



普通高等教育规划教材

统计学基础 第2版

S tatistics Essentials

徐江 李艳 纪晶华 主编



提供电子课件
www.cmpedu.com



机械工业出版社
CHINA MACHINE PRESS

普通高等教育规划教材

统计学基础

第2版

主 编 徐 江 李 艳 纪晶华

副主编 曹忠威

参 编 白 洁

机械工业出版社

本书根据经济管理类各专业统计学课程教学需要而编写,并且针对本门课程开设时间较早的特点,尽量避开烦琐的数学理论推导,本着实用、够用的原则,培养学生运用基本理论和基本方法来分析与解决实际问题的能力。在内容上以知识的实用性为主,理论为辅,注重基础;在形式上则注重图文结合和趣味性。全书共分九章,包括绪论、统计数据的搜集、统计数据的整理和展示、统计数据的指标描述、时间序列分析、统计指数、抽样推断、相关分析与回归分析、国民经济统计基础知识。每章章后配有丰富的思考题和练习题,并在书后附有练习题参考答案。

本书既可作为高等院校经济管理类各专业本科生教材,也可作为高职高专院校相关专业教材,还可作为企事业单位工作人员及工商管理硕士的参考读物。

图书在版编目(CIP)数据

统计学基础/徐江,李艳,纪晶华主编.—2版.—北京:机械工业出版社,2016.2

普通高等教育规划教材

ISBN 978-7-111-52291-1

I. ①统… II. ①徐… ②李… ③纪… 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(2015)第300631号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:曹俊玲 责任编辑:曹俊玲 马碧娟 商红云

责任校对:杜雨霏 封面设计:张静

责任印制:李洋

北京圣夫亚美印刷有限公司印刷

2015年12月第2版第1次印刷

184mm×260mm·15.75印张·387千字

标准书号:ISBN 978-7-111-52291-1

定价:32.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

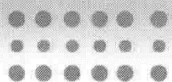
服务咨询热线:010-88379833 机工官网:www.cmpbook.com

读者购书热线:010-88379649 机工官博:weibo.com/cmp1952

教育服务网:www.cmpedu.com

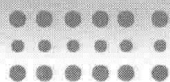
封面无防伪标均为盗版

金书网:www.golden-book.com



前

言



随着科技的日益发展,统计信息遍布社会生活的每一个角落,从国民经济总量、物价变动、人口数、出生率与死亡率、产品数量、商品销售额、自然灾害到房地产价格等,多有涉及。统计学是一门专门提供数据信息的学科,也是经济类、管理类专业的核心课程。学习统计学,掌握基本的统计理论与方法,具备一定的搜集、整理和分析统计数据的能力,是对经济类、管理类专业学生的基本要求。

本书主要根据经济类、管理类专业统计学课程教学的需要而编写,既可作为高等院校经济管理类各专业本科生教材,也可作为高职高专院校相关专业教材,还可作为企事业单位工作人员及工商管理硕士的参考读物。全书共分九章,包括绪论、统计数据的搜集、统计数据的整理和展示、统计数据的指标描述、时间序列分析、统计指数、抽样推断、相关分析与回归分析、国民经济统计基础知识。

本书针对本门课程开设时间较早的特点,尽量避开烦琐的数学理论推导,强调实际应用,让读者掌握常用的统计学原理和方法,提高依据实际问题进行统计设计和调查的能力,以及依据统计结果处理实际问题的能力。在写作思路,本书从统计数据出发,以统计数据的搜集、整理、分析和解释为主线,由描述统计到推断统计组织全书结构。在写作方法上,本书力求简明扼要,深入浅出,理论联系实际。为便于读者学习和掌握教材内容,本书每章后配有大量思考题和练习题,并在书后附录部分提供练习题参考答案。

本书由徐江、李艳、纪晶华主编。按篇章顺序,徐江编写第一~五章,并负责全书统稿;李艳编写第六、七章,并负责全书总体架构的设计;纪晶华负责编写第八章;曹忠威负责编写第九章和各章章后思考题和练习题;白洁负责编写练习题参考答案。在编写过程中,长春工业大学杨晶教授和吉林财经大学马敏娜教授提出了很多宝贵意见,在此表示由衷的感谢。

