

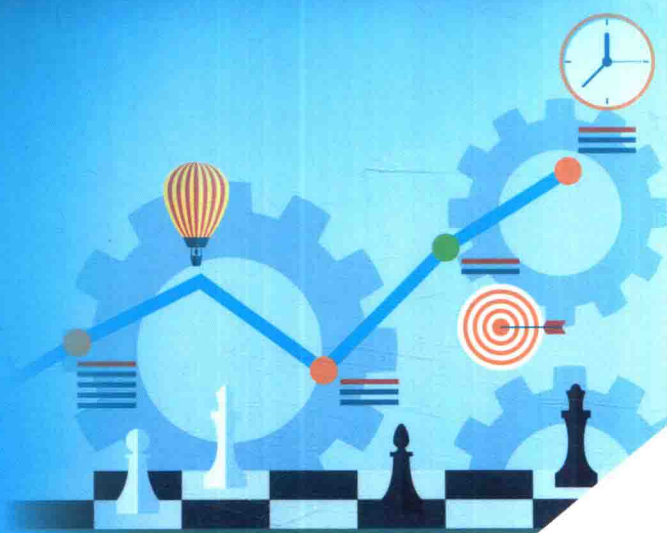
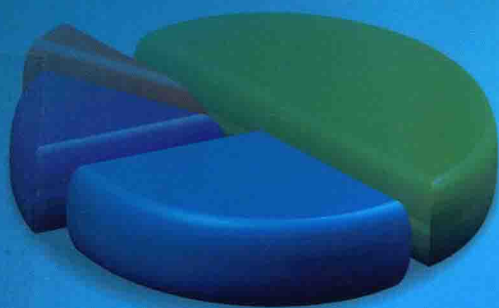
普通高等教育规划教材

统计学



Statistics

张雪飞 郭淑艳 李伟 赵福生 编



张外借

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

张雪飞 郭淑艳 李伟 赵福生 编



机械工业出版社

本书全面、系统地阐述了统计学的基本概念、原理和方法，并介绍了Excel在统计运算和分析中的应用。主要内容包括：绪论、统计数据的调查与整理、综合指标分析、时间序列分析、统计指数分析、概率统计基础、抽样推断与假设检验、相关与回归分析、统计决策与评价、国民经济生产统计和现代企业统计等。

本书可作为普通高等院校经济类、管理类各专业“统计学”课程的教材，也可作为统计工作者的参考用书。

图书在版编目（CIP）数据

统计学/张雪飞等编. —北京：机械工业出版社，2017.4
普通高等教育规划教材
ISBN 978-7-111-56460-7

I. ①统… II. ①张… 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（2017）第 064642 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：潘蕊 封面设计：张静

责任印制：李昂

中国农业出版社印刷厂印刷

2017 年 6 月第 1 版第 1 次印刷

184mm×260mm·19.5 印张·471 千字

标准书号：ISBN 978-7-111-56460-7

定价：43.00 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

统计是一种社会调查活动，不论是宏观社会的整体调查研究，还是微观事物的观察分析，都需要统计。

统计学在中国的发展有了很大的飞跃，呈现出多元化发展的趋势。统计学不仅自身的发展领域很宽广，而且在计算机科学、信息科学、经济学、管理学、金融工程等领域都有广泛的应用，并与之有力结合，共同发展。

统计学在社会科学与自然科学中应用的区别主要体现在统计试验上。在自然科学中统计试验是很容易做到的，而在社会科学中统计试验是很难进行的。人们对现实社会经济现象的一次观察，只能被看作是试验的一次观察，而对其进行重复观察几乎是不可能的。这就要求我们在为高等院校经济类、管理类各专业编写统计学教材时，要选择小样本、非重复试验等特有的统计方法为主要内容，以使对“统计学”课程的学习与其他专业课的学习和实践能够很好地结合，从而使得我们的统计学教材与概率统计等教材有明显的不同。

本书在各章采用本章要点的方式简要地介绍各章主要内容；以导入案例的方式引导学生学会综合应用本章的统计知识；以思考与练习的方式来巩固所学的知识。

本书对各类统计方法的含义、原理、应用原则以及 Excel 的应用例解等方面进行了系统的讲解。全书共 11 章，包括绪论、统计数据的调查与整理、综合指标分析、时间序列分析、统计指数分析、概率统计基础、抽样推断与假设检验、相关与回归分析、统计决策与评价、国民经济生产统计以及现代企业统计等。通过学习，学生能较好地掌握基本的统计思想和各种定量分析方法，帮助其提高分析问题的能力。本书由张雪飞统稿，具体编写分工如下：张雪飞编写第五、十、十一章；郭淑艳编写第一、三章，第七章第一、二、三、五节；李伟编写第二、四章，第八章第一、二节；赵福生编写第六章，第七章第四节，第八章第三、四节，第九章。

