

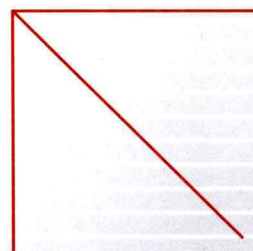


Economics

统计学

TONGJIXUE

主编 刘灵芝





统 计 学

TONGJIXUE

主编 刘灵芝



中国金融出版社

前言

在现代社会中，各行各业、各领域每天会产生丰富的数据信息，人们日常决策、管理工作都离不开它们。如何客观、准确地认识这些数据，并使其更好地服务于决策，必然离不开科学的方法，而统计学正好提供了这样一种适用于所有学科领域的收集数据、分析数据，并有助于认识现象规律性的方法和技术。

从统计工作过程来看，其主要包括数据收集、整理、显示和分析等环节，本书正是围绕这些环节来介绍统计方法和技术的。全书共有九章，第一章主要介绍统计学的基本原理，包括统计学的含义、统计学科分类以及统计学科所涉及的基本概念；第二章围绕统计数据的收集和整理，介绍相关的统计方法；第三章介绍统计整理的结果，即统计指标，这是进行数据分析的基础；第四章、第五章、第六章是推断统计学的主要内容，在介绍相关数理统计基本原理的基础上，重点介绍了抽样和参数估计的方法，以及假设检验的基本原理和应用；第七章相关与回归分析，介绍了研究变量之间数量关系的常用统计方法，在经济管理研究中有着重要的应用；第八章介绍了专门针对一类数据——时间序列数据的分析方法；第九章统计指数，介绍了指数的编制和应用，其中的因素分析法在经济研究中也有着重要的应用。

作为一本适用于高等院校经济管理类专业本科学生的统计学基础教材，本书在编写过程中，第一，尽量贴近学生实际，以培养学生应用统计方法解决实际问题的能力为出发点，对各种统计方法的介绍都做到以经济管理活动中的实际问题来引入，着重讲述统计思想，而避免过多的数学推导和证明。第二，为培养和提高学生的实际操作能力，对统计方法的应用，在各章附录中，都专门针对具体案例介绍了 Excel 软件的实践过程和结果，便于学生学习时对照使用。第三，为加强学生独立思考和自我学习能力的锻炼，本书在各章节内容安排上，设置了内容提要、导入案例、关键名词、重要公式、思考和练习等内容，有助于学生进行总结和复习，以加深对相关知识点和统计方法的掌握。

本教材由华中农业大学刘灵芝任主编并总纂、湖北经济学院黄小舟任第二主

编，华中农业大学杨明媚、武汉纺织大学泮传快任副主编，参加编写工作的还有华中农业大学熊巍、武汉科技大学黄之姣。其中，第二章由刘灵芝编写，第五章由黄小舟编写，第三章、第八章由杨明媚编写，第一章、第七章由泮传快编写，第九章由熊巍编写，第四章、第六章由黄之姣编写。

由于作者水平有限，书中难免存在缺点和错误，敬请读者批评指正！

编 者
2011 年 5 月

80	重要公式
81	思考和练习
84	第四章 概率基础
84	内容提要
84	导入案例
85	第一节 概率含义和古典概型
92	第二节 概率的基本运算
95	第三节 概率分布
105	第四节 常见的概率分布
110	Excel 操作与应用
117	关键名词
117	重要公式
117	思考和练习
119	第五章 抽样估计
119	内容提要
119	导入案例
119	第一节 抽样理论的一般概念
123	第二节 抽样误差
130	第三节 抽样估计概述
137	第四节 抽样设计与组织实施
149	Excel 操作与应用
151	关键名词
151	重要公式
152	思考和练习
155	第六章 假设检验与方差分析
155	内容提要
155	导入案例
156	第一节 假设检验的基本概念
160	第二节 总体均值和比例的假设检验
169	第三节 方差分析
186	Excel 操作与应用

277	Excel 操作与应用
278	关键名词
279	重要公式
279	思考和练习

第一章

统计学概论

内容提要

本章是对本书内容的概括、总结。通过学习，了解统计学的基本框架体系，系统把握统计学的含义、研究对象、学科性质、研究方法及统计活动的过程；在了解统计学产生和发展的基础上，深入理解统计学的基本概念，对统计认识活动的特点有一个全面的认识。

