

第 1 章

总 论

【学习目标】

- (1) 了解统计学的产生和发展过程、统计学科的种类；
- (2) 明晰统计工作过程，明晰统计学的研究对象和研究方法；
- (3) 掌握统计及统计学的含义、特点及作用；
- (4) 熟练掌握统计学中常用的统计总体、总体单位、标志、指标、变量、变量值等基本概念；
- (5) 明晰统计总体与总体单位、品质标志与数量标志、连续变量与离散变量、标志与指标的区别与联系；
- (6) 了解统计的工作过程和统计学的应用领域；
- (7) 掌握 Excel 加载“数据分析”工具库的方法。

【教学引例】学统计，耳聪目明

2024 年 1 月 17 日，国家统计局发布了《2023 年居民收入和消费支出情况》：2023 年，全国居民人均可支配收入 39 218 元，比上年名义增长 6.3%，扣除价格因素，实际增长 6.1%。分城乡看，城镇居民人均可支配收入 51 821 元，增长（以下如无特别说明，均为同比名义增长）5.1%，扣除价格因素，实际增长 4.8%；农村居民人均可支配收入 21 691 元，增长 7.7%，扣除价格因素，实际增长 7.6%。2023 年，全国居民人均可支配收入中位数是 33 036 元。

2023 年，全国居民人均食品烟酒消费支出 7 983 元，增长 6.7%，占人均消费支出的比重为 29.8%；人均衣着消费支出 1 479 元，增长 8.4%，占人均消费支出的比重为 5.5%；人均居住消费支出 6 095 元，增长 3.6%，占人均消费支出的比重为 22.7%；人均生活用品及服务消费支出 1 526 元，增长 6.6%，占人均消费支出的比重为 5.7%；人均交通通信消费支出 3 652 元，增长 14.3%，占人均消费支出的比重为 13.6%；人均教育文化娱乐消费支出 2 904 元，增长 17.6%，占人均消费支出的比重为 10.8%；人均医疗保健消费支出 2 460 元，增长 16.0%，占人均消费支出的比重为 9.2%；人均其他用品及服务消费支出 697 元，增长 17.1%，占人均消费支出的比重为 2.6%。

资料来源：2023 年居民收入和消费支出情况[EB/OL]. (2024-01-17). https://www.stats.gov.cn/sj/zxfb/202401/t20240116_1946622.html.

1.1 统计与统计学的产生和发展

人们在日常工作、生活、学习中,经常会接触到各类统计数据、统计图表,也常常要收集、处理、分析和使用各类统计数据,特别是在数字化时代,统计更是渗透于人类社会经济生活的方方面面,渗入人们日常和经济生活的每个角落。

1.1.1 统计的含义

人们一提到“统计”,首先想到的就是具体的统计工作。的确,统计工作是统计,但不是统计的全部。实际上,“统计”一词已被人们赋予了多种含义,在不同场合、不同的语言环境中有不同的解释。但普遍认为统计包含三层含义,即统计工作、统计资料和统计学。

统计工作,也即统计实践,它是对社会经济现象数量方面进行收集、整理、分析和提供数字资料工作的总称。例如,银行的统计科,每月要编制项目报表,这个过程就是统计工作;在我国,各级统计部门从事的统计数据收集、整理和分析工作也是统计工作。比如,负责全国人口、经济、社会等方面的统计调查。再比如,每10年进行一次的人口普查,人口普查时要经过方案设计、入户登记、数据汇总、分析总结和资料公布等一系列统计工作,并统计国家或地区的人口总量,同时也分类统计男性人口、女性人口,老年人口、中青年人口、儿童人口等数量,从而为国家或地区制定与人口相关的各项方针政策提供依据。又如,处理大量的经济数据,如国内生产总值(GDP)、月度居民消费价格指数(consumer price index, CPI)等,这些数据直接影响到国家的宏观经济政策,也是衡量通货膨胀的重要指标。此外,还要进行社会调查,如就业率、教育水平等,为制定政策提供参考依据。统计工作,也简称统计。

统计资料(统计信息),也即统计数据,它是统计工作的成果,是统计工作过程中所取得的各项数字资料和与之相关的其他实际资料的总称。

例如,2023年,我国国内生产总值超过126万亿元,比上年增长5.2%。四个季度GDP同比分别增长4.5%、6.3%、4.9%、5.2%,呈现前低、中高、后稳态势。^①这些由文字和数字共同组成的数字化的信息就是统计资料,是统计提供数据信息的基本表现形式,是统计工作的直接成果。

统计资料,包括原始资料和经过整理、分析形成的统计分析报告等次级资料。原始资料,如企业各车间的统计台账、人口普查时初次登记的资料等。而统计公报、全国房地产市场调查分析报告等现实和历史资料则是次级资料。统计资料的表现形式有统计表、统计图、统计分析报告、统计公报和统计年鉴等。统计资料,也可简称统计。

统计学,也即统计理论,是一门关于数据收集、整理、分析、解释和推断的科学。统计学是随着统计活动的不断发展和统计实践经验的日益丰富应运而生的。关于统计学的定义,已出版的国内外统计学教科书上有多种表述。

(1) 国家统计局(2022.11.29): 统计学是关于数据的一门学问。所有收集而来的数据

^① 2023年国民经济回升向好 高质量发展扎实推进[EB/OL]. (2024-01-17). https://www.stats.gov.cn/sj/zxfb/202401/t20240117_1946624.html.

都需要经过整理、分析才能得出结论,这就是统计学利用数据解决实际问题的全过程。……统计学,是一门找出统计规律的学问。

(2)《大不列颠百科全书》:统计学是收集、分析、表述和解释数据的科学。

(3)戴维·R.安德森(David R. Anderson)等(2006.1.商务与经济统计.北京,机械工业出版社):统计学是收集、分析、表述和解释数据的艺术和科学。

(4)吴喜之(2004.统计学:从数据到结论.北京,中国统计出版社):统计学是用以收集数据、分析数据和由数据得出结论的一组概念、原则和方法。

(5)道格拉斯·A.林德(Douglas A. Lind)等(2005.9.商务与经济统计技术.易丹辉等译.北京,中国人民大学出版社):统计学是对数据收集、组织、展示、分析和解释,从而帮助作出更为有效的科学的决策。

(6)曾五一,肖红叶(2006.统计学导论.北京,科学出版社):统计学是有关如何测定、收集、整理、归纳、分析反映客观现象总体数量的数据,以便给出正确认识的方法论科学。

(7)贾俊平(2006.9.统计学.北京,中国人民大学出版社):统计学是收集、处理、分析、解释数据并从数据中得出结论的科学。

(8)向蓉美(2017.1.统计学导论.成都,西南财经大学出版社):统计学,是一门收集数据、表现数据、分析数据、解释数据,从而认识现象数量规律、帮助人们更有效地进行决策的方法论科学。

(9)李金昌,苏为华(2019.统计学.北京,机械工业出版社):统计学是关于如何收集、整理和分析统计数据的科学。

综上所述关于统计学的定义,本书将统计学定义概括为:统计学是一门关于如何收集数据、整理数据、分析数据,并进行描述、推断和预测,从而正确认识现象总体数量方面的科学。当然,统计学也可以简称统计。

总之,统计是人们认识客观世界总体数量变动关系和变动规律活动的总称,是人们认识客观世界的一种有力工具。统计三种含义之间既有联系也有区别。统计资料是统计工作的成果;统计学是统计工作和统计资料的理论概括,而统计学形成的理论又指导统计工作有效进行,即统计工作一方面受统计理论的指导,另一方面又检验统计理论是否正确,并推动和促进统计理论向前发展;统计学与统计资料存在密切关系,统计学阐述的统计方法来源于对统计数据的研究,离开了统计数据,统计学就失去了存在的意义。统计学与统计工作、统计资料之间的关系表明:统计理论来源于统计实践,反过来又为统计实践服务,统计理论与统计实践是辩证统一的关系。

1.1.2 统计实践的产生与发展

统计,作为人类的一项社会实践活动,统计学,作为人们通过对事物的数量特征进行收集、整理、分析和推断预测,进而认识事物总体规律的一门科学,是随着人类的生产实践活动的产生和发展而产生和发展起来的。

统计实践萌芽于奴隶社会,当时的统治阶级为了治理国家,常常要进行征税、征兵、征劳役等统治活动,需要了解社会的基本情况,这就产生了统计。我国在原始社会末期,在奴隶制形成的过程中就已经出现了统计的萌芽。在距今4000多年前的夏朝,中国被分为九州,

人口 1 355 万,这是我国最早的土地与人口调查的资料。到了西周已建立统计报告制度;日成、月要和岁会,这项制度经春秋战国,至秦统一中国,形成了比较完善的“上计”报告制度,同时也逐渐形成了“举事必成,不知计数不可”和“强国知十三数”的统计思想。到了封建社会,统计不仅略具规模,而且已居当时世界先进水平。秦始皇最早建立了全国规模的人口调查登记制度,东汉进行了全国性土地测量,唐代有了计口授田制度,宋、明有田亩鱼鳞册,等等。

在西方,统计也有很长的历史。埃及在公元前 27 世纪进行过全国人口和财产调查。16 世纪初,伦敦出现“死亡公报”,公布死亡人数,其记录的人数,从开始的瘟疫死亡,扩充为受洗礼死亡,到世纪末的疾病死亡。17 世纪以来,资本主义国家由于工、商、农、贸、交通的发展,统计实践从国家管理领域扩展到社会经济活动的许多领域。18 世纪起,各资本主义国家都先后设立了专业的统计机关,收集各方面的统计资料,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,建立国际统计组织,召开国际统计会议。19 世纪中叶后,随着数学特别是古典概率的引入,统计活动有了飞跃发展。

