

C 第 1 章

CHAPTER ONE

绪 论

内 容 提 示

统计学在当今社会的经济管理等诸多领域正在发挥着重要的作用，因此应用统计已经成为每个经济管理者、研究者进行科学决策和研究的重要工具。

本章概括介绍统计学中的几个基本概念和问题。通过本章学习，应当掌握统计一词的含义；了解统计学的应用领域；掌握统计学中常用的基本概念；明确统计的研究过程与研究方法。

1.1 统计的含义与应用

1.1.1 统计的含义

统计活动起源很早,可以说自从有了国家就有了统计实践活动。最初的统计作为一种计数活动,为统治者治理国家而收集资料,为国家管理提供人力、物力和财力的依据。如今,在我们的日常工作与生活中,处处都有统计。例如,开会时主持人要统计到会的人数;体育比赛中裁判要统计各队的分数;企业要统计销售额、利润额,等等。但在不同场合“统计”一词却有不同的含义。具体来说,“统计”一词包含以下三种含义。

1. 统计工作

是指统计的实践活动,包括收集、整理、研究和提供统计数据的活动。这层含义的统计应用范围很广,尤其是在当今信息化、数字化的时代,每一个领域的管理者或研究者,都需要收集、整理和分析有关的统计信息,即从事相关的统计活动,以便加强对各领域的科学管理。实践证明,统计工作是管理工作的重要组成部分。

2. 统计数据

是指统计工作过程中收集到的各种数字资料和相关的其他资料,是统计工作的结果。人们可以在统计年鉴、报刊、各种媒体上见到大量的统计数据。准确、可靠的统计数据是进行有效的宏观管理和微观决策的重要依据。

3. 统计学

是指收集、整理和分析解释统计数据的方法与理论。统计学研究的对象主要是客观现象的数量方面,具体来说是研究总体现象的数量特征与规律性。由于各个领域都要研究事物的数量方面,所以统计理论和方法已经渗透到各个科学领域,成为科学研究的重要工具。

上述“统计”所指的含义各有不同,侧重点也不同。例如,如果某人在职业介绍中提到“曾做过统计、审计……”,这里的统计就是指统计工作;如果有人看到或听到“据统计……”等说法,这时的统计即指统计数据;如果学生在选择学习专业时填报“统计”专业,其中的统计即为统计学的含义。因此,我们应当结合具体的情况正确理解和使用“统计”一词。

1.1.2 统计学的应用

无论人们从事生产活动、科研活动还是社会活动,大多都离不开对数据资料的搜集、整理、分析和解释。目前,统计方法已应用到自然科学和社会科学的众多领域,统计学与各领域实质性学科交叉发展成为若干领域的应用统计学,如气象统计学、生物统计学、人口统计学、环境统计学、工程统计学、管理统计学、商务与经济统计学等。

我们这门课所讨论的统计学,主要强调统计学在工商管理领域中的应用。目前,统计学在工商管理中的应用越来越广泛,概括地讲,统计学在工商管理中的人、财、物三个方面的管理中有着极为广泛的应用。在人的管理方面,统计可以为人力资源管理及客户关系管理服务。通过员工满意度和顾客满意度的统计分析,企业决策层可以直接了解到企业员工和顾客的需求,从而更多地挖掘员工的潜力并更好地服务顾客。在财的管理方面,统计可以用于企业的财务分析,使企业决策层更清楚地了解企业的财务状况,并据此建立起企业的财务预警监测系统。对于重大的投资项目,统计可以用于投资风险分析。这些统计分析在保障企业财务安全方面能够起到至关重要的作用。在物的管理方面,统计分析可以用于对企业的产品质量进行监测与控制,从而保持产品质量的稳定性。以上三个方面是统计学在工商管理中的主要应用。此外,在工商管理中的其他许多领域也有统计应用的身影,而且这种身影还会被越来越多地发掘出来。

需要指出的是:尽管统计学非常有用,其应用领域极其广泛,但是单纯靠统计学并不能解决实际工作中的一切问题。在应用统计学研究实际问题时,需要人们具备两方面的知识,一方面要掌握统计学的基本理论与方法;另一方面要具有实际问题的知识背景。这样才能够把定量分析与定性分析结合起来,在应用统计方法进行定量分析的基础上,再应用各学科的专业知识对统计分析的结果做出合理的解释,从而得出令人满意的结论。例如,通过市场调研,利用统计方法可以得出某地区某种品牌家电的市场占有率是23%,但为什么会是这样的比例?要解释形成这一比例的原因则不能靠统计方法,而应由市场学或商品学的知识做进一步的研究和解释。

