

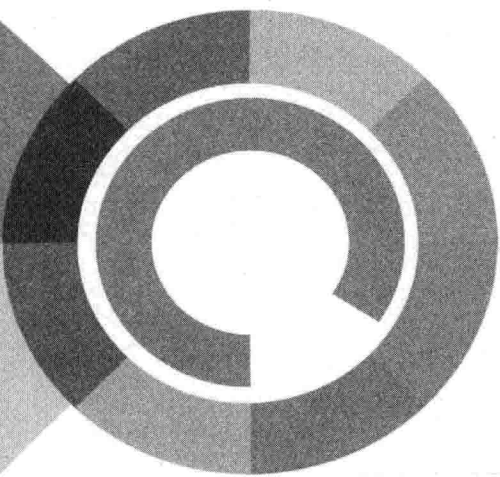


统计学

主 编：陈丽燕

副主编：路春艳 刘福香

 中国统计出版社
China Statistics Press



统计学

主 编：陈丽燕
副主编：路春艳 刘福香



中国统计出版社
China Statistics Press

图书在版编目(CIP)数据

统计学 / 陈丽燕主编. —北京 : 中国统计出版社,
2015.2

ISBN 978-7-5037-7348-8

I. ①统… II. ①陈… 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(2014)第 257536 号

统 计 学

作 者/陈丽燕 路春艳 刘福香

责任编辑/徐 颖

封面设计/李雪燕

出版发行/中国统计出版社

通信地址/北京市丰台区西三环南路甲 6 号 邮政编码/100073

电 话/邮购(010)63376909 书店(010)68783171

网 址/<http://csp.stats.gov.cn>

印 刷/河北三河市利兴印刷有限公司

经 销/新华书店

开 本/710×1000mm 1/16

字 数/230 千字

印 张/15.25

印 数/1—3000 册

版 别/2015 年 2 月第 1 版

版 次/2015 年 2 月第 1 次印刷

定 价/29.00 元

版权所有。未经许可,本书的任何部分不得以任何方式在世界任何地区以任何文字翻印、仿制或转载。

中国统计版图书,如有印装错误,本社发行部负责调换。

编写说明

21 世纪是信息时代,无论是自然科学领域、社会科学领域的研究,还是国家宏观管理和企业生产经营管理,甚至人们的日常生活,信息需求量日益增多,信息处理技术更加复杂,作为信息技术支柱的统计方法,越来越广泛地应用于各个领域。

本教材以普通高校本科教学的需求为第一目标。内容设置以实际应用为出发点,通过举例来讲解统计方法的含义、应用条件和应用技巧等内容,各章都配有小结、思考与练习题。

本教材是由黑龙江外国语学院和哈尔滨商业大学两所院校合作编写,各章执笔人是:第一章,陈丽燕;第二、三、四、五、八、十、十一、十二章,路春艳;第六、七、九章,刘福香。本书由陈丽燕任主编,路春艳、刘福香任副主编,王涛主审。陈丽燕和王涛对书稿进行了修改和完善。

在本书编写过程中,作者参考了其他相关书籍,选用了个别案例,引用了一些观点,在此谨向有关参考书的作者表示感谢。

由于作者水平有限,书中的错误和疏漏之处在所难免。敬请读者提出宝贵意见,以便进一步修订和改进。

目 录

第一章 导 论	(1)
第一节 统计学常识	(1)
第二节 统计学中的基本概念	(6)
第三节 统计学的主要研究内容及应用软件简介	(12)
本章小结	(14)
思考与练习	(15)
第二章 统计数据的获得、整理与表述	(16)
第一节 统计数据的获得	(16)
第二节 统计数据的整理	(19)
第三节 统计资料的表现形式	(24)
本章小结	(26)
思考与练习	(26)
第三章 数据分布特征的测度	(27)
第一节 集中趋势的度量	(27)
第二节 离散趋势的度量	(38)
第三节 偏态与峰度的度量	(43)
本章小结	(47)
思考与练习	(47)
第四章 概率基础	(49)
第一节 概率及其分布	(49)
第二节 正态概率分布	(58)
第三节 常用样本统计量的分布规律	(60)
本章小结	(64)
思考与练习	(64)

