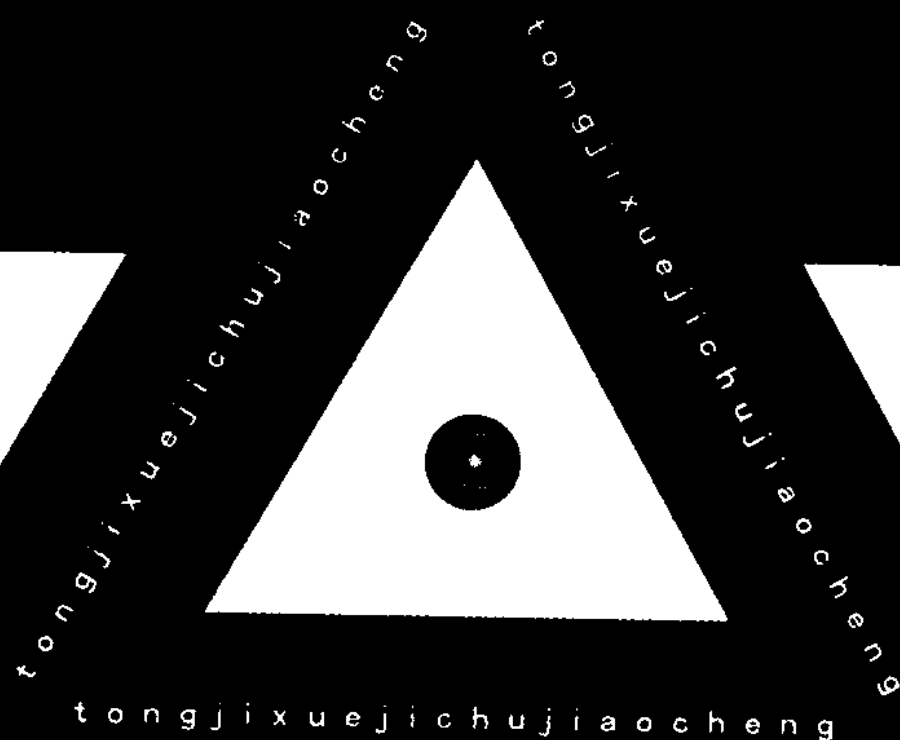


范义正 江华 编著

统计学基础教程

山西经济出版社



责 编:李慧平
复 审:宋晋平
终 审:赵建廷

图书在版编目(CIP)数据

统计学基础教程/范文正,江华编著. —太原:山西
经济出版社,2002.4
ISBN 7-80636-594-X

I.统... II.①范...②江... III.统计学-教材
IV.C8

中国版本图书馆 CIP 数据核字(2002)第 015230 号

统计学基础教程

范文正 江华 编著

*

山西经济出版社出版发行

030012 太原市建设南路 15 号 0351-4922102

<http://www.sxep.com.cn> E-mail: sxep@sx.cei.gov.cn

新华书店经销 山西科林印刷有限公司印刷

*

开本:850×1168 1/32 印张:8.625 字数:211 千字

2002 年 4 月第 1 版 2002 年 4 月太原第 1 次印刷

印数:1-4000 册

*

ISBN 7-80636-594-X

F·557 定价:15.00 元

目 录

第一章 导 言.....	(1)
第一节 事物的数量特征.....	(2)
第二节 统计学的源流.....	(7)
第三节 统计方法的基本思路	(11)
第四节 统计学的性质	(18)
第二章 统计资料的搜集与整理	(24)
第一节 统计资料及其搜集方法	(24)
第二节 调查方式与调查方案	(27)
第三节 调查误差	(33)
第四节 统计数据的整理	(37)
第三章 统计数据的描述与显示	(54)
第一节 绝对指标与相对指标	(54)
第二节 集中趋势的测定	(57)
第三节 离散程度的测定	(67)
第四节 探索性数据分析	(75)
第五节 统计数据的显示	(82)

第四章 抽样调查与推断	(92)
第一节 抽样调查与推断的基本问题	(93)
第二节 抽样误差	(99)
第三节 参数估计	(108)
第四节 抽样调查的组织形式	(115)
第五章 回归分析	(121)
第一节 变量间的相关关系	(121)
第二节 简单线性相关分析	(126)
第三节 一元线性回归分析	(132)
第四节 多元线性回归分析	(140)
第六章 时间序列分析	(147)
第一节 时间数列的描述方法	(147)
第二节 时间数列的因素分解	(153)
第三节 长期趋势分析	(156)
第四节 季节变动分析	(166)
第五节 周期波动分析	(170)
第七章 统计指数	(174)
第一节 指数的意义	(174)
第二节 指数的编制方法	(178)
第三节 指数体系与因素分析	(184)
第四节 常用指数举例	(188)
第八章 统计质量管理	(196)
第一节 产品质量抽样检验	(196)

