所包含的总体单位数称为样本容量。获取样本的方法包括随机抽样法和有意选择法。使用样本的目的是为了通过样本的信息来推断总体的有关特征。在抽取样本时,必须注意以下几个问题:

1. 样本单位必须是来自同一全体及总体,不许有总体外部单位的参加。

2. 从同一个总体中可以抽取多个样本,样本容量和样本个数的多少取决于所选取的抽样方法。

3. 抽取的样本必须具有代表性。由于抽取样本并不是目的,而是手段,它是为了解总体而服务的。样本的特征值是总体特征值的代表,由样本去推断所研究总体的特征总会存在一定的误差,而误差的大小与样本的单位数、抽样方法以及抽样形式密切相关。所以,要尽量抽取合适的样本,减少抽样误差,提高样本的代表性。

4. 抽取的样本必须具有客观性。

由上可知,我们可以进行随机抽样,也可以进行非随机抽样。不管是采取哪种抽样方式,都应该排除主观因素的影响,样本的选中与否不受主观偏好的制约。

1.3.2 标志和指标

标志是描述总体单位特征的名称,包括品质标志与数量标志。品质标志反映个体的属性特征,其标志表现只能是文字,如性别和产品等级等。数量标志描述现象的数量特征,其标志表现为数值,如年龄、体重和收入等。

尽管标志是各单位都具有的普遍属性,但是各单位有关标志的表现在不特定的时间和地点条件却可能表现出不同。当总体中各单位有关标志的具体表现都相同时,称之为不变标志,而表现出不同时,则为可变标志。例如,当我们以江西财经大学的学生为研究对象时,身份这一标志的具体表现都是学生,所以身份便是不变标志,而各学生的年龄存在差异,所以年龄是可变标志。要知道,一个总体一定有一个不变标志,否则各单位不能构成一个总体;同时,还必须要有可变标志,只有各单位间存在着差异,才需要进行统计研究。

指标是说明总体的综合数量特征的。一个统计指标包括指标名称和指标数值两部分。例如,经调查某校现有学生总数为3万人,这就是指标,是说明总体综合数量特征的,它包括指标名称——学生总人数、指标数值——3万。指标的名称是对所研究现象本质的抽象概括,也是对总体数量特征的质的规定性。指标数值则是反映所研究现象在具体点的时间、地点、条件下的规模和水平。总之,统计指标是统计研究对象的具体化,也是统计认

识客观事物的过程的起点。

指标按所反映的数量特点不同,又可以分为数量指标和质量指标。凡是反映现象总规模、总水平的统计指标称为**数量指标**,如人口总数、企业总数、职工总数、工资总数、国内生产总值等等,因为这些指标是反映现象或过程的总规模和水平,所以数量指标也被称为总量指标。而用来反映现象相对水平和平均水平的统计指标称为**质量指标**,例如职工平均工资、人口密度等。质量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以及对比情况。标志和指标的关系是个体和总体的关系。

1.3.3 数据及其类型

统计数据是对客观现象计量的结果,不同的客观现象能够予以计量或测度的程度是不同的。统计数据大体上分为两种类型:定性数据与定量数据。

定性数据也称品质数据,它说明的是现象的品质特征,只能用文字或数字代码来表示,不能用数值表示。定性数据具体可分为定类数据和定序数据。

定类数据是对现象进行分类的结果,表现为类别,由定类尺度计量而成。定类尺度也称为类别尺度或列名尺度,是最粗略、计量层次最低的计量尺度。定类尺度只能按照现象的某种属性对其进行平行的分类或分组。例如,人口按照性别分为男、女两类;企业按照经济类型分为国有经济、集体经济、股份制经济、外商投资经济等。定类尺度只能测度现象之间的类别差,不能反映各类现象之间的其他差别。定类数据是层次最低的数据,从数学运算的特性来看,定类数据只有等于或不等于的性质。

