



21世纪高等职业教育“十一五”规划教材

Yingyong Tongjixue

应用统计学

主编 钟新联 张海平



21世纪高等职业教育“十一五”规划教材

应用统计学

主编
副主编

钟新联
张建军
王海平
周宇
李锋

Yingyong Tongjixue

华中科技大学出版社

中国·武汉

图书在版编目(CIP)数据

应用统计学/钟新联 张海平 主编. —武汉:华中科技大学出版社,2009 年 12 月
ISBN 978-7-5609-5782-1

I. 应… II. ①钟… ②张… III. 应用统计学-高等学校:技术学校-教材 IV. C8

中国版本图书馆 CIP 数据核字(2009)第 205239 号

应用统计学

钟新联 张海平 主编

策划编辑:吴 晗
责任编辑:吴 晗
责任校对:刘 竣

封面设计:旻昊图文空间
责任监印:周治超

出版发行:华中科技大学出版社(中国·武汉)

武昌喻家山 邮编:430074 电话:(027)87542624

录 排:武汉正风图文照排中心

印 刷:华中科技大学印刷厂

开本:787mm×1092mm 1/16

印张:20.25

字数:486 000

版次:2009 年 12 月第 1 版

印次:2009 年 12 月第 1 次印刷

定价:34.00 元

ISBN 978-7-5609-5782-1/C · 122

(本书若有印装质量问题,请向出版社发行部调换)

总 序

世界职业教育发展的经验和我国职业教育发展的历程都表明,职业教育是提高国家核心竞争力的要素。职业教育这一重要作用和地位,主要体现在两个方面:其一,职业教育承载着满足社会需求的重任,是培养为社会直接创造价值的高素质劳动者和专门人才的教育。职业教育既是经济发展的需要,又是促进就业的需要。其二,职业教育还承载着满足个性需求的重任,是促进以形象思维为主的具有另类智力特点的青少年成才的教育。因此,职业教育既是保证教育公平的需要,又是教育协调发展的需要。

这意味着,职业教育不仅有着自己的特定目标——满足社会经济发展的人才需求以及与之相关的就业需求,而且有着自己的特殊规律——促进不同智力群体的个性发展以及与之相关的智力开发。

长期以来,由于我们对职业教育作为一种类型教育的规律缺乏深刻的认识,加之学校职业教育又占据绝对主体地位,因此职业教育与经济、与企业联系不紧,导致职业教育的办学未能冲破“供给驱动”的束缚;由于与职业实践结合不紧密,职业教育的教学也未能跳出学科体系的框架,所培养的职业人才,其职业技能的专深不够,职业工作的能力不强,与行业、企业的实际需求以及我国经济发展的需要相距甚远。实际上,这也不利于个人通过职业这个载体实现自身所应有的生涯发展。

因此,要遵循职业教育的规律,强调校企合作、工学结合,在“做”中“学”,在“学”中“做”,就必须进行教学改革。职业教育教学应遵循“行动导向”的教学原则,强调“为了行动而学习”、“通过行动来学习”和“行动就是学习”的教育理念,让学生在由实践情境构成的以过程逻辑为中心的行动体系中获取过程性知识,去解决“怎么做”(经验)和“怎么做更好”(策略)的问题,而不是在由专业学科构成的以架构逻辑为中心的学科体系中去追求陈述性知识,只解决“是什么”(事实、概念等)和“为什么”(原理、规律等)的问题。由此,作为教学改革核心的课程,就成为职业教育教学改革成功与否的关键。

当前,在学习和借鉴国内外职业教育课程改革成功经验的基础之上,工作过程导向的课程开发思想已逐渐为职业教育战线所认同。所谓工作过程,是“在企业里为完成一件工作任务并获得工作成果而进行的一个完整的工作程序”,是一个综合的、时刻处于运动状态但结构相对固定的系统。与之相关的工作过程知识,是情境化的职业经验知识与普适化的系统科学知识的交集,它“不是关于单个事务和重复性质工作的知识,而是在企业内部关系中将不同的子工作予以连接的知识”。以工作过程逻辑展开的课程开发,其内容编排以典型职业工作任务以及实际的职业工作过程为参照系,按照完整行动所特有的“资讯、决策、计划、实施、检查、评价”结构,实现学科体系的解构与行动体系的重构,实现于变化的具体的工作过程之中获取不变的思维过程完整性的训练,实现实体性技术、规范性技术通过过程性技术的物化。

近年来,教育部在高等职业教育领域,组织了我国职业教育史上最大的职业教育师资培训项目——中德职教师资培训项目和国家级骨干师资培训项目。这些骨干教师通过学习、了解、接受先进的教学理念和教学模式,结合中国的国情,开发了更适合我国国情、更具有中国特色的职业教育课程模式。

华中科技大学出版社结合我国正在探索的职业教育课程改革,邀请我国职业教育领域的专家、企业技术专家和企业人力资源专家,特别是接受过中德职教师资培训或国家级骨干教师

培训的高职院校的骨干教师,为支持、推动这一课程开发应用于教学实践,进行了有意义的探索——工作过程导向课程的教材编写。

华中科技大学出版社的这一探索,有两个特点。

第一,课程设置针对专业所对应的职业领域,邀请相关企业的技术骨干、人力资源管理者以及行业著名专家和院校骨干教师,通过访谈、问卷和研讨,由企业技术骨干和人力资源管理者提出职业工作岗位对技能型人才在技能、知识和素质方面的要求,结合目前我国高职教育的现状,共同分析、讨论课程设置存在的问题,通过科学合理的调整、增删,确定课程门类及其教学内容。

