

绪 论



应用案例

奥巴马在 2012 年美国大选中击败罗姆尼,成功连任。在这连任的过程中,有一个重要的工具就是统计数据。美国的《时代》杂志撰文指出奥巴马成功筹集数亿美元竞选资金、有效锁定目标选民、即时调整电视广告投放策略并最终赢得选举的秘诀就是大数据。

在总统竞选一开始,竞选活动经理吉姆·梅斯纳(Jim Messina)就已经打算搞一次完全不同的、以数据为驱动力的竞选活动,准备用数据去衡量这场竞选活动中的每一件事情。梅斯纳聘请了一个五倍规模于 2008 年竞选时的数据分析团队,并让雷伊德·加尼(Rayid Ghani)出任芝加哥竞选团队总部的首席科学家。这足以看出该竞选团队对数据统计分析巨大潜力的重视。下面简述数据统计分析在竞选过程中的应用。

在筹集资金方面,起初,竞选团队就创建了一个单一的巨大系统,数据分析团队可以利用此进行一些测试,分析哪些类型的人有可能被某种特定的事情说服。数据分析团队可以预测哪些人会在网上捐钱,哪些人会用邮件捐款,甚至对志愿者进行建模分析。竞选办公室的电话名单不仅列出了姓名和电话号码,并且按被说服的可能性和重要性进行了排序。

有相关人员称,民调数据与选民联系人数据每晚都在所有可能想象的场景下被计算机处理、再处理。每天晚上都模拟运行 66 000 次选举,这些模拟选举推算出了奥巴马在摇摆州的胜率,知道赢得这些州的机会在哪,并让他们得以通过这些数据来分配资源。

此外,竞选团队基于背后强大的内部数据库分析,通过建立复杂的模型定位不同的选民,购买了一些非传统类剧集之间的广告时间,替代了在当地新闻时段的传统做法。最后数据统计显示,广告效率相比 2008 年提高了 14%。

在竞选后期,通过数据分析发现社会化新闻网站 Reddit 中有许多动员的目标,奥巴马决定去社会化新闻网站 Reddit 回答问题。

由此可以看出,数据的挖掘分析为奥巴马的成功连任立下了汗马功劳,数据统计分析的地位在政治领域逐渐上升。一位官员说,“人们坐在屋里抽雪茄,高喊‘买《60 分钟》’的时代已经结束。”政治领域的大数据时代已经到来。案例中通过数据分析得出结论会涉及许多其他学科的支持,但是统计学作为一门基础学科,其发挥的重大作用理应引起大家的关注。^①

① 案例改编于数据挖掘助奥巴马连任 大数据左右美国政界. <http://it.sohu.com/20121108/n357017969.shtml>.

本章内容从讨论统计与统计学出发,对统计学进行概述性的分析,最后简要介绍Excel在统计分析中的应用。

第一节 统计与统计学

当今世界,人们在社会、经济、科技等各个领域都感受到信息的作用和影响,人类已步入了信息社会。而统计作为信息处理的一种重要手段,已成为广泛应用于经营管理、学术研究、社会生活等众多领域,了解统计知识对每个人来说也都是必要的。例如,在外出旅游时,你需要关心一段时间内的详细天气预报;在投资股票前,你需要了解股票市场价格的财务信息;在观看足球比赛时,除了关心进球的多少外,还要了解各个球队的技术统计结果。

通过统计,人们将会得到许多实用、有趣的东西,同时又会提出不少问题。例如,“统计”什么?为什么用商品零售物价指数可以反映社会商品零售价格总水平变动?为什么只调查少数人的人均收入就可以了解一地、一省乃至全国的人均收入?这些都反映了统计的奥秘。

一、统计的含义

“统计”在《现代汉语词典》(第7版)中有两个解释:一是总括地计算,二是指对某一现象有关的数据进行搜集、整理、计算和分析等。“统计”一词包含三个含义,统计工作、统计资料和统计学,即围绕研究目的和任务,运用科学的统计方法,获取真实的科学的统计数据,分析、了解和认识所研究事物的真相。

1. 统计工作

美国普渡大学的教授墨尔(David S. Moore)说过,统计是从数据中获取信息的科学。我们大家都很熟悉的统计,是体育、人口、经济和证券交易等领域数字资料的汇集。在日常生活工作中,“统计”有着多种含义。例如,上课老师统计上课的学生人数;NBA(美国职业篮球联赛)比赛中,人们所关注的球员的技术统计:得分、篮板、助攻等方面的数据;工厂在经过一定时间周期后,要对本厂该时期的效益进行统计,核算其当期完成的产量和产值;等等。此时,“统计”具有计数的含义,在这个意义上统计的起源可以追溯到远古时代。

在原始社会,人类的一般技术活动就蕴含着统计的萌芽。四五千年之前,居住在两河流域一带的人们就在石头上记录数字。从历史上看,早在奴隶制国家时代,由于政府赋税、徭役、征兵等需要,就对人口、土地、财产等方面进行调查统计。据记载,公元前27世纪,古埃及建造金字塔,为征集建筑费,就曾对全国的人口与财产进行调查。《旧约全书》中记载,公元前10世纪,犹太国王大卫和所罗门对全国进行了完整的人口和财产调查统计,以此作为实施统治措施的重要依据。我国春秋战国时代,统计已被认为是治理国家、创立基业的重要手段。齐国丞相管仲在《管子》一书中曾提出:“不明于计数,而欲举大事,犹无舟楫而欲经于水险也。”这就是说不善于利用计数而进行宏伟事业,犹如没有船和桨而想渡过激流险滩一样。秦国主持变法的丞相商鞅认为,“强国知十三数”,即他将十三

种数据作为反映国情国力的基本指标。由上可知,在“计数”的含义上,统计的应用十分广泛,而且成为人们治理国家的一项必要工作。发展至今,统计工作已成为国家进行宏观管理的一项专门职能工作,如我国从地方政府到中央都设有专门的统计部门,负责对国家、地区的经济、科技、文化、教育、社会等多方面进行统计信息的收集、整理、加工,统计工作成为信息、咨询、决策职能的重要载体。

2. 统计资料

在统计工作的基础上,会形成一系列的数字资料,这是统计工作的成果表现,也称统计资料或统计数据,这是“统计”的另一个含义。它与上述的统计工作是紧密相连的,是统计工作的直接结果,是反映了一定空间、时间条件下客观总体现象数量特征的数据型资料,因此也是很早就有的。根据历史记载,我国夏禹时代就开始有人口方面的统计数字。随着人类社会的不断进步,统计数字作为一种重要的信息,人们对它的需要越来越多、越来越细化。平时我们翻阅报刊,就可以看到各种统计数字,如股票指数、价格指数、失业率等。国家统计局、有关信息部门每年出版的统计年鉴和相关资料,反映国家或地区的宏观经济、人口、地理等方面基础资料以及科技、教育、文化等方面的发展等情况,这些都是在此意义上的统计。统计表、统计图、统计手册、统计年鉴等统计资料已成为国家的重要战略资源。