定序数据是对现象按照一定的排序进行分类的结果,表现为有顺序的类别,由定序尺度计量而成。定序尺度又称顺序尺度,是对现象之间等级差别和顺序差别的一种测度。它不仅可以测度现象之间的类别差,还可以测度次序差。例如,学生的考试成绩表示为优、良、中、及格、不及格;消费者对某产品的满意程度表示为很满意、满意、一般、不满意、很不满意等等。定序尺度不能测量类别之间的准确差值,只能比较大小,不能进行加、减、乘、除数学运算。定序数据的层次高于定类数据,从数学运算的特性来看,定序数据除了具有等于或不等于的性质以外,还有大于或小于之分。

定量数据也称数量数据,它说明的是现象的数量特征,是能够用数值来表示的。定量数据具体又分为定距数据与定比数据两种。

1. 定距数据。定距数据是既能反映现象所属的类别和顺序,又能反映现象类别或顺序之间数量差距的数据,由定距尺度计量而成。定距尺度也称间隔尺度,通常使用自然或物理单位作为计量尺度。定距尺度不仅能将现象区分为不同类型并进行排序,而且可以准确地指出类别之间的差距是多少。例如,甲、乙两位学生某门课程的考试成绩分别为88分和58分,不仅说明甲学生的成绩良好,乙学生的成绩为不及格,甲学生的分数高于乙学生,而且还说明甲学生分数比乙学生高30分。

定距数据的层次高于定序数据。从数学运算的特性来看,定距数据除了具有等于或不等于、大于或小于的性质以外,还适合于进行加减计算,但不适合于乘除运算。其原因是定距尺度中没有绝对的零点。定距尺度中的“0”是表示一个数值,即“0”水平,而不表示“没有”或“不存在”。例如,一个学生的统计学考试成绩为0分,表示他的统计学成

绩水平为 0, 并不表示他没有考试成绩或没有任何统计学知识; 一个地区的气温为 0°C , 表示的是温度的水平, 并不是没有温度。可见定距尺度中的“0”是一个有意义的数值。

2. 定比数据。定比数据是不仅能体现现象之间的数量差距, 还能通过对比运算, 即计算两个测度值之间的比值来体现相对程度的数据, 由定比尺度计量而成。定比尺度也称为比率尺度, 它有一个绝对“零点”, 在定比尺度中, “0”表示“没有”或“不存在”。比如, 一个人的收入为“0”, 表示这个人没有收入。因此, 定比尺度除了具有上述三种计量尺度的全部特性外, 还具有一个特性, 那就是可以计算两个测度值之间的比值。这也是它与定距尺度的唯一差别。现实生活中, 绝大多数的经济变量都可以进行定比测度。

定比数据是最高层次的数据。从数学运算的特性来看, 定比数据不仅可以加减运算, 还可以乘除运算。例如, 甲工资为 1200 元, 乙工资为 2400 元, 则乙的工资为甲的 2 倍。

定量数据按数值的表现形式不同, 可分为即绝对数、相对数和平均数三种形式。

(1) 绝对数, 揭示的是现象的总规模和总水平, 如一国的总人口、某一年度的国民生产总值等。它的计量单位一般为实物单位或者价值单位, 偶尔也使用复合单位。如人口数是用人计量, 这就属于实物单位; 价值单位是以货币形式进行计量, 如国民生产总值、出口总额; 复合单位是由两种或两种以上计量单位复合而成的, 如以“千瓦时”为电量的计量单位。

按绝对数所反映的时间状况不同, 绝对数可细分为时期数(流量)和时点数(存量)。时期数反映一段时期的总量, 如产品产量和出生人口数等, 并且可连续计数和累积, 数值大小与其所属的时间长短有直接的关系。时点数描述某一时点的总量, 如资产存量和人口数等。由于时点数的数值大小与时点相隔长短没有直接关系, 主要是反映现象在某一瞬间的水平, 因而只能间断计数且不能累积。