第五章 参数估计	(65)
第一节 参数估计的基本问题	(65)
第二节 点估计	(66)
第三节 区间估计	(70)
第四节 样本量的确定	(84)
本章小结	(86)
思考与练习	(86)
第六章 假设检验	(88)
第一节 假设检验的基本原理	(88)
第二节 总体参数的假设检验	(91)
第三节 非参数检验	(99)
本章小结	(103)
思考与练习	(103)
第七章 方差分析	(105)
第一节 方差分析的基本原理	(105)
第二节 单因素方差分析	(111)
第三节 双因素方差分析	(113)
本章小结	(119)
思考与练习	(120)
第八章 统计相关与回归分析	(121)
第一节 简单相关分析	(121)
第二节 线性回归分析	(127)
本章小结	(133)
思考与练习	(134)
第九章 时间序列分析	(135)
第一节 时间序列概述	(135)
第二节 时间序列的趋势分析	(145)
第三节 时间序列的结构分析	(152)
本章小结	(158)
思考与练习	(158)

第十章 指数	(160)
第一节 统计指数及其应用	(160)
第二节 总指数的编制	(164)
第三节 因素分析法	(174)
本章小结	(183)
思考与练习	(183)
第十一章 统计预测与决策	(185)
第一节 统计预测技术	(185)
第二节 统计决策的基本问题	(192)
第三节 风险型决策	(195)
第四节 不确定型决策	(198)
第五节 贝叶斯决策	(202)
本章小结	(205)
思考与练习	(206)
第十二章 国民经济统计基础知识	(207)
第一节 国民经济统计的基本内容和基本原则	(207)
第二节 国民经济统计的主要分类	(211)
第三节 国民经济统计的常用基本指标	(213)
第四节 国民经济统计的常用分析指标	(225)
本章小结	(229)
思考与练习	(230)
附表一:标准正态分布表	(231)
附表二:t 分布表	(232)
参考文献	(234)

第一章 导 论

统计是人类认识和探索未知领域的重要方法和途径,随着人们对定量研究的日益重视,统计方法已被应用到自然科学和社会科学的众多领域。本章将介绍统计学的一些基本问题,包括统计学的基本术语和常识性知识、主要研究内容及常用软件等。

第一节 统计学常识

一、统计学的学习目的和意义

(一)统计的含义与特征

1. 统计一词的含义

在科学研究、经济管理工作和日常生活中,人们经常使用“统计”一词,它的含义一般包括如下三个方面:

统计工作(statistical work)是对事物数量特征的有关资料进行搜集、整理和分析的活动过程;

统计资料(statistical material)是统计工作的对象和成果;

统计科学(statistical science)是对统计工作及统计资料的规律性进行总结和理论概括的系统学说。统计学是研究数据的科学,是方法论科学,它是由研究数据的方法构成的。研究数据的方法大体分为:搜集数据、整理数据、分析数据。

“统计”一词的英文为“statistics”,有两种含义:单数表示统计科学,复数表示统计数据。可见,统计一词在不同的场合,有不同的含义。

2. 统计的基本特征

虽然在不同场合统计一词的含义有所不同,但是人们总是要把它与总计、差异比较、大量的数据分析等现象联系起来。这说明数量性、整体性以及关联性的特性,已经固化到统计之中。

世界上的各类现象的发展变化规律,都表现为质与量的辩证统一,要认识其客



观存在的规律性,就必须认识其质与量的辩证关系,认识其数量关系的特征及度的界限,这一切都离不开统计,尤其是大数据时代的到来,将是统计学最为普及和应用最广的时期。

(二)学习统计学的现实意义

1. 统计与现实生活的关系

随着人类社会发展的不断信息化,如何测度社会经济活动,以及人们内心世界的各种潜在活动的量化信息,已经成为社会发展的共识。更是科学分析、政策制定和管理决策等社会活动的重要基础。

在现实的各行业的技术性工作中,越来越依赖数字化的、系统化的信息测量以及统计分析成果,因此统计工作被认为是一个最富挑战性的朝阳职业,能够运用定量分析手段和通信技术来处理各种难题,将是这个社会不断扩大的人才需求领域。这些工作虽表现为制图、制表、调查、计算以及分析报告的写作,但其背后都离不开统计学的支持。学习统计学及其方法的应用,将为你进入高层次的就业提供更多的机会。统计方法可用于创建解释工商业活动中的数量模型、市场的趋势模型,也可用于社会现象、物理活动和生物现象的规律总结,包括对数学原理的统计应用,以及对数据信息的收集、整理、分析和解释等的广泛运用。对科学规律有兴趣的人们都会发现,统计是一种有用的、有魅力的工具。