20 世纪 30 年代,各国普遍采用了随机抽样的方法。20 世纪 40 年代,电子计算机的应用,不仅提高了统计数据处理的效率和准确性,也为数据的积累、存储、更新、检索、加工及其一体化的自动化创造了物质技术基础。20 世纪 50 年代,以国民经济为整体的统计研究方法(国民经济账户体系及投入产出法)的推广应用,大大提高了世界各国对统计的认识和应用能力。

1.1.3 西方统计思想的形成与发展

封建社会生产的落后,限制了统计实践活动和统计理论的发展。资本主义经济的发展,特别是现代化大生产,促进了统计活动和统计科学的发展。统计学作为一门科学,其形成大体分为古典统计学时期、近代统计学时期和现代统计学时期。

1. 古典统计学时期

17 世纪中叶至 18 世纪是统计学形成初期,当时有政治算术学派、国势学派两大学派。其中,政治算术学派有统计学之实无统计学之名,国势学派有统计学之名无统计学之实。

政治算术学派的创始人是英国人威廉·配第(William Petty)(1623—1687)。配第首先提出了用数量方法科学地研究社会经济现象——政治算术,他的名著《政治算术》(1690)中的政治算术就是指数字和统计学方法。在序言中,配第明确指出:“我进行了这项工作,所使用的方法在目前还不常见。因为与只使用比较初级和最高级的词语以及单纯做思维的论证相反,我采用了这样的方法:用数字、重量和尺度等词汇来表示展望和论旨,都是真实的,即使不真实,也不会有明显的错误……”他在研究社会经济现象的规律时还应用推算法、分组法,编制原始数据的图表,计算一系列的总量指标、相对指标和平均指标,配第是最早估算国民收入的人。但是配第始终没有用“统计学”三个字。所以,政治算术学派有统计学之实,但无统计学之名。

国势学派的创始人是德国人赫尔曼·康令(Hermann Conring)(1606—1681)。康令开始定期地、系统地用对比的方法,讲述国家比较方面的知识,不仅讲述事实,而且试图探讨事实的因果关系,他把这门课程称为“欧洲最近国势学”,于是“国势学”由此产生。因为当时康令的学说在学术界影响很大,德国大学的许多教授都称赞并追随康令的学术思想,并把这门

课程定名为“统计学”，开始有了“统计学”这个名称。但是国势学派只是对各国情况做一般性的比较记载，如“某国人口众多”“土地辽阔”之类，而没有进行数量研究和描述。所以，国势学派有统计学之名，但无统计学之实。

2. 近代统计学时期

18世纪末至19世纪末是近代统计学时期，在这个时期，各种学派的学术观点已经形成，并且形成了两个主要学派，即数理统计学派和社会统计学派。

(1) 数理统计学派。18世纪，概率理论日渐成熟，为统计学的发展奠定了基础。19世纪中叶，概率论被引进统计学，从而形成了数理统计学派。数理统计学派的奠基人是比利时的阿道夫·凯特勒(Adolphe Quetelet)(1796—1874)。凯特勒在他的《社会物理学》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。凯特勒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，依此方法对自然现象和社会现象的规律进行观察，并认为要促进科学的发展，就必须更多地应用数学。总之，凯特勒把概率论引入统计学，为数理统计学的形成与发展奠定了基础。

(2) 社会统计学派。社会统计学派产生于19世纪后半叶，创始人是德国的经济学家、统计学家卡尔·古斯塔夫·阿道夫克尼斯(Karl Gustav Adolf Knies)(1821—1898)，以及恩斯特·恩格尔(Ernst Engel)、格奥尔·冯·迈尔(Georg von Mayr)等。他们融合了国势学派与政治算术学派的观点，在学科性质上认为：统计学是一门社会科学，是研究社会现象变动原因和规律性的实质性科学，以此与数理统计学派通用方法相对立。社会统计学派在统计对象上认为，统计学是研究总体的，而不是研究个别现象的，并且认为由于社会现象的复杂性和整体性，必须对总体进行大量观察和分析，研究其内在联系才能解释现象内在规律性，这是社会统计学派的实质性科学的显著特点。

数理统计学派与社会统计学派并存，争论了百年之久。在早期，由于数理统计学派尚未充分发展，社会统计学派在欧洲占有优势地位，对世界各国都有很大的影响。随着数理统计理论不断发展，其应用领域也不断扩大，数理统计学派开始在国际统计学界占据优势地位，但总的来说，二者出现了融合趋势。

3. 现代统计学时期

20世纪至今为现代统计学时期，其标志是推断统计学的问世。1907年，英国人威廉·西利·戈塞特(William Sealy 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。英国科学家弗朗西斯·高尔顿(Francis Galton)提出了相关与回归思想，并给出了计算相关系数的明确公式。英国统计学者卡尔·皮尔逊(Karl Pearson)发展了拟合优度检验，还给出了卡方统计量及其极限分布理论。波兰学者耶日·内曼(Jerzy Neyman)(1894—1981)创立了区间估计理论，并与皮尔逊一同发展了假设理论，等等。

20世纪50年代以来，统计理论、方法和应用进入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。无论是自然科学、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,因而可以说统计学与数学、哲学一样成为所有学科的基础。

总之,统计学大致经过三个发展阶段后,理论得到不断丰富和完善。目前,统计学越来越多地吸收数学方法,也越来越多地向其他学科领域渗透,形成各种以统计学为基础的边缘学科。随着统计学应用日益广泛和深入,特别是借助计算机等,统计学的功效越来越广泛,作用越来越强劲。

1.1.4 我国统计发展简史

新中国成立前,我国统计工作十分落后,统计学基本上照抄照搬西方统计理论,传播的主要是数理统计学派的观点。

新中国成立后,我国在学习苏联统计工作经验的同时,引进了苏联的统计学,即社会经济统计学。数理统计遭到批判和抛弃。1978年,党的十一届三中全会后,学术界百花齐放、百家争鸣。数理统计又重新受到人们的关注和重视,统计学者突破了狭隘观念的桎梏,认为:社会经济统计学、数理统计学和自然科学方面的统计学都是独立的统计学科,三者可以同时存在、互相借鉴、共同发展,形成了大统计观念。1996年10月,我国确立了统计学科体系的基本框架,肯定了统计学是包括社会经济统计学和数理统计学在内的一般方法论性质的科学,这为我国统计学的发展奠定了坚实的基础。

总之,随着大统计体系的建立,统计学作为一门独立的学科,其运用已经渗透到自然科学和社会科学的各个领域,统计科学工作者在总结我国统计实践经验的同时,不断吸收世界各国统计科学发展的成果,使我国现阶段统计学发展有三个明显的趋势:一是统计学依赖和吸收数学理论更多;二是以统计学为基础的边缘学科不断形成;三是与计算机技术相结合,借助大数据平台,统计学应用范围更广,统计学作用更强大。

【同步思考 1-1】

1. 什么是统计?统计三种含义之间有什么关系?
2. 在统计学发展史上主要有哪几种学派?其主要代表人物是谁?
3. 现代统计学与近代统计学最主要的区别有哪些?我国现阶段统计学发展的主要特点是什么?

1.2 统计学的研究对象和分科

1.2.1 统计学的研究对象及其特点

1. 统计学的研究对象

统计学的研究对象是指统计研究所要认识的客体,只有明确了统计的研究对象,才能根据统计研究对象的特点和性质,采用相应科学的研究方法,以达到认识对象客体规律性的目的。由统计学的发展史可知,统计学是从研究社会经济现象的数量开始的,随着统计方法的

不断完善,统计学得以不断发展。因此,统计学的研究对象是大量现象总体的数量方面。而社会经济统计学是在质与量的密切联系中,研究大量的社会经济现象总体的数量方面,即研究社会经济现象总体的数量特征和数量关系。

统计学在研究社会经济现象时,从定性研究开始,然后再进行定量分析,最后达到认识社会现象的本质、特征或规律,这就是质—量—质的统计研究过程和方法。

2. 统计学研究对象的特点

统计学研究对象的特点,最集中体现在总体性和数量性两个方面,而具体性、社会性和广泛性也是统计学研究有别于其他学科研究的特点,如图 1-1 所示。

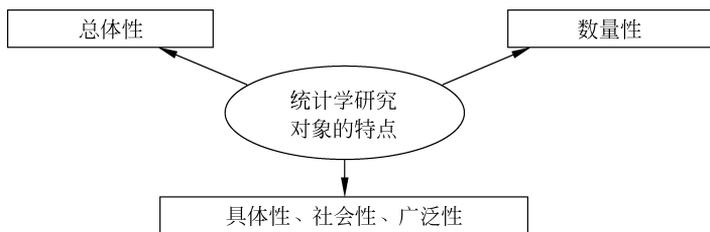


图 1-1 统计学研究对象的特点

(1) 总体性。统计工作研究的是总体的数量特征而非个体数量表现。因此,只有把大量的个体数量资料经过汇总、综合,才能表现出总体数量特征。从总体上研究现象的数量方面,是统计学的重要特点。社会经济现象是各种社会规律交错作用的结果,呈现出一种复杂多变的情景。统计对社会经济现象总体数量方面的调查研究,是从个体到总体,即必须对足够大量的个体(这些个体都表现为一定的差别、差异)进行登记、整理和综合,使它过渡到总体的数量方面,从而把握社会现象的总规模、总水平及其总发展变化趋势。例如,我国每 10 年进行一次的人口普查,要从登记每一个人的具体情况入手,但调查目的却不是要了解每个人的具体生活状况,而是要查清我国在人口数量、地区分布、构成和素质方面的变化,为科学制定国民经济和社会发展战略与规划,安排人民的物质、文化生活等提供可靠的资料。总之,统计研究最终要得到反映总体数量的指标,但统计工作是从研究个体开始的。只有从个体开始,才能对总体进行分析研究。