1.2 统计学中的基本概念

1.2.1 总体与样本

1. 总体

总体(population)就是某一特定研究中所有单元的集合,通常由具有相同性质的许多单元构成,而构成总体的每个单元被称为个体。个体是我们借以收集数据的单元,可以由人、单位或物充当。例如,要对全国工业企业的数量进行统计,全国工业企业是总体,其中每一个企业就是一个个体;若对某地区居民户的住房面积进行统计研究,则该地区所有居民户是总体,其中每一个居民户就是一个个体。

总体分为有限总体和无限总体。有限总体是指构成总体的个体数量是有限多的、可数的,如企业个数、职工人数等;无限总体是指构成总体的个体数量是无限多的、不可数的,如工厂流水线生产的产品、织物纤维等。在社会经济现象中,统计总体大多是有限的。总体是否有限,对统计调查方式的确定十分重要,显然对无限总体不能采用全面调查方式,而

对有限总体既可以用全面调查方式,也可以用非全面调查方式,这取决于统计研究的具体情况和目的。

2. 样本

样本(sample)是指从总体中抽取的部分个体构成的集合。统计研究的目的是认识总体的数量特征,但有时构成总体的个体数量很大,实际工作中不可能或不必要对每个个体的数量特征逐一调查,通常是以某种方式从总体中抽取一部分个体代表总体加以研究。例如,某外贸公司从供货方提供的一批电子元件中随机抽取100只,检验并推断该批电子元件的质量,这100只电子元件就构成一个样本。由此可见,样本是总体的代表,从统计中抽取样本的目的就是用样本的特征去推断总体的特征。

1.2.2 变量与数据

1. 变量

变量(variable)是指一个可以用两个或多个可能值表示的特征或属性。例如,人的性别可以用男或女两个值来表示,性别就是变量。变量通常是对总体中个体的特征或属性的反映,总体中每个个体可以有多种特征或属性,因此在一个总体中存在至少一个变量。例如,将企业中每个职工作为个体考察时,有职业、性别、民族、工种、年龄、文化程度、工资等多个特征或属性,这些特征或属性的名称都是变量。

按照变量的具体表现不同,变量分为品质型变量与数值型变量两种类型。

1) 品质型变量

品质型变量(qualitative variable)又称定性变量,是指用于反映事物品质特征的变量,其表现为类别,通常是用文字来表述的,如人口的性别、企业的经济类型等。

品质型变量按照其表现类别的情况不同,又分为分类变量和顺序变量两种形式。分类变量的表现只体现事物类别之间的不同,如性别的表现有男、女,企业经济类型的表现有国有企业、集体企业、私营企业、合资企业、独资企业等;顺序变量的表现不仅能体现事物类别的不同,还可以体现这些类别之间的优劣或顺序的不同,如考试成绩的表现可以是优、良、中、及格、不及格,人们对某种商品售后服务的态度可以表现为满意、一般、不满意,等等。

2) 数值型变量

数值型变量(quantitative variable)又称定量变量,是指用于反映事物数量特征的变量,通常使用自然或度量衡单位进行计量,其结果表现为具体的数值。如商品销售额、产品产量等。

数值型变量根据其取值是否连续,又可以分为离散变量和连续变量两种形式。离散变量的取值可以一一列举出来,各变量值之间都是用整数位断开的,如企业的人数、机器台数等,都能按整数算,不可能有小数,离散变量的数值只能用计数的方法取得。连续变量的取值是连续不断的,相邻两数值之间无限可分,如人的身高、体重、年龄等,都属于连续变

量,连续变量的数值要用测量或计算的方法取得。

2. 数据

数据(data)即为变量的具体表现。例如商品销售额可以表现为20万元、30万元、50万元;客户的性别可以表现为男、女等,这些变量的具体表现即是统计数据。

由于变量有品质型变量与数值型变量两种类型,因此相应的数据也可以分为品质型数据与数值型数据两种情况。

1) 品质型数据

品质型数据(qualitative data)又称定性数据,是用文字来表述的。如上面所列举的人口性别表现中的“男”“女”;企业经济类型表现中的“国有”“集体”“私营”;考试成绩表现中的“优”“良”等均为统计数据。虽然这些数据只是表现为某种类别,但在实际工作中为了便于统计处理,特别是为了便于计算机识别,我们可以对不同类别用不同的数字或编码来表示,比如用“1”表示男性人口,“0”表示女性人口;用“1”表示国有企业,“2”表示集体企业,“3”表示私营企业,等等。