导入案例

1941年，第二次世界大战正打得如火如荼。有一天，美国哥伦比亚大学著名统计学家沃德教授（Abraham Wald）遇到了一个意外的访客，那是英国皇家空军的作战指挥官。他说：“沃德教授，每次飞行员出发去执行轰炸任务，我们最怕听到的回报是‘呼叫总部，我中弹了！’请协助我们改善这个与飞行员生死攸关的难题吧！”沃德接下这个紧急研究案，他受委托分析德国地面炮火击中联军轰炸机的资料，建议机体装甲应该如何加强，才能降低被炮火击落的机会。但依照当时的航空技术，机体装甲只能局部加强，否则机体过重，会导致起飞困难及操控迟钝。沃德将联军轰炸机的弹着点资料描绘成两张比较表。

沃德的研究发现，机翼是最容易被击中的部位，而飞行员的座舱与机尾，则是最少被击中的部位。沃德详尽的资料分析，令英国皇家空军十分满意。但在研究成果报告的会议上，却发生一场激辩。负责该项目的作战指挥官说：“沃德教授的研究清楚地显示，联军轰炸机的机翼，弹孔密密麻麻，最容易中弹。因此，我们应该加强机翼的装甲。”沃德客气但坚定地说：“将军，我尊敬你在飞行上的专业，但我有完全不同的看法，我建议加强飞行员座舱与机尾发动机部位的装甲，因为那儿最少发现弹孔。”在全场错愕怀疑的眼光中，沃德解释说：“我所分析的样本中，只包含顺利返回基地的轰炸机。从统计的观点来看，我认为被多次击中机翼的轰炸机，似乎还是能够安全返航，而飞机很少发现弹着点的部位，并非真的不会中弹，而是一旦中弹，根本就无法返航。”指挥官反驳说：“我很佩服沃德教授没有任何飞行经验，就敢作这么大胆的推论，就我个人而言，过去在执行任务时，也曾多次机翼中弹严重受创，要不是我飞行技术老到，运气也

不错，早就机毁人亡了，所以，我依然强烈主张应该加强机翼的装甲。”这两种意见僵持不下，皇家空军部部长陷入苦思。他到底要相信这个作战经验丰富的飞行将军，还是要相信一个独排众议的统计学家？由于战况紧急，无法作更进一步的研究，部长决定接受沃德的建议，立刻加强驾驶舱与机尾发动机的防御装甲。不久之后，联军轰炸机被击落的比例果然显著降低。为了确认这个决策的正确性，一段时间后，英国军方动用了敌后工作人员，收集了部分坠毁在德国境内的联军飞机残骸，它们中弹的部位，果真如沃德所预料，主要集中在驾驶舱与发动机的位置。看不见的弹痕最致命，乍看之下，作战指挥官加强机翼装甲的决定十分合理，但他忽略了一个事实：弹着点的分布，是一种严重偏误的资料。因为最关键的资料，其实是在被击落的飞机身上，但这些飞机却无法被观察到，因此，布满了弹痕的机翼，反而是飞机最强韧的部位。空军作战指挥官差点因为太重视看得见的弹痕，反而作出错误的决策。这说明统计学家以数据为基础进行的决策往往比经验更可靠也更科学。

资料来源：<http://blog.eastmoney.com/xunzhaozhenli000>。

第一节 统计学的含义及其发展

一、统计的产生及其含义

“统计”一词的英文 statistics 最早源于现代拉丁文 statisticum collegium（国会）以及意大利文 statista（国民或政治家）。德文 statistik，最早是由德国政治学教授阿亨瓦尔（G. Achenwall, 1719—1772）所使用，代表对国家的资料进行分析的理论，也就是“研究国家的科学”。18 世纪中叶，辛克莱将英文中的 statistics 作为德文 statistik 的译文传入英国。“统计”一词为日本人所翻译，并在清朝光绪年间传入中国。

“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收集资料，弄清国家的人力、物力和财力，作为国家管理的依据。

