

高等财经类核心课程教材

统计学

马敏娜 王静敏 主编

高等学校财经类核心课程教材

统计学

马敏娜 王静敏 主编

TONGJIXUE



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容简介

本书是高等学校财经类核心课程教材,主要反映现代统计基本理论和方法,包括统计数据的采集、统计数据的整理、统计数据的描述、抽样与参数估计、参数假设检验、方差分析、相关与回归分析、时间序列分析、统计指数、国民经济统计基础知识。

本书内容系统全面,结构合理,脉络清晰,突出原理性和应用性相结合。在统计方法上尽量避免烦琐的数学推导和证明,通过大量的实例说明统计方法在经济管理中的应用。在应用叙述上深入浅出,文字简练。计算方法的实现采用了适用面较广的 Excel 软件和专业性较强的 SPSS 软件。

本书可作为高等学校财经类专业本科生的教材及人文、社会科学及其他学科的教材,也可作为高等学校和科研机构的教师、学者、管理工作者的学习与应用统计方法的参考书,还可作为用 Excel 软件和 SPSS 软件解决计算问题的实用手册。

图书在版编目(CIP)数据

统计学 / 马敏娜,王静敏主编. —北京:高等教育出版社,2011.12

ISBN 978 - 7 - 04 - 032877 - 6

I. ①统… II. ①马… ②王… 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 (2011) 第 221184 号

策划编辑 权利霞 施春花
插图绘制 尹文军

责任编辑 施春花
责任校对 胡晓琪

封面设计 王 睢
责任印制 韩 刚

版式设计 杜微言

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 高等教育出版社印刷厂

开 本 787 × 1092 1/16

印 张 22.75

字 数 510 000

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.landaco.com>

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

版 次 2011 年 12 月第 1 版

印 次 2011 年 12 月第 1 次印刷

定 价 33.20 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 32877 - 00

前 言

统计学是经济学类、工商管理类各专业的核心课程。统计学是关于数据采集、数据归纳整理、数据分析,并从中探寻客观对象内在规律性的方法论科学。随着我国社会主义市场经济的不断完善,人们对各国(地区)经济、管理、社会及企业产品的市场需求与预测的分析增加,对统计分析提出了更高要求,统计学在经济、管理及社会中的作用日益显著。

本书系统介绍了适用于各个应用领域的现代统计的基础理论和方法,包括:统计数据的采集与整理;描述统计、参数估计、假设检验、方差分析、相关与回归分析、时间序列分析、统计指数;国民经济统计的基础知识。

本书的四位作者为吉林财经大学统计学院教授,是吉林省统计学专业优秀教学团队的骨干教师,均有 20 多年的统计教学实践经验。吉林财经大学统计学课程 2005 年被评为吉林省精品课程,根据经济学类、管理学类专业的特点,我们在吸取国内外优秀统计教材精华的基础上编写了这本教材。与国内同类教材相比,本书具有以下特点:

1. 在内容组织上系统全面,结构合理,脉络清晰。由于在经济学类、管理学类各专业中,除统计学专业以外,其他各专业没有开设国民经济统计学课程,考虑到国民经济统计中国内生产总值(GDP)等主要指标应用的广泛性和计算的复杂性,大家在使用时经常发生一些错误,为此我们增加了国民经济统计基础知识一章,突出了原理性和应用性的结合。

2. 在统计方法上尽量避免烦琐的数学推导和证明,通过大量的实例说明各种统计方法及其适用条件,每章开篇引入统计实例,贴近现实的社会经济生活,提出问题,结合问题选用统计方法并对计算结果进行解释。

3. 在应用叙述上深入浅出,文字简练。本书选择的实例容易理解,具体生动,反映社会发展热点,引人深思,有助于培养学生理论联系实际的学风。

4. 在计算工具上我们选择了两种软件作为实现统计计算的工具:一是适用面广的 Excel 软件,二是专业针对性强的 SPSS 统计软件。

国内一般教材主要使用 Excel 软件。随着统计理论方法的普及应用,学生对 SPSS 统计软件应用的需求逐渐增加,有的院校还专门开设了 SPSS 统计软件应用课程,鉴于此,我们特意增加了 SPSS 统计软件的使用介绍。在计算量较大的部分章节,我们以课程实验的方式给出了两种软件的计算过程,学生可以根据自己专业的特点进行选择,有利于培养学生运用工具软件分析解决社会经济管理实际问题的能力。受篇幅所限和有利于及时更新,本书的课后习题和教学课件均放置在吉林财经大学精品课程网站上,读者可自行下载,也可通过封底所附网址下载。