(2) 数量性。统计学的研究对象是大量的社会经济现象的总体数量方面,包括社会经济现象数量方面的规模、水平、结构、速度、平均水平、平均发展速度以及数量关系、数量界限等。例如,某地区 2024 年上半年规模以上工业企业中,国有控股企业实现利润总额 38.4 亿元,同比下降 2.1%;股份制企业实现利润总额 510.5 亿元,同比增长 1.6%。所有这些指标都是从数量方面反映现象发展变化情况。统计的这一数量性特点将统计学与其他实质性社会科学,如历史学、哲学等学科区分开来,也将统计调查研究活动与那些以非数量性为主的调查研究活动,如社会、法律、考古等区分开来。

统计学研究的是大量的社会经济现象总体数量方面,但应该注意到统计的定量研究是建立在定性研究基础上的。例如,要想知道国内生产总值是多少,必须知道什么是国内生产总值,国内生产总值与国民生产总值有什么区别,国内生产总值与社会总产值有何区别,等等。因此,统计不是单纯地研究社会经济现象的数量方面,而是在质与量的密切联系中研究现象的数量方面。定性研究是基础,定量研究是目标。质—量—质,是一个完整的统计研究过程。

(3) 具体性。统计工作研究的总体数量是一个有具体时间、具体地点、具体条件限定的数量。如果单说“39 218 元”，这个数字没有任何意义，但如果说 2023 年，全国居民人均可支配收入 39 218 元，比上年增长 6.3%，扣除价格因素，实际增长 6.1%^①，这就是统计中所说的具体数量。因此，具体性是指在时间、地点、空间三方面都有明确的规定。

应该指出，虽然统计工作研究具体的数量，但为了进行复杂的定量分析，需要借助抽象的数学模型和数理统计方法，还要求遵循数学规则。因此，统计工作具体的数量研究需要密切联系抽象的数学方法，以抽象方法为手段，以具体数量为目的，体现统计工作中具体和抽象的辩证关系。

(4) 社会性。统计学的研究对象是人类社会活动的过程和结果。人类的社会活动都是人们有意识、有目的的活动，各种活动都贯穿人与人之间的关系，除了随机现象之外，还存在着许多确定性的因素。由于统计工作存在明显的社会性，各国政府都很重视统计工作的开展。因此，特别强调要克服统计工作中的主观随意性，抑制任意夸大或缩小统计数字，歪曲反映实际的现象出现。要强化统计法治建设，建立健全数据追溯和问责机制，确保数据可核查、可追溯、可问责。

《中华人民共和国统计法》(以下简称《统计法》)于 1983 年 12 月 8 日通过，1996 年 5 月 15 日修正，2009 年 6 月 27 日进行了再修订。统计立法的目的，就是科学、有效地组织统计工作，保障统计资料的真实性、准确性、完整性和及时性。

(5) 广泛性。统计学研究的是所有社会经济现象的数量方面。统计学既研究生产关系，又研究生产关系与生产力之间的关系；既研究经济基础，又研究经济基础与上层建筑之间的关系。同时，统计学还研究生产、流通、分配、使用等社会再生产的全过程以及社会、政治、经济、文化、教育等全部社会经济现象的数量方面。特别是随着互联网云技术、云服务和云应用在数据处理中的广泛应用，统计学在实践中的应用越来越普遍、越来越广泛。

1.2.2 统计学的分科

统计学是从研究社会经济现象开始，逐渐趋于成熟，并成为一门研究客观事物总体数量方面的方法论科学。这里所指的方法论包括：指导统计活动的原理和原则，统计核算和分析方法。这些方法是在统计实践中产生的，再经过理论概括，反过来又用于指导统计实践，为统计工作服务。由于统计学的研究对象既存在于自然领域，也存在于社会领域，因此，统计学是一门具有跨学科性质、有较高概括能力和较大适用范围的一般方法论学科。

按照统计方法的类型，统计学可分为描述统计学和推断统计学；按照统计研究方法和应用程度，统计学可分为理论统计学和应用统计学。

1. 描述统计学和推断统计学

(1) 描述统计学，是研究如何取得反映客观现象的数据，并通过图表形式对所收集的数据进行加工处理和显示，进而概括并分析得出反映客观事物的规律性数量特征的科学。它包括统计数据的收集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与

^① 2023 年国民经济回升向好 高质量发展扎实推进[EB/OL]. (2024-01-17). https://www.stats.gov.cn/sj/zxfb/202401/t20240117_1946624.html.

分析方法等。例如,2023年末全国人口140 967万,比上年末减少208万。从性别构成看,男性人口72 032万,女性人口68 935万;从年龄构成看,16~59岁的劳动年龄人口86 481万,占全国人口比重为61.3%;60岁及以上人口29 697万,占全国人口21.1%,其中65岁及以上人口21 676万,占全国人口15.4%。^①通过这些数字资料可以看出我国人口规模及内部构成情况。

(2) 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作出以概率形式表述的推断。例如,对一批出口草莓罐头进行产品质量检验,不能每一瓶都检验,只能抽取一部分进行检验,只要抽样合理,就可以根据抽检结果来估计和推断全部出口产品的质量。由于数据来源于概率抽样,因此,推断统计学的各种推算方法和推断结果的合理性与可靠程度都是以概率论为基础的。描述统计学与推断统计学之间的关系如图1-2所示。

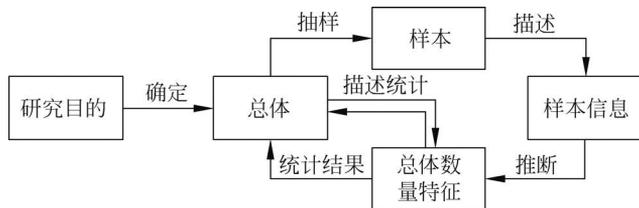


图 1-2 描述统计学与推断统计学之间的关系

描述统计学和推断统计学是统计方法的两个组成部分。描述统计学是整个统计学的基础,推断统计学是现代统计学的主要内容,并且它在现代统计学中的地位和作用越来越重要,已经成为统计学的核心内容。这是因为,在对现实问题的研究中,当总体不是很大时,我们采取全面调查来收集、整理和分析数据资料,而当总体很大时,通常要采用抽样调查来取得样本数据资料。但这并不等于描述统计学不重要,因为如果没有描述统计学收集可靠的统计数据并提供有效的样本信息,即使再科学的统计推断方法,也难以得出符合实际的结论。总之,从描述统计学发展到推断统计学,既是统计学发展的巨大成就,也是统计学走向成熟的重要标志。

2. 理论统计学和应用统计学

在统计学科的发展中,理论统计学和应用统计学是相互促进、共同提高的。

(1) 理论统计学,是论述统计学的基本理论、原理和统计方法的一门方法论科学。它广泛采用数理统计方法,集社会经济统计方法与数理统计方法之大成,既适用于社会经济现象的数量观察和研究,也适用于自然现象的数量观测和推断。理论统计学是统计学科的基础,理论统计学的特点是计量不计质,它具有通用方法论的理学性质。

(2) 应用统计学,是指研究如何应用统计方法去解决实际问题的科学。由于在自然科学及社会科学研究领域中,都需要通过数据分析来解决实际问题,因而,统计方法的应用几乎扩展到了所有的科学研究领域,如生物统计学、医学统计学、农业统计学、工业统计学、经济统计学和社会统计学等。需要指出,所有这些应用统计学的不同分支,都主要应用描述统

^① 2023年国民经济回升向好 高质量发展扎实推进[EB/OL]. (2024-01-17). https://www.stats.gov.cn/sj/zxfb/202401/t20240117_1946624.html.

