

第 1 章 数据统计凭什么成为最强的商业武器



引导案例

威尔斯的预言¹

威尔斯于 1903 年预言，统计思维和读写能力一样，终将成为良好公民所需具备的必要条件。

这段文字就写在哈佛大学医学院使用的统计学教材的开头。

作者 H.G.威尔斯人称“科幻小说之父”，也是一位思想家。时光机器及透明人等概念都是因其著作而闻名的。据说，他的科学知识之广博、眼光之高远，早就预言了核子武器和国际联盟的出现，甚至还预言了类似现在流行的维基百科全书的问世。

在现代统计学仍处于黎明期的 1903 年，我们无法确认威尔斯为何会如此预言。但经过了 100 年之后，正如读写能力的重要一样，我们也确实需要统计学的思维。读写能力视为教养或素养，因此，缺乏统计学方面的识读能力，即缺乏“统计学素养”，尤其对于现今社会发展与时代潮流所趋，是极为不智的。

不具备读写能力就无法了解合约及法律等内容，不具备统计学素养则无法了解概率及数据资料的运用。无论是以上何者，都可能陷入弱势族群的处境。



1.1 统计是什么

统计是对某一现象有关数据的收集、整理和分析。

统计是描述、总结某种现象的数据或其他信息。

统计学的目的：统计学有很多目的，而它最重要的目的是我们在面对不确定问题时做出决策的依据。

统计是这样一些数字：它们通过某种有意义的方式对原始事件和数字进行提炼，使得仅仅通过观察原始数据无法立即水落石出的一些理念得以昭示。这里的数据指的是我们能够据其做出结论的事实或数字。

“统计”一词从字面上理解，是指对大量事物的数量进行汇总和计数，即“统而计之”的意思。后来，从反映社会活动方面理解，“统计”一词延伸为用数字反映现象，表明统计还有用数字说明问题的作用。现

拓展 1



¹ 摘自日本西内启《统计学，最强的商业武器》2016 年 3 月。

代的“统计”一词,可以从三个不同方面理解,即统计工作、统计资料、统计学。

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对统计工作起着指导作用。统计工作一方面受统计理论指导,另一方面也检验着统计理论的正确与否,促进统计理论不断发展。统计学与统计工作、统计资料之间的关系表明,理论来源于实践,又反过来为实践服务,被实践检验,体现着理论与实践辩证统一的关系。

几乎所有学科都要依赖统计学的方法和数据分析。

拓展2



1.2 统计的研究对象、特点、方法和过程

1.2.1 统计的研究对象

统计、统计学是从研究社会经济数量方面开始的,随着统计实践的发展,统计方法的不断完善,统计学得以不断前进。这种理论与实践的关系,决定了统计学与统计工作的研究对象最终是一致的,都是社会经济现象总体的数量方面,即在质与量的辩证统一中研究大量社会经济现象总体的数量方面,反映社会经济现象发展变化规律性在具体时间、地点和条件下的数量表现,揭示事物的本质、相互联系、变动规律性和发展趋势。它们又有区别:统计工作是从实践上研究具体的社会经济现象发展规律;统计学则是从理论角度为统计工作提供数量研究和认识规律的科学方法,包括指导统计工作的原理原则,统计过程所应用的核算和分析方法,以及组织方法。其核心内容是数据的收集、整理、归纳、分析的原理和方法,这些方法论构成了统计学的科学体系。所以统计学就其性质来说,是研究如何收集数据、分析数据,以便从中做出正确推断的方法论学科。

1.2.2 统计的特点

1. 总体性

统计的研究对象不是个体现象的数量方面,而是许多个体现象构成的总体的数量方面。当然统计研究又必须是从个体入手然后经过分组、汇总,过渡到对总体特征的了解与认识。在认识和研究方法上遵循个体到总体的路径。

2. 数量性

统计的研究对象是社会经济现象的数量方面,即用大量数字资料说明现象的规模、水平、现象间的数量关系,以及决定现象质量的数量界限等。统计研究对象的数量性,是区别于其他社会经济调查研究活动的根本特点。必须指出,统计对现象数量方面的认识,是与现象的本质紧密结合在一起的,是以定性认识为基础,遵循定性一定量一定性的科学的认识规律。

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、具体条件限定的数值。具体性就是指统计指标在时间、地点、条件三方面有着明确的规定性。

统计固然是研究具体的数量，但为了进行复杂的定量分析，还需要借助抽象的数学模型和数理统计方法，遵循一定的数学规则。因此，在统计中对具体数量的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，体现了统计中具体和抽象的辩证关系。