2. 统计学在科学研究中的作用

统计方法主要包括描述和推断等领域,所以它必然成为科学研究的重要手段,是任何科学研究工作必备的知识。其主要作用体现如下:

- (1)统计描述系统是科学研究的前提;
- (2)统计推断是科学推算的依据;
- (3)对统计规律的观察和探索是科学研究的核心;
- (4)各门科学的专有统计技术和方法,常被称为专业统计,已形成各门科学体系中的重要组成部分。

二、统计学的发展简史

在人类文明的开始就有了初步的统计活动,可以说统计工作有着数千年的历史。而在这数千年的统计活动中,人们对统计规律的认识也在逐渐加深,并不断地总结出统计规律,使统计学成为一门独立的科学系统。

(一)古典统计学的产生(17世纪中期至19世纪初期)

在统计理论的形成过程中,由于各国的历史背景、经济水平与思想渊源的不同,对统计的研究内容和表达方式也相互存在差异。在17世纪中叶,欧洲各国相继进入资本主义工场手工业的经济迅速发展阶段,但是某些国家的封建制度尚未

解体,这使欧洲处于思想活跃的社会变革时期。为了适应各国经济发展的不同需要,欧洲各国不约而同的从不同领域开始了统计学的奠基工作。并相继形成了统计学的三大来源,即政治算术学派、国势学派和古典概率的应用,这三大来源到19世纪初都基本上形成了各自的理论体系,并为近代统计学奠定了理论基础,所以称17世纪中期至19世纪初的古典统计学为统计学的萌芽时期。

1. 政治算术学派

该学派源于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(Johan Graunt, 1620—1674)。

威廉·配第《政治算术》(1690年出版)一书是经济学和统计学史上的重要著作。该书采用“数字、重量、尺度”等定量的分析工具,对英国与荷兰、法国等当时主要发达国家的经济实力进行了比较分析。被马克思推举为英国古典政治经济学的创始人,同时也是统计学的创始人。政治算术中采用的研究方法及其所形成的理论体系,在统计学史上被称为政治算术学派。

约翰·格朗特在1662年发表了《关于死亡表的自然观察与政治观察》一书。书中通过大量观察发现了人口各年龄组的死亡率、性比例等重要数量规律,并对人口总数进行了较为科学的估计。因此,被认为是人口统计学的创始人。

德国的约翰·彼德·劳斯密尔希(Johan Peter Sussmilch, 1707—1767)深受政治算术学派的影响,并成为该学派的主要代表人物和继承者。其代表作《由人类之出生、死亡及繁殖证明在人类变动中所存在的神的秩序》(1741)一书,对大量不确定现象的比例规律进行研究。这部书虽把事物的规律性看做是神的安排,但它在某种程度上为概率论应用于人类生活的研究奠定了基础。

政治算术学派在当时的欧洲大陆广泛传播,并逐渐形成了两大支流。即以信奉配第为主的经济统计派、以信奉格朗特为主的人口统计派。18世纪人口统计派占主导地位,并以人口推算为其中心课题。

2. 国势学派

该学派产生于18世纪的德国,其创始人是黑尔姆斯塔特(Helmstadt)大学教授海尔曼·康令(Hermann Conring, 1606—1681)博士。他于1660年把国势学从法学、史学、地理学等学科中独立出来,在大学中讲授“实际政治家所必需的知识”。后由马丁·休姆采尔(Martin Schneitzel, 1679—1747)将其更名为《政治学·统计学讲义》。其中统计学一词(statisticum)的语意来源于拉丁语系的“状态”而后转化为“国家”,可见统计学在国势学派看来是指对国家政治状况的研究。

国势学派中最重要的继承人是休姆采尔的学生高特弗瑞德·阿亨瓦尔(Gottfried Achenwall, 1719—1772),他被当时德国誉为“统计学之父”,而他自己则推崇康令为统计学之父。并在1749年确定了统计学(statistic)这一学科的名称及有关

统计学的一些术语。

国势学派只是对国情及发展状态,使用比较级和最高级的词汇进行文字语言性描述。并不研究事物的计量分析方法,所以人们将其称为记述学派(或旧学派或德国学派),并认为国势学派有统计学之名,而无统计学之实。