计和推断统计的方法。

总之,作为应用统计工作者,不但要熟练掌握和应用各种统计方法,而且必须具备所研究和应用领域的专业知识,即专业知识和统计方法是做好统计应用的两项基本功。例如,要做好经济统计工作,其基本条件是具备经济理论和统计方法的知识能力,此外,还需具备计算机操作能力、写作能力和其他能力,即统计应用人才是一名复合型人才,这也是人文、社科、理、工、农、医各院校开设统计学课程的原因所在。统计学的分科如图 1-3 所示。

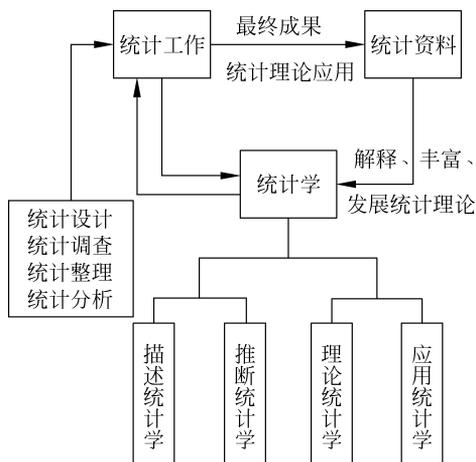


图 1-3 统计学的分科

1.2.3 统计的职能和统计学的作用

1. 统计的职能

在统计发展中,虽然其职能发生了很大变化,但反映国情国力、为国家管理提供基础信息的职能却始终没变,并且得到不断强化。

统计的职能,可以概括为三大方面,即信息职能、咨询职能和监督职能。

(1) 信息职能,是指根据科学的统计指标体系和统计调查方法,系统地收集、整理、传递、存储和提供大量以数量描述为基本特征的统计信息的职能。统计信息,是覆盖面最广、综合性最强的信息,因而是经济社会信息的主体,是党和政府以及社会各界了解情况、研究问题、进行科学决策和管理的重要依据。

(2) 咨询职能,是指统计机构和统计人员利用已经掌握的丰富统计信息,运用先进的技术手段和科学方法,深入开展综合分析和专题研究,通过去粗取精、去伪存真、由表及里、由此及彼的加工过程,透过表层数量关系,探求事物发展的内在联系和规律性,为党和国家提供各种咨询建议和对策方案,对科学决策和管理发挥参谋和助手作用的职能。

(3) 监督职能,是指根据统计调查和分析,及时、准确地从总体上反映经济、社会和科技的运行状态,并对其实行全面、系统的定量检查、监测和预警,以促使国民经济按照客观规律的要求,持续、稳定、协调地发展的职能。党的十九大报告从提高党的执政能力和领导水平、健全党和国家监督体系的角度,作出“完善统计体制”的重大部署,进一步强化了统计监督职能。统计部门作为我国重要的国民经济和社会发展评价和监督部门,将通过科学健全的统

计调查制度、监测制度和考核评价指标体系,为考核评估监督各地区、各部门国民经济和社会发展主要任务完成情况以及国家重大政策措施落实情况提供相关统计数据、监测数据和评价依据。

例如,2023年末全国全年出生人口902万人,出生率为6.39‰;死亡人口1110万人,死亡率为7.87‰;自然增长率为-1.48‰。而2013年末全国全年出生人口1640万人,出生率为12.08‰;死亡人口972万人,死亡率为7.16‰;自然增长率为4.92‰。^①由数据可见,2023年与2013年比,无论是新生人口总数还是人口出生率都有较大程度的减少和下降,人口数据信息,客观上要求国家相关部门及时出台相关鼓励生育的政策与措施,实现人口的良性发展。当然也可以在政策和措施实施后,通过实施前后的数据对比,评价政策和措施的实施效果。

统计信息职能、统计咨询职能、统计监督职能之间的关系如图1-4所示。

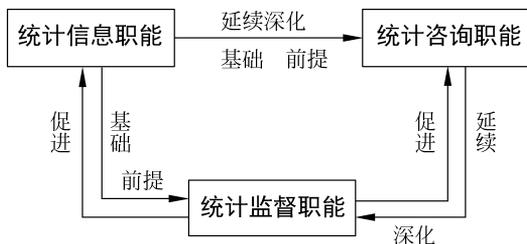


图 1-4 统计信息职能、统计咨询职能、统计监督职能之间的关系

2. 统计学的作用

统计学是一门应用性很强的学科。随着社会的发展,统计学的应用领域越来越广,并在认识事物、指导生产、经济管理和科学研究等方面都发挥出越来越重要的作用。

(1) 统计学为人们认识自然和社会提供了可行的方法和途径。人们要认识自然、认识社会,离不开各种各样的数据资料,并对这些资料进行分析研究,这就要用到统计理论和统计方法。例如,季节变动会对人们的日常生活和经济生活产生影响。微观经济领域的专家需要使用统计方法为决策者提供生产、消费和定价方面预测;宏观经济领域的专家可以利用统计方法描述居民家庭收入分布状况,对经济变量(如通货膨胀率、失业率、居民家庭平均收入水平等指标)的未来水平进行分析和预测,还可以对变量(如可支配收入与消费)之间的关系进行研究。统计是人们认识自然和社会的手段,统计学为人们利用数据进行认识和分析提供了理论和方法。

(2) 统计学对指导生产活动过程具有重要作用。生产的任务是要以尽量少的投入,产出数量多、质量好的产品。而影响产品数量和质量的因素很多,也有主次之分。如果我们对各种因素进行试验和观察,了解和分析各因素的影响方式和影响程度,找出最佳因素组合,就能少投入、多产出。此外,在各种情报信息、商品信息中,用统计方法对这些信息进行分析,能帮助企业对事物进行定量、定性分析,从而作出正确的决策。例如,企业市场营销部门运用网上零售统计信息,以及第三方支付数据信息,了解顾客对某一种商品喜爱的比例,以及他们为什么喜欢该种商品,用何种广告能让更多人知道、喜欢、购买该种商品等,从而增强

^① 根据 2014 年、2024 年国家统计局(<https://www.stats.gov.cn>)国民经济与社会发展统计公报数据整理。

企业竞争力、提高企业的经济效益。

(3) 统计学为经济管理提供了科学方法。统计数据是管理的重要基础。例如,生产定额的制定,既要使人们有可能完成,又要出效率,那么利用先进算术平均数不失为一种好方法。又如,可以通过抽样调查和样本分析等方法,对潜在风险进行预警和监测,及时采取相应的措施应对。当然,在宏观上,还可以通过对经济运行数据的实时监测和分析,及时发现经济发展中的问题,为政策制定者提供预警信息。这种预警机制有助于政府及时采取措施,防范和化解经济风险。也可以在政策实施后,通过对政策实施前后的数据进行对比分析,为政策效果评估提供科学依据。这不仅有助于政策的完善和优化,还能提高政策制定的针对性和有效性。

(4) 统计学为科学研究提供了有力手段。科学研究的任务是揭示客观事物的规律,而科学研究的方法大多是先根据若干观察或实验资料提出某种假设或猜想,然后再通过各种途径进行观察或实验加以验证,统计理论和方法就是这种科学研究的通用方法。统计学理论和方法,一方面有助于集中并提取观察实验中最本质的东西,从而有助于提供较正确的假设或猜想;另一方面,它又能指导研究人员如何去安排进一步的观察和实验,以判定提出的猜想或假说是否正确。例如,在医学界,人们利用统计方法来研究疾病的成因或影响因素,判断药物或医疗方案是否正确;在生物学界,人们用统计方法来研究基因分离定律、基因自由组合定律、基因稳定性定律等基因规律;在心理学界,人们用统计方法分析特定刺激的心理效应;在经济学界,人们用统计方法来研究经济运行状况及宏观决策、微观决策的正确与否,以及政策、决策的监督执行和调整。可以说,几乎所有的科学研究领域都离不开统计学。

总之,统计无时不有,统计无处不在。在市场调研与消费者行为分析、预测与决策分析、风险评估与控制、经济政策的评估和决策支持、效率评估和质量控制以及评估和监测经济政策效果等方面都大有用武之地。只要科学、准确、充分地应用统计,统计就能发挥出强大的作用。

【同步思考 1-2】

1. 怎样理解统计研究对象的数量性? 统计的数量性与数学的数量有什么不同? 总体数量特征与个体数量特征之间有什么关系?

视频 1 统计工作
利国利民



2. 统计学是如何分科的? 描述统计学与推断统计学、应用统计学之间有什么关系?

3. 统计学有什么作用? 怎样理解“统计无时不有,统计无处不在”?

1.3 统计数据类型与统计研究方法

1.3.1 统计数据类型

在信息时代,统计数据无处不在。从不同角度,统计数据有不同的分类方法。

1. 按照数据计量尺度不同,分为分类数据、顺序数据和数值型数据

(1) 分类数据。分类数据是指只能归于某一类别的非数字型数据,表现为类别,用文字

表述事物分类结果。例如,学生按专业分为统计学、会计学、金融学等,人口按性别分为男、女两类。但这种分类,也可以用数字代号来表示,以便统计整理。例如,用“1”代表统计学,用“2”代表会计学,用“3”代表金融学等;用“0”代表男性,用“1”代表女性。这时的数字只是代表不同类别的一个代码,没有任何程度上的大小之分,也没有数学运算意义。

(2) 顺序数据。顺序数据是指只能归于某一有序类别的非数字型数据,表现为有顺序的类别,用文字来表述。例如,产品质量分优、良、合格、不合格;教师职称分初级、中级、高级;学生对食堂服务满意度分满意、一般、不满意;等等。

在顺序数据中,虽然结果也表现为类别,但各类别之间却可以比较顺序。它具有“ $>$ ”或“ $<$ ”的数学特性。同样,顺序数据也可以用数字代码来表示,但其数字代码体现一种顺序或程度上的不同和数量上的差别。

(3) 数值型数据。数值型数据是指按数字尺度测量的观察值,是使用自然或度量衡单位对事物进行计量的结果,其结果表现为具体的数值,通常又分为定距数据和定比数据。

定距数据,是一种不仅能反映事物所属的类别和顺序,还能反映事物类别或顺序之间数量差距的数据,表现结果为数值。例如,两名学生统计学考试成绩分别为90分和52分,不仅说明前者成绩高于后者,还说明两者差38分。

定比数据,又称比率尺度或比较水平,也称比例尺度或等比尺度,是一种除了上述三种尺度的全部性质之外,测量不同变量(社会现象)之间的比例或比率关系的方法。它是在定距尺度基础上,确定相应的比较基数,然后对两种相关数进行对比而形成的相对数(或平均数),用来反映现象的结构、比重、速度、密度等数量关系。例如,将一个企业职工的工资总额与该企业的职工人数对比,计算出平均工资,并以平均工资反映该企业职工的平均收入情况。定比数据的主要数学特征是“ \times ”或“ \div ”。定比数据在统计对比分析中应用十分广泛。