由于品质型变量分为分类变量和顺序变量两种形式,相应的品质型数据也可以分为分类数据和顺序数据两种类型,如性别的“男”“女”,企业经济类型的“国有”“集体”“私营”“合资”等即为分类数据;而考试成绩的“优”“良”“及格”“不及格”,商品售后服务的“满意”“一般”“不满意”等则为顺序数据。

2) 数值型数据

数值型数据(quantitative data)又称定量数据,其结果表现为具体的数值。如人们的收入表现为700元、800元、1200元;学生的成绩表现为75分、82分、90分、100分;公司的销售额为90万元、110万元、120万元,等等。这些数据都是数值型数据。

由此可见,由品质型变量的表现可得到品质型数据,由数值型变量的表现可得到数值型数据。数值型数据总是表现为数字,而品质型数据既可以表现为文字,也可以表现为数字代码,但这种数字代码实际上是某特定个体的标记,切不要与数值型数据混淆使用。

就统计分析来说,区分品质型数据与数值型数据是很重要的。因为根据不同类型的数
据,我们将采用不同的整理技术和分析方法。例如,对于品质型数据我们主要是进行分类整理,并汇总计算出各类数据的频数或频率,分析方法主要是观察各类数据占总体的比重,以研究社会经济现象总体内部的构成;而对于数值型数据我们通常是按数值顺序进行分组整理,计算出各组数据的频数或频率,分析方法主要是对观察到的数据作各种统计运算处理,进行统计描述或统计推断等。

1.2.3 参数与统计量

1. 参数

参数(parameter)是用来描述总体特征的概括性值。例如,对某城市居民的收入情况进

行统计研究时,该城市居民的总平均收入(μ)、高收入者所占比例(p)等均为参数。

2. 统计量

统计量(statistic)是用来描述样本特征的概括性值。例如,对某城市居民的收入情况进行统计研究时,抽取2000户居民构成一个样本,则该样本的平均收入($\bar{x}$)、样本中高收入者所占比例(p)等均称为统计量。

统计量与参数是统计推断中的重要概念。在进行统计推断时,总体数据通常是不完全的,所以参数是一个未知的常数,但样本是经过抽样所确定下来的,所以统计量总是可以计算出来的。统计推断的目的就是要根据已知样本统计量去估计未知总体参数。

1.3 统计的研究过程与研究方法

1.3.1 统计的研究过程

统计研究的过程就是从收集个体现象的数据入手,在得到大量原始数据资料后,还要按照研究问题的需要对数据进行必要的整理,据此再对总体现象的数据特征进行分析和解释。因此可以将统计研究的过程概括为以下三个阶段。

1. 统计数据的收集

统计数据的收集是指对现象总体中全部或足够多的单位进行调查,收集大量的以数字为主的信息资料,借以反映总体的数量特征。统计数据的收集是统计工作的开始阶段,是进行统计整理和统计分析的基础,离开了统计数据,统计方法就失去了用武之地。因此,统计数据收集的质量如何,直接影响统计工作成果的质量,所以统计数据的收集是统计工作的起点和基础环节,它在统计工作中占有特别重要的地位。社会经济统计数据的收集方式主要是通过统计调查,如普查、抽样调查等。

2. 统计数据的整理

统计数据的整理是对零星、分散的统计数据进行系统化、条理化的加工处理过程,使数据更加符合统计分析的需要。数据整理是在数据收集之后于数据分析之前的一个中间环节,在统计研究的过程中承担着承前启后的作用。数据整理的主要内容包括:数据的审核,即查找数据中的错误,保证应用的数据准确可靠;数据分组,就是将数据按一定标准划分为若干组别,以利于研究总体事物的构成;数据的表述,就是将数据用图表等形式展示出来,使数据的分布特征表现得更加直观、生动、立体和形象。