本书配有电子课件,凡使用本书作为教材的教师,可登录机械工业出版社教育服务网(www.cmpedu.com)注册后下载。

由于编者水平有限,书中难免存在不足之处,敬请广大读者在使用本书的过程中给予批评指正,以便再版时修正。

编 者

目 录

前言		第六章 统计指数	119
第一章 绪论	1	第一节 指数的概念和分类	119
第一节 统计概述	2	第二节 综合指数	121
第二节 统计学的发展与分类	5	第三节 平均指数	127
第三节 统计学的基本概念	7	第四节 指数体系与因素分析	130
思考题	10	第五节 平均指标指数	134
练习题	10	第六节 常用的经济指数	137
第二章 统计数据的搜集	12	思考题	147
第一节 统计数据的来源	12	练习题	147
第二节 统计调查方案与问卷设计	14	第七章 抽样推断	151
第三节 统计数据的搜集方法	19	第一节 抽样推断的基本问题	151
第四节 课程实验	22	第二节 抽样误差	155
思考题	24	第三节 抽样估计	164
练习题	24	第四节 假设检验	168
第三章 统计数据的整理和展示	26	第五节 样本单位数的确定	179
第一节 统计数据的审核	26	第六节 课程实验	182
第二节 统计分组	27	思考题	185
第三节 次数分布	29	练习题	186
第四节 数据的图表展示	34	第八章 相关分析与回归分析	189
第五节 课程实验	41	第一节 相关关系的概念与分类	189
思考题	48	第二节 相关分析的内容与方法	191
练习题	48	第三节 回归分析	195
第四章 统计数据的指标描述	50	第四节 课程实验	202
第一节 总量指标与相对指标	50	思考题	209
第二节 数据的集中趋势	57	练习题	209
第三节 数据的离中趋势	68	第九章 国民经济统计基础知识	213
第四节 偏态与峰度	77	第一节 国民经济统计概述	213
第五节 课程实验	81	第二节 国民经济统计的基本分类	215
思考题	83	第三节 国民经济统计的主要指标	219
练习题	83	附录	225
第五章 时间序列分析	86	附录 A 练习题参考答案	225
第一节 时间序列的分析指标	86	附录 B 标准正态分布表	234
第二节 长期趋势的测定	98	附录 C 正态分布分位数表	235
第三节 季节变动的测定	105	附录 D t 分布表	237
第四节 循环变动和不规则变动的测定	109	附录 E χ^2 分布表	238
第五节 课程实验	112	附录 F F 分布表	239
思考题	115	参考文献	246
练习题	115		

第一章

绪 论

《2014 年国民经济和社会发展统计公报》显示：2014 年年末中国大陆总人口为 136 782 万人，比上年年末增加 710 万人，其中：城镇常住人口为 74 916 万人，占总人口比重为 54.77%。全年出生人口 1 687 万人，出生率为 12.37‰；死亡人口 977 万人，死亡率为 7.16‰；自然增长率为 5.21‰。

国民经济稳定增长。初步核算，全年国内生产总值 636 463 亿元，比上年增长 7.4%。其中，第一产业增加值 58 332 亿元，增长 4.1%；第二产业增加值 271 392 亿元，增长 7.3%；第三产业增加值 306 739 亿元，增长 8.1%。第一产业增加值占国内生产总值的比重为 9.2%，第二产业增加值占比为 42.6%，第三产业增加值占比为 48.2%。

就业继续增加。2014 年年末全国就业人员 77 253 万人，其中，城镇就业人员 39 310 万人。全年城镇新增就业 1 322 万人。2014 年年末城镇登记失业率为 4.09%。全国农民工总量为 27 395 万人，比上年增长 1.9%。其中，外出农民工 16 821 万人，增长 1.3%；本地农民工 10 574 万人，增长 2.8%。

劳动生产率稳步提高。全年国家全员劳动生产率为 72 313 元/人，比上年提高 7.0%。

价格水平涨幅较低。全年居民消费价格比上年上涨 2.0%，其中食品价格上涨 3.1%。固定资产投资价格上涨 0.5%。工业生产者出厂价格下降 1.9%。工业生产者购进价格下降 2.2%。农产品生产者价格下降 0.2%。