全书共分为 11 章,具体的编写分工如下:第一章、第十章、第十一章由马敏娜教授编

写；第二章、第三章、第四章由孙雅静教授编写；第五章、第九章由张平教授编写；第六章、第七章、第八章及各章的课程实验，由王静敏教授编写；全书由马敏娜教授和王静敏教授审定。

在本书的编写过程中，有关专家、学者提出了宝贵意见，同时也得到了高等教育出版社工作人员的大力支持，在此表示衷心的感谢。由于编者水平和精力所限，本书存在的不足敬请读者批评指正。

编者

2011年11月

目 录

第一章 总论 1

- 第一节 统计学的性质 1
- 第二节 统计学的发展与分类 5
- 第三节 统计学的基本概念 8

第二章 统计数据的采集 12

- 第一节 统计数据的来源 13
- 第二节 统计调查方案与问卷设计 15
- 第三节 统计数据的采集方法 24
- 第四节 课程实验 28

第三章 统计数据的整理 33

- 第一节 统计数据的审核 33
- 第二节 统计分组 34
- 第三节 次数分布 37
- 第四节 统计表 48
- 第五节 课程实验 51

第四章 统计数据的描述 65

- 第一节 总量指标与相对指标 65
- 第二节 集中趋势 71
- 第三节 离中趋势 87
- 第四节 偏态与峰度 97
- 第五节 课程实验 101

第五章 抽样与参数估计 106

- 第一节 抽样推断的基本问题 106
- 第二节 抽样分布 112
- 第三节 参数估计 119

第四节 其他抽样组织方式 134

第六章 参数假设检验 148

- 第一节 假设检验的基本原理 148
- 第二节 一个正态总体参数的假设检验 154
- 第三节 两个正态总体参数的假设检验 159
- 第四节 非正态总体大样本下参数的假设检验 164
- 第五节 总体比率的假设检验 (大样本) 167
- 第六节 假设检验中的其他问题 170
- 第七节 课程实验 173

第七章 方差分析 183

- 第一节 方差分析原理 183
- 第二节 单因素方差分析 188
- 第三节 课程实验 191

第八章 相关与回归分析 195

- 第一节 简单线性相关分析 195
- 第二节 一元线性回归分析 200
- 第三节 多元线性回归分析 213
- 第四节 课程实验 220

第九章 时间序列分析 239

- 第一节 时间序列的分析指标 239
- 第二节 长期趋势的测定 255
- 第三节 季节变动的测定 268

第四节 循环变动的测定	274	第十一章 国民经济统计基础知识	336
第五节 课程实验	277		
第十章 统计指数	306		
第一节 指数的概念和种类	307	第一节 国民经济统计的含义和意义 ...	336
第二节 综合指数	309	第二节 国民经济统计的基本分类	338
第三节 平均指数	315	第三节 国民经济统计的主要指标	342
第四节 指数体系与指数因素分析	318	附表	349
第五节 常用的经济指数	325	参考文献	357

第一章 总 论



统计实例

【庆祝新中国成立 60 周年系列报告】 1978 年，我国 GDP 总量居世界第 11 位。2000 年，我国 GDP 超过意大利，居世界第 6 位，是美国的 12.3%、日本的 25.7%、德国的 63.1%、英国的 83.1% 和法国的 90.2%。2005 年，我国 GDP 超过英国和法国，居世界第 4 位，是美国的 18.1%，比 2000 年提高 5.8 个百分点；是日本的 49.3%，接近一半，比 2000 年提高了近一倍；是德国的 80.5%，比 2000 年提高 17.4 个百分点。2008 年，我国 GDP 超过德国，跃居世界第 3 位，但已是位于世界第 2 位的日本的 78.6%，比 2005 年提高 29.3 个百分点。60 年多来，我国经济总量与主要发达国家之间差距不断缩小。