►► 1.2.3 统计的研究方法

统计是一门方法论科学，它提供对社会经济现象调查研究的理论、原则和方法。统计研究具有多阶段性，每个阶段都有许多特定的统计方法，归纳起来主要有以下几种基本方法。

1. 大量观察法

大量观察法是指对所研究现象总体中的足够多数的个体单位进行观察，以期认识具有规律性的总体数量特征的方法。

大量观察法常见的有统计报表、普查等，主要用于统计调查阶段。

2. 统计分组法

统计分组法是指根据统计研究的任务，将所研究的社会经济现象总体按照一定标志划分为若干组的方法。

由于所研究现象本身的复杂性、差异性及多层次性，需要我们对所研究现象进行分组或分类研究，以期在同质的基础上探求不同组或类之间的差异性。在统计分析阶段，可以通过统计分组法来划分现象类型、研究总体内在结构、比较不同类或组之间的差异（显著性检验）和分析不同变量之间的相关关系。

3. 综合指标法

综合指标法是指从总体范围角度研究现象数量特征和数量关系的范畴及其数值的方法。

统计研究现象的数量方面的特征是通过统计综合指标来反映的。常见的有总量指标、相对指标、平均指标和标志变异指标等。综合指标法用来描述统计学的核心内容。

4. 统计模型法

在以统计指标来反映所研究现象的数量特征的同时，我们还需要对相关现象之间的数量变动关系进行定量研究。在研究这种数量变动关系时，需要根据具体的研究对象和一定的假定条件，用合适的数学方程来进行模拟，这种方法就叫作统计模型法。

5. 统计推断法

统计推断法是指研究如何利用样本数据来推断总体特征的统计方法，包含参数估计和假设检验。

它可以用于总体数量特征的估计,也可以用于对总体某些假设的检验,所以在统计研究中应用很广泛,已成为现代统计学的基本方法。

以上五种方法相互联系,相辅相成,构成了一个完整的统计研究方法体系。

1.2.4 统计的研究过程

一次完整的统计活动过程,要经过统计设计、统计调查、统计整理和统计分析四个阶段。

1. 统计设计阶段

统计设计是对统计活动各个方面和各个环节所做的通盘考虑和合理安排,如确定调查对象,设计指标体系,编制分类目录,制订调查、整理和分析方案等。优良的统计设计是科学、有效地组织统计研究活动的前提。

2. 统计调查阶段

统计调查是根据一定的目的,通过科学的调查方法,收集社会经济现象的实际资料的活动,是认识客观经济现象的起点,也是统计整理和统计分析的基础。

3. 统计整理阶段

统计整理是对调查来的大量统计资料加工整理、汇总,使之系统化、条理化,成为能够说明总体特征的综合资料的过程。它既是统计调查的延续,又是统计分析的前提,起着承前启后的作用。

4. 统计分析阶段

统计分析是将加工整理好的统计资料加以分析研究,采用各种统计方法,计算分析指标,来揭示现象的本质及其发展变化的规律性,是统计工作的最后阶段。

统计研究过程的四个阶段并不是孤立、截然分开的。它们是紧密联系的一个整体,其中各个环节常常是交叉进行的。例如,小规模的投资,常把调查和整理结合起来,在统计调查过程中就有对事物的初步分析,在整理和分析过程中仍须进一步调查(见图1-1)。

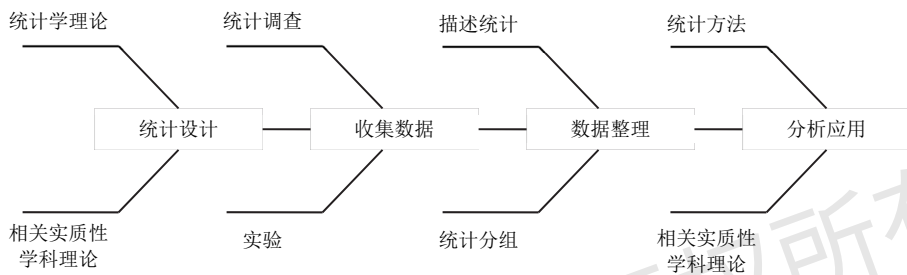


图 1-1 统计研究的过程

1.3 统计学中的一些基本概念

1.3.1 统计总体、总体单位和样本

1. 统计总体

统计总体是指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（集合体），简称总体。

在统计研究中，总体是所研究的人或事物的完整集合。

统计要研究客观现象总体的数量特征和数量关系，首先要对统计总体有明确的认识。

例如，我们要研究全国乡镇工业企业发展情况，全国的乡镇工业企业就组成一个总体。这些乡镇工业企业尽管资产规模、产品品种、技术力量、设备状况、经济效益等各不相同，但它们都是乡镇企业，都是工业生产单位，向社会提供工业产品或劳务服务，在这方面具有共同性，或称同质性，是统计总体赖以形成的客观基础。