第二,教学模式针对高职教育对象的智力特点,积极探讨提高教学质量的有效途径,根据工作过程导向课程开发的实践,引入能够激发学习兴趣、贴近职业实践的工作任务,将项目教学作为提高教学质量、培养学生能力的主要教学方法,把适度够用的理论知识按照工作过程来梳理、编排,以促进符合职业教育规律的新的教学模式的建立。

在此基础上,华中科技大学出版社组织出版了这套工作过程导向的规划教材。我始终欣喜地关注着这套教材的规划、组织和编写的过程。华中科技大学出版社敢于探索、积极创新的精神,应该大力提倡。我很乐意将这套教材介绍给读者,衷心希望这套教材能在相关课程的教学发挥积极作用,并得到读者的青睐。我也相信,这套教材在使用的过程中,通过教学实践的检验和实际问题的解决,不断得到改进、完善和提高。我希望,华中科技大学出版社能继续发扬探索、研究的作风,在建立具有我国特色的高等职业教育的课程体系的改革之中,做出更大的贡献。

是为序。

教育部职业技术教育中心研究所

学术委员会秘书长

《中国职业技术教育》杂志主编

中国职业技术教育学会

理事、教学工作委员会副主任

职教课程理论与开发研究会主任

姜大源研究员教授

2008年7月15日

前 言

通过对数据的统计，人们可以认识社会、了解社会。社会各界对统计信息的需求，从量上来说将日益增多，涉及的面将越来越广；从质上来说要求越来越精，期望也越来越高。统计信息作为一种政府的公共信息资源，是社会经济信息的主体，在社会信息资源中占有绝对的比重，这是其他任何社会信息资源无法相比的。在市场经济条件下，各种商务活动的竞争非常激烈，如何在激烈的市场竞争中取胜，是每一个市场主体都必须认真和冷静思考的问题。统计更是为国家制定政策、计划，进行宏观调控，提供信息、咨询、监督等多功能服务。市场经济越发达，统计的作用越广泛、越重要，统计工作的建设也就越需要加强，统计知识就越需要普及。

统计学是高职高专财经类专业的核心课程之一。为培养应用型人才，达到学以致用之目的，有必要掌握统计知识，并能应用统计方法进行数据分析与研究。

本书包括了统计学的基本理论、基本方法和技能，以及实践应用三个层次的内容，通过对统计数据、统计数据的采集、统计数据的整理与显示等内容的学习，培养学生敏锐的定量思维能力、分析能力和实际应用能力。本书力求用通俗易懂的语言、直观明了的统计图表和国家统计部门发布的最新统计数据为实例来讲述统计学的基本理论和方法，这既有利于教师组织教学，又有利于学生及时了解社会经济发展的基本情况。

本书由钟新联、张海平担任主编，张建军、周宇、王艳和李锋担任副主编。在本书的编写过程中，作者参考了不少同行的文献，引用了一些好的案例，在此谨向有关作者表示诚挚的感谢。

限于作者的水平和经验，本书内容难免有不妥和差错之处，敬请读者指正。

编 者

2009. 9

目 录

第一章 统计数据	(1)
第一节 统计数据概述	(1)
第二节 统计数据的计量	(8)
第三节 统计数据的质量	(11)
思考与练习	(17)
第二章 统计数据的采集	(21)
第一节 统计数据采集概述	(21)
第二节 统计数据采集方案的设计	(28)
第三节 统计数据采集的具体方法	(34)
第四节 统计数据采集的组织方式	(37)
第五节 统计调查误差与控制	(46)
第六节 统计调查报告	(48)
思考与练习	(61)
第三章 统计数据的整理与显示	(66)
第一节 统计数据的整理	(67)
第二节 统计分组	(69)
第三节 分配数列	(75)
第四节 统计汇总	(82)
第五节 统计数据的显示	(86)
思考与练习	(98)
第四章 静态数据	(103)
第一节 总量数据	(104)
第二节 相对数据	(107)
第三节 集中趋势指标	(115)
第四节 离中趋势指标	(134)
思考与练习	(144)
第五章 动态数列	(151)
第一节 动态数列概述	(151)
第二节 动态数列的水平指标	(154)
第三节 动态数列的速度指标	(162)
第四节 现象变动的趋势分析	(169)
思考与练习	(190)
第六章 统计指数	(199)
第一节 统计指数概述	(199)
第二节 总指数的编制	(204)

第三节 指数体系与因素分析·····	(217)
思考与练习·····	(236)
第七章 抽样推断·····	(242)
第一节 抽样推断概述·····	(242)
第二节 抽样方案的设计·····	(252)
第三节 抽样误差·····	(254)
第四节 抽样估计·····	(265)
第五节 样本容量的确定·····	(267)
思考与练习·····	(272)
第八章 相关分析与回归分析·····	(277)
第一节 相关分析概述·····	(277)
第二节 相关分析方法·····	(282)
第三节 简单线性回归分析·····	(292)
思考与练习·····	(307)
附录 正态分布概率表·····	(311)
参考文献·····	(313)