【世界经理人数数据】 美国投资银行美林公司和法国凯捷咨询公司公布的全球财富报告显示，2009 年中国富人总数增长三成，成为财富增长最快的国家之一，并带领亚洲在总财富方面超越欧洲。报告将净资产（不包括主要房产）价值 100 万美元以上的人士归为富人，按此标准，2009 年中国富人总数达到 47.7 万人，较 2008 年增加 31%，位居全球第四。

【中国人的每天消费】 目前按中国人口 13 亿计算，每天消费粮食约 80 万吨，相当于一个粮食基地县全年的总产量；猪肉约 6 万吨，相当于生猪 100 万头；食用植物油约 2 万吨，相当于 60 万亩^①油菜所榨的油；糖约 1.6 万吨，相当于 4.8 万亩甘蔗地所产的糖；鲜蛋 1 870 万千克；水产品 1 955 万千克，相当于云南省水产品年产量的一半；卷烟 2.2 亿盒，把这些烟三盒一叠排成直线，长约 6 266.6 千米，比中国东西之间的国土宽度还长 1 200 千米；酒 3.6 万吨，全年累计喝掉的酒可以装满 1.5 个杭州西湖；生活用布 3.6 万千米，几乎可以绕赤道一周；煤 60 万吨以上，相当于一个大中型矿井的全年产量；购买报纸 5 000 万份，约需 400 多辆中型货运汽车才能装载。

第一节 统计学的性质

一、统计的含义

什么是统计？在日常生活中，人们常常会提出许多问题。比如：今年全世界发生的自然灾害造成的损失是多少？自然灾害的发生是否有逐年递增的趋势？今年我国居民消费价格水平

^① 1 亩约等于 666.7 平方米。

(CPI)是上升,还是下降?一个城市竞争力的高或低是怎么计算出来的?不同机构对大学排名的结果不同,应当怎么理解?

一般来说,回答这些问题就要收集相关的资料(数据),对其进行整理分析,最后得出结论。比如,回答居民消费价格水平的涨跌问题,就要对消费品市场上的商品和服务价格进行调查,调查记录的结果(数据)就是统计数据,对其进行整理,采用科学的方法进行计算,得出结论,这个工作过程就是统计工作。在统计工作过程中,采用什么样的方法最科学有效,这是统计科学要解决的问题。

“统计”一词包含了三个层面的含义,即统计工作、统计数据和统计科学。

(一) 统计工作

统计工作也称统计实践,是指人们有目的地对客观现象的数量关系和数量特征进行的统计设计、统计调查、统计整理和统计分析,并据以揭示其内在数量规律性的工作过程。统计工作的每一个阶段,既有一定的独立性,又相互紧密联系着,形成一个完整的认识客观现象的过程。

1. 统计设计

统计设计是根据统计研究目的和研究对象的特点,对整个统计工作过程的各个方面和各个环节进行通盘考虑和安排。统计设计的结果是各种设计方案和统计方法制度。

2. 统计调查

统计调查是根据统计设计的要求,运用科学的方法收集数据资料的工作过程。统计调查是统计工作的基础阶段,为统计整理和统计分析提供基础数据。

3. 统计整理

统计整理是对统计调查的数据资料进行科学的加工,为统计分析准备系统化、条理化的综合资料的工作过程。统计整理是整个统计活动的中间环节,是统计调查的继续,也是统计分析的前提。

4. 统计分析

统计分析是以经过加工的统计资料为基础,利用专门的统计方法,分析现象的发展过程及规律性,对现象做出科学结论的工作过程。

综上所述,一个完整的统计工作过程是从统计设计(定性认识)到统计调查和统计整理(定量认识),最后通过统计分析(更高层次的定性认识),最终达到对现象本质和规律性的认识过程。

(二) 统计数据

统计数据也称统计资料或统计信息,是通过统计工作所取得的反映客观现象的各项数据及与之相关的资料的总称,表现为各种统计图、统计表、统计公报、统计年鉴等。

(三) 统计科学

统计科学是指导人们如何进行统计设计、统计调查、统计整理、统计分析的理论和方法的科学,也是指导人们如何从事统计工作的科学。

统计的三层含义既有区别又有联系。统计数据是统计工作的成果；统计工作与统计科学是实践与理论的关系，统计理论源于统计实践，是统计实践经验的总结，反过来又指导实践，在实践中不断升华，使统计理论方法体系不断完善和发展。