另外,“统计”一词还有“统计学”的含义。作为一门科学,统计学的出现要比统计工作和统计资料晚得多。它是本书要探讨的主要内容。

二、统计学

统计学是关于数据的科学,是一门收集、整理、总结、描述和分析数据,并在分析的基础上得出正确结论和合理决策方法的科学,它所提供的是一套有关数据收集、处理、分析以及解释并从中探索数据的内在规律性的方法。

对统计学的定义,需要注意以下几点。

(1) 统计学的研究对象是客观事物现象的数量方面,是总体现象的大量数据,而不是个别的单个的数据。因为单个数据在没有其他数据作为参考对照时,是说明不了什么问题的。例如,某学生在英语考试中取得 80 分的成绩,但仅通过这么一个数字是无法说明该学生的成绩如何的,因为不知道其他学生的考试成绩。只有其他学生的考试成绩知道了,才能对该学生的成绩进行评价,得出他是处于中等水平,还是较高水平,抑或是较低水平的结论。另外,统计学所研究的客观现象十分广泛,既包括自然现象,也包括社会现象。如前所述,早期统计所研究的问题有人口调查、保险精算等,后来逐渐扩大到社会经济和生物实验等方面。随着人类的实践活动的需要和信息化的发展,各领域都需要研究事物的数量方面,需要联系数量方面的特征来研究事物的本质。因此统计学的应用范围越来越宽,凡能以数量来表现的现象特征均可作为统计学的研究对象。统计方法已渗透到多个科学领域,成为目前最活跃的学科之一。萨维奇曾说:“统计学基本上是寄生的。靠研究其他领域内的工作而生存。这不是对统计学的轻视,这是因为对很多寄主来说,如果没有寄生虫就会死。对于有的动物来说,如果没有寄生虫就不能消化它们的食物。因此,人类奋斗的很多领域,如果没有统计学,虽然不会死亡,但一定会变得很弱。”

(2) 统计学是一门方法论科学,它是由收集、整理、描述和分析数据资料的统计方法组成的。统计方法来源于对数据资料的研究,其目的是研究数据资料以揭示出数据内在的规律性。统计方法有多种,主要包括描述统计方法、推断统计方法。如平均法、统计分组法、时间序列分析法、指数法、假设检验法、方差分析法、相关分析法、回归分析法、国民经济核算方法等。过去在我国统计学界有实质性科学和通用性科学之争,随着实践的发展,人们普遍认为统计学是一门实用性很强的方法论学科。统计方法是来源于实践的,是从现实问题中产生的,统计方法的发展是动态的。随着统计方法的应用日益广泛,其内容不断充实,尤其是近代概率论的发展为统计方法提供了理论基础,发展到目前,统计方法已相对独立地形成了自己的科学体系,其内容包括如何收集资料,如何对收集的资料加以整理、概括和表现,以及如何对取得的数据进行分析和推断等一系列方法。现在统计方法已成为宏观及微观经济管理、科学决策和科学研究的重要工具,统计学学科体系日趋完善,成为一门充满活力的学科。

(3) 统计学的研究目的是探索和研究客观现象数量的内在规律性。什么是数据内在的数量规律性呢?下面我们用新生儿性别例子进行说明。我们知道,对每个家庭中新生婴儿的性别来说,既有可能是男孩,也有可能是女孩,是无法事先确定的。在过去未实现计划生育的时候,某个家庭可能连续生几个男孩,另一个家庭可能连续生几个女孩,甚至出现过七朵金花的情形。因而,新生婴儿的性别从个别家庭来看,似乎是没有特殊规律的。但如果对多个新生婴儿进行观察,即观察几千个或几万个新生婴儿,就会发现男孩会比女孩稍多一些,大致为每生 107 个男孩,就会有 100 个女孩出生。 $107:100$ 就是新生婴儿性别比例的稳定的数值,即新生婴儿性别比例的数量规律性。之所以将 $107:100$ 称为数量规律性,是因为古今中外新生婴儿的性别比例都大致如此。这种数量规律性是由人类遗传规律和人类社会发展规律所决定的。众所周知,人类社会自身的协调发展,需要男女性别结构保持在 $1:1$ 的水平。在少儿时期,男孩比女孩略多,但男孩的死亡率比女孩的死亡率要高一些。到中年时期,男女人数基本平衡。到老年时期,由于生理结构的原因,男性的平均预期寿命比女性的要低,从而男性老人比女性老人相对要少一些。这样,从总体上看,人类社会男女总人数还是基本平衡的,在中青年结婚生育时,也是大致平衡的。当然,近些年来我国新生婴儿性别比例较高,一般在 $110:100$ 以上,这主要是由农村人口愿意要男性婴儿的缘故造成的。通过“男女平等”的宣传和计划生育工作力度的加大,这一趋势又有所改善。总的说来,新生儿性别比例是人类长期进化的结果,是一种和谐的平衡。

那么为什么统计方法能通过对数据的大量观察进而探索和揭示数据的内在规律性呢?这是由客观事物的本身特点和统计方法的特性所共同决定的。根据唯物辩证法的原理,客观事物是必然性和偶然性的对立统一,当然反映数量特征的数据也是受必然性和偶然性两种因素的支配,是两者的对立统一。必然性反映了事物本质的联系,是比较稳定的,它决定了事物的表现形式是有规律可循的。如果客观事物只有必然性的一面,那么事物的表现形式就会简单化,也就不需要用统计方法了。但由于受偶然性的影响,事物的表现形式千变万化,反映事物数量特征的数据千差万别,这样必然性和数量规律性就被掩盖在千差万别的事物的表现形式中了。而统计学提供了一系列的统计方法,对表面上杂乱

无章的数据加以整理和处理,根据研究目的和客观事物的特点探索出数据的内在规律性来,然后结合专业知识对这些数量规律性进行解释和分析。

三、统计工作、统计资料与统计学

统计工作、统计资料与统计学是一个事物的三个方面,它们之间既存在着显著的区别,也存在着密切的联系。统计工作和统计资料之间是过程和结果的关系,统计工作以统计资料的形式表现出来,统计资料的质量直接取决于统计工作的情况;统计学与统计工作之间是理论与实践的关系,统计学的发展基于统计工作实践,是统计工作的经验总结,同时统计学的发展又能够为统计工作的开展提供理论和方法的指导。这三个方面构成了统计的完整概念,如图 1-1 所示。

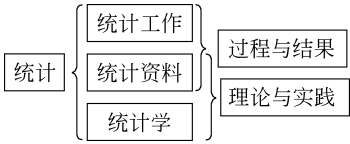


图 1-1 统计工作、统计资料与统计学的关系

第二节 统计学的发展历程

统计学是一门很古老的科学,一般认为统计学的学理研究始于古希腊的亚里士多德时代,迄今已有 2 300 多年的历史。但是将其作为一门学科系统地研究和发展,只有不到 400 年的历史。德国的斯勒兹曾说过:“统计是动态的历史,历史是静态的统计。”统计学是在不断发展和完善的,回顾统计学的发展历程,可以分为古典统计学时期、近代统计学时期以及现代统计学时期三个阶段。

一、古典统计学时期

古典统计学时期主要是指 17 世纪中叶至 18 世纪中叶的统计学时期,这段时期成为统计学的创立萌芽时期。在这一时期,统计学理论初步形成了一定的学术派别,主要有国势学派和政治算术学派。