第二节 统计过程控制.....	(203)
第九章 统计决策.....	(219)
第一节 统计决策的意义、种类和原则	(219)
第二节 确定统计决策方法.....	(223)
第三节 风险型决策方法.....	(228)
第四节 统计决策例析.....	(244)
各章主要知识点.....	(253)
附录.....	(256)
一、希腊字母表	(256)
二、统计报表举例	(257)
三、随机数表(局部)	(259)
四、正态分布概率表	(261)
五、参考及推荐阅读书目	(263)
后记.....	(265)

目 录

第一章 导 言	(1)
第一节 事物的数量特征	(2)
第二节 统计学的源流	(7)
第三节 统计方法的基本思路	(11)
第四节 统计学的性质	(18)
第二章 统计资料的搜集与整理	(24)
第一节 统计资料及其搜集方法	(24)
第二节 调查方式与调查方案	(27)
第三节 调查误差	(33)
第四节 统计数据的整理	(37)
第三章 统计数据的描述与显示	(54)
第一节 绝对指标与相对指标	(54)
第二节 集中趋势的测定	(57)
第三节 离散程度的测定	(67)
第四节 探索性数据分析	(75)
第五节 统计数据的显示	(82)

第四章 抽样调查与推断	(92)
第一节 抽样调查与推断的基本问题	(93)
第二节 抽样误差	(99)
第三节 参数估计	(108)
第四节 抽样调查的组织形式	(115)
第五章 回归分析	(121)
第一节 变量间的相关关系	(121)
第二节 简单线性相关分析	(126)
第三节 一元线性回归分析	(132)
第四节 多元线性回归分析	(140)
第六章 时间序列分析	(147)
第一节 时间数列的描述方法	(147)
第二节 时间数列的因素分解	(153)
第三节 长期趋势分析	(156)
第四节 季节变动分析	(166)
第五节 周期波动分析	(170)
第七章 统计指数	(174)
第一节 指数的意义	(174)
第二节 指数的编制方法	(178)
第三节 指数体系与因素分析	(184)
第四节 常用指数举例	(188)
第八章 统计质量管理	(196)
第一节 产品质量抽样检验	(196)

第二节 统计过程控制·····	(203)
第九章 统计决策·····	(219)
第一节 统计决策的意义、种类和原则 ·····	(219)
第二节 确定统计决策方法·····	(223)
第三节 风险型决策方法·····	(228)
第四节 统计决策例析·····	(244)
各章主要知识点·····	(253)
附录·····	(256)
一、希腊字母表 ·····	(256)
二、统计报表举例 ·····	(257)
三、随机数表(局部) ·····	(259)
四、正态分布概率表 ·····	(261)
五、参考及推荐阅读书目 ·····	(263)
后记·····	(265)

第一章 导 言

在介绍统计学知识之前,先请读者回忆一些经常接触到的反映现实情况的数字,如人口数、产品销售量、品牌知名度、天气预报等,它们通过对某些现象数量特征的描述使我们了解了一些事物,我们可以把这些数量信息称为统计数据。再请读者回忆一些“规律”,如“吸烟有害健康”、“着凉后容易感冒”、“恩格尔系数(食品支出占家庭总支出的比重)低于 25% 的家庭生活很富裕”,它们描述了一些“关系”——我们常把这样比较稳定的“关系”视为“规律”,这些“规律”是存在的,但并不十分准确。请读者从这两类现象中总结一些你认为是共同的东西。

我们每天都要接触许多数字或数量化的现象,有些是日常生活中反映“小”问题的数字,如去某处要花多少时间、今天会不会下雨,有些是反映一些“大”现象的数字,如股票指数、失业率,这些现象都与统计学有一定关系。人们从对数量的感知、计数开始,发展到现在的数学、统计学等学科,都是对数量关系、数量规律的认识与探索。

统计学是一门方法论或工具性学科,从方法上看,它是收集、整理、描述、显示和分析统计数据的科学;从功能上看,它是对大量个别现象的数量特征进行较可靠的概括性判断的科学。