二、统计学的定义

统计学是一门收集、整理和分析统计数据的科学。统计学是属于方法论的科学，其目的在于探索客观现象内在的数量规律性，从而认识客观现象内在的质的规律性。

中国香港著名统计学家胡孝绳在其所著的《统计学》中说：“统计学也可称为统计方法，即就研究对象的数量资料(quantitative data)，从事收集(collecting)、显示(presenting)、分析(analyzing)及解释(interpreting)，借以表现其真相的一种学问。”

《大不列颠百科全书》把统计学定义为，“统计学是收集、分析、表述和解释数据的科学”。

三、统计学的研究对象及其特点

(一) 统计学的研究对象

统计学的研究对象是客观现象的数量方面，包括数量多少、数量关系和数量界限。

1. 数量多少

这是指现象的数量规模、总量水平。如 2009 年年末中国总人口为 133 474 万人，全年社会消费品零售总额 132 678.4 亿元，全年粮食产量 53 082 万吨，全年国内生产总值 340 506.9 亿元，分别说明中国 2009 年年末的人口规模、全年的消费总量、全年粮食生产总量、全年经济发展总量。

2. 现象之间的数量关系

现象之间的数量关系包括发展速度、结构、比例、差异程度、平均水平等。如 2009 年中国全年国内生产总值比上年增长 8.4%。第一、二、三产业增加值占国内生产总值的比重分别为 10.3%、46.3% 和 43.4%。2009 年中国人口出生率为 12.13‰，死亡率为 7.08‰，自然增长率为 5.05‰，出生人口性别比为 105.9:100。

3. 质量互变的数量界限

质量互变的数量界限如营销中的盈亏平衡点，居民生活贫困线，企业规模中大、中、小型企业的数量界限，学习成绩中的及格、中等、良好、优秀的数量界限等。

应当注意的是统计学在现象数量方面的研究应当和性质方面联系起来，在质与量的密切联系中研究现象的数量方面。统计学需要数学运算，但统计学不是数学，数学撇开事物的具体内容研究抽象的数量关系和空间形式，而统计学则是研究具体客观现象本身的数量关系，总结客观现象的内在数量规律。

现象的内在数量规律是什么呢？比如，我们都做过掷硬币的小游戏，知道我们向上抛一枚硬币，硬币落下后哪一面朝上本来是偶然的，但当我们上抛硬币的次数足够多后，达到上万次甚至几百万次以后，我们就会发现，硬币每一面向上的次数约占总次数的 50%。这种情况下，

偶然中包含着必然。这就是我们所说的数量规律性。

在进行农作物实验时，如果其他实验条件相对固定，那么我们会发现某种农作物产量与耕地的密度和深度有关，密度和深度达到多少时，产量会达到最高呢？这一数量关系就是我们要探索的数量内在规律。

（二）统计学研究对象的特点

统计学研究对象的特点主要有：数量性、差异性、总体性。

1. 数量性

数量性是统计研究对象的基本特点，客观现象本身是可以用数字度量的，才属于统计学的研究范畴，这也是统计学科与其他学科的重要区别。

2. 差异性

由于客观现象的复杂性，使各种现象的数量表现不尽相同，表现为个体千差万别，且个体数量表现具有不确定性或随机性。因此，才要求统计采用其特有的方法去解释现象的本质及规律性。差异是统计存在的前提条件，没有差异就没有统计，差异为统计提供用武之地。

比如，居民消费价格指数(CPI)是反映一定时期内城乡居民所购买的生活消费品价格和服务项目价格变动趋势及程度的物价指数。市场上与老百姓生活水平相关的消费品有成千上万种，各种商品价格不同，同一种商品价格也会因为规格、花色、产地、品牌不同而不同，有的价格是上升的，如食品类，有的是下降的，如电子产品、通信等，也有的价格是不变的，所以国家要对其进行统计。按照国家统计局公布的数据，2009年全年CPI比上年下降0.7%。

3. 总体性

从总体上研究客观现象的数量方面，是统计学区别于其他学科的一个重要特点。统计学从整体的观点出发，从大量个体现象入手，进而上升到对整体的数量特征的认识，反映现象发展变化的内在规律性。