本书内容通俗易懂、简明实用，可作为普通高等院校经济类、管理类及其他人文社科类相关专业“统计学”课程的教材，也可作为统计工作者的参考用书。

本书在编写过程中参阅了大量文献资料，在此对有关专家和作者表示由衷的感谢。

由于作者水平有限，本书中难免有错误和疏漏之处，恳请广大读者批评指正。

编 者

目 录

前 言

第一章 绪论	1
本章要点	1
导入案例	1
第一节 统计的产生与发展	2
第二节 统计学的含义和研究对象	3
第三节 统计学研究方法	6
第四节 统计学的基本概念	8
第五节 统计的职能	13
思考与练习	14
案例分析	14
第二章 统计数据的调查与整理	18
本章要点	18
导入案例	18
第一节 统计数据的调查	19
第二节 统计数据的整理	32
第三节 统计数据的显示	38
思考与练习	42
案例分析	43
第三章 综合指标分析	50
本章要点	50
导入案例	50
第一节 总量指标	51
第二节 相对指标	55
第三节 平均指标	62
第四节 标志变异指标	74
第五节 Excel 的应用例解	78
思考与练习	81
案例分析	83

第四章 时间序列分析	86
本章要点	86
导入案例	86
第一节 时间序列概述	87
第二节 水平分析	90
第三节 速度分析	94
第四节 长期趋势分析	96
第五节 季节变动分析	102
第六节 Excel 的应用例解	106
思考与练习	107
案例分析	108
第五章 统计指数分析	109
本章要点	109
导入案例	109
第一节 统计指数的概念和种类	110
第二节 综合指数的编制和计算	113
第三节 平均指数和平均指标指数的因素分析	116
第四节 指数体系和因素分析	121
第五节 统计指数的应用	126
第六节 Excel 的应用例解	133
思考与练习	135
案例分析	137
第六章 概率统计基础	143
本章要点	143
导入案例	143
第一节 概率基础	143
第二节 随机变量及其分布	150
第三节 抽样分布	157
第四节 大数定律和中心极限定理	161
第五节 Excel 的应用例解	163
思考与练习	165
案例分析	166
第七章 抽样推断与假设检验	168
本章要点	168
导入案例	168

第一节 抽样推断概述	169
第二节 抽样误差	174
第三节 参数估计	178
第四节 假设检验	188
第五节 Excel 的应用例解	194
思考与练习	196
案例分析	197
第八章 相关与回归分析	200
本章要点	200
导入案例	200
第一节 相关与回归分析的基本概念	200
第二节 相关分析	204
第三节 回归分析	207
第四节 Excel 的应用例解	217
思考与练习	220
案例分析	221
第九章 统计决策与评价	222
本章要点	222
导入案例	222
第一节 统计决策概述	222
第二节 统计决策的方法	226
第三节 统计综合评价概述	241
第四节 统计综合评价方法	242
第五节 Excel 的应用例解	257
思考与练习	258
案例分析	260
第十章 国民经济生产统计	261
本章要点	261
导入案例	261
第一节 国民经济生产统计概述	265
第二节 国内总产出统计	267
第三节 国内生产总值统计	269
第四节 国民经济生产统计分析	272
思考与练习	273
案例分析	273

第十一章 现代企业统计	274
本章要点	274
导入案例	274
第一节 企业投入统计	275
第二节 企业产出统计	284
思考与练习	288
案例分析	290
附 录	292
附录 A 标准正态分布表	292
附录 B t 分布表	294
附录 C χ^2 分布表	295
附录 D F 分布表	298
参考文献	302

第一章

绪论



本章要点

1. 明确统计的含义、方法及职能。
2. 重点理解统计的基本概念及各概念之间的区别与联系。

导入案例

《2015 年国民经济和社会发展统计公报》摘录

在信息高度发达的今天，我们几乎每天都能从电视、网络、报纸、杂志等各种渠道接触到大量的统计数据。例如，国家统计局 2016 年 2 月最新发布的《2015 年国民经济和社会发展统计公报》显示：