在原始社会时期，人类简单的计数活动孕育着统计的萌芽，随着社会生产力的发展，人类社会到了奴隶社会以后，奴隶制国家组织的人口、财富和军事统计得到了长足的发展，统计被认为是维护阶级统治、兴邦安国的重要手段。

在中国，统计有几千年的历史。进入封建社会以后，中国的户籍统计和田亩统计都有很大的发展，不论是统计方法、统计制度还是统计组织，都在世界上居于先进水平。

资本主义以前的统计活动，多半是在赋税、征兵工作中进行的，深深地打下了阶级的烙印；另一方面，由于自然经济、封建割据的束缚，统计范围、统计制度和统计方法都是比较落后的。资本主义生产方式在人类历史上确立以后，对统计工作提出了新的要求，也大大促进了统计活动的发展，为统计科学的产生奠定了物质基础。

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收集活动,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。

(一) 统计工作

统计工作,是收集、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、企业、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。

(二) 统计资料

统计资料是统计工作活动进行收集、整理、分析和研究的对象及最终成果。不管是个人、集体和社会,还是国家、部门、企业及科研机构,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收集和分析,以指导自己的学习、工作和生活;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,这就需要对有关的市场资料、生产资料、质量数据、人员资料、成本资料等进行收集、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GNP比2000年翻一番,就需要我国有关GNP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初公布的《国民经济与社会发展统计公报》等即是统计数据资料,也可称为统计。

(三) 统计学

一般来说,统计学是对目标数据资料进行收集、整理、分析和研究,以揭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。统计学是以收集、整理、分析和研究等统计技术为手段,对所研究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。

统计工作和统计学是实践和理论的关系。统计理论来源于统计实践,它是统计工作经验的总结和概括。反过来,统计理论又是指导统计工作的原则和方法。

总之,统计工作、统计数据、统计学三者之中,统计工作是基础,是源。没有统计工作,就不会产生统计数据,没有统计工作,缺少这个实践基础,统计学也就不可能形成和发展。

二、统计学的发展

在有人类文明时就有了统计活动,即统计工作有着数千年的历史,在这数千年的统计活动中,人们对统计规律的认识逐渐加深,并不断总结、规范为统计学。而统计学成为系统和独立的科学只有三百多年的历史。按统计方法及特征的历史演变顺序,一般可将统计学的发展史分为三个阶段,具体如下。

（一）古典统计学时期

古典统计学时期是从17世纪70年代至19世纪初期，是统计学的萌芽时期。在这个时期，统计学逐步由以定性分析为主转为以定量分析为主，在该期间产生了两个有代表性的学派：国势学派和政治算术学派。

1. 国势学派。国势学派产生于17世纪的德国，其创始人是海尔曼·康令（Hermann Conring, 1606—1681）教授。国势学派只是对国情的记述，未能进一步揭示社会经济现象的规律，也不研究事物的计量分析方法，只是用比较级和最高级的词汇对事物的状态进行描述。所以，人们也把它叫做记述学派（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。到了18世纪，阿亨瓦尔对康令教授的理论进行了延续和发展，其本人对康令教授甚为推崇，誉其为“统计学之父”，但后人鉴于阿亨瓦尔教授的贡献及因其提出“统计”一词，而把他尊为“统计学之父”。

2. 政治算术学派。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（Johan Graunt, 1620—1674）。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和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。

（二）近代统计学时期

近代统计学时期是从19世纪初至20世纪初，其间概率论的引进使统计学获得重大的发展和突破。这个时期产生了两个统计学派：数理统计学派和社会统计学派。

1. 数理统计学派。数理统计学派的代表人物是比利时的凯特勒博士（A. Quetelet, 1796—1874），其深受古典概率学者拉普拉斯的影响，在其所著的《社会物理学》一书中将概率论引入统计学。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，他被称为“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。该学派的创始人是克尼斯（K. G. A. Knies, 1821—1898），代表人物有恩格尔（C. L. E. Engel, 1841—1896）以及梅尔（G. V. Mayr, 1841—1925）等人。他们认为统计学是研究社会现象的具体应用学科，并以此与数理统计学派相对立。