3. 古典概率论的应用

古典概率论的研究虽始于16世纪的意大利,但17世纪中期才得到一般化的解法。并在18世纪的法国、瑞士等国得到广泛发展,最终于19世纪初叶由法国数学家、统计学家拉普拉斯(P. S. Laplace, 1749—1827)在总结前人成果的基础上,出版了名著《概率论分析理论》一书,从而形成完整的应用理论体系。他对统计学的贡献主要在如下几个方面:(1)总结了古典概率论研究的成果,初步奠定了数理统计学的理论基础;(2)将大数定律作为概率论与政治算术的桥梁,为社会经济统计奠定了数量分析基础;(3)用自然科学的方法研究社会现象,为数理统计产生提供了必要的理论依据。

(二)近代统计学体系的形成时期(19世纪初至20世纪初)

近代统计学的主要贡献是建设和完善统计学的理论体系,并逐渐形成了以随机现象的推断统计为主要内容的数理统计学和以传统的政治经济现象描述为主要内容的社会统计学两大学派。

1. 数理统计学派

19世纪前半叶,资本主义制度在欧洲许多国家中已经成熟,机械唯物论的世界观和自然科学的成果,已否定了所谓的神的秩序,证实了世界存在着自然规律,这为数理统计的建立创造了充分条件。

比利时的凯特勒博士(Lambert Adolphe Jacques Quetelet, 1796—1874)深受拉普拉斯影响,在其《社会物理学》中将概率论引入统计学。认为概率论是适于政治及道德科学中以观察与计数为基础的方法,并对自然现象和社会现象的规律性进行观察,认为要促进科学的发展,就必须更多地应用数学。他的统计学著作有56部之多,按其贡献可以认为他是古典统计学的完成者,近代统计学的先驱,也是数理统计学派的奠基人;同时,他还是第一届国际统计会议(1853年)的召集人,因此,他被称之为“近代统计学之父”。

2. 社会统计学派

社会统计学派产生于19世纪后半叶的德国。因德国的资本主义产生较晚,所以为之服务的社会统计学派,较英国的政治算术学派晚了近二个半世纪。但由于当时数理统计学尚未充分发展,社会统计学派便在欧洲大陆占有优势地位,并向世界各国广泛传播。

该学派的创始人是克尼斯(K. G. A. Knies, 1821—1898),他认为统计学是一

门独立的具有政治算术内容的社会科学。

另一位有影响的创始人是乔治·逢·梅尔(Georg von Mayr, 1841—1925)。他把统计学作为实质性研究的社会科学。并认为统计学是以社会集团的规律性为其独立的研究对象,以大量观察法为其特殊的研究方法,初步建立了社会统计的学科体系。

社会统计学派的代表人物是厄恩斯特·恩格尔(Christian Lonrenz Ernst Engel, 1841—1896),他通过工人家庭生活费用调查发现“恩格尔法则”,并用一定消费单位“凯特”表示整个家庭的消费能力等。

(三)现代统计学的大发展时期(20世纪初到现在)

1. 二十世纪的两大对立学派

(1)欧美数理统计学。自19世纪末叶以来,欧洲自然科学飞跃发展,促进了数理统计学的快速发展。进化论和能量守恒定律的出现促进了描述统计的完善,是描述统计学派发展的顶峰。20世纪20年代以后,在细胞学的发展推动下,统计学迈进了推断统计的新阶段,直到50年代是推断统计学派发展最迅速的时期。

在这期间有影响的理论和大师很多,例如:20世纪初英国统计学家戈赛特(William Sealy Gosset, 1876—1937)的 T 分布理论;20年代英国统计学家费希尔(R. A. Fisher, 1890—1962)的 F 分布理论;30年代波兰统计学家奈曼(Jerzy Splawa Neyman, 1894—1981)等人的假设检验理论及置信区间估计等理论;40年代的美国统计学家瓦尔德(A. Wasld, 1902—1950)等学者的统计决策理论、多元分布理论等。到了50年代,经过几代大师的努力,推断统计的基本框架已经建成,并逐渐成为20世纪的主流统计学。这是统计学全面发展的阶段,由于受计算机和新兴科学的影响,使统计学越来越依赖于计算机技术,成为数量分析的方法论科学。这一时期统计学的研究和应用范围越来越广,使得在现代统计学史中很难找到权威性的代表人物。当今的统计学家只能限制在有限的专业领域内从事某方面的研究,这是现代统计学的主要特点。如科克伦(W. G. Cochran, 1909—1980)的实验设计理论、安得森(Th. W. Anderson)的复变数分析等。