2. 按照来源不同,分为观测数据、实验数据与大数据

(1) 观测数据。观测数据是通过统计调查或观测的方式而获取的反映研究现象客观存在的数量特征的数据,这类数据是在没有人为控制的条件下获得的。有关社会经济现象的统计数据几乎都是观测数据,比如前面提到的各种统计数据。

(2) 实验数据。实验数据是在人为控制的条件下,通过实验的方式获得的关于实验对象的数据。自然科学研究中的数据大都属于实验数据,如生物实验数据、产品性能实验数据、药物疗效实验数据等。随着实验方法在经济等领域的应用,逐步形成了实验经济等学科,在经济等领域出现了许多实验数据。

(3) 大数据。大数据是基于现代信息技术而产生的数字化信息,其本意是那些体量巨大、类型多、结构复杂、增速快,以至于依赖现有的主流软、硬件工具无法及时加以消化处理的数据,它通常具有体量巨大、形式多样、快速增加和蕴含信息价值四大特点。

3. 按照加工程度不同,分为原始数据与次级数据

(1) 原始数据。原始数据是指直接向调查对象收集的、尚待加工整理、只反映个体特征的数据,或通过实验采集的原始记录数据。原始数据是统计数据收集的主体。

(2) 次级数据。次级数据也称加工数据或二手数据、三手数据甚至更多手数据,是指经过加工整理、能反映总体数量特征的各种非原始数据。次级数据又包括直接根据原始数据整理而来的汇总数据,以及根据各种已有数据推算而来的推算数据。如果次级数据满足有

关分析和研究需要,就没必要去收集原始数据,以免造成财物及时间的浪费。

4. 按照时间或空间状态不同,分为时序数据和截面数据

(1) 时序数据。时序数据是时间序列数据的简称,是对同一现象在不同时间上收集到的数据(空间状态相同,时间状态不同),描述的是现象某一方面(或某几方面)的数量特征随时间变化的情况。例如,把2000年以来我国国内生产总值数据按时间先后顺序加以排列,就形成了国内生产总值的时序数据。

(2) 截面数据。截面数据是指对同类现象在相同或近似时间上收集到的数据(空间状态不同,时间状态相同),描述的是在相同时间状态下同类现象数量特征在不同空间状态下的差异情况。例如,2023年我国各省、区、市的地区生产总值数据,就是截面数据。

有时,时序数据与截面数据可以结合起来,成为平行数据(时间状态、空间状态都不同),如列出2000—2023年各省、区、市的地区生产总值数据,就是平行数据。

统计数据的分类方法,如图1-5所示。

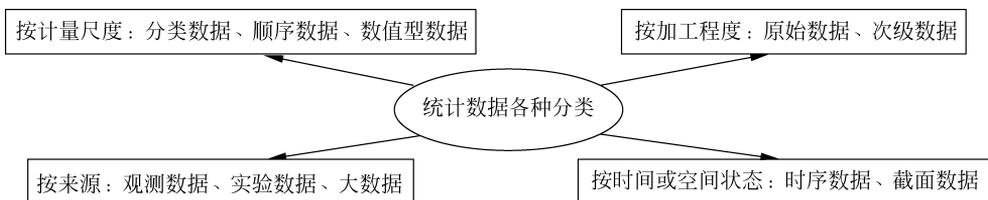


图 1-5 统计数据的分类方法

1.3.2 统计数据研究过程

统计工作是一种对社会经济现象总体数量进行调查研究的活 动,它是对事物的表面、本质及其规律性的认识活动。这一活动是由浅入深的过程,通常这个过程可以概括为统计设计、统计调查、统计整理、统计分析四个阶段。

1. 统计设计

统计设计是统计工作的第一个阶段,它根据统计研究对象的性质和研究目的,对统计工作的各个环节和各个方面进行统筹安排。统计设计的结果表现为各种统计设计方案,如统计指标与统计指标体系、分类目录、统计报表制度、调查方案、汇总或整理方案等。

统计设计包括对统计活动全过程设计和单项统计设计两个方面。对统计活动全过程设计是指针对一项统计研究任务,对收集、整理、分析数据的工作全过程所做的设计。单项统计设计是指对收集、整理、分析数据的某一个环节所做的进一步的设计。对统计活动各个方面的设计,主要指的是统计研究对象的各个组成部分,它们是统计工作横的方面。例如,工业企业统计,包括:工业企业经营的内部条件和外部条件;人力、物资、资金等生产要素;生产、供应、销售等生产经营环节。再如,整个社会经济统计,包括:人口、环境、资源等社会发展的环境条件;物质资料的生产、分配、流通、消费的扩大再生产过程;政治、文化、教育、科学、卫生、体育等社会活动;人民的物质和文化生活状况;国际以及其他各国的经济和社会状况。对统计各个环节的设计,主要是指统计工作实际进行时的各个阶段,它们是统计工作

纵的方面。这些阶段包括：统计资料的收集；统计资料的汇总整理；统计分析；统计资料的提供、保存、公布；等等。

2. 统计调查

统计调查也称数据收集，它是统计工作过程的第二个阶段。统计调查的方式方法主要有统计报表制度、普查、抽样调查、重点调查、典型调查等。这一阶段是统计实践活动的开始，属于表层和感性认识阶段。但因为统计是要用数字说话的，而统计数字来源于统计调查，所以，“没有调查就没有发言权”。统计调查属于定量认识阶段，它的工作质量如何，直接关系到影响到以后各阶段的工作质量。

3. 统计整理

统计整理，是按照一定的目的和要求，对统计调查收集到的大量的零乱的资料进行科学的加工和分类，使之系统化、条理化，成为能够说明总体特征的综合资料。统计调查阶段收集的资料既丰富也零乱，既大量也粗糙。因此，需要统计整理去粗取精、去伪存真。这一阶段是对事物由表层认识到深层认识的连接点，对统计分析的质量有举足轻重的作用，是一个承上启下的中间环节。

4. 统计分析

统计分析，是在统计整理的基础上，对统计资料进行多种多样的定量分析和定性分析或评价、论证，由表及里、由浅入深、由此及彼，得出科学的结论，达到对事物本质和规律的认识。这一阶段是认识活动上升为深层次和理性认识的研究阶段。

通过统计整理和统计分析，可以得到许多有用的统计资料。统计资料的提供并不意味着统计研究的终结。统计的目的在于认识客观世界的规律。对于已经公布的统计资料需要加以积累，同时还可以进行加工，结合相关的实质性学科的理论知识去分析和利用，从而更好地将统计数据 and 统计方法应用于相关领域中，使统计更好地发挥信息、咨询、监督的职能。

总之，统计工作的四个阶段是一个统一体，无论哪个环节出了偏差，都会背离统计认识活动的规律，从而歪曲反映事物。统计调查出现偏差，会直接影响统计整理的质量和统计分析结果的正确性；统计分析出现偏差，会造成统计调查和统计整理两个阶段前功尽弃。统计工作各个阶段的工作质量和效果是密切相关的，因此，要注意它们之间的衔接和协调。

需要指出，也有学者认为统计工作过程除了以上四个阶段以外，还包括统计预测和决策阶段、统计资料提供和保管阶段。应该说这两个方面工作确实都是统计工作的重要内容，因为整个统计工作决不能仅仅满足对统计数字的简单加工和初步分析，统计工作的一个重要内容是利用过去和现在的资料对现象的长期趋势做判断和预测，并以此作为国家或企业决策的依据。因此，统计信息的披露与保管都是统计工作不可或缺的组成部分。

1.3.3 统计数据研究方法

统计学的基本研究方法有大量观察法、统计分组法、综合指标法、归纳推断法，如图 1-6 所示。



图 1-6 统计学的基本研究方法

1. 大量观察法

大量观察法,是指为了对现象整体的数量规律有所了解,必须对所研究现象的全部或足够多的个体进行调查的方法。只有通过大量观察法才能消除偶然的、次要的因素影响,以反映主要的、共同起作用的因素所呈现的规律,形成对现象总体数量规律的认识。例如,一个袋子里有编号为 1、2、3 三个球,如果我们只拿出来少数几次,可能拿到哪一个号码会多一些,但只要我们拿出来的次数足够多,那么拿到 1 号、2 号、3 号的概率会大体相同,即基本上为 $1/3$ 。又如,我们掷硬币和掷骰子,每掷一次出现哪一面或哪个点是不确定的,但当我们掷 N 次硬币或骰子时,就会发现掷硬币出现正面和反面的可能性几乎各为 50% ,掷骰子出现 1、2、3、4、5、6 点的可能性几乎各为 $1/6$ 。

大量观察法实际上不是一种具体方法,而是一种统计思想,它强调观察的个体要充分多,只有这样才能将现象的个别偶然差异充分抵消,从而准确地揭示出所研究现象的数量特征和规律。否则,就可能以偏概全,得出片面的或错误的结论。

2. 统计分组法

统计分组法,是根据统计研究的目的和任务,按照一定的标志将总体划分成不同的类型或组的一种统计研究方法。通过统计分组,突出组与组之间的差异、抽象组内各单位的差异,以便划分现象的类型,反映总体的内部结构和现象之间的相互关系,从而达到正确运用指标来表明事物本质与规律性的目的。统计分组法主要适用于统计整理工作阶段,但在统计调查、统计分析等阶段,也都有独特意义。因此,统计分组法既是统计调查研究中的一种基本方法,也是统计整理和统计分析阶段的重要方法,它贯穿于统计研究和统计工作的全过程。

3. 综合指标法

综合指标法,是指运用各种经过科学分类汇总的综合指标和分析方法,研究和说明现象与本质的综合数量特征的方法。常用的综合指标主要有总量指标、相对指标、平均指标等。常用的统计分析方法有时间序列分析、统计指数分析、相关与回归分析等,通过各种分析方法,得出反映现象总体在一定时间、地点、条件下的规模、水平、对比关系、集中趋势、差异程度、依存关系、发展趋势和变化规律。综合指标法主要应用于统计分析工作阶段。

4. 归纳推断法

归纳推断法,也称统计推断法,它是以一定的置信标准,根据样本数据来推断总体数量特征的一种方法。在统计研究中,我们所观察的单位常常是部分单位或少数单位,而要判断总体对象却是大量的。这就需要根据样本资料,对全部总体数量的特征作出具有一定置信度的判断。例如,我们调查万分之一的城市居民户的收入水平,推断出城市全部居民户的收入水平;调查 1% 农田的收获量,推断出上万亩农田的收获量等。随着市场调查在市场经济

济中发挥越来越重要的作用,归纳推断法也越来越被广泛重视,并成为统计研究的基本方法。

【同步思考 1-3】

1. 统计数据有哪几种分类方法?按数据计量尺度和数据时空不同,其各是怎样分类的?