全年国内生产总值 676 708 亿元，比上年增长 6.9%。其中，第一产业增加值 60 863 亿元，增长 3.9%；第二产业增加值 274 278 亿元，增长 6.0%；第三产业增加值 341 567 亿元，增长 8.3%。第一产业增加值占国内生产总值的比重为 9.0%；第二产业增加值占比为 40.5%；第三产业增加值占比为 50.5%，首次突破 50%。全年人均国内生产总值 49 351 元，比上年增长 6.3%。全年国民总收入 673 021 亿元。年末全国总人口 137 462 万人，比上年年末增加 680 万人，其中城镇常住人口 77 116 万人，占总人口比重（常住人口城镇化率）为 56.10%，比上年年末提高 1.33 个百分点。全年出生人口 1 655 万人，出生率为 12.07‰；死亡人口 975 万人，死亡率为 7.11‰；自然增长率为 4.96‰。全国人户分离的人口 2.94 亿人，其中流动人口 2.47 亿人。人均预期寿命 76.34 岁。年末全国就业人员 77 451 万人，其中城镇就业人员 40 410 万人。全年城镇新增就业 1 312 万人。年末城镇登记失业率为 4.05%。全国农民工总量 27 747 万人，比上年增长 1.3%。其中，外出农民工 16 884 万人，增长 0.4%；本地农民工 10 863 万人，增长 2.7%。

.....

当听到或阅读到这些统计数据时，你是否会思考以下问题：统计数据对我们的生活有用吗？这些统计数据是如何得来的？统计数据与将要学习的统计学之间有着怎样的关系？等等。所以我们需要了解“什么是统计”以及“统计能解决什么问题”。

第一节 统计的产生与发展

统计学是统计实践活动中产生和发展起来的,它将在今后的统计实践中进一步完善和发展。

一、统计实践的产生和发展

从历史上看,统计实践活动远远早于统计学的产生,它是随着人类社会经济的发展,随着治国和管理的需要而产生和发展起来的,至今已有四五千年的历史。统计活动起源于原始社会末期,当时人们就用结绳记事、结绳计量的方法来对狩猎品和采集到的野果的数量等进行简单计数。

在奴隶社会,统治者为了维护自己的统治地位,实现对内统治、对外战争的需要,进行征兵、徭役、征税,开始了人口、土地和财产的统计活动。据晋皇甫谧(215—282)的《帝王世纪丛书》记载,我国早在四千多年前的夏朝,为了治国治水的需要,就进行了初步的国情统计,全国分为9个州,人口为1 355万人,土地为2 438万顷。在国外,古希腊、古罗马时代,已开始了人口和居民财产的统计活动。在埃及,早在建造金字塔时,为了征集所需财物和征用劳力,就在全中国进行了人口、劳力和财产的调查。

封建社会,统计已略具规模。封建君主和精明的政治家日益意识到统计对于治国强邦的重要作用。

封建社会末期,特别是进入资本主义社会以后,社会生产力迅速发展,社会分工日益深化,交通、航运、贸易日趋发达,国际市场逐步形成。这时,不仅政府需要有包括人口、土地、财富、赋税、军事等方面国情国力的统计,各类企业主、商人为了经营管理和争夺市场,也需要各种商业情报和市场信息。统计逐步扩展到工业、农业、贸易、银行、保险、交通、邮电、海关等部门,并且出现了专业的统计机构和研究组织。统计逐步成为社会分工中的一个独立部门和专业。

二、统计理论的产生和发展

统计实践经过封建社会末期的丰富和发展,客观上需要从理论上加以概括和总结,于是统计学应运而生。从17世纪下半叶开始,欧洲出现了一些统计理论著述。在资本主义社会,随着统计实践活动的发展,众多学者开始总结丰富的统计实践经验,纷纷著书立说,使得统计学在理论和方法上不断丰富。由于这些统计学者所处的历史条件不同,研究的领域不同,因此形成了不同的统计学派。其中主要的统计学派有:

1. 国势学派

国势学派也称记述学派,产生于17世纪的德国,代表人物是康令(H. Cornig, 1606—1681)和阿亨瓦尔(G. Achenwall, 1719—1772),代表作是《近代欧洲各国国势学概论》。他们在德国的大学开设了“国势学”课程。国势学派把统计学理解为,国家重要事项的记述,他们搜集大量的实际材料,分门别类地记述国家组织、土地、人口、军队、居民职业、宗教、资源、财产等社会经济情况,注重事件的文字记述,但缺乏数量的分析。由此对比后人所认为的统计学,国势学派所理解的统计学是不符合要求的,存在着名不副实的缺陷。然