(2)东方社会经济统计学。十月革命胜利后,苏联的大多数统计学家受社会统计学派的影响,主张统计学是一门实质性的社会科学。1954年3月,由苏联科学院、中央统计局、教育部联合召开了统计科学讨论会,并把统计学定义为:统计学是在质与量的密切联系中研究大量社会现象的数量方面,研究社会发展规律在具体地点及时间条件下的数量表现的社会科学。这一定义对我国及东欧社会主义国家的影响都很大,在这些国家中形成了以马克思政治经济学为理论基础的社会经济统计学派。该学派是在反虚无论、反消亡论、反万能科学论和反数学形式主义中形成的,并以物质产品为核算范围,建立了物质产品平衡表核算体系(the system of

material product balances,简称 MPS)。直到 1993 年以后,该核算体系才逐渐被改革掉,取而代之的是国民经济核算体系(即 SNA),但 MPS 对我国的影响至今还是很大的。

2. 二十一世纪统计学的发展趋势

21 世纪是新科学不断产生,各学科广泛融合的时期,所以现代主流统计学的发展趋势和特点如下:

(1)统计学的家族越来越庞大和独立。统计学在传统的理学和经济学中逐渐独立,必将成为一个新的科学门类。

(2)统计学可以分为方法论和实质性科学两部分内容,即通用的方法论是以统计学的知识和方法应用为主要研究内容,如本书所讲述的知识;而实质性科学,常被称之为专业统计学,它是在不同的行业领域所产生的专门性研究内容的学科,与其专业有着密切的联系。

(3)统计学的发展规律。第一,随着现代数学的发展,更广泛地应用数学方法;第二,统计学与其他新科学新理论的结合,不断产生新的边缘科学或新的统计分支;第三,借助电子计算机,使大量数理方法得以普及应用,并已成为实证分析的主要工具,同时专业统计方法在不断产生;第四,统计的作用是从描述向推断、预测及决策方向发展。

在最近半个世纪统计活动的发展中,国民经济账户体系的建立、概率论和其他数学方法的广泛应用、统计机构的进一步完善、信息处理手段的自动化,被称为“现代统计”的四大标志。

第二节 统计学中的基本概念

一、有关统计对象的基本术语

(一)总体与总体单位

1. 总体

总体(population)是统计总体的简称,指统计所研究的事物全体,它是客观存在的具有某种共同性质的许多单位组成的集合体。例如,在全国人口普查中,所有“具有中华人民共和国国籍且在中国境内常住”的人所构成的集合,就是一个统计总体。这是我们研究中国人口状况所需调查的全部对象,它是客观存在的,是由许多人(总体单位)组成的集合体,其中的每个人都具有“中华人民共和国国籍且在中国境内长期居住”这种共同的属性。

2. 总体单位

总体单位(population unit)是指构成总体的个别事物,简称个体。例如,上面所说的人口总体,其总体单位就是每一个具有“中华人民共和国国籍在中国境内常住”的人。总体是一个集合的概念,总体单位则是集合中的元素。

3. 统计研究总体的特征

总体根据其总体单位数是否有限,分为有限总体和无限总体。有限总体是指总体中包含的总体单位数量是有限的;无限总体是指总体单位数目是无限的。

总体和总体单位是相对而言的,是根据统计研究的目的来确定的。同一事物,在一定条件下是总体,在另一条件下可以是总体单位。所以在统计研究中要注意统计总体具有如下三个基本特征:

同质性(homogeneity)是指构成总体的每一个单位都必须具有某种共同性质,该性质是形成总体的客观依据,也是我们确定总体范围的标准。各单位必须具有这种共同性质,是由统计研究的目的决定的。

大量性(large quantity)是指总体必须由许多单位组成。一个或少数单位不能形成总体,因为统计研究的目的是要揭示大量事物的普遍规律性,所以统计研究的对象必须包括足够多的个体。

差异性(heterogeneity)指总体的各单位除了某一方面的共同性外,在其他方面必须是有差异的,这些差异是统计研究的基础和前提。

作为总体必须同时具备上述三个特征,才能进行一系列的统计计算和分析研究,三者缺一不可。现象不同质,无法将其结合在一起;少量单位,反映不出现象的规律性;各个单位都一样,无差别就无须进行统计研究。