2. 统计工作分为哪几个阶段?

3. 大量观察法是统计分析的基本方法。统计研究的过程中为什么要对研究事物的全部或足够多单位进行观察?



拓展阅读 1 统计的作用

1.4 统计学中的基本概念

统计是关于数据的科学,统计要研究社会经济现象总体的数量特征,那么,统计总体就是统计学的基本范畴,而统计总体又是由总体单位构成的,总体单位又通过许多标志来表现它的特征,其中数量标志可汇总成统计指标,各种相关联的指标构成指标体系(当总体很大时,有时又用少数单位的数量特征来推断和反映总体的特征)。因此,统计总体、总体单位、样本总体、标志、总体指标、样本指标、指标体系等就是统计学的基本范畴。弄清这些基本范畴,有助于把握统计的基本内涵。

1.4.1 统计总体、样本总体、总体单位

1. 统计总体

统计总体简称总体,它是根据一定目的确定的所要研究事物的全体。它是由客观存在的、具有某种共同性质的许多个别事物构成的整体。例如,我们要研究全国城镇居民的收支情况,就以全国城镇居民作为一个总体。有了这个总体,我们就可以研究全国城镇居民的各种数量特征,如人均可支配收入、人均消费等指标。

统计总体的范围随着统计研究目的不同可大可小,如上例中统计总体可以是某市所有的城镇居民家庭,也可以是某区或全省乃至全国所有的城镇居民家庭。

统计总体必须同时具有同质性、大量性和变异性三个特点。

(1) 同质性。同质性是形成统计总体的一个必要条件,作为总体的一个重要特征。同质性是指构成统计总体的每一个单位在某一方面必须具有共同的属性。它是根据统计研究目的而确定的。研究目的不同,所确定的总体也不同,其同质性的意义也随之变化。例如,研究城镇居民户的生活状况,所有城镇居民户构成了统计总体,农民户就不在统计范围之列,所有的城镇居民户在这一点上都是同质的,无论是收入低还是收入高。如果研究的是城镇居民贫困户的生活状况,那么,贫困线下的城镇居民户就构成了统计总体,贫困线下的城镇居民户在这一点上都是同质的,而贫困线上的城镇居民户就是非同质的了。

(2) 大量性。大量性是指总体中应包括足够多的个别事物,每一个个别事物常常可能受偶然因素的影响,表现出各种各样的差异。因此,少数个别事物的特征往往不能说明总体的特征,只有研究多数单位形成的总体,才能使偶然因素的作用相互抵消,从而显示出事物

的本质特征。

(3) 变异性。变异性是指构成总体的每一个个别单位,在某一方面性质是相同的,但在其他方面必定有差异。例如,同是某股份制企业的职工,也有工种、工资、工龄、文化程度等方面的差异。统计研究实质上就是研究总体各单位某种品质或数量变异的程度、趋势等,从而寻找出规律。

2. 样本总体

样本总体简称样本,它是根据一定的研究目的和研究任务,在全部研究总体中随机抽取少数单位组成的小总体,并通过研究小总体的数量特征,来推断大总体的可能数量变化规律。例如,我们要研究全国城镇居民的收支情况,可以在全国城镇居民的总体中,随机抽取1万个或10万个居民,组成小总体,然后通过研究这个小总体居民的收支情况,来推断和了解全部总体的数量特征;再如,要了解某批5000台电视机质量情况,可以随机抽取500台组成一个样本总体,通过对500台电视机质量检查,如画面清晰度、分辨率、外观、发热、使用时长等指标,来推断和了解全部5000台电视机的质量情况。

3. 总体单位

总体单位,简称单位,也称个体,它是构成统计总体的每一个个别事物,是各项统计特征的承担者。比如,上例中的每一台电视机。根据统计研究目的不同,总体单位可以是人或物,可以是企业单位或地区、部门,也可以是时间、质量或长度等。有时总体单位以自然计量单位表示,如设备以“台”表示,产品以“件”表示等。有时总体单位以物理计量单位表示,如时间用“小时”“分”“秒”表示,质量用“吨”“千克”“克”表示,长度用“米”“厘米”“毫米”表示,面积用“平方千米”“平方米”表示,等等。例如,某市每一个股份制企业就是全市股份企业总数这个统计总体的总体单位;每一个工业企业的每一台设备则是该市工业企业设备总量这个总体的一个基本单位,即总体单位。

总体和总体单位的概念是相对而言的,随着研究目的不同,总体范围可相互变化。同一个研究对象,在一种情况下是总体,在另一种情况下可能就变成总体单位了。例如,要研究全国人口情况,全国为总体,每一个人就是总体单位;而要研究全国各省份的人口情况,全国为总体,各省份为总体单位;同样,要研究某省各县人口状况,则该省变成总体,而各县又成了总体单位。

统计总体、样本总体、总体单位之间的关系如图1-7所示。

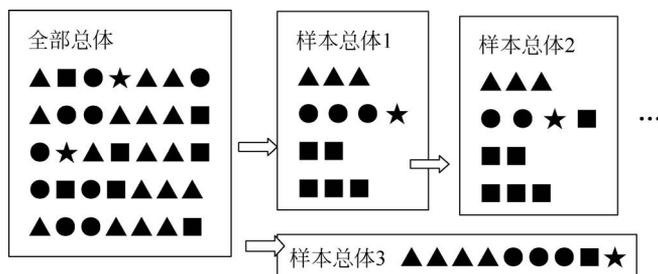


图 1-7 统计总体、样本总体、总体单位之间的关系

需要指出的是,当一个总体所包含的单位数是无限的,无法一一计数,称为无限总体。例如,连续大量生产的某种小件产品中,总产量是无限的;当总体单位数是有限的,是可以一一计数的,称为有限总体。例如,我国人口总数约14亿,但总归是可数的,是有限的,可以计数。区分无限总体和有限总体是确定科学的调查研究方法的前提条件。通常,对于无限总体,无法进行全面调查,只能进行非全面调查,也即对样本总体进行调查;对于有限总体,既可进行全面调查,也可进行非全面调查。

1.4.2 标志和变量

1. 标志

标志也称标识,是说明总体单位属性或特征的名称。每个总体单位从不同的角度和要求观察,可以有多个属性特征。标志依附于总体单位,总体单位是标志的直接承担者。例如,上例中的每一台电视机的质量情况。每个工人都具有性别、工种、文化程度、技术等级、年龄、工龄、工资等属性和特征,这些就是工人作为总体单位的标志。标志的具体表现是在标志名称之后所表明的属性或数值。统计研究是从登记标志状况开始的,并通过对标志的特征进行综合反映出总体的数量特征。

标志有不同的种类。按说明现象的性质不同,标志可分为品质标志和数量标志。品质标志表明单位属性方面的特征,品质标志的表现只能用文字来描述。例如,某工人的性别是“男”,民族是“汉族”,则“性别”和“民族”是品质标志的名称,而“男”和“汉族”是品质标志属性的具体表现。数量标志是表明总体单位数量方面特征的,用各种不同的数值表示。例如,职工的“年龄”是数量标志名称,其标志具体表现,如“38岁”“42岁”为标志值;职工的“薪资”是数量标志名称,其标志具体表现,如“6 100元”“8 160元”为标志值。

标志按变异情况可分为不变标志和可变标志。当一个标志在各个单位中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中的具体表现不完全相同时,如“性别”,有的人表现为“男”,有的人表现为“女”,再如“薪资”,有的表现为“6 100元”“8 160元”,也有的表现为“11 200元”或“4 150元”等,那么,这个标志称为可变标志。在划分统计总体时,可以选定某一标志的具体表现,把它固定下来,把所有具备这种标志的表现结合在一起,形成一个统计总体。所谓统计总体的同质性,实际上就是统计总体中的各个单位都具有某一共同的标志表现。

无论是品质标志还是数量标志,当它们在总体各单位中表现不完全相同时,就称为变异。变异性是统计总体必须同时具备的三个特点之一,也是统计研究的前提。

2. 变量和变量值

变量是统计中常用的重要概念。变量是对可变化的量而言的。统计中的变量是指可变的数量标志。变量的数值表现就是变量值或标志值。例如,工人的身高和年龄不会每个人都相同,那么,“年龄”和“身高”是数量标志的名称,也即变量,而“年龄”和“身高”数量标志的具体表现数值,如38岁、42岁以及6 100元或8 160元等是数量标志值,也称变量值。