70 个大中城市新建商品住宅销售价格月同比上涨城市个数上半年各月均为 69 个，下半年月同比上涨城市个数逐月减少，12 月份为 2 个，月同比价格下降城市个数增加至 68 个。

财政收入稳定增长。全年全国一般公共预算收入 140 350 亿元，比上年增加 11 140 亿元，增长 8.6%，其中税收收入 119 158 亿元，增加 8 627 亿元，增长 7.8%。

外汇储备略有增加。年末国家外汇储备 38 430 亿美元，比上年年末增加 217 亿美元。全年人民币平均汇率为 1 美元兑 6.142 8 元人民币，比上年升值 0.8%。

当看到或听到这些统计数据时，人们常常会思考许多问题：统计数据对人们的生活有用吗？这些统计数据是如何获得的？我们所获得的数据准确吗？如果数据不准确，那么误差是怎么产生的呢？

本章将介绍统计学的一些基本问题，包括统计学的含义、统计数据及统计学的分类，以及统计学中常用的基本概念等。

第一节 统计概述

一、统计的含义

统计是社会生产力发展的必然产物,是适应经济发展、国家管理的需要而产生和发展起来的。说到“统计”一词,首先人们想到的通常是一系列的数字、各种报表。但实际上并不止于此。“统计”一词一般包含三个层面的含义,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,并据以揭示其内在的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,形成了一个完整的认识过程。

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,运用科学的方法搜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,它是统计调查的继续,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,对认识的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的过程。

(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,统计理论来源于统计实践,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,统计理论方法体系不断完善和发展。

二、统计学的定义

统计学是一门搜集、整理和分析统计数据的科学。统计学属于方法论的科学,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。

我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,是就研究对象的数量资料(quantitative data),从事搜集(collecting)、显示(presenting)、分析(analyzing)及解释(interpreting),借以表现其真相的一种学问。”

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

数量多少是指现象的数量规模、总量水平。例如,2014年年末我国总人口为136 782万人,2014年全年社会消费品零售总额为262 394亿元、2014年我国全年粮食产量为60 709.9万t、全年国内生产总值为636 463亿元,分别说明了我国2014年年末的人口规模,2014年我国的消费总量、粮食生产总量、经济发展总量。

2. 现象之间的数量关系

例如,发展速度、结构、比例、差异程度、平均水平等,如2014年我国全年国内生产总值比上年增长7.4%,第一、二、三产业增加值占国内生产总值的比重分别为9.2%、42.6%和48.2%,2014年我国人口出生率为12.37‰,死亡率为7.16‰,自然增长率为5.21‰,出生人口性别比为115.88:100。

3. 质量互变的数量界限

例如,营销中的盈亏平衡点,居民生活贫困线,企业规模大、中、小的数量界限,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?

例如,我们都做过掷硬币的小游戏,知道向上抛一枚硬币,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,我们会发现某种粮食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、差异性和总体性。

1. 数量性

数量性是统计研究对象的基本特点，只有客观现象本身是可以数字度量的，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，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，差异为统计提供了用武之地。

例如，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，各种商品的价格不同，同一种商品也会因为规格、花色、产地、品牌的不同而不同，有的价格是上升的，如食品类，有的是下降的，如电子产品、通信类等，也有的价格是不变的，所以国家要对其进行统计。

3. 总体性

从总体上研究客观现象的数量方面，是统计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，一个家庭生男孩还是生女孩是不确定的，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了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、经济学，要掌握和分析大量有关信息就必然会用到统计方法。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。这些资料和信息取得，需要运用统计手段去搜集有关数据，经过分析研究，得出科学结论。例如，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。

在经济生活和日常生活中，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。在网络购物中，我们点击某个商家的店铺，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量子统计,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,客户可以比较、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学去解释周围所发生的事情,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 H. G. 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,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,人口有 1 355 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,由于工业、农业、贸易、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以下两条主线展开:

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 年)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、质量、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 年)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,如出生的男婴多于女婴,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并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,概率论日益成熟。19世纪中叶,比利时的阿道夫·凯特勒(Adolphe Quetelet, 1796—1874),在其《社会物理学》一书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描述统计学已转向推断统计学,1907年,英国人戈塞特(William Sealey Gosset, 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,迅速成为估计参数的重要方法。英国科学家弗朗西斯·高尔顿(Francis Galton, 1822—1911)提出了相关与回归方法。英国统计学者皮尔逊(Karl Pearson, 1857—1936)发展了适合度检验,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,并与他人合作发展了假设理论。美国学者瓦尔德(A. Wald, 1902—1950)提出了决策理论和序贯抽样方法。美国化学家威尔科克森(Frank Wilcoxon, 1892—1965)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。由于受计算机、信息论等现代科学技术的影响,一方面,统计学的研究领域进一步扩大,如多元统计分析、时间数列分析、贝叶斯统计、非参数统计、数据挖掘等。另一方面,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

1. 描述统计学与推断统计学

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两大类,即描述统计学和推断统计学。

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。描述统计学是表述、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,我们所搜集或遇到的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。这就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。

尽管描述统计学能够完成统计的基本任务,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,使人们不可能掌握总体的全部资料。例如,要了解水库里有多少条鱼、原始森林里有多少根木材、某台电视机的使用寿命是多少、某项投资是否能盈利等。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的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,则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,描述统计学是整个统计学的基础和统计研究工作的起点,它为统计研究工作搜集可靠的数据和有效的样本信息。而推断统计学则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。推断统计学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。

描述统计学与推断统计学虽然方法不同,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,推断统计学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。

2. 理论统计学与应用统计学

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,用于探索各个领域的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。例如,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、时间数列分析、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

统计学的概念很多,其中有几个概念是经常要用到的,有必要单独加以介绍。

一、总体与样本

(一) 总体

总体是我们所研究的全部对象的全体。构成总体的每一个个体单位,称为个体或者总体

单位。例如，我们要研究某班学生的学习情况时，则这个班级的所有学生就构成统计总体，每个学生便是总体单位；进行全国人口普查时，全国人口是总体，其中的每个人是总体单位；某地进行工业企业调查时，该地所有工业企业是总体，其中每个工业企业是总体单位；对工业品出厂价格变化情况进行调查时，所有出厂工业产品是总体，每一种出厂的工业产品是总体单位。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。

总体根据其所包含的单位数目是否可数，分为有限总体和无限总体。总体中含有有限的个体称为有限总体，如全国的人口、工业企业等。总体中含有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。

统计总体具备三个基本特征，即同质性、大量性和差异性。

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。例如，研究我国人口状况时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：

第一层含义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的研究，而不是对个体的研究。当然，只有在对许多个体的分析、综合、研究的基础上，才能揭示出总体的数量特征。

第二层含义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做出具有一定可靠性的描述和估计。

3. 差异性

差异性是指总体中个体单位的具体表现不同，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的变化而变化。当研究工业企业时，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，也就是说总体中的个体单位应当具有足够多的数量，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。

（二）样本

样本也称子样，是指从总体中抽取的一部分个体，样本是总体的缩影，是总体的代表。以样本的调查或观察结果来推断总体的数量特征，是抽样调查的目的。例如，日常生活中的抽血化验。

一个样本所包含的单位数称为样本容量。

二、标志与指标

1. 标志

标志是指说明总体单位(个体)特征或属性的名称。例如,某市一家民营企业作为总体单位,其“企业性质”“企业类型”“生产能力”“职工人数”“工资总额”等都是标志,工人作为总体单位,则他们的性别、工种、文化程度、技术等级、工资、婚姻状况等都是标志。

标志按是否可以用数值表示来划分,可分为品质标志和数量标志两种类型。品质标志是说明总体单位属性特征的名称,如“企业性质”“企业类型”等,品质标志只能用文字而不能用数值表示。数量标志是说明总体单位数量特征的名称,如“生产能力”“年产值”“销售收入”“职工人数”等。

2. 指标

指标是统计指标的简称,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。例如,国内生产总值、社会商品零售额、人均国民收入、人口密度、居民可支配收入等都是比较严格的概念。

指标数值说明现象在具体时间、地点、条件下在该特征方面的具体表现。例如,2014年年末全国就业人员 77 253 万人,其中城镇就业人员 39 310 万人;全年城镇新增就业 1 322 万人;城镇登记失业率为 4.09%。全国农民工总量为 27 395 万人,比上年增长 1.9%。其中,外出农民工 16 821 万人,增长 1.3%;本地农民工 10 574 万人,增长 2.8%。