(2) 相对数, 是两个具有相互联系的数据的比值。具体可分为比例和比率两种数据形式。比例是总体各个部分的数量占总体该数量的比重, 通常反映总体的内部结构。比率是不同类别的数量之比, 既可以是同一个总体中各不同部分的数量对比(如人口统计的性别比), 也可以是同一现象在不同时间或空间上的数量之比(如本年度的 GDP 与上一年度的 GDP 之比, 本国 GDP 与其他国家 GDP 之比等)。常用的相对数有结构相对数、动态相对数、比较相对数、强度相对数、计划完成相对数、比例相对数。

(3) 平均数, 是反映现象总体的一般水平或分布的集中趋势, 有算术平均数、调和平均数、几何平均数、幂平均数和位置平均数等。

§ 1.4

统计研究方法

统计学的研究对象和性质决定了统计学有自己专门的研究方法。这些研究方法有大量观察法、统计描述法、统计推断法。

1.4.1 大量观察法

大量观察法是指对所研究的事物的全部或其中的大量个体进行逐一的调查,以达到对现象总体数量特征及其规律性的认识。它是以大数定理为依据的。由于社会现象和自然现象都受到各种社会规律或自然规律相互交错作用的影响,在所研究事物的总体中,个体往往受到偶然因素的影响,它们之间会存在数量上的差异性,如果只是选取其中之一进行观察,其结果是不能代表总体的一般特征的;只有对全部或者足够多的单位进行观察并加以综合,才能够抵消那些影响个别单位的偶然因素,现象的一般特征才能显示出来,从而得到合乎实际的结论。例如,每个学生的月消费金额有多有少,但是随着我们观察人数的增加,调查的结果就越具有代表性,越合乎实际。

大量观察法广泛应用于我国的统计实践中,例如各种基本的、必要的统计报表,普查、重点调查和抽样调查等等。

1.4.2 统计描述法

统计描述法是指通过了解客观实际,并对搜集到的数据进行加工整理和综合分析,得到各种能够反映总体数量特征的综合指标,以此反映现象总体的总量规模、结构比例、速度快慢等实际状况。统计描述是统计研究的基础,它为统计推断、统计咨询、统计决策提供必要的事实依据。它通过对零散无序的原始资料的整理归纳,运用分组法和综合指标法,得到现象总体的数量特征,揭示客观事物的内在数量规律性,达到对事物认识的目的。

统计描述包括以下三个方面的内容:

1. 统计分组法。根据统计研究的目的及任务,将调查得到的大量资料按照一定的标志划分为若干个组,使组内的单位具有相对的同质性,组间的单位具有明显的差异性,研究它们的分布情况。例如,要研究江西财经大学学生的月消费额情况,可以选择“消费额”为标准进行分组,就可以得到该大学学生消费额的大概分布情况。

2. 综合指标法。综合指标法是运用各种统计指标来反映总体的一般数量特征和数量关系的方法。通常使用的综合指标主要有总量指标、相对指标、平均指标、变异指标等。这些指标各自从不同的角度刻画了总体的特征,将其综合使用,可以显示出现象在具体时间、地点条件下的总量规模、相对水平、集中趋势、变异程度,并进一步从动态上研究现象的发展趋势和变化规律。除此之外,通过综合指标的对比分析,还可以分析现象的各个因素,分析它们之间的相互关系,分析各种因素的变动及其对总体的影响。综合指标法常与大量观察法一起使用。

3. 统计模型法。它是根据一定的理论和假设前提,用数学方程去模拟显示客观现象间相互关系的一种研究方法。此方法可以对客观现象的发展变化过程中存在的数量关系进行比较完整和近似的描述,反映出所研究的综合指标之间的关系,简化了客观存在的其他复杂关系,便于用模型对所关心的现象变化进行数量上的评估和预测。

1.4.3 统计推断法

统计推断法就是以一定的置信水平要求,根据样本数据来判断总体特征的归纳推理方