而,“统计学”一词就是从国势学派演变而来的。

2. 政治算术学派

政治算术学派产生于17世纪的英国,代表人物是威廉·配第(W. Petty, 1623—1687),代表作是《政治算术》。该书运用大量的数据资料,对英国、荷兰、法国的政治事项、社会结构、经济状况、军事力量等国情国力首次进行了解剖分析。书中运用了具体的数字、重量、尺度等方法,对社会经济等现象及其相互关系做了系统的数量运算与对比分析,为统计学的创立奠定了方法论基础。该学派成功地将经济理论和统计分析方法结合在一起,形成了既不同于数学,又不同于政治经济学的新学科。因此,马克思称威廉·配第为“政治经济学之父”,在某种程度上说,他也是统计学的创始人。政治算术学派的另一代表人物是约翰·格朗特(John Graunt, 1620—1674)。在他的论著《对死亡率公报的自然观察和政治观察》中,首次通过大量观察研究发现了一系列人口统计规律:新生儿性别比例;死亡率男性高于女性;一般疾病和事故的死亡率较稳定,而传染病的死亡率波动较大;编制了初具规模的“生命表”,对年龄死亡率与人口寿命进行了分析。尽管该学派的学者运用了统计学的理论与方法,但都没有使用“统计学”这个名称。

3. 数理统计学派

数理统计学派产生于19世纪的比利时,代表人物是阿道夫·凯特勒(A. Quetelet, 1796—1874),著有《社会物理学》《概率论书简》等著作。他将法国古典概率理论引入到统计学中,认为无论是自然现象还是社会现象都有规律可循,一切事物都受大数定律支配。统计学是可应用于任何学科研究的方法,并开创性地应用了许多统计方法。到了19世纪60年代,凯特勒把统计学发展过程中的三个主要源泉,即国势学派、政治算术学派和古典概率学派加以结合,使之成为近代应用数理统计学。其后,由葛尔登(F. Galton, 1822—1911)、皮尔生(K. Pearson, 1857—1936)、戈塞特(W. S. Gosset, 1876—1937)和费雪(R. A. Fisher, 1890—1962)等统计学家,经过多方面的研究,提出并发展了相关与回归、假设检验、 χ^2 分布、 t 分布等理论,使数理统计学逐渐成为一门独立的学科。

4. 社会统计学派

社会统计学派产生于19世纪后半叶,其先驱者是德国大学教授克尼斯(K. G. A. Knies, 1821—1897),主要代表人物为恩格尔(C. L. E. Engel, 1821—1896)和梅尔(G. V. Mayer, 1841—1925)。该学派认为,统计学的研究对象是社会现象的数量方面,描述社会现象内部的联系、相互关系以及发展规律。统计应当包括资料的搜集、整理以及对其的分析研究。社会统计学派认为,全面调查,包括人口普查和工农业调查,居于重要地位;以概率论为理论基础的抽样调查,在一定的范围内具有实际意义和作用。社会统计学派在理论上比政治算术学派更加完善,在时间上比数理统计学派更早成熟,因此在国际统计学界有较大的影响。

第二节 统计学的含义和研究对象

一、统计与统计学的含义

统计作为一种社会实践活动已有悠久的历史。在外语中,“统计”一词与“国家”一词来自同一词源。因此,可以说自从有了国家就有了统计实践活动。最初,统计只是为统治者

管理国家的需要而搜集资料,弄清国家的人力、物力和财力,作为国家管理的依据。

今天,“统计”一词已被人们赋予多种含义,因此很难给出一个简单的定义。在不同场合,“统计”一词可以具有不同的含义。它可以是指统计数据的搜集活动,即统计工作;也可以是指统计活动的结果,即统计数据资料;还可以是指分析统计数据的方法和技术,即统计学。

1. 统计工作

统计工作,是搜集、整理、分析和研究统计数据资料的工作过程。统计工作在人类历史上出现得比较早。随着历史的发展,统计工作逐渐发展和完善起来,使统计成为国家、部门、事业和企业、公司、个人及科研单位认识与改造客观世界和主观世界的一种有力工具。统计工作,可以简称为统计。例如,某统计师在回答自己的工种时,会说我是干统计的。这里所说的统计指的就是统计工作。