第一章

统 计 数 据

▶▶ 能力目标

- ◆ 识别统计数据的表现形式、统计数据的内容质量和统计数据的表述质量;
- ◆ 区分总体与总体单位,区分统计标志与统计指标、变量;
- ◆ 熟悉统计数据及其类型,熟悉各种数据的计量尺度及其特点、关系。

▶▶ 重点与难点

本章重点:统计和统计数据的基本概念;统计数据的表现形式和计量尺度及其特点;总体与总体单位;标志与指标。

本章难点:标志、指标、变量及其分类;定类尺度、定序尺度、定距尺度和定比尺度的特点及其关系。

▶▶ 任务导入

在人们的日常生活中,“统计”无时不有,无处不在。我们几乎每天都能从实际工作、生活中感悟到或从各种媒体上了解到不同国家、地区、一般市场主体等方面的相关信息,这些信息都是相关人员深入客观实际进行采集、整理和分析研究的结果。如第29届北京奥运会进行期间每天奖牌榜上公告的奖牌信息,沪、深证券交易所提供的证券交易信息等;又如经济普查、人口普查、“5.12”汶川地震灾情调查、贫困大学生状况调查、市场物价调查等发布的信息;车市、房市,世界金融危机对经济发展的影响、索马里海盗劫持过往商船的相关信息等。这些信息资料主要是通过不同的统计数据的形式表现出来的,因此,我们有必要对什么是统计数据、统计数据的类型、统计数据的表现形式、统计数据的计量及统计数据的质量等有所了解。

阅读2009年2月26日国家统计局发布的《中华人民共和国2008年国民经济和社会发展统计公报》,进一步领会到统计数据的作用及魅力,据此就能了解我国经济和社会发展中主要的、具体的信息。

第一节 统计数据概述

一、统计的含义

按《大不列颠百科全书》的定义:“统计学是收集、分析、表述和解释数据的科学。”这一定义揭示了统计学是一门处理数据的方法和技术的科学。《中华人民共和国统计法》第二条规定:

“统计的基本任务是对国民经济和社会发展情况进行统计调查、统计分析,提供统计资料和统计咨询意见,实行统计监督。”这是我国法律对统计基本任务的规定,也是统计法律对我国统计职能的规定。

在我国,“统计”一词最早仅具有“汇总”的意思。早在原始公社时期就有了结绳记事的计量方法。《周易正义》中记载:“事大,大结其绳;事小,小结其绳;结之多少,随物众寡。”这可谓统计之萌芽,但它仅是一种总量的计算。从字面上理解,统计就是统而计之。可见,统计最基本的含义就是计算某类事物的总量。现代之统计已远远超出了原来的范畴,一般认为,“统计”一词包括三个方面的含义,即统计工作、统计数据和统计学。

(1) 统计工作,即统计实践活动。它是指利用各种科学方法,对社会、经济及自然现象的总体数量特征进行采集、整理和分析等工作过程的总称。例如,国家统计局进行国情国力与社会发展情况的调查(如人口普查、经济普查等),各级统计部门对所属地区的工业、农业与进出口贸易等方面的数据资料进行采集、整理、分析,营销人员为推销产品需要对该产品的市场供求状况进行调查与分析,等等,这些方面的工作都属于统计实践活动。

需要指出的是,统计实践活动作为一项具体的经济工作,是与现代管理密切联系在一起的。只要从事现代化的经济活动,并强化科学、高效的管理,就必然存在统计实践活动,而且管理越规范、越科学,对统计工作的要求也就越高。管理者只有掌握准确、及时、全面的统计信息,才能做出可行的计划、科学的决策,从而实施有效的管理。如果没有准确、及时、全面的信息做支撑,实施科学管理、提高经济效益将无从谈起。

(2) 统计数据,即统计信息。它是统计工作者通过统计调查所采集的各项原始数据及其派生数据,是统计工作的成果。

(3) 统计学,即统计科学。它是一门认识社会和自然的方法论科学,是关于如何采集、整理和分析统计数据的理论、原则、方式方法的科学,包括社会经济统计学和数理统计学等。

显而易见,统计工作、统计数据和统计学之间既有区别,又有联系。三者之间的联系主要表现为以下三点。

第一,过程和结果的关系。统计实践活动的目的是为了获取统计数据;统计数据的取得必须依靠统计实践活动,而且,统计工作的好坏直接影响统计数据的数量和质量。

第二,实践与理论的关系。一方面,统计学是统计实践活动的经验总结和理论概括;另一方面,统计学又指导统计实践活动如何有效进行。统计学与统计工作存在一种理论与实践的双向作用关系。

第三,统计实践活动先于统计学发展起来。自从有了国家以后,统计实践活动就伴随着社会、政治、经济的发展和国家的管理需要而发展起来;而统计学的出现却只有近四百年的历史。