例如，一个家庭生男孩生女孩是不确定的，但从总体上看，新生婴儿的男女性别比例约为106:100，古今中外这一比例大致相同，这是由人类自然发展的内在规律所决定的。尽管从新生婴儿来看，男性略多于女性，但男性的平均寿命比女性低。因此，人口在性别比例上保持大体平衡，保证了人类社会的进化和发展。

上面的例子说明，仅就个体观察，其结果往往是随机的、不确定的，但通过大量观察，利用统计方法是可以探索出现象总体内在的数量规律性的。

四、统计学的作用

统计学为研究客观现象的内在规律提供指导原则和方式方法，它是认识现象内在规律的手段。统计是认识社会最有力的武器之一，这个基本性质决定了统计方法在社会实践中的重要作用。

在科学研究领域中，不论是工程技术、农学、医学、军事科学，还是社会科学、经济领域，要掌握和分析大量有关信息就必然要用到统计方法。由于客观现象千差万别，它们相互联系、相互制约的关系非常复杂，因此更需要以大量的实际资料为依据。这些资料和信息

得,需要运用统计方法去收集有关数据,经过分析研究,得出科学结论。比如,科学家如何评价新理论的有效性?医学人员如何设计实验来测定新药的疗效?人口学家如何预测未来某一时间世界人口数量?经济学家如何对国家的宏观形势进行分析预测?审计师如何确定账目的差错率?投资公司如何确定一项投资的风险大小?营销管理者如何根据市场状况制定营销决策?这些都离不开统计。

在经济领域和日常生活中,人们在不同的行业,从不同的角度,以不同的方式应用统计方法去认识世界、改造世界。例如,在网络购物中,我们点击某个商家,要看商家的信誉、货品的出售量、顾客的购买评价等。很多网络店铺都采用了统计方法,可以看出最近7天的客户访问量、商品被访排行、商品的价格区间、回头客访问量,从而进行比较、分析、评价、决策。气象预报员利用大片地区的气象资料进行预报。企业根据产品信息决定产品的生产。银行利用统计数据确定货币投放量。教练员运用统计方法评价运动员的水平。股票投资者根据股价变动趋势决定股票的买卖。农民利用过去和现在的观察估计收成好坏。人们总是有意无意地使用统计方法去解决周围所发生的事情,使统计变成认识世界的工具。

统计学在发现和试验新产品、新方法、新思想方面也有着广泛的应用。100多年前英国作家和历史学家韦尔斯(H. G. Wells)指出:“统计思想有一天会像读写能力一样成为追求效率的公民的需要。”

第二节 统计学的发展与分类

一、统计学的产生与发展

(一) 统计实践史

统计学的产生与统计实践活动是密不可分的,统计作为一种社会实践活动,已有四五千年的历史。

人类的统计实践是随着计数活动而产生的。因此,对统计实践发展的历史可追溯到人类社会初期的打绳结、画道道计数。统计实践的真正萌芽是在古代奴隶社会。统治阶级为了征税、征兵、服劳役的需要进行统计。我国在公元前21世纪的夏朝,就有了人口与土地数字的记载,当时全国分为九州,人口1355万人。古代埃及也有类似的记载。统计实践的广泛发展始于资本主义社会。17世纪以来,由于工、商、农、贸、交通的发展,资本主义国家先后设立专业的统计机关,收集各方面的统计资料,定期或不定期地举行人口、工业、农业、贸易、交通等项调查,出版统计刊物。

(二) 统计学说史

统计学的理论和方法则是在长期统计实践活动的基础上发展起来的,统计学至今已有300多年的历史。统计学产生于17世纪中叶,其发展过程主要沿着两条主线展开。

1. 政治算术学派

这是以政治算术学派为开端形成和发展起来的以社会经济问题为主要研究对象的社会经济统计。政治算术学派产生于17世纪中叶的英国,主要代表人物是威廉·配第(William Petty, 1623—1687)和约翰·格朗特(John Graunt, 1620—1674)。威廉·配第在其代表作《政治算术》(1676)一书中,运用大量的数字资料对英国、法国、荷兰三国的经济实力进行比较。他在书中运用数字、重量、尺度等进行数量对比分析的方法,为统计学的产生奠定了基础。