2. 统计数据资料

统计数据资料,是统计工作活动进行搜集、整理、分析和研究的主体及最终成果。不管是个人、集体和社会,还是国家、部门和事业、企业、公司及科研机构,都离不开统计数据资料。个人要进行学习、工作和家政管理,需要对有关的统计数据资料进行搜集和分析,以指导自己的学习、工作和生活;公司和企业要管理好生产和销售,必须进行市场调研、生产控制、质量管理、人员培训、成本评估等,这就需要对有关的生产资料、市场资料、成本资料、人员资料、质量数据等进行搜集、整理、分析和研究;国家要进行经济建设和社会发展,更离不开有关国民经济和社会发展的统计资料,像我国的十年规划,2010年的国民生产总值比2000年翻一番,就需要我国有关国民生产总值的历史数据资料和相关数据资料,需要有关各国的相关统计数据资料,以此为基础进行分析和决策。还有像国家统计局编辑、中国统计出版社出版的每年一册的《中国统计年鉴》,以及国家统计局每年年初公布的《国民经济与社会发展统计公报》等都是统计数据资料,也可称为统计。例如,电视台、电台和报纸杂志所说的“据统计”中的统计指的就是统计数据资料。

3. 统计学

一般来说,统计学是对研究对象的数据资料进行搜集、整理、分析和研究,以显示其总体的特征和规律性的学科。统计学的研究对象是客观事物的数量特征和数据资料。统计学是以搜集、整理、分析和研究等统计技术为手段,对所研究对象的总体数量关系和数据资料去伪存真、去粗取精,从而达到显示、描述和推断被研究对象的特征、趋势和规律性的目的。统计学,亦可简称为统计。例如,我们所学的课程——统计课,实际指的就是统计学课程。

早期统计学的学派之一,政治算术学派的创始人威廉·配第和约翰·格朗特,首先在其著作中使用统计数字和图表等方法来分析研究社会、经济和人口现象,这不仅为人们进一步认识社会提供了一种新的方法和途径,也为统计学的发展奠定了基础。

目前,随着统计方法在各个领域的应用,统计学已发展成为具有多个分支学科的大家族。因此,要给统计学下一个被大家普遍接受的定义是十分困难的。在本书中,我们对统计学做如下解释:统计学是一门搜集、整理和分析统计数据的方法科学,其目的是探索数据的内在数量规律性,以达到对客观事物的科学认识。

统计数据的搜集是取得统计数据的过程,它是进行统计分析的基础。离开了统计数据,统计方法就失去了用武之地。如何取得所需的统计数据是统计学研究的内容之一。

统计数据的整理是对统计数据的加工处理过程,目的是使统计数据系统化、条理化,符合统计分析的需要。数据整理是介于数据搜集与数据分析之间的一个必要环节。

统计数据的分析是统计学的核心内容,它是通过统计描述和统计推断的方法探索数据内在规律的过程。

可见,统计学是一门有关统计数据的科学,统计学与统计数据有着密不可分的关系。在英文中,“statistics”一词有两个含义:当它以单数名词出现时,表示作为一门科学的统计学;当它以复数名词出现时,表示统计数据或统计资料。从中可以看出,统计学与统计数据之间有着密不可分的关系。首先,统计学是由一套搜集和处理统计数据的方法所组成的,这些方法来源于对统计数据的研究,目的也在于对统计数据的研究。统计数据不用统计方法去分析也仅仅是一堆数据而已,无法得出任何有益的结论。其次,统计数据不是指单个的数字,而是由多个数据构成的数据集。单个的数据显然用不着统计方法去进行分析,仅凭一个数据点,我们也不可能得出事物的规律,只有经过对同一事物进行多次观察或计量得到大量数据,才能利用统计方法探索出其内在的规律性。

二、统计学的研究对象

一般来说,统计学的研究对象是自然、社会客观现象总体的数量关系。正是因为统计学的这一研究的特殊矛盾,使它成为一门万能的科学。不论是自然领域,还是社会经济领域,客观现象总体的数量方面,都是统计学所要分析和研究的。

统计学研究对象的特点有如下几点:

1. 数量性

统计学的研究对象是自然、社会经济领域中现象的数量方面,这一特点是统计学(定量分析学科)与其他定性分析学科的分界线。数量性是统计学研究对象的基本特点,因为,数字是统计的语言,数据资料是统计的原料。一切客观事物都有质和量两个方面,事物的质与量总是密切联系、共同规定着事物的性质。没有无量的质,也没有无质的量。一定的质规定着一定的量,一定的量也表现为一定的质。但在认识的角度上,质和量是可以区分的,可以在一定的质的情况下,单独地研究数量方面,通过认识事物的量进而认识事物的质。因此,事物的数量是我们认识客观现实的重要方面,通过分析研究统计数据资料,研究和掌握统计规律性,就可以达到我们统计分析研究的目的。例如,要分析和研究国民生产总值,就要对其数量、构成及数量变化趋势等进行认识,这样才能正确地分析和研究国民生产总值的规律性。