统计的含义如图 1-1 所示。

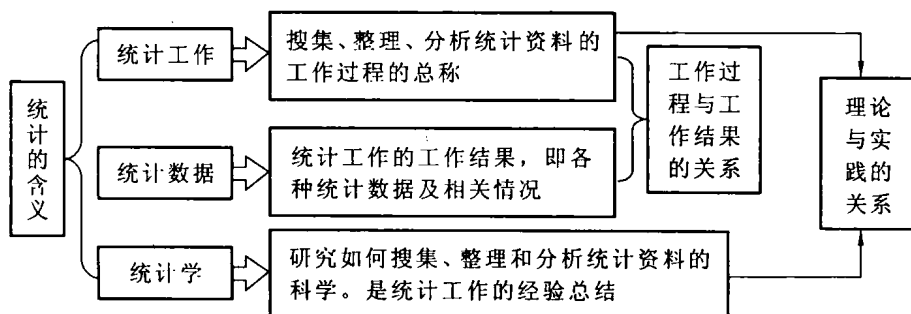


图 1-1 统计含义示意图

二、数据的概念

数据是对客观事物的符号表示,是用于表示客观事物的规模、水平和质量等未经加工的原始素材,如数字、图形、表格、符号、字母等。或者说,数据是通过对事物的观察而得来的事实和概念,是关于现实世界中的事件、行为、其他对象或概念的描述。

在计算机科学中,数据是指所有能输入到计算机并被计算机程序处理的符号的介质的总称,即用于输入计算机进行处理,具有一定意义的数字、字母、符号和模拟量等的统称。

三、统计数据的概念

统计数据是活动中所取得的反映国民经济和社会现象的规模、水平和结构等的各种数字资料以及与之相联系的其他情况的总称,统计数据包括通过统计调查直接获取的原始数据和经过整理、分析后所形成的次级资料,是统计工作的成果或“产品”。统计研究的是客观事物的数量信息,所以它离不开统计数据,统计数据是对客观现象进行计量的结果。统计数据的形式是多种多样的,具体表现为统计表、统计图、统计公报、统计年鉴、统计手册、统计分析报告等。从内涵上讲,包括为管理国家事务、管理经济和文化事务、管理社会事务等所需要的宏观和微观资料;从外延上讲,包括描述过去、剖析现在、预测未来的资料等。

就所采集的统计数据而言,一般要求满足以下六点要求。

(1) 准确性。即所采集的统计数据必须真实地反映客观事实,其误差不能超过统计要求所允许的误差范围。这就要求具有报送统计资料义务的公民、法人和其他组织,不得虚报、瞒报、伪造、篡改统计资料;地方、部门、单位的领导人不得强令、授意统计机构和统计人员实施统计违法行为。

(2) 及时性。即统计数据应及时采集、及时整理、及时报送。

(3) 全面系统性。即所采集的统计数据从“量”和“面”上要符合统计工作的基本要求,能真实地反映客观现象之间的内在联系和发展变化的规律性。

(4) 连续性。即所采集的统计数据在时间上能进行动态对比,不是偶然的、孤立的,更不能是臆造的。

(5) 尊重隐私性。《统计法》第十五条规定:“属于私人、家庭的单项调查资料,非经本人同意,不得泄露。”“统计机构、统计人员对在统计调查中知悉的统计调查对象的商业秘密,负有保密义务。”法律如此规定,主要是基于两方面的考虑:一是个人或家庭的隐私、调查对象的商业秘密,若泄露可能损害被调查者的合法权益;二是这样规定有利于消除被调查者的后顾之忧,增进被调查者对调查者的信任感,使被调查者能够如实提供统计资料。这也是统计工作的国际惯例。

(6) 可分享性。除属于国家机密、商业秘密和个人或家庭的单项调查资料外,统计机构应将采集、整理的统计数据、统计报告、统计公报刊载于公共媒体上,或放置于开放场所中,以便各级党政机构和社会公众查询,以及根据各级党政机构和社会公众的特殊需要,对统计数据进行处理、加工、整理、汇总后供其使用。

四、统计数据的类型

统计数据的表现形式多种多样,如统计表、统计图、统计公报、统计年鉴、统计手册、统计分析报告等。统计数据是采用某种计量尺度对事物进行计量的结果,采用不同的计量尺度会得

到不同类型的统计数据。由于现象中个体的特征存在着不同的类型,其测度的方法也不一样。根据计量学的分类方法,按照对事物计量的精确度不同,可将不同计量尺度的计量结果从低到高、从粗到精分为定类数据、定序数据、定距数据和定比数据四种类型。

(1) 定类数据。定类数据是由定类尺度计量形成的,它表现为类别,即说明事物类别的一个名称,但不区分顺序,这类数据的具体表现就是定类变量值。如人的“性别”具体表现为“男”或“女”,这里的“男”或“女”就是定类数据;又如企业的“经济类型”也是一个定类变量,其变量值具体表现为“国有经济”“集体经济”“个体经济”等,这里的“国有经济”“集体经济”“个体经济”就是定类数据。

(2) 定序数据。定序数据是由定序尺度计量形成的,它表现为类别,即说明事物有序类别的一个名称,但有顺序,这类数据的具体表现是定序变量值。如“产品等级”就是一个定序变量,其具体表现为“一等品”“二等品”“三等品”等,这里的“一等品”“二等品”“三等品”等就是定序数据;又如职工的“文化程度”也是一个定序变量,其具体表现为“小学”“中学”“高中”“大学”等,这里的“小学”“中学”“高中”“大学”就是定序数据。