按变量值是否具有连续性,变量又分为连续变量和离散变量两种。连续变量的数值是连续的,相邻两值之间可做无限分割、无限取值,既可用整数表示,也可用小数表示。如以质量、长度、面积等物理单位计量的都是连续变量。连续变量的数值要用测量或计算的方法取

得。离散变量是指相邻两值只能以整数位断开、只能以整数表示的变量。如人数、企业数、设备台数等都只能按整数计算,有了小数则令人不可思议。离散变量的数值只能用计数的方法取得。

1.4.3 统计指标和指标体系

1. 统计指标

统计指标简称指标,是反映社会经济现象总体某一综合数量特征的范畴,由指标名称和指标数值两部分构成。指标名称反映现象所属的范畴,也即指标概念;指标数值反映现象在具体时间条件下达到的规模、水平及比例关系。例如,2023年,初步核算,全年国内生产总值1260582亿元,按不变价格计算,比上年增长5.2%。全年全国居民人均可支配收入39218元,比上年名义增长6.3%,扣除价格因素,实际增长6.1%。^①

统计指标一般具有三个特点:①数量性,即统计指标都是用数字表示的。②综合性,即统计指标是总体单位同质数量综合的结果。③具体性,即统计指标是现象在不同时间、地点、条件下的具体反映。如上例中的全国居民人均可支配收入39218元,“39218元”体现了数量性,也体现了综合性,即2023年全国居民人均可支配收入不断增长,经济发展水平和居民生活水平在提高。统计指标可以按其不同的研究目的进行分类。

统计指标按其反映时间特点不同,可分为时点指标和时期指标。时点指标是反映现象在某一时点上的数值表现,如人口数、黄金储备量等;时期指标是反映现象在某一时期内的数量表现,如产值、产量等。有关这两类指标的特点和计算,将在总量指标、时间序列分析中做详细论述。

统计指标按其反映总体特征不同,可分为数量指标和质量指标。数量指标是指反映总体规模大小、水平高低的指标,一般以绝对数表示,如人口总数、企业总数、职工总数、工资总额、国内生产总值、商品流转额、商品进出口总额等;质量指标是指反映总体的强度、密度、效果等的指标,一般用相对数和平均数表示,如劳动生产率、职工平均工资、人口密度、工人出勤率、流动比率、速动比率等。有关数量指标和质量指标的特点及计算,将在总量指标与相对指标一章中详细论述。

统计指标按其作用和表现形式不同,可分为总量指标、相对指标和平均指标。总量指标是反映现象总体规模、水平的指标,如工资总额、产品产量、销售额、利税总额等。相对指标是两个有联系的总量指标进行对比的结果,用来说明现象总体的结构、发展程度、比例、强度、密度等,如产品的优质品率、平均增长率、人口密度等。平均指标是按总体内各单位某一数量标志值计算的反映总体一般水平的指标,如平均工资、工人劳动生产率等。

2. 标志与统计指标的关系

标志与统计指标之间既有联系又有区别(图1-8)。

二者的区别主要表现为:①指标是说明总体特征的,而标志是说明总体单位特征的。②指标都是用数字表示的,而标志既可用数字表示,也可用文字表示。

^① 中华人民共和国2023年国民经济和社会发展统计公报[EB/OL]. (2024-02-29). https://www.stats.gov.cn/sj/zxfb/202402/t20240228_1947915.html.

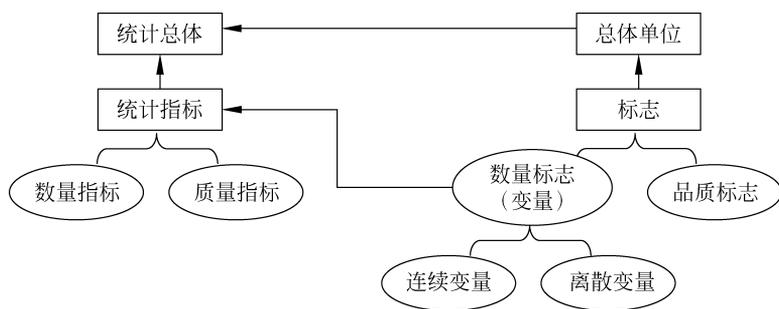


图 1-8 标志与统计指标之间的关系

二者的联系主要表现为：①指标数值都是由总体单位的数量标志值汇总而来的。②指标与数量标志存在一定的变换关系，随着研究目的的变化，总体和总体单位发生了变化，指标和标志也会发生相应的变化。例如，如果所要研究的是全国工业企业情况，则各企业的职工人数、固定资产、工业增加值等都是总体单位（即各个企业）的标志，而如果研究目的变成研究某一企业的职工状况，则该企业变成一个总体，企业职工人数变成了统计指标，每个职工的文化程度、技术等级、性别、年龄等就成为标志。

3. 样本指标与总体指标

(1) 样本指标。样本指标，也称统计量，是根据样本总体计算出来的数据，如样本均值、方差或标准差、是非比率等。样本是从总体中抽取的部分单位构成的整体，是实际进行调查登记的对象，因此，样本单位的数据是可计量、可测算、可知的，统计量也是可知的。

由于样本来源于总体，因此样本统计量必然在一定程度上可以用来推断有关的总体参数。比如，可用样本均值去估计总体均值、用样本方差或标准差估计总体方差或标准差、用样本是非比率去估计总体是非比率等。

(2) 总体指标。总体指标，也称参数，是反映总体某种特征的量。参数是根据总体所有数据计算出来的，如总体均值、方差或标准差、是非比率等。当总体是有限总体、总体规模不是特别大时可以直接计算。但由于实际工作中，常常不可能收集齐总体所有数据，所以总体指标的真实数值往往是未知的，就需要利用样本数据去估计或推断总体数据。

在用样本统计量去估计或推断总体参数时，往往需要通过一定的函数关系式对样本统计量进行再加工，这就形成了构造统计量。常用的构造统计量有 Z 统计量、 t 统计量、 χ^2 统计量、 F 统计量等。我们将在第 6 章“抽样推断”中再详细阐述和计算。

4. 统计指标体系

一个统计指标只能反映特定现象的一个侧面或一个侧面的某一特征，但任何客观现象都具有多个相互联系、彼此制约的方面和特征。比如，对一个企业的综合评价，包括财务指标、市场指标、创新指标、风险指标等。我们如果要全面、深入地反映客观事物，就必须将各种相互联系的指标构成一个整体，用以反映所研究对象各方面的相互依存和制约关系，反映总体的全貌。因此，统计指标体系就是指由一系列相互联系、相互制约的统计指标构成的整体。在社会经济生活中，统计指标体系中各个指标之间的联系表现为两种形式。

(1) 统计指标之间存在的客观联系是通过严密的数学公式表现的，比如：

$$\text{农作物收获量} = \text{播种面积} \times \text{单位面积产量}$$

国民总收入 = 国内生产总值 + 来自国外的净要素收入

期末库存量 = 期初库存量 + 本期购进量 - 本期销售量

资产 = 负债 + 所有者权益

(2) 各统计指标之间存在着相互补充的关系。例如,考核工业企业的九项指标:产量、品种、质量、原材料、燃料、动力消耗、成本、利润、流动资金占用。再如,为了反映工业企业生产经营的全貌,需要设立产量、产值、品种、质量、职工人数、劳动生产率、工资总额、原材料、设备、财务成本等多项指标组成工业企业统计指标体系。以上两例中,多个指标所构成的指标体系就属于这种情况。

在设立统计指标体系用以全面、综合反映现象的状态时,应该遵循客观性、科学性、可行性、预见性原则。指标体系的建立不但要遵循指标之间内在的客观联系,还要考虑获取资料是否可能以及指标体系的设置是否可行;不但要考虑指标体系是否能反映实际问题,还要使新设立的指标体系具有一定的超前意识,从而更好地适应不断变化的需要。

【同步思考 1-4】

1. 标志是说明总体单位特征的名称。其中,品质标志不能用数值表示,数量标志只能用数值表示。那么,文化程度是品质标志还是数量标志? 试以一个班同学为例,分别列三个不变标志和可变标志的例子。

2. 统计指标是反映现象总体数量特征的概念及数值。统计指标由指标名称、指标数值、空间范围、时间范围、计量单位、计算方法六要素构成。那么,人口出生率 6.39‰,这个指标完整吗? 还缺少哪几个要素? 应该怎样明确?