2. 总体性

统计学的研究对象是自然、社会经济领域中现象总体的数量方面,即统计的数量研究是对总体普遍存在着的事实进行大量观察和综合分析,得出反映现象总体的数量特征和资料规律性。自然、社会经济现象的数据资料和数量对比关系等一般是在一系列复杂因素的影响下形成的。在这些因素当中,有起着决定和普遍作用的主要因素,也有起着偶然和局部作用的次要因素。由于种种原因,在不同的个体中,它们相结合的方式和实际发生的作用都不可能完全相同。所以,对于每个个体来说,就具有一定的随机性质,而对于有足够多数个体的总体来说又具有相对稳定的共同趋势,显示出一定的规律性。例如,对工资的统计分析,我们并不是要分析和研究个别人的工资,而是要反映、分析和研究一个地区、一个部门、一个

企业事业单位的总体的工资情况和显示出来的规律性。统计研究对象的总体性,是从个体的实际表现的研究过渡到对总体的数量表现的研究的。例如,工资统计分析,要反映、分析和研究一个地区的工资情况,就要先从每个职工的工资开始统计,然后再综合汇总得到该地区的工资情况,只有从个体开始,才能对总体进行分析研究。研究总体的统计数据资料,不排除对个别事物的深入调查研究,但它是为了更好地分析研究现象总体的统计规律性。

3. 具体性

统计学的研究对象是自然、社会经济领域中具体现象的数量方面。即它不是纯数量的研究,是具有明确的现实含义的,这一特点是统计学与数学的分水岭。数学是研究事物的抽象空间和抽象数量的科学,而统计学研究的数量是客观存在的、具体实在的数量表现。统计学研究对象的这一特点,也正是统计工作必须遵循的基本原则。正因为统计的数量是客观存在的、具体实在的数量表现,它才能独立于客观世界,不以人们的主观意志为转移。统计资料作为主观对客观的反映,必然是存在是第一性,意识是第二性,存在决定意识,只有如实地反映具体的已经发生的客观事实,才能为我们进行统计分析研究提供可靠的基础,才能分析、探索和掌握事物的统计规律性。虚假的统计数据资料是不能成为统计数据资料的,因为它违背了统计学研究对象的这一特点。

4. 变异性

统计学研究对象的变异性是指构成统计学研究对象的总体各单位,除了在某一方面必须是同质的以外,在其他方面又要有差异,而且这些差异并不是由某种特定的原因事先给定的。也就是说,总体各单位除了必须有某一共同标志表现作为它们形成统计总体的客观依据以外,还必须在所要研究的标志上存在变异的表现。否则,就没有必要进行统计分析研究了。例如,高等院校这个统计对象,除了都是从事高等教育的教学活动这一共同性质之外,各高等院校在隶属主管部门、院校性质、招生规模、专业设置等各方面又有差异。工人作为统计数据资料对象,每个工人在性别、年龄、工龄、工作性质、工资等方面是会有不同表现的。这样,统计分析研究才能对其表现出来的差异探索统计规律性。

第三节 统计学的研究方法

统计学根据研究对象的性质和特点,形成了自己专门的研究方法,这些基本方法是试验设计法、大量观察法、统计描述法和统计推断法。

一、试验设计法

统计是要分析数据的,但首先需要考察的是,数据的来源是否合适,试验采集的数据是否符合分析的目的要求。由于安排不科学,使试验数据不能反映现象的真实情况,或不能用以估计总体的数量特征,那么接着一系列分析工作也就会是白费工夫了。例如,要比较某农作物 A 品种和 B 品种的收获率高低,分别在两地段播种 A 品种和 B 品种,结果获得 A 品种的单位面积产量高于 B 品种的数据。如果根据这个数据判断 A 品种优于 B 品种,那么这个结论就太不可靠了。原因是影响收获率高低的因素不但有种子品种的差异,还有土地区位、肥沃程度等差异,所以我们需要事先做出安排,使试验结果数据的差异中排除可控因素(土地)的差异,而显示不可控因素(品种)的差异。所谓试验的统计设计,就是指设计试