定类数据和定序数据是说明事物的品质特征,它不能用具体的数值表示,其结果只能表现为类别,称为定性数据或品质数据。

(3) 定距数据。定距数据是由定距尺度计量形成的,它表现为数值,可进行加、减运算。

(4) 定比数据。定比数据是由定比尺度计量形成的,它表现为数值,可进行加、减、乘、除运算。

定距数据和定比数据说明的是现象的数量特征,它能够用数值来表现。由于定距尺度和定比尺度属于同一测度层次,所以可以把该两种数据看作是同一类数据,统称为定量数据或数值型数据。

区分测量的层次和数据类型是十分重要的,因为对不同类型的数据将采用不同的方法来处理和分析。需要指出的是,适用于低层次测量数据的统计方法,也适用于较高层次的测量数据,因为后者具有前者的数学特性。比如,在描述数据的集中趋势时,对定类数据通常是计算众数,对定序数据通常是计算中位数,但对定距和定比数据既可以计算众数和中位数,还可以计算其他统计量。反之,适用于高层次测量数据的统计方法,则不一定能用于较低层次的测量数据,因为低层次数据不具有高层次测量数据的数学特性。例如,对于定距和定比数据可以计算平均数,但对于定类数据和定序数据则不能计算平均数。理解这一点,对于选择统计分析方法是十分有用的。

五、统计数据的承担者

(1) 总体单位。总体单位是指构成总体的各个基本单位,简称个体,它是各项统计数据最原始的承担者。虽然统计研究的最终目的是为了揭示现象总体的运动过程 and 变化规律,但统计研究都是从对个体单位的相关数据的采集入手的,然后根据统计研究的目的和要求进行加工汇总和分析。如果没有大量个体单位的原始数据为基础,统计研究将成为“无源之水”“无本之木”。根据不同的研究目的和要求,总体单位可以是人、物或事件等。例如,在人口普查中,普查范围内的每个人就是总体单位;在调查我国高校快速发展的情况时,我国的每一所高校就是总体单位。

(2) 统计总体。统计总体是统计所要研究的全体事物,是由客观存在的性质相同的许多个体单位所构成的整体,简称总体。例如,在我国人口普查中,普查范围内的每个人是总体单位,所有接受普查的人口便形成一个统计总体,这个总体中每个人都是有生命的、客观存在的,

而且都是具有中国国籍的人,有了这个总体,我们就可以研究全国人口的各种统计指标,如人口的规模、性别、年龄结构、地区分布等;又如,要研究我国高校快速发展的情况,则全国所有的高校便组成一个总体,其中每所高校都是客观存在的,且都是教育教学单位,据此可以计算出我国高校相应的教职工人数及其结构、在校学生人数、专业结构等指标。这些指标都具有综合性,它是同质总体各单位某一数量标志值的差异综合,即指标数值均由总体中各单位的相应数量标志值汇总而来,且指标数值的大小受各单位相应标志值大小及其变化的影响。可见,凡统计指标都是综合指标,都是说明总体综合数量特征的。

要弄清总体与总体单位所代表的具体内容,还必须明确研究对象与研究内容的关系,研究对象是研究内容的承载体。例如,要研究某校学生的学习情况,研究对象则是该校的全部学生,因而总体也就是该校的全部学生,总体单位则是该校的每一个学生,而其研究内容则是该校学生的学习、生活、纪律、卫生情况和社会实践活动等。

需要明确指出的是,统计研究的社会现象是指总体现象,是对总体的综合数量特征作出科学的归纳和分析。然而,这种研究又是以个别单位为基础的,总是从搜集个体单位的数据开始,然后将每个总体单位的数据进行汇总综合,以反映总体的综合数量特征。这就要求统计总体必须同时具备以下三个方面的特征。

第一,同质性。同质性就是要求构成总体的各个具体单位在所研究的标志方面要有共同的性质,也就是说,总体是由具有某一相同标志表现的基本单位所组成。例如,在研究我国高校快速发展的情况时,全国所有的高校便构成总体,该总体各单位的共性是“从事教育教学工作”,凡从事非教育教学的单位(如企业单位、行政单位等)不能成为该总体的基本单位。可见,同质性是构成总体的基础,是统计研究的前提条件,如果没有这个基础和前提,统计汇总和数据分析就没有任何科学意义。

第二,大量性。总体一定要由大量的个体单位所组成,因为统计研究的目的是要揭示现象发展变化的趋势和规律性,而这种趋势和规律性只能在大量事物的普遍联系中才能表现出来,只有通过大量的偶然性才能表现出必然性来。以2009年为例,不能以个别商品物价的涨跌幅度(如石油价格、汽车价格等)来说明和评估整个商品物价的变动水平,因为影响个别商品价格涨跌的因素是不一样的,带有一定的特殊性和偶然性,只有组成总体的个体单位足够多,这种特殊性和偶然性因素的影响就趋于相互抵消,才有可能显示出必然性来。总体的大量性和个体的差异性紧密联系在一起的,个体单位之间的数据差异越大,构成总体的单位数就应该越多;调查研究对数据要求的精确度越高,调查或观察的总体单位数就要求越多。