约翰·格朗特在其《对死亡表的自然观察和政治观察》(1662)一书中,通过大量观察的方法,研究并发现了一系列人口统计规律,如男婴出生多于女婴,两者之比为14:13。各年龄组死亡率男性高于女性等。因此,他被认为是人口统计学的创始人。

2. 概率论数理统计学派

统计学的第二个源头是以概率论的研究为开端、以概率论为基础形成和发展起来的、以随机现象为主要研究对象的数理统计。概率论研究起源于意大利文艺复兴时代。最初的研究是为赌徒们找出掷骰子取胜的一套办法,合理地算出取胜的机会。在18世纪,概率理论日益成熟。19世纪中叶,比利时的阿道夫·凯特勒(Lambert Adolphe Jacques Quetelet, 1796—1874),在其《社会物理学》中将古典概率论引入统计学,使统计学进入一个新的发展阶段。

20世纪以后,统计理论方法和应用进入了一个全面发展的阶段。这一时期的主要特征是由描述统计学转向推断统计学。1907年,英国学者戈塞特(W. S. Gossett, 1876—1937)提出了小样本 t 统计量理论,丰富了抽样分布理论,为统计推断奠定了基础。英国的费舍尔(R. A. Fisher) (1890—1962)提出了极大似然估计量的概念,迅速成为了估计参数的重要方法。英国科学家弗朗西斯·高尔顿(Francis Galton)提出了相关与回归方法。英国统计学者皮尔逊发展了拟合优度检验,还给出了卡方统计量及其极限分布。波兰学者奈曼(J. Neyman, 1894—1981)创立了区间估计理论,与他人合作发展了假设理论。美国学者瓦尔德提出决策理论和序贯抽样方法。美国化学家威尔科克森(Frank Wilcoxon)发展了一系列非参数统计方法,构建了现代统计学基本框架。统计学在计算机、信息论等现代科学技术的影响下,研究领域进一步扩大,如多元统计分析、时间序列分析、贝叶斯统计、非参数统计、数据挖掘等。同时,统计的应用领域也在不断扩大,几乎所有的自然科学、社会科学研究都离不开统计方法。

二、统计学的分类

统计学作为认识、揭示大量现象数量规律性的方法和工具,由于其研究目的不同、处理资料的条件不同、使用的分析方法不同而分成不同类别。

(一) 描述统计与推断统计

描述统计学是研究如何取得统计数据,利用一些综合性的指标来描述所研究现象的数量关系和数量特征。描述统计是表述、分析、解释全部总体资料的数量特征的科学方法。其目的是通过适用、简单、明确、易于理解的形式去揭示现象的数量规律性。在社会、经济和日常生活中,我们所收集或遇到的原始数据资料大部分都是杂乱无章、无规可循的,而人们却可以通过统计描述的过程,用更加有意义或更加适用的形式,遵循描述数据资料的某些规则和程序,使

现象的数量规律性呈现出来。如同我们对一个物体进行画像一样,把形状、结构和大小画出来。

尽管描述统计能够完成统计的一些基本任务,但是,由于人们认识能力的局限性和自然、社会经济现象的复杂性,人们不可能掌握一些总体的全部资料。例如,不可能准确了解水库里有多少尾鱼,原始森林有多少木材,某台彩电的使用寿命是多少,某项投资是否能赢利等。在这种情况下,我们不能进行描述而只能对现象的数量特征进行推断。

推断统计学是研究如何根据样本数据去推断总体数量特征的方法,它是在对样本数据进行描述的基础上,对统计总体的未知数量特征或未知总体分布形式做出具有一定把握程度的推断。

我们知道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,如果我们收集到的是总体数据,则经过描述统计之后就可以达到探索内在数量规律性的目的了。但如果我们所获得的数据只是研究总体的一部分数据,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做出科学的推断。显然,描述统计是整个统计学的基础和统计研究工作的起点,它为统计研究工作收集可靠的数据和有效的样本信息。而推断统计是现代统计学的核心和统计研究工作的关键环节,因为统计最终能否科学准确地探索到总体内在的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。推断统计在实践中具有相当重要的意义。它既可以节省费用支出,又可以节省时间,是非常有效、实用、科学的统计方法。