总体具有三个特征：①大量性；②同质性；③变异性。

2. 总体单位

总体单位是构成总体的每个事物或基本单位，也称个体（元素）。

原始资料最初就是从各个总体单位取得的，所以总体单位是各项统计数字最原始的承担者。例如，研究某个工业部门的生产情况时，该工业部门的所有工业企业可以作为一个总体，每个工业企业是总体单位，将每个工业企业的某些数量特征加以登记汇总，就取得该工业部门的统计资料。

总体和总体单位是相对而言的，随着研究目的不同，同一事物在不同情况下，可以作为总体，也可以作为总体单位。

3. 样本

样本是从某个总体中抽取部分单位所组成的整体。样本中个体的数目称为样本容量。样本是总体的一个子集，它的数据是进行实际测量而获得的。

一般情况（不同领域标准不一）下，样本中个体的数目大于 30 称为大样本；样本中个体的数目小于或等于 30 称为小样本。在有些情况下，大样本和小样本所用的统计处理方法不一样。

▶▶ 1.3.2 标志和标志表现

1. 标志

标志是反映总体各单位属性和特征的名称。

每个总体单位从不同方面考察都具有许多属性和特征。例如，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，这些就是工人总体单位的标志。

2. 标志表现

标志表现是总体各单位的某一标志之后所表明的属性或数值。标志分为品质标志和数量标志两种。

拓展 3



凡是只能用文字表示的、表明单位属性方面的特征,如工人的性别、设备的种类、企业的经济类型等称为品质标志。性别标志具体表现为男女;设备的种类标志具体表现为车床、铣床、冲床等。凡是可用数值表示的、表明单位数量方面的特征,称为数量标志,如工人的工龄、工资,企业生产设备的能力,职工人数,产品产值等。工龄标志具体表现为不同的年数,某工人工龄是15年,则工龄是数量标志,15年是该标志的数值表现。

不论是品质标志还是数量标志,如果在一个总体的各单位具体表现都相同,就称为不变标志。例如,在工人总体中,职业这一标志各单位表现都是工人,所以职业便是不变标志。在一个总体中,当一个标志在各单位的具体表现有可能不同时,这个标志便称为可变标志。例如,在工人总体中,各单位的工龄可能表现不同,所以工龄便是可变标志。可变标志的属性或特征的具体表现由一种状态变到另一种状态,统计上称为变异,所以可变标志也称为变异标志。在一个总体中不变标志和变异标志各自发挥重要的作用。一个总体至少要有一个不变标志,才能使各单位结合成一个总体。

▶▶ 1.3.3 统计指标和指标体系

1. 统计指标

统计指标是反映统计总体数量特征的概念和数值。例如,2016年江苏省常州市人均国内生产总值122 721元,反映了常州市经济发展的状况和水平。

由此可见,完整的统计指标是由5项基本要素构成的:时间范围、空间、指标名称、指标数值及计量单位,其中指标名称和指标数值是核心要素。

指标概念(名称)是统计所研究的社会经济现象的科学概念,表明社会经济现象的质的规定,反映某一社会现象内容所属的范围。

指标数值是统计所研究现象的具体数量综合的结果,对某一社会经济现象总体特征从数量上加以说明。

统计指标必须包括时间状态、空间范围、计量方法等限定,不能随意变动,同时必须注意由于上述条件的变化而引起数值的可比性问题。总之,统计指标是统计名称及其指标数值的有机结合,是统计研究对象的具体化,也是统计对客观事物认识过程由质到量、质量结合的起点。

2. 统计指标的分类

统计指标按照指标属性差异,分为数量指标和质量指标。

凡是反映现象总规模、总水平和工作总量的统计指标都称为数量指标,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。这些指标不论是总体单位总量还是总体标志值总量,都反映现象或过程的总规模 and 水平,所以数量指标也称为总量指标,用绝对数来表示。

凡是反映现象相对水平和工作质量的统计指标都称为质量指标,如粮食平均亩产量、职工平均工资、人口密度、出生率、死亡率、工人出勤率等。质量指标是总量指标的派生指标,用相对数或平均数来表示,以反映现象之间的内在联系和对比关系。