第三,变异性。即在同质条件下,要求组成总体的各单位在其他方面的表现又必须不同或不完全相同。例如,在了解我国高校快速发展的情况时,在我国高校这个总体中,每一所高校除了具有“从事教育教学工作”的共同性质外,其他如校园面积、教职工人数、在校学生人数、专业设置、教学设施、教学与科研成果等方面的数量表现是不相同的,也是不可能完全相同的。总体各单位的同质性是进行统计研究的前提条件,而总体各单位之间的变异性才是统计研究的内容。也正是总体各单位之间存在着程度不同的差异,才需要通过统计研究来发现现象变化的原因、趋势和规律。

六、统计数据的主要表现形式

统计数据的表现形式很多,从统计数据承担者的角度可以把统计数据概括为两大类——标志值和统计指标。它们是统计数据的主要表现形式。

1. 标志值

标志是反映总体单位属性或特征的名称。例如,对人口普查中的个体单位“每一个人”来说,有“民族”“性别”“年龄”“文化程度”“职业”等方面的特征,这里的“民族”“性别”“年龄”“文化程度”“职业”等名称就是标志。显而易见,总体单位是标志的承担者,标志是依附于总体单位的。每个总体单位有许多方面的标志,每个标志是从某一特定方面来说明总体单位的属性或特征的。标志按其性质不同可分为品质标志和数量标志。品质标志表明总体单位属性方面特征,具体表现为品质标志表现不能用数值表示,而只能用文字陈述。例如,职工的性别、工种、文化程度等;设备的种类、用途、性能等。数量标志是表明总体单位数量特征的,具体表现为数量标志表现能用数值表示。例如,职工的身高、体重、年龄、工龄、工资等;设备的价值、使用年限等。

标志值就是标志在总体单位上的具体体现,即标志表现。如人口普查中的个体单位某人的“性别”为“男”或“女”、“年龄”为“28岁”或“55岁”等。这里的“性别”“年龄”等是标志名称,而“男”或“女”、“28岁”或“55岁”等就是标志表现,即标志值,它分别回答某人的性别是什么、年龄是多少。任何一项统计工作都要掌握总体单位在特定的时间、地点、条件下实际发生的情况,因此标志值是统计最为关心的问题。品质标志的标志表现即品质标志值是通过定类尺度或定序尺度测度的,它不能用数值表示,只能用文字陈述,其结果只能表现为类别,称为定性数据或品质数据。数量标志的标志表现即数量标志值是通过定距尺度或定比尺度测度的,它能用数值表示,是定量数据的重要组成部分。

2. 统计指标

(1) 统计指标的概念。统计指标是反映总体综合数量特征的范畴,它表明现象总体在具体时间、地点、条件下的综合数量表现。统计指标一般由指标名称和指标数值两个基本部分组成。指标名称反映现象所属的一定范畴,指标数值反映现象在具体环境下所达到的规模、水平及比例关系等。例如,《中华人民共和国2008年国民经济和社会发展统计公报》显示,2008年底全国总人口为132 802万人;全年出生人口1 608万人,出生率为12.14‰;死亡人口935万人,死亡率为7.06‰。其中,“总人口”“出生人口”“出生率”“死亡人口”“死亡率”等都是指标名称,相应的“132 802万人”“1 608万人”“12.14‰”“935万人”“7.06‰”等都是指标数值。指标数值是通过对总体各单位进行调查,然后进行综合汇总和计算而得到的,它是通过定距尺度或定比尺度测度的,其结果只能表现为数值,是定量数据的重要组成部分,人们日常生活中所称的统计数据基本上就是指这类数据,它大量存在于社会经济活动中。

统计指标的基本作用就是用现象的具体数量来反映总体事物存在的客观事实,为统计认识社会,进行定量分析提供坚实的客观基础。统计指标是统计工作的中心问题之一,被形象地称为“统计的语言”。

(2) 统计指标的特征。统计指标具有可量性、综合性、具体性等特征。

① 可量性。统计指标是对现象的某种综合数量特征进行概括而形成的科学范畴,但不是所有作为概括现象的范畴都能形成统计指标,只有能用数字加以计量的范畴才有可能被称为统计指标。例如,进出口贸易总额、从业人数、财政收入或支出额等。对于无法用数字加以计量的范畴,就不可能成为统计指标。例如,股份制企业的组织形式、筹资方式等就不是统计指标。可见,凡是统计指标都能用数字加以计量,所以可量性是统计指标的基本特征,也是社会经济范畴转化为统计指标的前提条件。

② 综合性。统计指标是反映总体综合数量特征的,其数值既是同质总体各单位某一数量

标志值的总计,又是各单位某一数量标志值的差异综合。所以,凡统计指标都是综合指标,都是说明总体综合数量特征的。

③ 具体性。统计指标是反映具体现象在具体时间、地点、条件下的具体数量特征,而不是抽象的现象、概念和数字。可见,统计指标是指特定的现象,包含着特定的时间、空间、内容和计量单位,因而不存在脱离具体内容的统计指标。