验的合理程序,使得搜集的数据符合统计分析方法的要求,以便得出有效的客观的结论。它主要适用于自然科学研究和工程技术领域的统计数据搜集。

试验的统计设计要遵循以下三个基本原则:

1. 重复性原则

重复性原则即允许在相同条件下重复多次试验。如果只能一次试验,则所得的数据作为总体的估计量精度就很差,这时试验的误差等于观察的误差,观察误差可能是试验误差的结果,很难用观察的数据来代表总体情况。多次重复试验的好处是显然的:其一可以获得更加精确的效应估计量;其二,可以获得试验误差的估计量。这些都是提高估计精度或缩小误差范围所需要的。

2. 随机性原则

随机性是指在试验设计中,对试验对象的分配和试验次序都是随机安排的。这种安排可以使可控的影响因素作用均匀化,突出不可控影响因素的作用。例如,在种子品种的试验中如果不是将A品种固定在甲地段、B品种固定在乙地段,而是两地段随机地选择不同品种多次重复试验,则可以断定这种安排在不同品种收获率的差异中,土地因素的影响大大减少,而品种因素的影响大大提高了。所以随机化原则是试验设计的重要原则。

3. 区组化原则

区组化原则即利用类型分组技术,对试验对象按有关标志顺序排队,然后依次将各单位随机地分配到各处理组,使各处理组组内标志值的差异相对扩大,而处理组组间的差异相对缩小,这种试验设计安排称为随机区组设计。这样就可以提高处理组的估计精度。

二、大量观察法

大量观察法是统计学所特有的方法。所谓大量观察法,是指对所研究的事物的全部或足够数量进行观察的方法。社会现象或自然现象都受各种社会规律或自然规律相互交错作用的影响。在现象总体中,个别单位往往受偶然因素的影响,如果任选其中之一进行观察,则其结果不足以代表总体的一般特征;只有观察全部或足够的单位并加以综合,影响个别单位的偶然因素才会相互抵消,现象的一般特征才能显示出来。大量观察的意义在于可使个体与总体之间在数量上的偏误相互抵消。

大量观察法的数学依据是大数定律。大数定律是随机现象的基本规律。大数定律的一般概念是:在观察过程中,每次取得的结果不同,这是由偶然性所致的,但大量、重复观察结果的平均值却几乎接近确定的数值。狭义的大数定律,就是指概率论中反映上述规律性的一些定理,表述平均数的规律性与随机现象的概率关系。

大数定律的本质意义在于经过大量观察,把个别的、偶然的差异性相互抵消,而必然的、集体的规律性便显示出来。例如,当我们观察个别家庭或少数家庭的婴儿出生情况时,生男生女的比例极为不同,有的是只生男不生女,有的是只生女不生男,有的是女多男少,有的是男多女少,然而经过大量观察,男婴、女婴的出生数则趋向均衡。也就是说,观察的次数越多,离差的差距就越小,或者说频率出现了稳定性。这就表明,同质的大量现象是有规律的,尽管个别现象受偶然性因素的影响出现偏差,但观察数量达到一定程度时就呈现出规律性,这就是大数定律的作用。

三、统计描述法

统计描述是指对由试验或调查而得到的数据进行登记、审核、整理、归类、计算出各种能反映总体数量特征的综合指标,并加以分析从中抽出有用的信息,用表格或图像把它表示出来的方法。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。统计描述也是对客观事物认识的不断深化过程。它通过对分散无序的原始资料的整理归纳,运用分组法、综合指标法和统计模型法得到现象总体的数量特征,揭露客观事物内在数量规律性,达到认识的目的。

分组法是研究总体内部差异的重要方法,它通过分组可以研究总体中不同类型的性质以及它们的分布情况,如产业的经济类型及其行业分布情况;可以研究总体中的构成和比例关系,如三次产业的构成、生产要素的比例等;可以研究总体中现象之间的相关依存关系,如企业经营规模和利润率之间的关系等。

综合指标法是指运用各种统计指标来反映和研究客观总体现象的一般数量特征和数量关系的方法。通过综合指标的计算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。

统计模型法则是综合指标法的扩展。它是指根据一定的理论和假定条件,用数学方程去模拟客观现象相互关系的一种研究方法。利用这种方法,可以对客观现象和过程中存在的数量关系进行比较完整和全面的描述,凸显所研究的综合指标之间的关系,从而简化客观存在的复杂的其他关系,以便利用模型对所关心的现象变化进行评估和预测。