3. 指标和标志的联系与区别

两者的主要联系：

- 许多统计指标的数值是从总体单位的数量标志值汇总而来的。例如，我国钢产量（统计指标）是由我国每个钢铁企业（总体单位）的钢产量（数量标志）的具体数值汇总而来的。所以，数量标志是统计指标的基础。
- 有些统计指标与数量标志之间在一定条件下存在着变换关系。由于研究目的不同，原来的统计总体可以变为总体单位，相应的统计指标就变成数量标志。例如，当把某企业作为统计总体时，其产量、职工人数、工资总额都是统计指标；当把该企业作为总体单位时，其产量、职工人数就成了数量标志。

两者的主要区别：

- 说明的对象不同。指标是说明总体特征的，标志则是说明总体单位特征的。
- 表示方法不同。标志有不能用数值表示的品质标志与能用数值表示的数量标志两种，而指标都是用数值来表示的，没有不能用数值表示的指标。

4. 指标体系

指标体系是由一系列相互联系的统计指标所组成的有机整体，用以反映所研究现象各方面相互依存、相互制约的关系。

单个统计指标只反映总体某个数量特征，说明现象某一侧面情况。但是客观现象是错综复杂、具有多方面联系的。要反映客观现象的全貌、描述现象发展的全过程，只靠单个统计指标是不够的，应该设立统计指标体系。

例如，工业企业是在一定生产经营主体的组织下，由资本金、劳动力、物资、技术、设备、生产、供应、销售等相互联系的整体活动。为了反映企业生产经营的全貌，就应设立产量、产值、品种、质量、职工人数、劳动生产率、工资总额、原材料、设备、财务成本等指标群，来组成工业企业统计指标体系，其中，产品产量、总产值、增加值、品种、质量指标又构成企业生产统计指标体系，而固定资金、流动资金、生产费用、产品成本、销售利润又构成企业财务指标体系等。

社会经济统计指标体系大体上可以分为两大类，即基本统计指标体系和专题统计指标体系。基本统计指标体系是由社会指标体系、经济指标体系、科技指标体系三个子系统构成的，每个子体系又可以设若干门类指标或指标群，以反映其基本情况和相互联系。专题统计指标体系是指针对社会经济专门问题而制定的专项指标体系，如经济效益指标体系、国际收支指标体系、小康生活水平指标体系等。

指标体系的设置不但是客观现象的反映，而且是人们对客观认识的结果。随着客观形势的发展变化及实践经验和理论研究的积累，指标体系也将不断更新，臻臻完善。

拓展 4





本章小结

1. “统计”一词有三种含义:统计工作、统计资料和统计学。统计工作是对客观事物数量方面的调查研究。在社会经济领域,通过统计调查、资料的整理和分析,以综合指标的形式反映社会经济现象在一定时间、地点、条件下的数量特征和数量关系,揭示现象发展变化的规律,为社会和各级政府管理部门提供信息和咨询,对国民经济和社会发展进行监测。在当今信息时代,统计信息是社会信息的主体。统计不只是认识社会的有力武器,也是科学决策和管理的重要工具。

2. 统计研究对象:在质与量的辩证统一中研究大量社会经济现象总体的数量方面规律。具有总体性、数量性和具体性的特征。

3. 统计研究的基本方法。从统计工作过程看,统计在各个不同阶段有着不同的工作内容和要求,相应地,就需要运用各种不同的统计研究方法,其中基本方法有大量观察法、统计分组法、综合指标法、统计模型法和统计推断法。

4. 统计研究过程。一个完整的统计研究过程一般包括统计设计、统计调查、统计整理及统计分析四个阶段。

5. 统计学中最基本的概念是统计总体和样本、标志和指标。对这些概念必须正确理解,根据具体的研究任务和对象加以确定。



复习思考题

一、名词解释

统计 统计学 总体 总体单位 样本 标志 标志表现 统计指标 指标体系 数量标志 品质标志 数量指标 质量指标

二、简答题

- (1) 怎样理解统计的含义?它们之间有什么关系?
- (2) 统计研究对象是什么?它有哪些特点?
- (3) 统计研究过程分为哪几个阶段?
- (4) 统计研究的基本方法包括哪些?
- (5) 什么是统计总体和总体单位?试举实际例子说明。
- (6) 什么是指标?指标和标志有什么区别和联系?
- (7) 什么是指标体系?