1. 国势学派

最初的统计学是来源于对国家重大事项的记述,持这种观点的被称为国势学派,又称为记述学派。国势学派起源于 17 世纪的德国,其创始人是德国西尔姆斯特大学的海尔曼·康令(H.Conring,1606—1681)和哥廷根大学的阿亨瓦尔(G.Achenwall,1719—1772)。

1660 年,康令第一个在德国西尔姆斯特大学以“国势学”为题讲授政治活动家应具备的知识。他将国势的记述上升至理论化和系统化的高度,内容是关于各国行政、人口、土地、财政等方面的国家显著事项。阿亨瓦尔在哥廷根大学开设“国家学”课程,继承和发展了康令的观点,认为统计学是研究国家显著事项的学问。他在 1749 年出版的著作《近代欧洲各国国势学纲要》中首先提出并且使用了统计学这一名词,书中讲述了“一国或多数

国家的显著事项”,主要用对比分析的方法研究了解国家组织、领土、人口、资源财富和国情国力,比较了各国实力的强弱,为德国的君主政体服务。由于他最早提出统计学这一名词,所以有人称其为“统计学之父”。

总体来说,国势学派认为统计学是对政治事项的记述,偏重事物性质的解释,而不注重数量对比和数量计算,这与现代统计学相差甚远,但却为统计学的发展奠定了经济理论基础。但随着资本主义市场经济的发展,对事物量的计算和分析显得越来越重要,该学派后来发生了分裂,分化为图表学派和比较学派。

2. 政治算术学派

政治算术学派是用计量的方法来研究计量经济问题,运用大量观察法、分类法以及对比、综合、推算等方法对社会经济生活问题进行分析。政治算术学派起源于英国,产生于17世纪中叶,其代表人物有威廉·配第(William Petty,1623—1687)和约翰·格朗特(John Graunt,1620—1674)。

1662年,约翰·格朗特在其著作《对死亡表的自然观察和政治观察》一书中对当时英国人口情况的分析揭示出一系列的数量规律关系。如男婴出生多于女婴(14:13),男性死亡多于女性,一切疾病和事故在全部死亡原因中占有稳定的百分比等。他在该书中指出,为了找出一个要在多年内形成的规律,需要进行多次观察。因此格朗特虽然未用到统计学这一概念,但他在实践中却应用了大量观察方法去研究分析规律性的问题,其独特的方法给人以启迪。他是最早利用统计资料查明统计规律性的人。随后,威廉·配第在著名的《政治算术》一书中,用大量的数字来比较分析英国、法国、荷兰三国的经济实力,采用了与传统方法不同的分析方法,用数字、重量和尺度来表达他自己想说的问题和观点。马克思对威廉·配第的评价很高,认为他“在某种程度上也可以说是统计学的创始人”。

由上可知,政治算术学派是以数量分析为特征,以客观现象数量关系为研究对象,在研究内容和方法上可以说是统计学的正统起源。但是毕竟它只是应用了数据,而并未对数据的收集、如何利用数据进行统计分析等进行系统论述,因此它并不是现代意义上的统计学。



扩展阅读 1-1 威廉·配第对统计的贡献

二、近代统计学时期

近代统计学时期主要是指18世纪末到19世纪末的这一段时间,这时期是统计学的形成时期。这段时期的统计学分为数理统计学派和社会统计学派。

1. 数理统计学派

数理统计学派是从19世纪逐步形成的,由于它主要是在英美等国发展起来的,所以又称为英美数理统计学派。该学派的代表人物主要有拉普拉斯和比利时统计学家、数学家以及天文学家阿道夫·凯特勒(Adolphe Jacques Quetelet,1796—1874)。

拉普拉斯是最早将概率论方法引入统计学的,他在1802年利用抽样调查方法对法国人口进行了调查,估算了全国的人口数,并给出了误差区间。其后,凯特勒融合了国势学派和政治算术学派的已有成果,并把概率论的原理和大量观察法引入了统计研究的领域,

他不仅在人口问题、寿命问题等自然领域运用该统计方法,在犯罪问题等社会领域也运用概率论的方法进行研究,取得了丰硕的成果。凯特勒将概率论引入统计学这一举动极大地提高了统计计量的准确性,使统计方法发生了重大的发展,使统计学产生了质的飞跃,为近代统计学奠定了基础,因此有人推崇其为“近代统计学之父”。

后来,统计学的发展也相应地要求概率论的方法不断地发展,如高尔顿(F. Galton, 1822—1911)提出了相关和回归的概念;皮尔逊发展了相关和回归的理论,提出了假设检验和卡方分布。这都极大地丰富了数理统计学派的发展,使得它渐渐地从统计学中分离出来成为一门独立的应用数学。

2. 社会统计学派

社会统计学派主张统计学是研究社会现象内部的联系以及它们变动的规律的社会科学,它形成和发展于 19 世纪后半叶的德国。由于欧洲各个国家之间的联系,所以社会统计学派形成后在欧洲大陆占据着很大的优势地位,因此该学派也被称为大陆学派。该学派的奠基人为克尼斯(Karl Gustav Adolf Knies, 1821—1898),代表人物主要有德国的统计学家、经济学家恩格尔(Ernst Engel, 1821—1896)和梅尔(Georg Mayr, 1841—1925)。

1850 年,克尼斯在《作为独立科学的统计学》书中,提出了“国家论”与“统计学”科学分工的主张。他认为国家论是用文字记述的国势学的科学命名,统计学则是用数值解读大量现象的一门科学。恩格尔提出了统计调查、整理和分析三阶段的统计方法,并且除了创立为人所知的恩格尔系数,恩格尔还提倡在人口调查中用个人调查卡片法,并且在消费计量中,首创了消费权数。梅尔认为“统计学是在对总体现象大量观察的基础上,对人类社会生活实际状态及其所产生的规律性,作出的系统性的表述和说明”。此外,他认为统计方法是统计学中不可缺少的一部分。

社会统计学派认为统计学的研究对象是社会总体,而不是个别的社会现象,在研究方法中使用大量观察法。社会经济的不断丰富与发展,以及社会科学本身的不断细化以及定量化发展,也要求统计方法能够不断改进、不断完善,满足社会调查的需要。所以,社会统计学派的研究逐步由实质性科学向方法论转变,但是仍然首要强调现象的实质。在 20 世纪,人类社会经历了两次世界大战,国际政治风云几番突变,这些都对社会统计学造成了一定的影响。

三、现代统计学时期

现代统计学时期主要是指 20 世纪初到现在,这段时期称为统计学的形成时期。现代统计学沿袭了国势学派统计学的名称,内容上除了对国家重要事项的调查外又扩大了研究的范围,同时又吸取了政治算术学派、概率论对客观现象进行数字计量、大量观察和研究分析的方法,在三者基础上逐渐融合而形成。现代统计学的发展具有以下几个特征。