3. 统计数据的分析和解释

统计数据的分析和解释是统计学的核心内容,它是通过各种统计方法描述现有数量的状况、推断未知总体的数量特征,以达到探索数据内在规律的目的。这些有关问题有待于

以后章节详细介绍。

1.3.2 统计的研究方法

由上述统计研究过程各阶段的内容可以看出，统计研究的起点是数据，目的是探索大量数据内在的数量特征与规律性，而贯穿于这一过程的统计研究方法主要有描述统计和推断统计。

1. 描述统计

描述统计，就是运用各种表格、图形和数字来概括总体数量特征的方法。在对观察到的大量原始数据进行整理汇总后，可以绘制统计表和统计图来显示总体数量的分布特征，也可以计算得出平均数、变异指标、时间序列水平、指数等一系列统计值，应用这些统计值可以描述性地揭示现象在一定条件下的集中趋势、离散趋势、动态变动状况等。

近年来，随着计算机技术的普及和发展，描述统计的各种方法都可以借助计算机来实现，人们也对统计的计算机应用怀有浓厚的兴趣，有关内容将在本书以后的章节中予以介绍。

2. 推断统计

推断统计，就是在人们可以控制的范围内，根据样本资料的特征，对总体的特征做出估计和预测的方法。当面对的总体范围很大，甚至是无限总体，或者由于经费、时间有限或破坏性检验等各种情况时，会使得我们只能在客观上根据局部观察的结果来推断总体的特征。例如，要考察一批灯泡的平均使用寿命，只能从该批灯泡中抽取一小部分进行检验，推断这一批灯泡的平均使用寿命。在研究现象的总体数量关系中，推断统计是现代统计学的基本方法，它既可以用于对总体参数的估计，也可以用于对总体某些假设的检验。

描述统计与推断统计在统计研究过程中的关系可以通过图 1-1 得到反映。

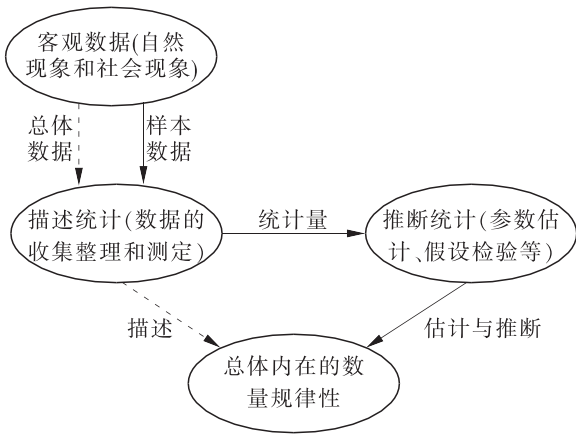


图 1-1 统计方法在统计研究过程中的关系

由图 1-1 可以看出: 统计研究过程的起点是个体数据, 终点是客观现象总体内在的数量规律性。如果我们收集到的是总体数据(如普查), 则经过描述统计之后就可以达到统计研究的目的了; 如果我们所获得的是样本数据, 就必须运用概率论与数理统计的理论并根据样本统计量的信息对总体做出科学的推断, 显然推断统计是在描述统计基础上得以进行的。

尽管描述统计可以在获得总体数据时直接推测出总体数量规律性来, 但这种情况在实际工作中很少见到。大量的管理和研究工作不可能组织普查, 例如城市居民家庭每月的收入支出调查、某种商品的市场调查、某个事件的民意测验等都只能是通过抽样调查方法然后对总体的数量规律进行科学的推断。另外在实践中总体单位很多, 加之要考虑时间、费用和人力的投入, 难以一一进行调查, 所以推断统计被广泛应用于管理的各个领域。