三、判断题(把“√”或“×”填在题后的括号里)

- (1) 统计调查过程中采用的大量观察法,是指必须对研究对象的所有单位进行调查。
()

- (2) 社会经济统计所研究的领域必须是社会经济现象总体的数量方面。()
- (3) 对某市中小学教师的收入状况进行普查,该市中小学教师的工资水平是数量标志。()
- (4) 品质标志说明总体单位的属性特征,质量指标反映现象的相对水平或工作质量,二者都不能用数值表示。()
- (5) 由女学生组成的总体中“性别”这个标志是不变标志,不变标志是构成总体的基本条件。()

四、单选题

- (1) 社会经济统计的主要特点是()。
- A. 社会性、总体性 B. 抽象性、数量性
C. 具体性、总体性 D. 数量性、总体性
- (2) 社会经济统计现象形成总体的必要条件是()。
- A. 总体单位的大量性 B. 总体单位间的同质性
C. 总体单位间的差异性 D. 总体单位的社会性
- (3) 对某城市工业企业未安装设备进行普查,总体单位是()。
- A. 工业企业全部未安装设备 B. 工业企业每台未安装设备
C. 每个工业企业的未安装设备 D. 每个工业企业
- (4) 已知每位教师的工龄是15年,这里的“工龄”是()。
- A. 变量 B. 指标 C. 品质标志 D. 数量标志
- (5) 以产品等级来反映某种产品的质量,则该产品的等级是()。
- A. 数量标志 B. 数量指标 C. 品质指标 D. 品质标志
- (6) 总体的两大特点是()。
- A. 同质性和大量性 B. 同质性和数量性
C. 大量性和数量性 D. 同质性和变异性
- (7) 下列属于品质标志的是()。
- A. 工人年龄 B. 工人性别 C. 工人体重 D. 工人身高
- (8) 统计一词的基本含义是()。
- A. 统计调查、统计整理、统计分析 B. 统计设计、统计分析、统计计算
C. 统计方法、统计分析、统计计算 D. 统计科学、统计工作、统计资料
- (9) 要了解某市工业企业生产设备情况,则统计总体是()。
- A. 该市全部工业企业 B. 该市每个工业企业
C. 该市工业企业的某一台设备 D. 该市工业企业的全部设备
- (10) 数量指标一般表现为()。
- A. 平均数 B. 相对数 C. 绝对数 D. 众数
- (11) 统计指标的特点是()。
- A. 综合性、具体性 B. 准确性、及时性、全面性
C. 同质性、差异性 D. 科学性、客观性
- (12) 下列指标属于质量指标的有()。

A. 合格品数 B. 职工总数 C. 资金产值率 D. 工资总额

(13) 调查某大学 5 000 名学生学习情况, 则总体单位是 ()。

A. 5 000 名学生 B. 5 000 名学生的学习成绩
C. 每名学生 D. 每名学生的学习成绩

五、多选题

(1) 统计学的研究方法主要有 ()。

A. 实验设计法 B. 大量观察法 C. 统计分组法
D. 综合指标法 E. 统计推断法

(2) 在全国人口普查中 ()。

A. 全国人口总数是统计总体 B. 每户是总体单位
C. 某人的具体年龄是变量 D. 性别男或女是品质标志表现
E. 人口的平均寿命是统计指标

(3) 在下列各项中, 属于统计指标的有 ()。

A. 某企业总资产额 2 000 万元 B. 某学生统计学考试成绩 85 分
C. 某地区国民生产总值 150 亿元 D. 某居民户的人均支出 500 元/月
E. 某市年末就业人口数

(4) 在下列统计指标中, 属于质量指标的有 ()。

A. 工资总额 B. 单位产品成本 C. 出勤人数
D. 人口密度 E. 合格品率

(5) 有一统计报告如下: 某市国有商业企业 650 家, 职工总数 41 万人, 上月的商品零售总额 90 亿元, 职工平均工资额为 1 500 元。其中, A 企业的零售额为 39 万元, 职工人数 820 人, 则报告中出现有 ()。

A. 统计总体 B. 总体单位 C. 标志 D. 指标 E. 变量

(6) 品质标志表示事物的质的特征, 数量标志表示事物的量的特征, 所以 ()。

A. 数量标志可以用数值表示 B. 品质标志可以用数值表示
C. 数量标志一般不用数值表示 D. 品质标志不可以用数值表示
E. 两者都可以用数值表示

(7) 指标按其数值形式不同, 可以分为 ()。

A. 总量指数 B. 相对指标 C. 平均指标
D. 描述指标 E. 评价指标

(8) 指标按其反映的数值特点不同, 可以分为 ()。

A. 综合指数 B. 数量指标 C. 质量指标
D. 全及指标 E. 样本指标