1. 推断统计的发展

进入 20 世纪以后,随着科学技术尤其是计算机技术和新兴学科的发展,统计学的研究和应用范围越来越广,在随机抽样的基础上建立了推断统计学的理论和方法,即通过随



扩展阅读 1-2 高尔顿简介

机样本来推断总体数量特征。推断统计起源于英国数学家戈赛特(W.S.Cosset, 1876—1937)的小样本 t 分布理论,经过英国统计学家费雪(R.A.Fisher, 1890—1962)的充实,费雪提出了重要的随机性原则,然后由美国统计学家内曼(J. Neyman, 1894—1981)等人进一步发展,并建立了统计假设理论。从20世纪中叶开始,推断统计学逐步成为主流的统计学。



扩展阅读 1-3 戈赛特与小样本 t 分布理论

2. 统计学研究领域更加广泛,在现代管理中地位日益重要

在20世纪之前,统计学主要应用于生物统计、社会统计、经济统计这几个领域。随着社会和科学技术的发展,统计学目前已经应用到社会生活的各个领域,成为通用的方法论科学。它被广泛用于研究社会和自然界的各个方面,并发展成为有着许多分支学科的科学。统计也不仅仅局限于对过去现象的总结,开始向统计预测和统计决策发展。在现代管理中,统计面向社会、面向人民的服务领域和内容不断拓展,提供的信息范围也在不断地延伸,在国家管理、企业预测、家庭投资决策中发挥着越来越不可替代的作用。

3. 信息论、控制论、系统论与统计学的相互渗透和结合

统计学与信息论、控制论、系统论不断结合,相互借鉴方法、思想,这改变了世界的科学图景和科学家的思维方式,也使统计科学和统计工作中吸取了营养、拓宽了视野、丰富了内容,使统计科学不断发展、不断完善。

4. 计算机技术在统计中的应用

计算机技术的发展,为统计数据的收集、处理以及储存的过程提供了很多便利,为统计学发展的理论基础提供了有力的支持,这极大地加速了统计学的发展和统计学在现代生活中的应用。在数据形式多样、容量巨大的当今社会,统计学也要求将计算机技术应用到统计研究中,所以计算机技术的应用使得统计学和统计工作发生了革命性的变化。

第三节 统计学分科及与其他学科的关系

一、统计学的分科

随着实践的不断发展和对科学研究的不断深入,科学的学科体系也不断趋于完善。由于人们分析的角度和研究的重点有所区别,必然会出现各个相互联系而又有区别的分支子科学。统计学也不例外,大致有以下两种分类方法。

1. 描述统计学和推断统计学

统计学可分为描述统计学和推断统计学,主要是反映统计学发展的两个历史阶段,同时在研究方法上各有不同的侧重点。

描述统计学是研究如何对客观现象数量特征进行数字的计量、概括和表示的方法。在客观现象中,有些现象的数字描述是比较简单的,如参加会议人数、火车站数量等;而在某些领域如社会经济方面就比较复杂,如要分析可持续发展的程度或对不同国家国际竞争力进行比较,就要涉及多个方面。因此需要确定一些反映现象数量特征的范畴,即统计

指标。研究某一个问题,要选择恰当的统计指标,对问题进行全面的反映、刻画和系统的认识。有的问题只需要一个统计指标就可以了,然而对复杂的问题,一个指标往往只能说明某一方面的问题,用一个指标来全面评价复杂的问题是不现实的。对于社会经济现象的发展状况,由于是较复杂的,仅用一个统计指标来反映显然是片面的和行不通的。为了比较全面系统地认识社会经济现象,就需要用多个相互联系的指标来反映所研究问题的各个侧面,这些相互联系的统计指标就构成了统计指标体系。通过统计指标体系来对现象进行刻画,有利于我们全面地把握问题的本质。有了统计指标体系后,就可以根据不同的研究对象和研究目的,确定相应的收集数据的方法。在一些自然科学中人们通常是根
据实验观察来获得数据,而在社会科学中往往通过实地调查和访问来取得。收集来的原始数据,不经整理,是杂乱无章的,难以直接看出什么问题。对数据整理、加工后,要按各种分组方法整理,并综合成一些统计图、统计表醒目简便地表达出来,并对事物数量的集中趋势和离散程度进行分析。以上所述的内容就是描述统计学的主要内容。描述统计学就是以便利化和信息化的方式对数据进行整理、汇总、显示,表达形式包括图示法,或者通过数值法、平均值法等。

在 20 世纪之前统计学基本上处于描述阶段,进入 20 世纪后,随着概率论的发展,推断统计学不断发展进步并成为统计学研究的主流。推断统计学又称为归纳统计学,它是研究如何根据部分总体单位数据去推断总体的情况,是在概率论的基础上发展起来的。由于客观事物范围广泛、数量很多,在对客观现象收集数据时,对所有单位去做调查是不可行的或不现实的。例如,要研究我国某品牌电器产品的市场占有率情况,由于我国地域宽广,如果用全面调查的方法,显然不现实。由于电器产品的顾客群是较难识别的,抽样总体是无法确定的。因而通常采取抽取部分顾客进行研究,从而对总体情况作出推断。当然,由于只调查了部分顾客,包含的信息不全面,而且抽样时是随机的,必然会出现抽样误差,使推断结论产生了一定的不确定性。但是推断统计学根据概率论的原理可以使归纳推断所产生的不确定性得到度量,因而抽样调查也成为推断统计研究的重要内容。推断统计学是利用样本数据信息对总体特征作出推断的一系列方法。由上所述,描述统计学是统计学的基础,而推断统计学则是近代统计学的核心。

2. 理论统计学和应用统计学

理论统计学是指扎根于概率论的统计学的数学原理。把研究对象一般化、抽象化,以概率论为基础,从理论的角度,对统计方法加以推导论证,其中心内容是以归纳方法研究随机变量的一般规律。从广义来说,统计学理论是包括概率论的,另外还包括随机化原则的理论、各种估计的原理、假设检验的原理以及一般决策的原理等一些并不属于传统概率论的内容,这些原理可以看成概率论公理的拓宽。通常在统计活动实践中遇到一些新问题,用原有的统计方法无法解决时,就需要统计工作者针对新出现的问题去建立一个与实际情况相适应的统计模型,构造新的统计方法去研究分析,这些活动就要靠统计理论来指导,才能取得一定成果。

应用统计学是从所研究的领域或者特殊问题出发,是采用恰当的指标体系和统计方法研究对象的性质,以解决所需研究的问题。统计学是应用性很强的一门学科,统计学的基本原理方法应用于不同专业领域就形成各种各样的专业统计学。应用统计学形成了一

整套统计分析方法体系,具体内容包括适用于各个领域的如参数估计(parameter estimation)、假设检验(hypothesis testing)、方差分析、相关与回归等一般性的统计方法,还包括专有分析方法,如经济统计学中的指数分析法等。近些年来,由于统计研究的范围越来越广,统计方法也日趋复杂化和专门化,在应用统计方法时需要对于统计模型和实际情况的不一致而引起的各种误差的性质和大小作出判断,或提出改进的措施。由于统计工具的专门化,其通用性就受到限制。可以说,现在一个统计学家要熟悉所有的专门统计工具方法已不现实。为适应社会实践发展的需要,既熟悉统计知识又熟悉某一领域业务的应用统计人员就应运而生,同时也产生了相应的应用统计学。这类统计学的特点不侧重于统计学原理的推导,而是侧重于阐明统计的思想,并将理论统计学的结论作为工具广泛应用于各个具体领域。本书主要阐述应用统计学,特别侧重于工商管理方面的应用。

统计学经过几百年的发展,目前已经成为横跨社会科学和自然科学领域的多科性的科学。横向来看,各种统计学都具有统计的本质特点。从纵向来看,统计学方法应用于各种实质性科学,同它们相结合,产生了一系列专门领域的统计学,如表 1-1 所示。

表 1-1 统计学在各领域的应用

学 科	应 用 领 域	应用统计学
统计学	经济	经济统计
	管理	管理统计
	社会	社会统计
	教育	教育统计
	物理	物理统计
	生物	生物统计
	医学	医学统计
	...	