描述统计与推断统计虽然方法不同,用途各异,但两者的目的却是一致的,即揭示大量现象的数量规律性。在实践工作中,推断统计常常以描述统计为基础,使用描述统计方法的测定结果进行推断。因此,人们使用统计方法去探索现象的数量规律性,不仅要掌握描述统计方法,还要掌握推断统计方法。

(二) 理论统计与应用统计

随着统计学被广泛地应用到社会科学和自然科学的各个领域当中,统计学也逐渐发展成为由若干个分支学科组成的学科体系。根据统计学研究的侧重点不同,可将统计学分为理论统计学和应用统计学。

理论统计学是指统计学的数学原理,是抽象地研究统计学的一般理论和方法。它包括描述统计学与推断统计学。

应用统计学是将理论统计学的基本原理应用于各个实质性科学领域,用于探索各个领域内在数量关系和数量规律。统计方法与相应的实质性科学相结合,产生了相应的统计学分支,如统计方法在社会学中的应用形成了社会统计学;统计方法在经济学领域中的应用形成了经济统计学;统计方法在人口学中的应用形成了人口统计学;统计方法在教育中的应用形成了教育统计学;统计方法在生物学中的应用形成了生物统计学;统计方法在医学中的应用形成了医学统计学;统计方法在物理学中的应用形成了物理统计学等。以上这些应用统计学的不同分支所应用的基本方法都是一样的,即描述统计和推断统计的主要方法,但由于各应用领域都有其特

殊性，统计方法在应用中就具有了不同的特点。

事实上，很难划分开理论统计学与应用统计学的界限，因为在统计的发展史中，每一个统计方法的突破都是以应用为前提的，如恩格尔系数、时间序列分析、相关分析与回归分析等。然而，正是由于统计的方法性与工具性，所以作为一个从事应用统计的工作者，不仅应当精通统计理论，而且还要精通其所从事的领域的专业知识，这样才能更好地应用统计。

第三节 统计学的基本概念

为了有利于以后各章的学习，首先要了解几个统计学的基本概念。

一、总体与个体

（一）总体与个体的含义

凡是客观存在的，在同一性质基础上结合起来的许多具有差异的个体单位构成的整体，称为总体，它也是我们的研究对象。构成总体的每一个个体单位，称为个体。

最常见的总体是由自然物体组成的。例如，我们要研究某班学生的学习情况，因为每个学生是客观存在的，都具有是这个班级的学生这一共同性质，而每个学生的学习成绩等方面各不相同，所以，由这个同一班级的许多学习成绩不同的学生所构成的班级便是统计总体，每个学生便是个体。

进行全国人口普查，全国人口是总体，其中的每个人是个体；某地进行工业企业调查，该地所有工业企业是总体，其中每个工业企业是个体；对工业品出厂价格变化情况进行调查，所有出厂的工业产品是总体，每一种出厂的工业产品是个体。

在有些情况下，我们的研究对象可能是具有相同性质的个体的数据资料，而不是个体本身。因此，从全部有关个体所获得的数据资料的集合，也可以视为总体，如股市一周的交易次数、我国高速公路一年的交通事故次数、产品质量检查时生产线上某一时段的产品数量等。

总体中含有限的个体称为有限总体，如全国的人口、工业企业等。总体中含无限的个体称为无限总体，如原始森林中的树木、海洋中的鱼、宇宙中的星球等。

（二）总体的特征

统计总体具备三个基本特征，即同质性、大量性、差异性。

1. 同质性

同质性是指各个个体在某一方面必须具有共同的表现，否则对众多个体的综合便无法描述或揭示其总体特征。如研究我国人口状况时所针对的个体都必须具有中华人民共和国国籍，否则所概括的人口特征便不能描述中国的人口情况。

2. 大量性

大量性具有两层含义。一是指总体必须由数量众多的个体单位所组成。这是因为统计的目的是对总体进行研究，而不是对个体进行研究。当然，只有在对许多个体分析、综合、研究的

基础上,才能揭示出总体的数量特征。二是指要观察总体中足够多的个体单位。在统计研究中,只要达到目的要求,为了节约成本和时间,也可以不必观察全部个体单位,但必须在一定误差的前提下观察到足够多的个体单位,否则也无法对总体特征作出具有一定可靠性的描述和估计。