案例 2014 年中国内地上映票房排名前十的电影数据

表 1-1 示出了 2014 年中国内地上映票房排名前十的电影, 其中包括排名、电影名称、总票房、总人次、总场次、制片地、类型 7 个统计指标。

表 1-1 2014 年中国内地上映票房排名前十的电影

排名	电影名称	总票房 /亿元	总人次 /万人	总场次 /万次	制片地	类型
1	变形金刚 4: 绝迹重生	19.79	4743	93	美国/中国内地	动作/科幻
2	心花路放	11.69	3398	92	中国内地	喜剧/爱情
3	西游记之大闹天宫	10.46	2492	60	中国内地/香港	动作/奇幻
4	智取威虎山 3D	8.83	2149	78	中国内地/香港	动作/战争/冒险
5	星际穿越	7.51	2070	60	美国/英国/ 加拿大	剧情/科幻/悬疑/ 冒险/家庭
6	X 战警: 逆转未来	7.24	1936	63	美国/英国	动作/科幻/冒险
7	美国队长 2	7.21	1840	58	美国	动作/科幻/冒险
8	猩球崛起 2: 黎明之战	7.11	1932	66	美国	剧情/动作/科幻
9	爸爸去哪儿	6.96	2187	51	中国内地	儿童/喜剧/家庭
10	分手大师	6.66	2001	56	中国内地	喜剧/爱情

数据来源: <http://58921.com/alltime/2014>;

http://www.douban.com/doulist/1295618/?start=25&sort=seq&sub_type=

(检索时间: 2015 年 11 月 3 日)

问题

(1) 该资料中有多少个变量?

- (2) 哪些是品质型变量, 哪些是数值型变量?
- (3) 该资料中有多少个数据?
- (4) 哪些是品质型数据, 哪些是数值型数据?
- (5) 若以该资料作为样本, 分别指出总体、个体。
- (6) 结合本例, 你能具体列举几个参数和统计量吗?

解 析

(1) 变量有 7 个, 即第一行中各项目: 排名、电影名称、总票房、总人次、总场次、制片地、类型。

(2) 品质型变量包括: 排名、电影名称、制片地、类型。

数值型变量包括: 总票房、总人次、总场次。

(3) 数据共有 70 个, 即对应于第一行的各变量的具体表现。

(4) 品质型数据包括: 对应于品质型变量下的每个数据。

数值型数据包括: 对应于数值型变量下的每个数据。

(5) 总体: 2014 年中国内地上映的电影。

个体: 2014 年中国内地上映的每部电影。

(6) 参数: 2014 年中国内地上映电影的总票房、总人次、总场次等。

统计量: 2014 年中国内地上映的电影票房排名前十位的电影总票房、总人次、总场次等。

本章小结

本章介绍了统计的概念。统计一词通常有三种含义: 统计工作、统计数据和统计学, 各种含义通常用在不同的场合。统计学的应用范围非常广泛, 本书主要强调统计学在工商管理领域中的应用。

统计学中的常用概念包括总体与样本、变量与数据、参数与统计量等。总体是指某一特定研究中所有元素的集合, 通常由具有相同性质的许多个体构成。样本是指从总体中抽取的部分元素构成的集合。变量是指一个可以取两个或多个可能值的名称或概念。数据即为变量的具体表现, 也称为变量值。统计数据可以分为品质型数据与数值型数据两种类型, 不同类型的数据通常要用不同的统计方法进行整理与分析。参数是反映总体特征的概括性值。统计量是用于反映样本特征的概括性值。参数与统计量是统计推断中的重要概念。

统计研究的过程可以概括为数据的收集、数据的整理、数据的分析和解释几个主要阶段。在整个统计研究过程中所使用的研究方法主要有统计描述法和统计推断法。

重要概念

统计	总体
样本	变量
参数	统计量
描述统计	推断统计



思考与练习

思考题

1. 统计一词有哪几种含义?
2. 总体与样本有何关系?
3. 什么是变量? 变量有哪些分类? 变量与数据有何关系?
4. 举例说明参数和统计量的关系。
5. 简述统计研究的过程。

练习题

一、判断题(正确的标注 T; 错误的标注 F)

- () 1. 若对某地区居民的收入情况进行统计研究, 则关注的变量是该地区居民的收入。
- () 2. 贯穿统计研究过程的方法主要包括描述统计和推断统计。
- () 3. 总体是样本的代表。
- () 4. 统计数据可以分为品质型数据与数值型数据两种基本类型。
- () 5. 统计量是用于反映总体特征的概括性值。

二、单项选择题

1. 一家商户接待了 10 万名客户, 从中抽取了 300 名客户进行满意度调查, 则调查的样本容量是()。
- A. 10 万 B. 300 C. 10 300 D. 99 700
2. 某咨询公司对我国 15 个城市 500 户家庭进行电话调查, 了解到家庭年均收入约为 4.5 万元, 则这个统计研究关注的总体是()。
- A. 我国 15 个城市 B. 我国城市家庭