(3) 统计指标的设计。统计指标是统计的语言。设计统计指标的目的是为了取得一定的指标数值,使其能客观地反映现象总体的数量特征。在市场经济体制逐渐完善的今天,统计工作如何发挥其应有的作用,关键是看人们如何去发挥统计指标作用,而如何去发挥统计指标的作用又取决于所设计的统计指标是否科学、是否合理、是否可行。一个完整的科学的统计指标的设计,要符合以下几点要求:

第一,指标所反映的总体特征、概念要有科学的理论依据;

第二,指标要有明确的计算口径和计算范围;

第三,指标要有科学的计算方法;

第四,指标要有确切的计量单位。

(4) 统计指标的分类。社会经济现象纷繁复杂、变化多端,用以反映其发展变化的统计指标千差万别,且随时间的变化而变化。为了求得指标的核算和分析符合客观实际和科学性原则,保证统计结果准确无误,有必要对统计指标进行分类。统计指标的分类方法很多,下面介绍其中几种主要的分类方法。

第一,按反映现象的性质不同,指标可分为实体指标和行为指标。实体指标所反映的是具有实物形态的、客观存在的总体的数量特征,如煤炭产量、生猪存栏头数等;行为指标所反映的是某种行为的数量特征,如工伤事故指标、犯罪行为指标、影视收视率等。

第二,按数据的依据不同,指标可分为客观指标和主观指标。客观指标是指可以通过实际度量或计数来取值的、具有具体性和客观性的指标,如上述煤炭产量、生猪存栏头数等实体指标;主观指标是指不可能或难以直接度量或计数取值,而只能凭人们的主观估计、评价取值的指标,如文艺演出、部分奥运会比赛项目中的评委亮分、民意测验结果等。必须指出的是,主观指标虽然不是实际度量或计数得出来的,带有模糊性,但是该结果的得出仍然有一定的标准、规则,并非完全主观臆造。如第 29 届北京奥运会中的体操、跳水等项目的评分,应该说这类指标仍然具有一定的客观性。

第三,按反映的现象的时间状态不同,指标可以分为时期指标和时点指标。

第四,按计量单位的特点不同,指标可分为实物指标和价值指标。实物指标是以实物单位加以计量的指标。所谓实物单位,是根据事物的实物形态及性能特点,由国家统一规定计量单位,主要有自然单位、度量衡单位和标准实物单位等,如米、千克、千瓦·时等计量的指标。价值指标是以货币单位计量,反映现象价值量的指标,如国内生产总值、进出口总额等指标。另外还有一种用劳动时间单位计量的指标,如工时、工日等,这类指标一般在企业内部使用,宏观核算一般不用。

第五,按表现形式不同,指标分为总量数据、相对指标和平均指标。

第六,按反映总体特征的不同,指标可分为数量指标和质量指标。数量指标是指反映现象总体规模大小、水平高低的指标,一般表现为统计绝对数,如人口总数、产品产量等。质量指标是反映现象总体的强度、密度、效果、工作质量等方面的指标,这类指标一般用相对数、平均数等表示,如人口密度、劳动生产率、资金利润率等。

统计中把标志和指标统称为变量,统计数据是统计变量的具体表现。变量可以分为:定类变量,其结果表现为定类数据;定序变量,其结果为定序数据;数值型变量,其结果为定距数据和定比数据。

数值型变量可以表现为不同的具体数值。如人的身高不可能都一样,因此“身高”这一数量标志就成为一个数值型变量,其变量值在不同人身上可以分别表现为 1.72 米、1.81 米、1.68 米等;又如我国各地县的 GDP 是不一样的,则“GDP”也是一个数值型变量,其具体表现可以分别为 2 460 亿元、4 139 亿元、8 287 亿元、6 598 亿元等。这些具体数据就是变量值,一般情况下人们所说的变量与变量值就是指这类变量与变量值。

数值型变量按其取值是否连续,可分为连续变量和不连续变量。连续变量指变量在其取值范围内可取任意值,且任意两个变量值之间可以有无穷多个具体值,不能一一列举。如人的身高、体重;企业的销售收入、利润等。不连续变量又称离散变量,是指变量值只能取整数的变量,相邻两个整数之间不可能有小数连接。如企业的职工人数、设备台数;某地区的人口数、学校数、企业数等。需要说明的是,有些性质上属于连续变量的现象但通常按整数取值,可以把它当作离散变量看待。如年龄、成绩等虽然属连续变量,但一般按整数计数,实际工作中把它当作离散变量来处理。

第二节 统计数据的计量

对统计数据的属性、特征进行分类、标示和计算,称为统计计量或统计量度。例如,对进出口贸易额进行的统计,对工业企业经济效益进行的统计,对居民生活水平进行的统计,也分别可以说是对进出口贸易额进行的计量,对工业企业经济效益进行的计量,对居民生活水平进行的计量。