二、统计学与其他学科的关系

1. 统计学与数学的关系

统计学是研究客观现象数量方面的学科,其理论基础是数学理论——概率论,因此,可以说统计学是应用数学的一个分支,与数学学科的关系十分密切,与其他的应用数学有一定的共性。因而统计学,尤其是理论统计学,需要有良好的数学基础。

但统计学与其他的数学分支又不一样,它有如下特点。

(1) 统计学处理的数据是受偶然性的影响而导致区别和差异的数据,它是研究随机现象的一门学科。这一点与其他数学分支存在一定区别。

(2) 在研究方法上,数学常常是用演绎的方法,即从一般到特殊,从一些假设命题、已知的事实出发,按一定的逻辑推理去推断相应的结论,主要是通过严格的定义、假设的命题以及条件进行论证。而统计学在本质上是用归纳的方法,即从特殊到一般,它是根据观

察到的样本的特征,“归纳”起来去推断总体的特征,这一点与概率论的方法也有区别。因此目前国际上把统计学看成与数学相互独立的一门学科。

(3) 统计学与数学在研究基础上也存在着差异。数学是抛开具体的对象,以最一般的形式研究数量的联系和空间形式。而统计学特别是应用统计学则总是与客观事物联系在一起,从客观事实出发,归纳得出其数量规律,得出研究结论。从客观现象出发,最终还要回归到客观对象中去。

2. 统计学与其他专门学科的关系

统计学的应用性很强,其一般的数据分析方法适用于其他学科中的偶然现象,因此它与很多专业学科都有关系。但是统计方法只是从事物的外在数量表现去推断该事物可能的规律性,它本身并不能说明为何会存在规律性。至于为什么会存在这种规律性,则是各专门学科领域内的任务。例如,我们用统计方法分析一些资料得出结论,吸烟与某些人体疾病有关,这是通过吸烟者的发病率和吸烟者的发病率的对比得出的结论,但它只能揭示出这种关系,并不能解释吸烟何以会增加患病的概率,如何去解释这种关系则是医学专业学科的任务。所以统计方法只是一种工具,进行定量分析时必须和定性分析结合起来。

需要指出的是,将统计方法应用于社会经济领域更应注意定性和定量分析的结合,因为社会经济现象的影响因素比自然现象的影响因素更为错综复杂,且往往是不可控制的。统计分析方法虽然是一个强有力的工具,但必须慎重使用,不了解其背景而随意使用统计分析方法是十分危险的做法。例如近年来我国经济发展水平和刑事案件的数量都在提高,能否由此说明经济发展水平越高,刑事案件发生率越高呢?显然,这样的结论是荒谬的,刑事案件的增加是受多种因素影响的结果,因此有必要在应用统计方法时全面周密地考虑全部有关情况,同时把统计学知识和其他有关专业知识结合起来,定性和定量分析方法相结合,才能更好地发挥统计工具的作用。

因此,统计学现在与其他学科也有着密切的关系。在统计工作中要继续深入实际采集数据,研究时不仅要用到统计的方法,还要掌握该领域的知识,才能得到有意义的成果。统计学可以帮助人们认识有关的数量规律,检验其他专门学科理论,也提高了其他学科的精确程度。同时在对专门学科进行分析的过程中,也对统计学提出了一些挑战,这样也有助于统计学的不断发展。统计学只是利用现有的数据对数量规律进行分析,要分析背后的机理还需要专门领域的知识。

第四节 统计学的一些基本概念

统计学中有许多概念,最常见的、基本的概念有:总体与总体单位,样本,标志与统计指标,统计量,变量,这些概念在统计学习中经常用到,所以,需要准确理解这些基本概念。

一、总体和总体单位

总体是所关注研究对象的全体,即根据一定研究目的所包含全部研究个体的集合。凡是客观存在的,在某些相同性质基础上结合起来的许多个别事物的全体,都可以称为总

体。例如,要研究我国铁路客运列车情况,全国所有的铁路客运列车就是总体,它是由许多铁路客运列车次数组成,而且每次列车都是铁路客运的基本单位,其承担的职能是相同的,具有同质性。

组成总体的个体,称为总体单位,在数理统计中称为样本点。例如上面提到的铁路客运列车,它们聚集在一起,就构成全国铁路客运列车这个总体,每次铁路客运列车即为总体单位。总体单位的总个数,称为总体单位数,常用字母 N 表示。根据总体中包含的单位数的范围可以将统计总体分为有限总体和无限总体两种。有限总体是指总体包含的单位数是有限的。无限总体是指总体范围不能明确确定的。数理统计中的总体,一般是无限的。

当我们进行抽样调查或研究时,统计总体就称为全集总体,简称总体。总体与总体单位,不是固定不变的,而是随着研究的目的与任务的不同,以及地位的不同变换位置。例如如果我们的研究任务变了,要研究我国铁路客运列车的发车准时性,则全国所有的铁路客运列车的发车时间成为统计总体,而每次铁路客运列车的发车时间就是总体单位。

综上所述,统计总体和总体单位的确定,取决于统计研究的目的和要求。而形成统计总体的必要条件,亦即统计总体必须具备的三个特性:大量性、同质性和变异性。统计总体一般是由许多总体单位组成的,这就是所谓总体的大量性。总体的同质性是其中的各个单位在某个方面或某几个方面具有共同的性质。此外,总体中各个单位除了具有某种或某些共同的性质以外,在其他方面则各不相同,具有质的差别和量的差别,这种差别称为变异或差异。正因为有变异才用到统计,没有变异也就用不到统计了。

二、样本

统计研究是为了能够确定总体的某些特征,但是当总体的数量单位较多,或者某些研究的总体本来就是无限总体时,对总体进行逐个研究就显得不必要或者不可能了,所以,采取一定方式,用从总体抽取的部分单位代表总体进行研究。从总体中抽取的部分单位组成的集合称为样本总体,简称样本。样本都是有限总体,其单位数(样本容量)常用字母 n 表示。从我国铁路客运列车中挑选出 100 条有代表性的车次,这 100 条车次就构成了样本,可以通过这 100 条列车车次的情况来推断全国的铁路客运列车情况。

三、标志与统计指标

表明总体单位特征的概念或名称,称为标志。例如,每条铁路客运线路的起点、终点、票价、运行速度等不同的特征,都是每次列车(总体单位)的标志。标志按其特征的性质的不同,可以分为品质标志与数量标志。品质标志是表明总体单位属性的特征,如每条铁路客运线路的起点、终点等都属于品质标志。而数量标志是表明总体单位数量的特征,每条铁路客运线路的票价、运行速度都属于数量标志。数量标志亦称为变量,其具体数值又叫标志值或变量值。