（三）现代统计学时期

现代统计学时期是从20世纪初至今。在该时期，所有学科都得到了迅猛的发展，统计学概莫能外，从理论和应用技术上都获得了重大的发展。

在该时期，为现代统计学的发展作出重大贡献的主要统计学家和理论有：20世纪初英国的戈赛特（W. S. Gosset, 1876—1937）的T分布理论，20世纪20年代英国的费雪（R. A. Fisher, 1890—1962）的F分布理论，30年代波兰的尼曼（J. S. Neyman, 1894—1981）等人的假设检验理论及置信区间估计等理论，40年代美国的瓦尔德（A. Wald, 1902—1950）等学者的统计决策理论、多元分布理论等。

20世纪50年代以来，统计理论、方法和应用进入了一个全面发展的新阶段。一方面，统计学受计算机科学、信息论、混沌理论、人工智能等现代科学技术的影响，新的

研究领域层出不穷,如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计、线性统计模型、探索性数据分析、数据挖掘等。另一方面,统计方法的应用领域不断扩展,几乎所有的科学研究都离不开统计方法。因为不论是自然科学、工程技术、农学、医学、军事科学,还是社会科学都离不开数据,要对数据进行研究和分析就必然要用到统计方法,现在连纯文科领域的法律、历史、语言、新闻等都越来越重视对统计数据进行分析,国外的人文与社会学科普遍开设统计学的课程,因而可以说统计方法与数学、哲学一样成为所有学科的基础。

三、统计学在我国的应用

在科学技术飞速发展的今天,统计学广泛吸收和融合相关学科的新理论,不断开发应用新技术和新方法,深化和丰富了统计学传统领域的理论与方法,并拓展了新的领域。今天的统计学已展现出强有力的生命力。在我国,社会主义市场经济体制逐步建立,实践发展的需要,对统计学提出了新的更多、更高的要求。随着我国社会主义市场经济的成长和不断完善,统计学的潜在功能将得到更充分的开发和挖掘。

第一,对系统性及系统复杂性的认识将为统计学的未来发展增加了新的思路。随着科学融合趋势的兴起,统计学的研究触角已经向新的领域延伸,新兴起了探索性数据的统计方法的研究。研究的领域向复杂客观现象扩展。21世纪统计学研究的重点将由确定性现象和随机现象转移到对复杂现象的研究,如模糊现象、突变现象及混沌现象等新的领域。可以这样说,复杂现象的研究给统计开辟了新的研究领域。

第二,定性与定量相结合的综合集成法将为统计分析方法的发展提供新的思想。定性与定量相结合的综合集成方法是钱学森教授于1990年提出的。这一方法的实质就是将科学理论、经验知识和专家判断相结合,提出经验性的假设,再用经验数据和资料以及模型对它的确实性进行检测,经过定量计算及反复对比,最后形成结论。它是研究复杂系统的有效手段,而且在问题的研究过程中处处渗透着统计思想,为统计分析方法的发展提供了新的思维方式。

第三,统计科学与其他科学渗透将为统计学的应用开辟新的领域。现代科学发展已经出现了整体化趋势,各门学科不断融合,已经形成一个相互联系的整体。模糊论、突变论及其他新的边缘学科的出现为统计学的进一步发展提供了新的科学方法和思想。将一些尖端科学成果引入统计学,使统计学与其交互发展将成为未来统计学发展的趋势。统计学也将会有一个令人振奋的前景。今天已经有一些先驱者开始将控制论、信息论、系统论、图论、混沌理论以及模糊理论等方法和理论引入统计学,这些新的理论和方法的渗透必将对统计学的发展产生深远的影响。

统计学产生于应用,在应用过程中发展壮大。随着经济社会的发展、各学科相互融合趋势的发展和计算机技术的迅速发展,统计学的应用领域、统计理论与分析方法也将不断发展,在所有领域展现它的生命力和重要作用。

第二节 统计学的特点和分类

一、统计学的特点

统计学区别于其他学科的特点主要有：数量性、总体性、具体性和不确定性。

（一）数量性

任何现象总表现为质和量两个方面，而统计学研究的是现象量的方面。统计学研究的基本思路是，从现象外在的量的方面入手，对现象的量的方面进行观察、研究以寻求现象的本质和规律。应当指出，由于统计研究都是基于定量研究，这很容易进入另外一个误区，那就是脱离现象的质去研究现象的量，这往往会得到错误甚至是荒诞的结论，因此在统计研究中必须注意定量研究和定性研究的结合。