统计学是关于统计数据的采集、整理、分析和推断的科学。关于社会经济统计学的性质,概括地说,它是一门关于社会经济现象总体综合数量特征和数量关系的方法论科学。统计实践活动的研究对象是统计总体的综合数量特征及其相互关系,但人们所关注的主要是对总体的某方面的特征或某些数量特征的计量,而对总体的综合数量特征的计量却取决于对总体中各个个体的相应数量特征的计量。在采集统计数据前,人们需要对现象进行计量或测度,这就涉及数据的计量尺度问题。不同的事物,其计量的尺度是不一样的。个体单位的品质属性即其品质标志表现,如人口的性别、职业、爱好,设备的性能、种类等是不能用数值计量的。只有个体单位的数量标志表现才能用数值计量或测度,如人口的身高、体重,设备的价格或价值等。同时,由于客观事物有的比较简单,有的比较复杂,有的特征和属性是可见的(如人的外貌特征、校园的自然环境等),有的则是不可见的(如人的偏好、欲望和理想等,学校的管理水平、人文底蕴等),有的表现为数量差异,有的表现为品质差异。因此,统计计量也就有定性计量和定量计量的区别,并且可分不同的层次。美国社会学家、统计学家史蒂文斯(S. S. Stevens)1968 年按照变量的性质和数学运算的功能特点,将统计计量划分为四个层次或四种计量尺度。

一、定类尺度

定类尺度又称名义尺度,它是按研究对象的某种属性将其划分为不同组别或类型的一种测度。对于个体的某些特征,既不能用通常的物理或化学的计量尺度对其进行测度,又不能用经济计量尺度进行测度,也不能对其进行排序分等,而只能根据其特征的具体表现进行分类。

定类尺度在日常生活中的应用最为广泛。例如,将人口的性别分为男、女两类,将公司分为股份有限公司和有限责任公司等,将会计人员按工作岗位分为会计、出纳、稽核等。这种既不能直接用数值大小测度,又不能用等级顺序计量,而只能划分类别的测度就称为定类尺度,它主要用于对个体单位的品质属性进行测度。为便于统计处理,尤其是便于计算机识别,一般也可以将不同类别的个体单位用不同的数字或编码表示。例如,将男性人口用“1”表示,将女性人口用“0”表示;将会计岗位用“1”表示,出纳岗位用“2”表示,稽核岗位用“3”表示等。不同代码反映同一水平的各类(组)别,并不反映其大小顺序。各类中虽然可以计算它的单位数,但不能反映第一类的一个单位可以相当于第二类的几个单位等等。同理,这些数字或编码本身并不能代表现象规模的大小或数量的多少,它仅仅是不同类别的代码。该尺度是最粗糙的一种计量尺度,也是最基本的一种计量尺度,它是其他计量尺度的基础。

定类尺度只能测定事物的类别差异,不对类别之间的关系做任何假定,其特征可概括为以下三点。

(1) 只能区别事物的类别,无法区别事物的大小、优劣。这就是说,类别之间不存在是否相等的关系,也不存在谁先谁后的关系,它们是平等的并列关系,类别之间的位置或顺序是可以互换的,互换后其性质或数量并不因此而改变。例如,人口的性别分为男、女两类,但不存在男优于女或女先于男之说;又如,企业某车间已安装的某种机械设备的编号是从左到右,还是从右到左并不改变设备本身的价值与性能;再如,学生上体育课时按高矮秩序排队,从左到右报数与从右到左报数,虽然各学生两种报数所报数字有所不同,但并不因此而改变各学生的性格、学习成绩和爱好等,它只是排列顺序或编号的不同而已。

(2) 对事物的分类必须遵循穷尽原则和互斥原则。穷尽原则是指在对总体进行分类时,必须保证总体中的每个单位归属于其中的某一组或某一类别中,不能游离于所分的组或类别之外,这就是通常所说的分组要遵循完整性的原则;互斥原则是指总体中的每个单位在某一分组标志条件下只能归属于其中的某一组或某一类别中,不能在同一分组标志条件下同时出现在其他组,这就是通常所说的归属的唯一性。

(3) 对定类尺度计量的数据进行分析的统计指标主要是频数或频率。例如,将某校学生按专业分类后,可分别计算出各专业的学生人数,也可以计算出各专业学生人数占总人数的比重。

二、定序尺度

定序尺度又称为顺序尺度,它是对个体单位进行分类后所确定的现象之间等级差别或顺序差别的一种测度。对于个体单位的某些特征,虽然不能用通常的物理或化学的计量尺度对其进行测度,也不能用经济计量尺度进行测度,但可以对其进行排序分等。例如,对合格产品按其性能的好坏,可分为一等品、二等品、三等品等。这种尺度虽然不能表明一个单位的一等品等于几个单位的二等品,但一般情况下能明确表示一等品的性能优于二等品,二等品的性能又优于三等品等等。又如,会计人员的职称可分为高级、中级、初级等;学生的体育、美术等成绩可分为优、良、及格、不及格等;公务员年度考核结果可表现为优秀、称职、一般、不称职等。需要指出的是,各个等级之间虽有高低、好坏之分,但相邻等级之间没有明确的数量界限,无法说清某一等级比其相邻等级多多少或少多少、优多少或劣多少,这种不能用数字的大小计量,而只能用顺序或等级来说明个体单位特征表现位次的测量计量尺度就称为定序尺度或顺序尺度。可见,定序尺度虽然比定类尺度精确,但相邻顺序等级