统计指标是说明统计总体特征的数量表现。它反映某现象总体的质的规定性与量的

确定性。这种数量表现可以是表明总体单位属性特征的单位数加总,这种指标也称为质量指标,其也可以是总体单位的标志值加总,这种指标也称为数量指标。不管是数量指标还是质量指标,都是用数量表现的。统计指标,特别是总量指标,一般都包括五个部分,即指标的名称、时间、范围(空间)、指标数字以及相应的计量单位。例如,2018 年我国铁路完成旅客发送量 33.7 亿人次,这在一定程度上就反映了我国铁路客运运输的基本情况。

统计指标与标志两者既有区别,又有密切的联系。它们的主要区别如下。

(1) 统计指标是说明总体特征的,而统计标志是说明总体单位特征的。

(2) 统计指标都是用数量表示的,而统计标志可以用数量表示(即数量标志),也可以不用数量表示(即品质标志)。

(3) 统计指标是许多个体现象的数量综合的结果,而统计标志未经过任何综合,只代表某一个具体现象。

统计指标与标志两者又有密切的联系。首先,统计指标是建立在统计标志的基础上的。其次,它们可以随着研究的任务与目的不同而互相变换关系。例如,如果我们的任务是研究我国铁路客运列车情况,则每次铁路客运列车的客运量、营业里程、准时性等是总体单位(即每次列车)的标志,客运量、营业里程这些标志值汇总起来的铁路客运列车的客运量、营业里程,就是统计指标。但是,由于研究的目的不同,所以统计指标和标志也都是相对的,如果研究某条铁路客运线路的情况,那么该条铁路线路就成了总体,该线路的客运量、营业里程、准时性就是统计指标。

从不同角度反映统计总体特征并且相互联系的统计指标构成的整体,在统计中称为统计指标体系。统计指标体系根据其反映的对象的不同,可分为基本统计指标体系和专业统计指标体系;根据其内容的不同,可分为社会、经济、科技的统计指标体系;根据其功能的不同,又可分为描述、评价、监测、决策的统计指标体系。上述各种统计指标体系是各自独立、自成系统的,但它们之间又是有联系、相互交叉的,各自从不同的方面反映有关的情况。

四、统计量

标志是表示总体单位特征的概念或名称,统计指标是说明统计总体特征的数量表现,通常是未知的。对于描述样本特征的数量表现,称为统计量。为了研究的方便,通过一定的抽样方法抽取样本,对于这个样本,它的统计量都是可以知道的。选取样本的目的就是运用样本统计量来估计总体参数。

对于总体、样本、总体单位、标志、统计指标及统计量的关系可以用图 1-2 来表示。

五、变量

在一个总体中,不管是品质标志或数量标志,当某标志在每个总体单位上的具体表现都相同时,称此标志为不变标志。当某标志在每个总体单位的具体表现不同时,称为可变标志。变量泛指一切可变标志,既包括可变的数量标志,也包括可变的品质标志。变量的

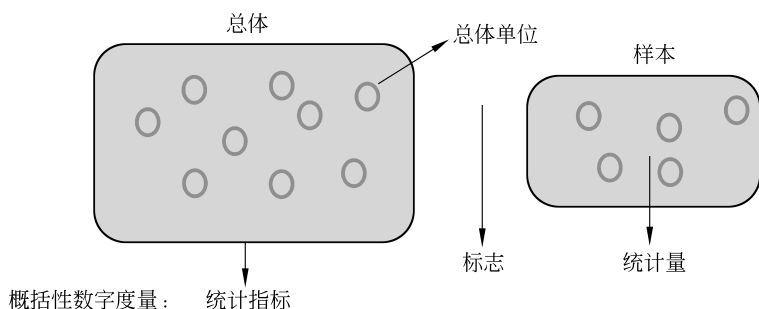


图 1-2 总体、样本、总体单位、标志、统计指标及统计量的关系

具体数值称为变量值,也称标志值。例如,每次铁路客运线路的票价是不同的,票价就是一个变量,票价有可能为 1 748 元、146 元、87 元。这些具体的数值,就是(数量)标志值,也叫变量值。

按照变量的性质,变量主要可以分为分类变量、顺序变量以及数量变量几种类型。用分类型数据表示的变量称为分类变量,我国铁路客运列车车次类型分为高铁、动车、直达、特快、快速等,这就是分类变量。顺序变量是用顺序类型数据表示的变量,高铁铁路客运列车座位等级就是一个顺序变量,其变量值可以为商务座、特等座、一等座、二等座。数量变量是用数量型数据表示的变量。数量变量按其值是否连续出现,可以分为离散变量与连续变量。离散变量的变量值只能取整数表示,如票价、客运量等。连续变量的变量值是连续的,相邻两个值之间可以取任意多个值。例如列车运行时速等,它们的计量可以取到小数点以后的任何位数,但为了便于统计资料的整理,一般的总量指标都取整数或取到小数点后面 2 位或 3 位,对于用百分数、千分数表示的相对数,一般最多取到小数点后面 1 位或 2 位。当然也有例外,如银行存贷款利率、不同币种的兑换率等常取到小数点后 4 位,甚至更多位。

变量按其性质的不同,可以分为确定性变量与随机性变量。所谓确定性变量,是指其变化方向及结果变量值是可以确定的。例如函数关系 $y=8x$ 中的变量 y ,它与变量 x 的变化方向是一致的,而且其结果是变量 x 的 8 倍,即如果变量 x 取值 2,变量 y 就是确定值 16。所谓随机性变量,是指它的取值是不确定的。例如农作物的收获量与其施肥量的关系,一般地说,当施肥量增加时,其收获量也增加,但增加多少,是不确定的,这个收获量变量是随机性变量。但如果就变量 x 和变量 y 说,如果变量 x 限定在 1、2、3、4、5、6 这 6 个整数,如同一颗骰子的 6 个点,变量 y 是变量 x 的 8 倍数,变量 x 的取值由掷骰子决定,则变量 x 的取值是随机的,即为随机性变量,此时变量 y 的取值也是随机的。在概率统计中有:随机性变量的函数仍是随机性变量。随机性变量的取值带有偶然性,在统计学中是一个很重要的概念。

第五节 统计软件

一、Excel 软件

Microsoft Excel 是微软公司的办公软件 Microsoft Office 的组件之一,是由 Microsoft 为 Windows 和 Apple Macintosh 操作系统的电脑而编写和运行的一款试算表软件。它可以进行各种数据的处理、统计分析和辅助决策操作。从 1982 年只用于 Mac 系统的 Excel 1.01 版本问世以来,微软公司不断在原来的基础上更新完善 Excel 功能,发展到目前已经到了 Excel 2019 版本,功能越来越丰富,使用也越来越方便,目前广泛地应用于管理、统计财经、金融等众多领域。

1. 数据整理功能

在 Excel 2010 版本中,每个工作表由 1 048 576 行、16 384 列组成。行号(如 1、2、3,范围:1~1 048 576)和列标(如 A、B、C,范围:A 至 XFD)分别用数字和字母表示。在 2007 版本之前的一个 Excel 工作簿(book)可显示 255 个工作表(sheet),每一工作表由 65 536 行 $\times$ 256 列的表格组成。每张工作表最多可容纳 65 536 个观察个体、256 个变量。