（二）总体性

统计研究是基于量的研究，但是统计研究的是总体的量，而单个的量构成了统计的研究对象。只有基于总体的研究，得出的统计规律才是正确可靠的，而仅对个量的分析得到的结论往往是偏颇、不可靠的。比如，我们想知道某产品在一特定市场上的满意度，那么我们的研究对象就要基于这个市场上的所有消费者构成的总体，而不是基于道听途说的几个人对这个产品的评价。诚然，在实际中我们往往难以做到对这个市场上所有消费者进行调查，但我们会利用科学的方法在所有消费者中选择部分有代表性的消费者作为样本调查，但最后我们仍然要用这个样本信息去推断总体，也就是说统计的研究对象始终为总体。

（三）具体性

具体性是指统计研究的量是具体的量而非抽象的量，这点是统计学与数学的主要差别。数学中的量是抽象的，比如2可以指2个人也可以指2斤大米，因为数学更注重的是理论基础而不是应用性。而统计学则更注重应用性，统计学中的量都是具体的量，因此数学中的简单运算可能在统计学中也是不可行的。比如，在数学中 $2+3=5$ ，但如果这2是指2斤大米而3是指3个苹果呢？这就遇到了一个可加性的问题。

（四）不确定性

不确定性是指统计的结论往往存在误差，统计研究往往难以得到一个完全精确的结论。误差始终伴随着统计研究的整个过程，抑或说统计研究的一个重要任务就是不断地降低统计误差。形成统计误差的原因有很多，有些是客观存在的无法避免的，比如计量误差、无回答误差、抽样误差，有些是由于人的疏忽主观造成的，比如登记误差，但即便是主观造成的误差也是难以避免的，只能通过努力去降低。

二、统计学的类别

统计方法已被应用到自然科学和社会科学的众多领域，统计学也发展成为由若干分

成了一些不同的特点。

第三节 统计学中的基本概念

一、总体和样本

(一) 总体

统计总体简称总体，是指由特定的研究目的确定的一定范围内的所有单位构成的集合体。构成总体的这些个别单位称为总体单位。例如，我们要研究我国的工业企业的利润状况，那么我国所有的工业企业就是一个总体。对于该总体来说，每一个工业企业就是一个总体单位。

总体和总体单位的划分也不是绝对的，随着研究目的的不同可以相互转换。比如，当我们研究某地区的一个企业的时候，这个企业是一个总体，但当我们研究该地区的所有企业的时候，这个企业就是一个总体单位。

总体中包含总体单位的数量，叫总体单位数，也叫总体容量。根据一个总体的单位数大小，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，如某地的人口数、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。应当指出在统计实践中，当一个总体非常大的时候，哪怕从数量上讲是有限的，我们也会把它当做无限总体，如全国人口数，理论上是有限的，但这个数目实在过于庞大，我们就把它当做无限总体看待。

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。

(二) 样本

广义而言，样本是指从总体中抽取的部分单位构成的集合，但为了符合统计理论的需要，统计中所指的样本特指从总体中按照随机原则抽取的部分单位构成的集合。构成样本的单位叫样本单位。总体往往是唯一的，但样本不是唯一的，一个总体可以抽出很多个样本。总体和样本的这种关系可以参见图 1-1。

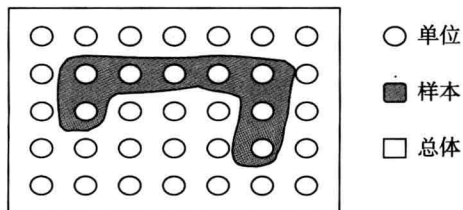


图 1-1 总体和样本的关系

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,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28岁,民族为汉族等,这里“女”、“28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。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,国籍和在国境内常住是不变标志,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. 标志按其性质可以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,如职工年龄、工资、工龄等。品质标志主要用于分组,将性质不相同的总体单位划分开来,便于计算各组的总体单位数,计算结构和比例指标。数量标志既可用于分组,也可用于计算标志总量以及其他各种质量指标。