统计指标按其反映总体现象的内容不同而分为质量指标和数量指标。前者反映现象间的内在联系和比例关系,后者则用来说明现象的总量或规模。例如,平均工资、平均学习成绩、单位产品成本、商品流转次数、利润率、劳动生产率等就是质量指标,其表现形式为相对数或平均数;而社会总产值、工资总额、商品销售额、利润总额等就是数量指标,其表现形式为绝对数。

3. 标志与指标的关系

统计指标与统计标志既有明显的区别,又有密切的联系。两者的主要区别在于:

(1) 指标是说明总体特征的;而标志则是说明总体单位特征的。
(2) 统计指标必须是可量的;统计标志却未必都是可量的。因为品质标志的标志表现就不是用数值,而是用文字表示的;数量标志的标志表现是用数值表示的。

(3) 统计指标具有综合性,如许多人的平均年龄、平均工龄、平均工资、工资总额等;而统计标志一般不具有综合性,例如,一个人的年龄、工龄、工资等,不是指标而是标志。

指标与标志的联系,主要有以下两点:

(1) 许多统计指标的指标数值是从总体单位的数量标志值汇总而来的。例如,一个市的工业总产值是从它所属的各工业企业的总产值汇总而来的,一个省的粮食总产量是从各个县的粮食总产量汇总而来的。

(2) 指标与标志之间存在着变换关系。由于研究目的不同,原来的统计总体若变成总体单位,则相应的统计指标也就变成标志了。反过来亦然。例如,要了解我国粮食生产情况,则我国的粮食总产量是指标,而某省的粮食总产量是标志。现在要了解某省的粮食生产

情况,则该省的粮食总产量是指标,而其中某县的粮食总产量就是标志。现在改为研究某县的粮食生产情况,则该县的粮食总产量是指标,而其中某乡(镇)的粮食总产量就是标志。

三、变异、变量与变量值

所谓变异,就是总体各个单位中标志的具体表现。例如,经济类型有国有经济、集体经济和私营经济等,性别表现为男和女,工人数可以有 500 人、1 000 人等。变异是普遍客观存在的,并且决定了总体的数量特征。有变异,才需要统计。人们就是通过总体的变异实现对总体数量特征的认识的。

所谓变量,是指总体单位的数量标志或者总体的统计指标。变量的数值表现就是变量值,即可变数量标志的不同取值。例如,某工厂工人的月工资额有 2 560 元、3 650 元、3 780 元等,这里的“月工资额”为变量,各个工人工资额的具体数量是变量值。

变量可分为连续变量和离散变量。连续变量的各个变量值是连续不断的,相邻的两个数值之间可以做无限分割。例如,工人的身高、体重等都是连续变量,其数值可以细算到若干位小数,如某人身高 1.772 5m,也可以只取三位小数,即 1.773m。离散变量的各个变量值则以整数断开,而且只能用整数表示。例如,企业数、职工人数、牲畜头数以及设备数等。



思考题

1. 什么是统计?一般有几理解?
2. 为什么说有变异才会有统计?
3. 什么是标志和指标?两者有何区别与联系?
4. 统计研究的基本方法有哪些?



练习题

一、单项选择题

1. 要了解某企业职工的文化水平情况,则总体单位是()。
 - A. 该企业的全部职工
 - B. 该企业每一个职工的文化程度
 - C. 该企业的每一个职工
 - D. 该企业全部职工的平均文化程度
2. 总体与总体单位不是固定不变的,是指()。
 - A. 随着客观情况的变化发展,各个总体所包含的总体单位数也是在变动的
 - B. 随着人们对客观认识的不同,对总体与总体单位的认识也是有着差异的
 - C. 随着统计研究目的与任务的不同,总体和总体单位可以相互转化
 - D. 客观上存在的不同总体和总体单位之间,总是存在着差异
3. 下列变量中,属于连续变量的是()。
 - A. 大中型企业个数
 - B. 大中型企业的职工人数
 - C. 大中型企业的利润额
 - D. 大中型企业拥有的设备台数
4. 下列标志中,属于数量标志的是()。
 - A. 学生性别
 - B. 学生年龄
 - C. 学生专业
 - D. 学生住址
5. 下列标志中,属于品质标志的是()。