Excel 在数据整理方面,具有自动填充、数据编辑、数据查询、筛选、排序等功能,方便处理,并且其他软件如 SPSS、Eviews 等统计软件都可以读取 Excel 的文件数据。所以,统计软件所用到的数据可以通过 Excel 录入或者进行一些简单的处理,从而方便统计软件数据的整理及录入。

在数据处理方面,Excel 还拥有丰富的图形类型可供用户选择,它提供了十几种图表类型,包括柱形图、条形图、折线图、饼图、散点图等,用户可以根据自己的需求自行设置图表格式。通过图表,使用者不仅能够丰富数据的表达形式,同时图表表示也可以使得数据之间的关系、发展趋势等较为直观地表现出来,如图 1-3 所示。

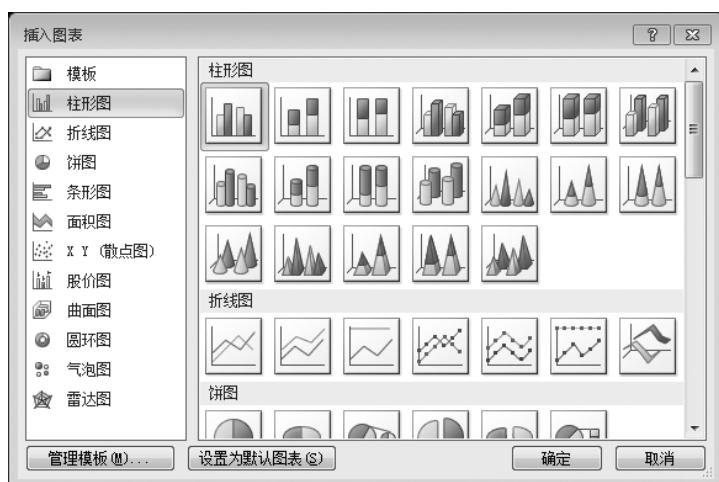


图 1-3 Excel 插入图表界面

2. 数据函数功能

函数作为 Excel 处理数据的一个最重要手段,功能是十分强大的,通过函数功能能够对数据进行自动处理和计算,在生活和工作实践中可以有多种应用,将各个领域经常用的一些计算使用函数表示,可以方便地进行数据处理和分析,甚至可以用 Excel 来设计复杂的统计管理表格或者小型的数据库系统。

Excel 内置函数有近 400 个,包括数学与三角函数、统计函数、日期与时间函数、文本函数、逻辑函数、查找与引用函数、数据库函数、信息函数、工程函数、财务函数、用户定义函数 11 类,如图 1-4 所示。

通过函数和运算符,可以自行设置 Excel 公式,进行统计计算。在建立公式时,首先应该输入“=”,表示编辑公式的开始,如图 1-5 所示。



图 1-4 Excel 插入函数界面

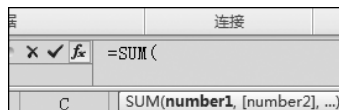


图 1-5 Excel 编辑栏输入函数界面

一般可使用下列操作步骤建立公式。

- (1) 将光标定位在要建立公式的单元格。
- (2) 输入“=”,或者如果直接利用 Excel 内置函数建立公式,可在编辑栏左侧单击插入函数“fx”按钮。
- (3) 输入公式的内容。如果输入的是内置函数,如求和函数 SUM,Excel 会自动弹出函数的格式文本框,提示使用者正确的函数输入格式。
- (4) 完成公式的编辑后,可按 Enter 键或单击编辑栏左侧的“√”按钮,表示确定;如果要取消编辑的公式,可单击编辑栏的“×”按钮。

3. 数据分析功能

Excel 除了能够进行数据的整理以及相关的运算之外,还具有一项强大的功能,即进行数据分析。通过数据分析,可以利用数据分析工具简单地完成常用的统计学分析,如数据的统计学描述、制作直方图、进行 t 检验、方差分析、相关与回归分析等等。

Excel 提供了一组数据分析工具,称为“分析工具库”。只需为每一个分析工具提供必要的数据和参数,该工具就会使用适宜的统计或数学函数,在输出表格中显示相应的结

果。其中的一些工具在生成输出表格时还能同时产生图表。

按照 Excel 2010 版本,“数据分析”在“数据”菜单下,“数据分析”的界面如图 1-6 所示。

在之前的 Excel 版本中,“工具”菜单下一般没有“数据分析”这一选项,需要自行进行安装。

安装“数据分析”的步骤如下。

(1) 单击“工具→加载宏”,弹出加载宏界面。

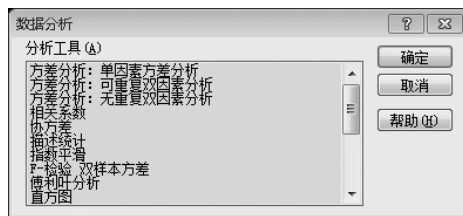


图 1-6 “数据分析”的界面

(2) 选取“分析工具库”等,根据屏幕提示进行安装。安装完“加载宏”后,打开“工具”菜单,即可发现“数据分析”选项。

此外,在数据分析中,数据透视图也是一种比较实用的工具,通过它可以较快地将数据呈现到表格或者图形中去。Excel 还可以进行辅助决策,如单变量求解、模拟运算器等,这些功能的设置都大大地方便了使用者,促进了有效的决策。

二、R 软件

R 是属于 GNU 系统的一个自由、免费、源代码开放的软件。R 是 S 语言的一种实现。S 语言是由 AT&T(美国电话电报公司)贝尔实验室开发的一种用来进行数据探索、统计分析、作图的解释型语言。最初 S 语言的实现版本主要是 S-PLUS。S-PLUS 是一个商业软件,它基于 S 语言,并由 MathSoft 公司的统计科学部进一步完善。后来 Auckland 大学的 Robert Gentleman 和 Ross Ihaka 及其他志愿人员开发了一个 R 系统。R 的使用与 S-PLUS 有很多类似之处,两个软件有一定的兼容性。S-PLUS 的使用手册,只要经过不多的修改就能成为 R 的使用手册。所以有人说:R 是 S-PLUS 的一个“克隆”。

R 是一套完整的数据处理、计算和制图软件系统。其功能包括:数据存储和处理系统;数组运算工具(其向量、矩阵运算方面功能尤其强大);完整连贯的统计分析工具;优秀的统计制图功能;简便而强大的编程语言:可操纵数据的输入和输出,可实现分支、循环,用户可自定义功能。

与其说 R 是一种统计软件,还不如说 R 是一种数学计算的环境,因为 R 并不是仅仅提供若干统计程序,使用者只需指定数据库和若干参数便可进行一个统计分析。R 的思想是:它不仅可以提供一些集成的统计工具,而且可以提供各种数学计算、统计计算的函数,从而使使用者能灵活机动地进行数据分析,甚至创造出符合需要的新的统计计算方法。