(二)标志、变异与变量

1. 标志

标志(mark)常写为标识,是指总体单位各种特征的名称,有数量标志与品质标志之分:数量标志表明事物的数量属性,它可以用数值表示;而品质标志表明事物的品质属性,即所有的非数量属性,它常用文字和符号来直观地表示。例如,在工人总体中,某工人的性别是男,民族是汉族,年龄是35岁,工资是2800元。在这里的“男”和“汉族”分别是品质标志“性别”和“民族”的属性,即这类标志的具体表达没有使用数字;而“35岁”和“2800元”则分别是数量标志“年龄”和“工资”的数值表述,它们使用了数字。

2. 变异

变异(variation)是指标志的具体表现在总体各单位间的差别。包括质的差别和量的差别。如人口的性别标志可以表现为男和女;年龄标志可以表现为30岁、40岁、50岁等。这种差别就叫变异。



在一个总体中,每个总体单位都具有至少一个不变标志与许多可变标志。不变标志是指对所有总体单位都有完全相同的具体表现的标志,正因为具有这个不变标志,才使它们集合在一起构成统计总体;可变标志是指在总体各单位之间具有不同表现的标志。如对某地区所有工业企业构成的这个总体来说,其不变标志是某个地区、工业、企业等,它对总体各单位进行了具体的界定,构成总体的同质性;而每个企业的职工人数、产量、产值等都各不相同,这些都是可变标志,它表现为总体的变异性。总体的同质性和总体单位的变异性是进行统计核算的必备条件。

3. 变量

变量(variable)是指使用数字来表述所研究的各单位的某数量标志的数列。而某单位数量标志的具体取值是变量值。

按变量值是否连续可把变量分成离散型变量和连续型变量;离散变量是指标志表现为有限的或可数的数值;连续变量的数值是连续不断的,只能用区间表示的不可数的数值,如人的身高、体重等。

按变量值的确定性,可把变量分为确定性变量和随机性变量。确定性变量是指变量受某种或几种确定性因素的影响,其数值是具体确定的。事后统计得到大多数成果多属于确定性变量,如一次调查所得到的某标志的统计数列等。随机性变量是受某种或几种不确定性因素影响,其取值是带有很大偶然性的变量。例如,抽取部分零部件,检查其质量是否符合规定,其误差的大小带有一定的偶然性,所以误差的大小就是一个随机性变量,它的表现是不确定的。

(三) 统计指标与指标体系

1. 统计指标

统计指标(statistical indicator)是指反映统计研究总体某一综合数量特征的概念和具体数值。如人口数 13 亿、土地面积 960 万平方千米等,都属于统计指标。一项完整的统计指标应该由两个基本要素构成,即指标名称和指标数值。除了这两个基本要素外,还应有时间限制、空间限制、计算方法和计量单位四个要素。统计指标是一种统计结果,是反映确定的统计对象综合数量特征的概念及其具体数值。统计指标具有数量性、综合性、具体性和客观性等主要特点。

通过统计指标,可以反映社会经济现象的规模、水平、比例关系和速度等。研究社会经济发展规律的数量表现,检查国民经济和社会发展规划以及各项政策的执行情况,衡量生产经营活动的经济效益。因此,统计指标成为人们认识社会、管理经济、科学研究的基本依据之一,起着社会指示器和反映数量规律性的作用。

统计指标按照不同标准可以分为不同类别:

(1)按统计指标所说明的总体现象的性质不同,可以将统计指标分为数量指标和质量指标。

数量指标(quantitative indicator)是反映总体绝对量多少或规模大小的统计指标,常以总量数值的形式来表示,所以也常叫总量指标。如人口数、企业数、商品销售额等。数量指标的数值大小往往随着总体范围的大小而增减变动。

质量指标(qualitative indicator)是表明总体本质的属性,并说明事物属性程度的统计指标。虽然统计指标都是以数量形式出现,但是它通过数量关系来反映其总体内在的品质特征的统计指标。质量指标通常是由数量指标派生出来的,用相对数或平均数来表示。如反映统计对象质量的指标产品合格率,反映利用程度的指标设备利用率,反映经营效益的指标资金利润率等。