3. 统计总体、样本总体、总体单位之间有什么关系? 标志与统计指标之间又有什么联系与区别? 试各举一例说明。

1.5 Excel 在统计数据处理中的功能概述

作为美国微软公司推出的 Office 套装办公软件的主要成员,Excel 是运行在 Windows 环境下的电子表格系统。Excel 因具有电子表格管理、数据清单管理、统计图表处理以及数据分析与决策功能,在统计应用中有着举足轻重的地位,得到了广泛的应用。本书以 Excel 2019 为工具进行统计分析。

1.5.1 Excel 数据分析工具的加载及其使用

作为 Office 电子表格文件处理工具的 Excel,不仅具有进行相关电子表格处理的功能,而且带有一个可以用来进行统计数据处理分析的宏程序库——“分析工具库”。通常计算机安装了 Office 后,其 Excel 电子表格系统并不能直接使用“分析工具库”来进行统计数据的处理分析,需要加载宏,启动“数据分析”宏“分析工具库”系统后,才能运行统计数据的数理分析工具。

打开 Excel 2019 电子表格系统后,如果在“数据”选项卡中没有“分析”组,则说明 Excel

系统尚未加载分析工具宏程序,必须在 Excel 中加载并启动“分析工具库”宏程序。

打开 Excel 2019 后,依次单击“文件”选项卡、“选项”和“加载项”类别。在“管理”框中,单击“Excel 加载项”,然后单击“转到”。此时将显示“加载项”对话框。在“加载项”框中,勾选“分析工具库”复选框,然后单击“确定”(图 1-9)。

完成了 Excel 数据分析程序宏的加载后,单击“数据”选项卡“分析”组中的“数据分析”命令,即会弹出 Excel 的统计分析工具对话框(图 1-10),如果选中其中的某一个统计分析工具,并单击“确定”按钮,就会弹出该分析工具的运行对话框,然后通过运行对话框的对话,可以进入该统计分析工具宏程序的运行过程。在整个分析工具宏程序库中设有各种数据处理分析的工具宏程序,包括用于描述统计分析的描述统计和直方图分析工具宏等,也包括可以进行推断统计分析的方差分析、相关和回归分析、统计推断和检验以及时间序列指数平滑法等分析工具宏,具体的统计分析工具所包含的内容见表 1-1。

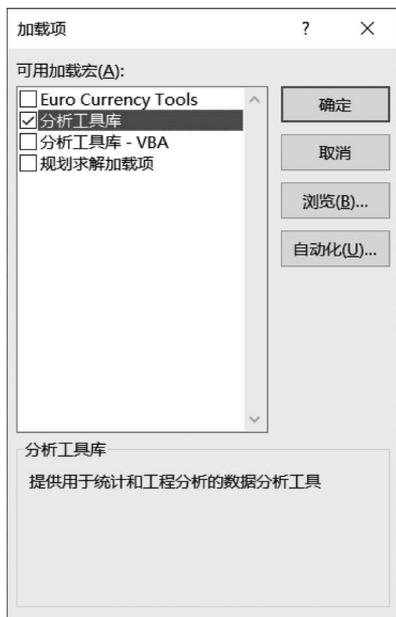


图 1-9 加载分析工具库

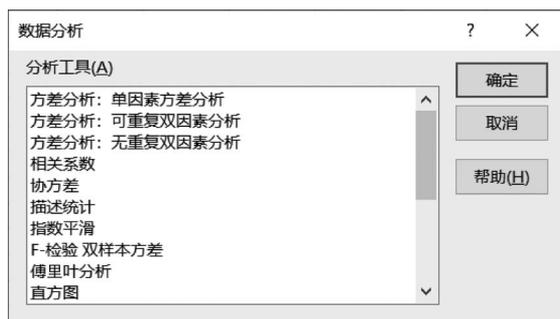


图 1-10 数据分析工具

表 1-1 Excel 统计分析工具

序号	分析工具	类别
1	方差分析: 单因素方差分析	方差分析
2	方差分析: 可重复双因素方差分析	
3	方差分析: 无重复双因素方差分析	
4	相关系数	相关与回归
5	协方差	
6	回归	
7	直方图	数据整理
8	移动平均	时间序列预测
9	傅里叶分析	
10	指数平滑	

续表

序 号	分 析 工 具	类 别
11	抽样	抽样设计
12	随机数发生器	
13	排位与百分比排位	参数估计
14	描述统计	
15	F -检验：双样本方差	假设检验
16	t -检验：平均值的成对二样本分析	
17	t -检验：双样本等方差假设	
18	t -检验：双样本异方差假设	
19	Z -检验：双样本平均差检验	

1.5.2 Excel 统计函数及其使用

Excel 具有大量的内置函数,如财务函数、日期和时间函数、数学和三角函数以及统计函数。其中统计函数的功能简介,请参见附录 C。通常在应用 Excel 进行数据处理分析时,应尽量使用这些内置函数,其一方面可以减少因计算公式的输入带来的麻烦;另一方面,还可以根据处理分析的需要,在电子表格上编辑出由多种内置函数组合而成的复杂运算公式,以适应某些多步骤运算过程的特殊处理分析需要。

在 Excel 运行过程中调用统计函数主要采用两种方法:其一,在工作表的单元格中直接输入统计函数的名称(必须在统计函数名称前加“=”),就会立即弹出该函数的初始输入对话框,只要在有关的参数选项内填入确定的参数就能得到函数的计算结果值。其二,在工作表的单元格内输入“=”后,查找工作表左上方的“名称”显示格内出现的函数选择表,选择某个函数名称,同样会得到该函数的初始输入对话框(图 1-11)。

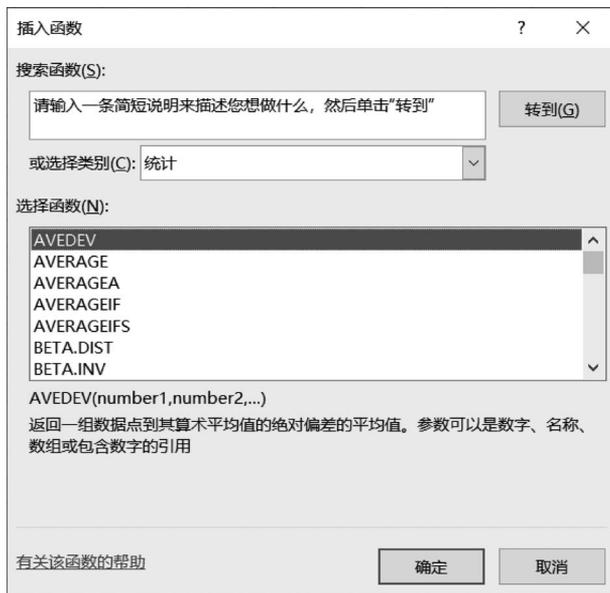


图 1-11 插入函数对话框

本章小结

统计学是一门关于如何收集数据、整理数据、分析数据,并进行描述、推断和预测,从而正确认识现象总体数量方面的科学。

(1) 统计包括统计工作、统计资料、统计学三层含义,三层含义之间既有联系也有区别。统计资料是统计工作的成果;统计理论来源于统计实践,反过来又为统计实践服务,统计理论与统计实践是辩证统一的关系。

(2) 统计学发展史上的主要学派有政治算术学派、国势学派、社会统计学派和数理统计学派。

(3) 统计研究具有总体性、数量性两大突出特点,此外,也有具体性、社会性和广泛性的特点。统计总体具有同质性、大量性和变异性三大特点。统计指标具有数量性、综合性、具体性三个特点。

(4) 统计研究的基本方法有大量观察法、统计分组法、综合指标法、归纳推断法。其中大量观察法是普遍遵循的方法,但不是一种具体的方法,而是一种统计思想。

(5) 统计的基本概念有:统计总体、样本总体、总体单位、标志、变异、统计指标、样本指标、总体指标、统计指标体系等。标志按说明现象性质不同,分为品质标志和数量标志;按变异情况分为不变标志和可变标志。按变量值是否具有连续性,变量可分为连续变量和离散变量。

(6) 标志与统计指标之间既有联系又有区别;统计总体、样本总体、总体单位既有联系又有区别;样本指标与总体指标之间既有联系又有区别。

(7) 利用 Excel 的数据分析工具和函数,可以提高处理统计数据的效率。

思考与练习

● 知识题

- (1) 什么是统计?简述统计工作、统计资料、统计学三者之间的关系。
- (2) 什么是总体?什么是总体单位?总体与个体有什么关系?
- (3) 什么是标志?什么是统计指标?二者有什么联系与区别?
- (4) 统计研究常用的方法有哪几种?为什么要用大量观察法?
- (5) 统计的基本职能是什么?简述统计信息职能、咨询职能和监督职能三者之间的关系。

● 实务题

为了研究某学校教职员工的工资情况,根据此目的,明确下列概念。

- (1) 总体是()。
 - A. 该学校
 - B. 该学校全部教职工
 - C. 该学校教职员工的工资
 - D. 该学校教职工人数
- (2) 总体单位是()。
 - A. 该学校
 - B. 该学校全部教职工

C. 每一个教职工

D. 该学校教职工工资的总额

(3) 标志是()。

A. 该学校的教职工人数

B. 学校的工资金额

C. 每个教职工的工资

D. 每个教职工的平均工资

(4) 指标是()。

A. 该学校的教职工人数

B. 该学校教职工的平均工资

C. 该学校的教职工的工资金额

D. 每个教职工的工资额

● 实训题

实训一

(1) 实训目的：通过本题练习，熟悉各种统计基本概念。

(2) 实训资料：某企业为了解组装车间职工的人数、日产量和各组总日产量的关系，调查了2024年3月该企业职工人数和产量情况，经过分组、汇总和计算，得到表1-2。

表 1-2 某企业组装车间职工人数及产量资料

按日产量分组/件	工人数/人	每组日产量/件
15~25	4	80
25~35	25	750
35~45	15	600
45 以上	6	300
合计	50	1 730

(3) 实训要求：在上述资料中，总体是什么？总体单位是什么？什么是总体单位的数量标志？什么是总体单位的品质标志？本例中反映总体数量特征的指标有哪几个？

实训二

(1) 实训目的：通过本题练习，加深对总体、总体单位、品质标志、数量标志、数量指标、质量指标的认识。

(2) 实训资料和实训要求：调查某数码商场销售的全部手机情况，试指出总体、总体单位是什么，试列举出若干品质标志、数量标志、数量指标、质量指标。

即测即练