- A. 工人性别
- B. 工人年龄
- C. 工人体重
- D. 工人工资
- 6. 下列属于数量指标的是 ()。
- A. 劳动生产率
- B. 废品量
- C. 单位产品成本
- D. 资金利润率
- 7. 下列属于质量指标的是 ()。
- A. 销售总量
- B. 工资总额
- C. 平均工资
- D. 上缴利润额
- 8. 变异是指 ()。
- A. 标志的具体表现不同
- B. 标志和指标各不相同
- C. 总体的指标各不相同
- D. 总体单位的标志各不相同

二、多项选择题

- 1. 下列标志中,属于数量标志的有 ()。
- A. 性别
- B. 工种
- C. 工资
- D. 民族
- E. 年龄
- 2. 下列属于质量指标的有 ()。
- A. 产品合格率
- B. 废品量
- C. 单位产品成本
- D. 资金利润率
- E. 上缴利润额
- 3. 在说明和表现问题方面,其正确的定义有 ()。
- A. 标志是说明总体单位特征的
- B. 指标是说明总体特征的
- C. 变异是可变标志的差异
- D. 变量是可变的数量标志
- E. 标志值是变量的数值表现
- 4. 变量按其取值是否连续可分为 ()。
- A. 确定性变量
- B. 随机性变量
- C. 连续变量
- D. 离散变量
- E. 常数
- 5. 下列有关品质标志和数量标志的说法,正确的有 ()。
- A. 数量标志可以用数值表示
- B. 品质标志也可用数量表示
- C. 数量标志不可以用数值表示
- D. 品质标志不可以用数值表示
- E. 两者都可以用数值表示

三、综合题

要调查某商场销售的洗衣机情况,试指出总体、总体单位是什么?试举若干品质标志、数量标志、数量指标和质量指标。

第二章

统计数据的搜集

统计数量规律性的确认与识别依赖于大量的统计数据，它是统计分析与推断的基础。那么统计数据如何取得呢？这便是统计数据搜集的内容。

我们常常会关心一些社会问题，比如今年全世界发生的自然灾害造成的损失是多少？今年我国居民消费价格指数是上升还是下降？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？人口老龄化的社会对人们生活的影响有哪些？以及人们对延迟退休的态度等。

为了回答这些问题，就要搜集相关数据进行分析。本章将对上述问题加以讨论。

第一节 统计数据的来源

一般来说，统计数据主要来源于两种渠道：一是通过直接调查和科学实验，我们称之为数据的直接来源；二是引用公开出版的统计数据，即数据是由别人通过调查或者实验的方式搜集的，使用者只是找到它们并加以利用，我们称之为数据的间接来源。

一、统计数据的直接来源

统计数据的直接来源是指通过专门组织的调查和科学实验搜集数据。科学实验是取得自然科学数据的主要手段，如医学、生物学等的数据都是通过实验和观察取得的。本章主要介绍专门组织的调查，专门调查是取得社会经济数据的重要手段。在计划经济中，主要是政府统计部门的统计调查；在市场经济中，还有一些专门进行调查的机构做市场调查等。

（一）统计调查的意义与要求

统计调查是指根据统计研究的目的和要求，利用各种统计调查的方式方法，采集统计数据的工作过程。统计调查所采集的统计数据，既包括原始数据，也包括次级数据。原始数据也称初级数据、直接数据或第一手数据，是指向调查单位直接搜集的尚待汇总整理的、需要由个体过渡到总体的统计数据。例如，为了了解某市某一时期内工业生产情况及企业的经济效益，我们必须对该市所有工业企业的产品产量、产值、产品质量、劳动生产率、原材料、能源消耗量、成本、流动资金和利润额等资料进行登记和采集，这些数据就是原始数据。次级数据也称间接数据或第二手数据，是指已经经过加工整理，由个体过渡到总体，能够在一定程度上说明总体特征的统计数据。如上例中，已经有人对其进行一定的加工整理，但还不能完全满足要求，还须做进一步的加工整理，这样的数据就是次级数据。一切次级数据都是