数量指标和质量指标是有着密切联系的相对概念,有的指标在某种条件下是数量指标,而条件改变就可能成为质量指标。所以在经济分析时,要根据环境情况区别对待,作具体处理。

(2)按统计指标形成的方式不同,可以将统计指标分为总量指标、相对指标和平均指标。

总量指标(total amount indicator)是反映社会经济现象总的规模、水平或总量的统计指标,其数值的取得往往是对总体各单位具体数值的汇总。

相对指标(relative indicator)是反映两种现象的差距相对量的统计指标,其数值的取得是由两个有联系的统计指标数值的对比结果。

平均指标(average indicator)是同质总体内标志总量与总体单位数相除的结果,用以表明总体单位标志值的一般水平的统计指标。

2. 统计指标体系

每一个统计指标都是对统计总体的某一侧面的特征反映,所以只用单一统计指标来反映总体现象,很难避免其片面性。为此,将若干相互联系的统计指标组成一个整体,形成反映统计总体内在联系的有机体系,就形成了统计指标体系(statistical indicator system)。它是根据统计任务的需要,由能够全面反映统计对象的数量特征和数量关系的互相联系的一套指标构成。

二、有关统计数据的基本术语

(一)统计数据的计量类型

统计指标的可量性决定了在对于社会经济现象的数量方面进行研究时,必须予以量化,从数据量化的抽象程度不同分为以下几个层次:

1. 定类尺度

定类尺度(nominal scale)或称作列名尺度,就是将研究对象按某种特征划分成若干部分,并给每一类别定名,但不对类别之间的关系做任何假定。定类尺度是最粗略、精度最低的计量尺度,也是最基本的尺度。例如,在人口统计中按地区分



组、民族分组,并用数字作为代号,如北京为01,河北为02等。在形式上,定类尺度具有对称性和传递性两种属性,对称性说明各类之间彼此相对称,传递性则表示运算上各类量值只具有相等与不相等的性质。这种测定尺度和分组在实际统计活动中使用得很广泛,主要用于计算各组数值占总体数值的比重和众数等特征值,以及进行逻辑运算等,但不能对各类编号进行加减乘除计算。

2. 定序尺度

定序尺度(ordinal scale)或称为顺序尺度,它是把各类事物按一定特征的大小、高低、强弱等顺序排列起来,构成定序数据。例如,将产品按其质量高低列成一等品、二等品、三等品,学生的成绩排列为优、良、中、及格、不及格等,这种测定尺度的量度层次要比定类尺度高一些,它不仅可以分类,而且可以确定这些类别的顺序,各类之间还能比较等级和次序上的差别。在运算上,各类量值除了具有等与不等的特征外,还有大于或小于之分,但其序号仍不能进行加减乘除计算。定序尺度除了可用来计量比重(频率)外,还可进行累计频数(率)、中位数等特征值的计算。

3. 定距尺度

定距尺度(interval scale)或称间隔尺度,它是把定序排列的各类事物间的差距,以一定的度量单位明确起来,构成定距的数据。这是更精确的计量尺度,一般要求建立某种物理的量度单位。如考试成绩以“分”计量;长度以“米”计量;等等。成绩每分之间的间隔是相等的,80分与90分的差距等同于90分与100分的差距。在运算上,除了等于、不等于、大于、小于之外,还可进行加减运算,但不能进行乘除运算。例如可以说 30°C 与 25°C 相差 5°C ,且它与 10°C 与 5°C 之间的差距相等,却不能说 10°C 比 5°C 热一倍。

4. 定比尺度

定比尺度(ratio scale)或称比率尺度,是量度层次最高的数据测定尺度。它是在定距尺度的基础上增加了一个绝对零点,并抽象掉事物的度量差异的测定尺度。换言之,定距尺度中的“0”只表示某一个值,即0值;而定比尺度中的“0”是绝对零点,表示没有。定距尺度与定比尺度的差别,在于是否存在绝对零点,“0”在两者间的意义是不同的,如某人数学考试得0分,只能表示他的数学成绩是0分,不等于说他完全没有数学水平,但如说某人的身高为0米,则表示此人是不会存在的。在运算上,定比尺度可以用于任何统计运算和比较。所以应用广泛,许多统计的最终结果是以定比尺度给出的。

在测定尺度的应用中,同类事物用不同的尺度量化,就会得到不同的尺度数据。如农民收入数据按实际值填写就是定距尺度;按高、中、低收入水平分别记载就是定序尺度;而按有无收入计量则成为定类尺度了;如果说某人的收入是另一人的两倍,则是定比尺度。又如,学生成绩若具体打分就是定距尺度,用优、良、中、及