3. 差异性

差异性是指总体中的个体单位的具体表现不同。例如,人口有性别方面的差别、年龄方面的差别等。正是这种差异的存在,才需要统计对其进行综合概括、抽象分析,从而揭示总体的数量特征。

总体与个体不是一成不变的,它可以随着研究目的发生变化。当研究某地区工业企业的生产情况时,则所有工业企业是总体,个体单位是各个工业企业,如果研究某工业企业各种设备的情况,则每台设备是个体单位,而所有设备是总体。不论怎样变化,总体的三个基本特征不变,就是说总体中的个体单位应当具有足够多的个数,个体之间至少在某一个方面具有共性,在其他方面具有差异。

对总体与个体的区别关键取决于对个体单位的认识和判别,然而,在实践中的许多场合,也可能对总体中的个体不容易区分,如检查玉米的含杂率、某河流有害物质的含量等,其个体不明确,这时一般将一个观测单位看成一个个体,如10千克玉米或100毫升水。

二、变量与数据

(一) 变量与数据的含义

在统计研究中,我们对总体数量特征的认识是从对个体单位的属性特征或数量特征的调查或测量开始。我们把这些特征称为变量。变量是个体单位所具有的特征,一般也称其为标志。

数据是调查或测量的结果。数据由变量和其测量值组成。变量无论是用文字表示的属性特征,还是用数值表示的数量特征,一般统称为数据。当然,从调查的角度看,数据不仅仅是直接的调查结果,还可以包括其各种衍生结果。例如,性别有男、女之分;所有制形式有全民所有制、集体所有制之分等;学习成绩有50分、60分、90分;学生出勤人数有100人、105人等。

(二) 变量的分类

1. 定量变量与定性变量

变量按其表现特征是否为数值型可分为定量变量与定性变量。变量用非数值表示的是属性特征,称为定性变量或属性变量,如性别、观点等。变量用数值表示的是数量特征,称为定量变量,如商品价格、学生出勤人数等。

2. 连续变量和离散变量

定量变量按其取值是否连续可分为连续变量和离散变量。连续变量的数值从理论上说是连续不断的,在任意两个数值之间可能有无限个数值,如产值、身高、体重等。离散变量的数值只能以整数计量,如企业数、设备台数、职工人数等。

3. 定类变量、定序变量、定距变量、定比变量

变量按所反映的对象不同划分,可以分为以下四种类型。

(1) 定类变量(数据)。定类变量是用于区分个体属性差别的变量,表现为类别,不能区分顺序,如民族、颜色、行业等。

(2) 定序变量(数据)。定序变量不仅反映现象类别,还可反映个体之间等级差别或顺序差别。例如,产品质量有一等、二等、三等,消费者满意度有满意、比较满意和不满意,对改革方案有赞成、比较赞成和反对等。

(3) 定距变量(数据)。它是对个体类别或次序之间间距进行的一种测度结果,它不仅能将事物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。它可以进行加减运算,但不能计算乘除。没有绝对零点,即可以以一个任意数为0点。如山峰相对于地面的高度、天气温度等。

(4) 定比变量(数据)。它是对个体数量特征绝对量大小的测度结果,除能区分类别、排序、比较大小、求出大小差异、可采用加减运算以外,还可以进行乘除运算,具有绝对零点,即“0”表示“没有”或“不存在”,如企业利润、职工人数、银行存款等。

变量之间有时可以转换的。如学生成绩可以用及格、不及格表示,及格、不及格是定序变量;也可以用50分、60分、72分等具体分数表示,称为定比变量。

表 1-1 对某种常用消费品使用情况满意度调查的数据

被访者编号	满意度	性别	月收入(元)
1	满意	男	3 000
2	不满意	男	5 500
3	比较满意	男	3 100
4	比较满意	女	1 800
5	不满意	女	4 700
...	...	...	...
1 000	比较满意	女	1 900

表 1-1 中有三个变量,其中满意度为定性变量、定序变量;性别为定性变量、定类变量;月收入为定量变量、定比变量、连续变量。

三、指标与指标体系

指标是统计指标的简称,它是用来说明现象总体数量方面特征的。一个完整的统计指标应包括指标名称和指标数值两个部分。指标名称是说明现象总体特征的科学概念,它给统计总体某一方面特征以质的规定,如国内生产总值、社会商品零售额、人均国民收入、人口密度等。