三、统计指标

(一) 统计指标及其构成要素

对统计指标的含义,一般有两种理解和两种使用方法:

1. 统计指标是指反映总体现象数量特征的概念,如人口数、商品销售额、劳动生产率等。它包括三个构成要素:指标名称、计量单位、计算方法。这是统计理论与统计设计上所使用的统计指标含义。

2. 统计指标是反映总体现象数量特征的概念和具体数值。例如,2009年我国国内生产总值为340 507亿元。这个概念含义中包括了指标数值。按照这种理解,统计指标除了包括上述三个构成要素外,还包括时间限制、空间限制、指标数值。这是统计实际工作中经常使用的统计指标的含义。

一般认为,对统计指标的这两种理解都是成立的。在作一般性统计设计时,只能设计统计指标的名称、内容、口径、计量单位和方法,这是不包括数值的统计指标。然后

经过收集资料、汇总整理、加工计算可以得到统计指标的具体数值,用来说明总体现象的实际数量状况及其发展变化的情况。从不包括数值的统计指标到包括数值的统计指标,在一定意义上反映了统计工作的过程。

(二) 统计指标的特点

1. 数量性。即所有的统计指标都是可以用数值来表现的。这是统计指标最基本的特点。统计指标所反映的就是客观现象的数量特征,这种数量特征是统计指标存在的形式,没有数量特征的统计指标是不存在的。正因为统计指标具有数量性的特点,它才能对客观总体进行量的描述,才使统计研究运用数学方法和现代计算技术成为可能。

2. 综合性。这是指统计指标既是同质总体大量个别单位的总计,又是大量个别单位标志差异的综合,是许多个体现象数量综合的结果。例如,某人的年龄、某人的存款额不能叫做统计指标,一些人的平均年龄、一些人的储蓄总额和人均储蓄才叫做统计指标。统计指标的形成都必须经过从个体到总体的过程,它是通过个别单位数量差异的抽象化来体现总体综合数量的特点的。

3. 具体性。统计指标的具体性有两个方面的含义。一是统计指标不是抽象的概念和数字,而是一定的具体的社会经济现象的量的反映,是在质的基础上的量的集合。这一点使社会经济统计和数理统计、数学相区别。二是统计指标说明的是客观存在的已经发生的事实,它反映了社会经济现象在具体地点、时间和条件下的数量变化。这一点又和计划指标相区别。统计指标反映的是过去的事实和根据这些事实综合计算出来的实际数量,而计划指标则说明未来所要达到的具体目标。

(三) 标志与指标的区别和联系

它们的主要区别是:

第一,标志是说明总体单位特征的,指标是说明总体特征的。例如,一个工人的工资是数量标志,全体工人的工资总额是统计指标。

第二,标志有用文字表示的品质标志和用数值表示的数量标志,指标则都是用数值表示的,没有不能用数值表示的指标。

它们的主要联系是:

第一,统计指标的数值多是由总体单位的数量标志值综合汇总而来的。例如,工资总额是各个职工的工资之和,工业总产值是各个工业企业的工业总产值之和。由于指标与标志的这种综合汇总关系,有些统计指标的名称与标志是一样的,如上例中的工业总产值。

第二,标志与指标之间存在着变换关系。如果由于统计研究目的的变化,原来的统计总体变成总体单位了,则相对应的统计指标也就变成了数量标志。反过来,如果原来的总体单位变成总体了,则相对应的数量标志也就变成了统计指标。

四、变异、变量和变量值

统计中的标志和指标都是可变的,如人的性别有男女之分,各时期、各地区、各部门的工业总产值各有不同等,这种差别叫做变异。变异就是有差别的意思,包括质的差

别和量的差别。变异是统计的前提条件。

变量就是可以取不同值的量，这是数学上的一个名词。在社会经济统计中，变量包括各种标志和统计指标。变量就是各种标志的名称或指标的名称，变量的具体数值表现则称为变量值或者观察值。例如，职工人数是一个变量，因为各个工厂的职工人数不同。某工厂有 852 人，另一工厂有 1 686 人，第三个工厂有 964 人，等等，都是职工人数这个变量的具体数值，也就是变量值。要注意区分变量和变量值。如上例，852 人、1 686 人、964 人三个变量值的平均数，不能说是三个变量的平均数，因为这里只有“职工人数”这一个变量，并没有三个变量。