该语言的语法表面上类似 C,但在语义上是函数设计语言(functional programming language)的变种并且和 Lisp 以及 APL 有很强的兼容性。特别的是,它允许在“语言上计算”(computing on the language)。这使得它可以把表达式作为函数的输入参数,而这种做法对统计模拟和绘图非常有用。R 软件界面如图 1-7 所示。

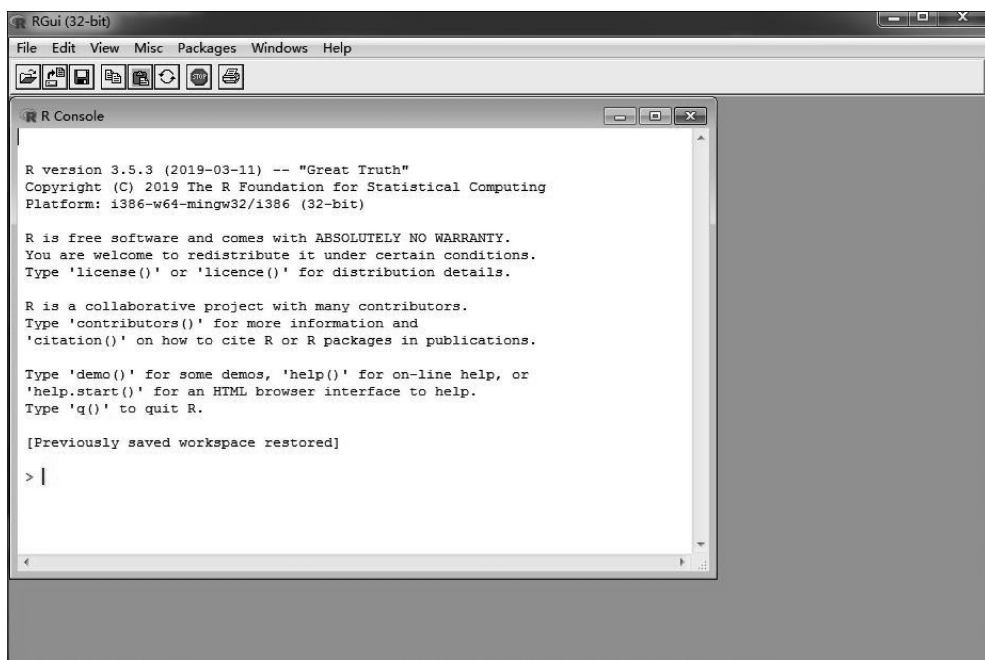


图 1-7 R 软件界面

R 是一个免费的自由软件,它有 Unix、Linux、MacOS 和 Windows 版本,都是可以免费下载和使用的。在 R 主页可以下载到 R 的安装程序、各种外挂程序和文档。在 R 的安装程序中只包含了 8 个基础模块,其他外在模块可以通过 CRAN(综合 R 档案网络)获得。

R 软件具有以下特点。

(1) 免费性。R 是自由开源软件。这意味着它是完全免费,开放源代码的。可以在它的网站及其镜像中下载任何有关的安装程序、源代码、程序包及其源代码、文档资料。标准的安装文件自身就带有许多模块和内嵌统计函数,安装好后可以直接实现许多常用的统计功能。

(2) 制图功能强。在可视化方面,R 的效果特别好,输出的图形可以直接保存为 JPG、BMP、PNG 等图片格式,还可以直接保存为 PDF 文件。

(3) 统计分析能力突出。R 内嵌了许多统计分析函数,一些中间结果既可保存到专门的文件,也可直接用于进一步的分析。

(4) 交互性强。除了图形输出是在另外的窗口处,它的输入输出都是在同一个窗口进行的,输入语法中如果出现错误会马上在窗口中得到提示,对以前输入过的命令有记忆功能,可以随时再现、编辑修改以满足用户的需要。

(5) 更新性。所有 R 的函数和数据集是保存在程序包里面的。只有当一个包被载入时,它的内容才可以被访问。一些常用、基本的程序包已经被收入了标准安装文件中,随着新的统计分析方法的出现,标准安装文件中所包含的程序包也随着版本的更新而不断

变化。在安装文件中,已经包含的程序包有: base——R 的基础模块、mle——极大似然估计模块、ts——时间序列分析模块、mva——多元统计分析模块、survival——生存分析模块等。

(6) 帮助功能良好。可通过 help 命令随时了解 R 所提供的各类函数的使用方法或例子,便于用户学习和使用。

思考与练习

1. 试述统计与统计学的含义。
2. 试述统计工作、统计资料与统计学之间的关系。
3. 试述总体、样本、总体单位三者的概念。
4. 试述标志、统计指标以及统计量的概念。

5. 为了调查 2018 年我国民营企业的基本情况,选取了目前在上海证券交易所和深圳证券交易所上市交易的 A 股 12 家公司的情况,所调查公司的基本情况如表 1-2 所示。数据来源于国泰安数据库。

- (1) 该次研究的总体、样本以及总体单位分别是什么?
- (2) 该次研究包含了哪几个变量?



习题 5 数据

即测即练



表 1-2 所选取的 12 家民营公司基本情况

公司中文全称	员工人数 /人	流动资产/ 百万元	流动负债/ 百万元	总资产/ 百万元	股东权益(不含 少数股东权益)/ 百万元	营业收入/ 百万元	净利润/ 百万元	净资产 收益率/元	每股收益/ 元
信雅达系统工程股份有限公司	6 891	11 881.65	4 540.64	16 120.77	11 383.33	12 228.50	-45.99	0.00	-0.01
浙江康恩贝制药股份有限公司	9 334	47 541.74	35 419.78	107 134.05	57 925.87	67 866.45	8 083.39	0.14	0.30
安徽皖江物流(集团)股份有限公司	8 048	34 764.20	36 086.27	167 772.83	101 393.36	110 879.46	5 698.25	0.06	0.15
浙江祥源文化股份有限公司	200	9 703.67	1 710.94	20 627.44	18 687.24	7 205.86	127.29	0.01	0.02
铜陵精达特种电磁线股份有限公司	3 114	37 933.77	18 792.79	56 240.21	36 879.56	118 977.96	5 149.89	0.14	0.26
卧龙电气集团股份有限公司	14 887	96 080.60	91 073.63	184 546.54	64 686.29	110 760.33	6 880.22	0.11	0.53
江苏长电科技股份有限公司	23 600	109 053.81	184 048.77	344 274.01	122 951.64	238 564.87	-9 266.40	-0.08	-0.58
山东金晶科技股份有限公司	3 157	38 616.32	48 629.53	96 546.78	42 781.23	51 246.03	840.99	0.02	0.06
用友网络科技股份有限公司	16 079	83 401.06	73 481.28	152 208.97	76 519.63	77 034.95	8 101.87	0.11	0.42
广东榕泰实业股份有限公司	733	33 306.73	19 982.57	54 049.21	31 878.73	17 001.06	1 535.85	0.05	0.22
泰豪科技股份有限公司	3 267	91 887.40	67 761.18	131 459.33	43 829.06	61 273.85	2 741.64	0.06	0.32
贵州益佰制药股份有限公司	8 977	31 209.99	19 835.40	64 755.68	39 105.94	38 828.63	-6 844.87	-0.18	-0.86