四、统计推断法

统计在研究现象的总体数量关系时,需要了解的总体对象的范围往往是很大的,有时甚至是无限的,而由于经费、时间和精力等各种原因,以致有时在客观上只能从中观察部分单位或有限单位进行计算和分析,根据局部观察结果来推断总体。例如,要说明一批灯泡的平均使用寿命,只能从该批灯泡中抽取一小部分进行检验,推断这一批灯泡的平均使用寿命,并给出这种推断的置信程度。这种在一定置信程度下,根据样本资料的特征,对总体的特征做出估计和预测的方法称为统计推断法。统计推断是现代统计学的基本方法,在统计研究中得到了极为广泛的应用,它既可以用于对总体参数的估计,也可以用作对总体某些分布特征的假设检验。从这种意义上来说,统计学是在不确定条件下做出决策或推断的一种方法。

第四节 统计学的基本概念

统计科学和其他科学一样,在论述本门科学的理论与方法时,要运用一些专门的概念,有些是基本的、常用的,有些是局部的,在论述专门问题时使用局部的概念,局部的概念在后面章节讲解。本节只就几个基本的、常用的概念加以阐述。

一、统计总体与总体单位

统计总体简称总体,是指客观存在的、在同一性质基础上结合起来的许多个别单位的整体。构成总体的这些个别单位称为总体单位。例如,所有的工业企业就是一个总体,这是因

为在性质上每个工业企业的经济职能是相同的,即都是从事工业生产活动的基本单位,这就是说,它们是同性质的。这些工业企业的集合就构成了统计总体。对于该总体来说,每一个工业企业就是一个总体单位。

总体可以分为有限总体和无限总体。总体所包含的单位数是有限的,称为有限总体,如人口数、企业数、商店数等。总体所包含的单位数是无限的,称为无限总体,如连续生产的某种产品的生产数量、大海里的鱼资源数等。对有限总体可以进行全面调查,也可以进行非全面调查。但对无限总体只能抽取一部分单位进行非全面调查,据以推断总体。

确定总体与总体单位,必须注意以下两个方面:

(1) 构成总体的单位必须是同质的,不能把不同质的单位混在总体之中。例如,研究工人的工资水平,就只能将靠工资收入的职工列入统计总体的范围。同时,也只能对职工的工资收入进行考察,对职工由其他方面取得的收入就要排除,这样才能正确反映职工的工资水平。

(2) 总体与总体单位具有相对性,随着研究任务的改变而改变。同一单位可以是总体也可以是总体单位。例如,要了解全国工业企业职工的工资收入情况,那么全部工厂是总体,各个工厂是总体单位。如果旨在了解某个企业职工的工资收入情况,则该企业就成了总体,每位职工的工资就是总体单位。

二、统计标志

(一) 标志和标志表现

统计标志简称标志,是指统计总体各单位所具有的共同特征的名称。从不同角度考察,每个总体单位可以有許多特征。例如,每个职工可以有性别、年龄、民族、工种等特征。这些都是职工的标志。

标志表现是标志特征在各单位的具体体现。职工的性别是女,年龄为32岁,民族为汉族等,这里“女”“32岁”“汉族”就是性别、年龄、民族的具体体现,即标志表现。

(二) 标志的分类

1. 标志按变异情况可分为不变标志和变异标志

当一个标志在各个单位的具体表现都相同时,这个标志称为不变标志;当一个标志在各个单位的具体表现有可能不同时,这个标志称为可变标志或变异标志。例如,中国第五次人口普查规定:“人口普查的对象是具有中华人民共和国国籍并在中华人民共和国国境内常住的人。”按照这一规定,在作为调查对象的人口总体中,国籍和在国境内居住是不变标志,而性别、年龄、民族、职业等则是变异标志。不变标志是构成统计总体的基础,因为至少必须有一个不变标志将各总体单位联结在一起,才能使它具有“同质性”,从而构成一个总体。变异标志是统计研究的主要内容,因为如果标志在各总体单位之间的表现都相同,那就没有进行统计分析研究的必要了。

2. 标志按其性质可分为品质标志和数量标志

品质标志表示事物的质的特性,是不能用数值表示的,如职工的性别、民族、工种等。数量标志表示事物的量的特性,是可以数值表示的,如职工的年龄、工资、工龄等。品质标志主要用于分组,将性质不同的总体单位划分开来,便于计算各组的总体单位数,计算结