变量按照是否可以连续取值可分为连续变量与离散变量两种。在一定区间内可任意取值的变量叫连续变量，其数值是连续不断的，相邻两个数值可作无限分割，即可取无限个数值。例如，生产零件的规格尺寸和人体测量的身高、体重、胸围等为连续变量，其数值只能用测量或计量的方法取得。可按一定顺序一一列举其数值的变量叫离散变量，其数值表现为断开的。例如，企业个数、职工人数、设备台数、学校数、医院数等，都只能按计量单位数计数，这种变量的数值一般用计数方法取得。

关键名词

统计学 统计工作 统计资料 总体 总体单位 样本 样本单位 统计标志
不变标志 可变标志 品质标志 数量标志 统计指标 变异 变量 变量值

思考和练习

思考题

1. 简述统计的含义。
2. 统计学区别于其他学科的特点是什么？
3. 总体和样本之间的关系是什么？
4. 什么是连续变量和离散变量？如何判断？

问题是：从表中数据是否能够看出2008年我国人口分布的规律性？如何才能找到人口分布的规律性？

第一节 统计数据的收集

客观现象是错综复杂的，要认识这种复杂现象的规律性，就必须从客观实际中收集大量的数据资料，并通过一定的方法进行加工和计算。而统计资料的收集正是根据统计研究的目的和任务，运用科学有效的方法，有计划、有组织地向客观实际收集数据资料的过程。从统计工作阶段来看，数据收集是统计工作的起点，处于统计工作过程的基础阶段，是统计整理和分析的前提。

一、数据来源

任何数据的初始来源基本上都始于调查或实验。但从数据使用者的角度来看，数据主要来源于两个渠道。

一个渠道是已经存在的数据，往往被称为二手数据或间接数据。所谓二手数据或间接数据，是指由研究者以外的机构或他人已经收集得到的数据。在一些情况下，如果发现与研究相关的所需要的数据已经存在，为了节省费用和时间，则可以直接引用。目前随着互联网的持续、快速发展，网络已成为二手数据和统计信息的一个重要来源。比如反映上市公司基本状况和财务信息的数据，各行业发展的相关数据，与居民生活相关的各种价格数据，宏观经济和社会发展的投资、消费、进出口、外汇储备等数据，都可以通过相关网站进行搜索和应用。也可以通过行业或统计部门公开出版的刊物和年鉴获得各种统计数据，如《中国统计年鉴》、《中国农村统计年鉴》、《中国工业经济统计年鉴》、《中国金融年鉴》、《中国保险年鉴》、《世界经济年鉴》、《世界发展报告》等。

可见二手资料来源广泛，既经济又方便，但在使用时一定要注意对其进行鉴别和选择，注意数据的含义、计算口径和计算方法，避免误用或滥用。

数据来源的另一个渠道是需要研究者亲自动手直接收集才能获取的数据，被称为一手数据或直接数据。其取得方式又有两种，即实验和非实验两种方法。

实验法是在所设定的特殊实验场所、特殊状态下，对研究对象进行实验以获取数据资料的一种方法。比如医学上对新药疗效的临床实验，可以获得新药品对疾病的治愈率能达到百分之多少的数据，或者通过实验的方式获得新药的剂量影响病人血压变动的数据等。实验方法主要应用于自然现象，目前也被越来越多地应用于心理学、教育学、社会学和经济管理学科中。比如一种新的商品在大量投产前，可以通过小范围的让消费者免费试用的方式，来得到消费者对新产品偏好的数据，以帮助生产者完善产品的设计，更好地获得市场份额，这实际上就是一种针对产品的市场实验。

但是总的来说，对于社会经济现象的研究，获取数据运用最多的方法还是非实验性的方法。常用的方法有直接观察法、访问法、问卷法、电脑辅助电话调查法等。