许多人简单地认为统计就是搜集数字,加加减减,形成诸如“旅游人数”、“客运量”之类的统计指标,其实这只是较规范的计数活动,是千百年来人类计数活动的延续与扩展。统计学无论在方

法和理论上还是应用范围上,已远远超出了这个范围,它已经发展成为广泛应用于自然科学、政治法律、经营管理、社会生活等广阔领域的科学而有效的分析方法。在最近的一百年内,由于统计学有了很大的进步,又由于统计学的广泛运用,一定程度上使物理学、生物学、医学、经济学、管理学、人类学、气象学、军事科学、语言学等许多学科都获得了实质性的进步。在色彩斑斓的现代社会,统计学是不可或缺的思维与分析利器。

本书介绍统计学的基本方法和内容,阐述一些基本统计思想。多数统计方法看起来很平易,只涉及加减乘除、开方等简单运算,但其思想方法,不仅与数学不同,且在一定程度上,与一些日常生活养成的思考方式也不同。掌握这些统计思想和统计方法有助于我们在看待事物时具有“大局观”,在叙述时能正确地使用统计数字,正确地理解和分析别人所作的统计表述,或者帮助自己阐明观点;同时,统计方法的合理运用,能使我们区分“直觉”和客观的见识。也许,由于统计学自身的特点,它比其他课程更宝贵的是,统计学能帮助我们用较明确而可靠的依据评价自己的直觉印象和简单判断。

需要提前指出的是,如同其他课程一样,统计方法绝不是万能的,它不能代替其他课程,不能代替你自己的常识和经验,也不能代替良好的经营意识。或者说,统计学有很强的“技术”特色,对你性格的形成不会有很大的帮助和影响,但它能使你目光锐利、思虑周全。

第一节 事物的数量特征

一、认识数量的习惯方法

人们对世界的认识是逐步加深和拓宽的,认识逐渐丰富的过

程,实际上是对“差异”——可以是属性上的差异,也可以是数量上的差异——了解更多的过程。

人们在观察自然现象的同时,自觉或不自觉地发现了事物之间的差异,有些是性质上的明显不同,如一棵树和一条蛇,一座山和一条河,又在有些相似或相同的事物中发现了数量的差异,如一棵小树和一棵大树。不同的物种表现质的差异,不同的大小表现量的差异,这些都需要人们用不同的方法去表达、标记它们,也就是说,不同的事物有不同的性质和表现,为了反映这种表现上的差异性,人们找到了许多方法,其中把同类事物之间的差异用不同的数量来表现是普遍而有效的方法,这就是量化的方法。

度量是人的主观行为,他肯定要用自己习惯的器物和方法去度量,“人是万物的尺度”。人们总是用自己习惯的事物为单位来度量自己所接触的物体。人们对数量的认识习惯主要有以下特征:

1. 计量单位与计量结果简单化

计量单位是通用的标准,如市制的尺在中国古代是手掌尽力伸开时拇指尖到中指尖的距离,英尺的原文(foot)就是脚,源自古巴比伦时人们用成年男子的脚掌来计量尺度,英制的码是一位英国国王的鼻尖到大拇指的距离,都是以人体上的某种距离为度量标准,这种标准的好处是“随身带着准尺”,使用方便。与计量单位相对的是计量结果,比较简单的结果数目更容易使用,如计量身高可用尺或码,计量路程却需要使用里。

2. 以平均数为中心形成一个数量区间,但认识的精度不高

如在描述身高时,通常说“一米七左右”,多数情况下,精确到厘米就够了(使用英制单位,一般都精确到英寸。1英寸=2.54厘米),没有人说“一到二米”或“1.69852米”。这与人的肉眼观察能力有一定关系:肉眼能在近距离分辨1.6米和1.7米的区别,所以不能把身高简化为“一到二米”,但很难分辨1.695米和1.705米,

所以也不必细致到“1.69852 米”。其精度以人的认识能力和经验为尺度。相同的道理,如果某篇新闻报道说市民中有 70% 的人同意某一种意见,人们一般都会把这一数字还原为 65% ~ 75%,或者 60% ~ 80%,但很少会有人认为是 69% ~ 71%,或者 40% ~ 100%。

3. 数量常被当做间隔较大的离散型变量

人们认识数量往往只认识其大略,最简单的是按照程度副词来排列级别,如很、极、几乎、完全等,较复杂一点的约在二个数量级以内,如七十六、四百八、五万二等,只是在较特殊的场合才使用精确数字。这符合人们的认识习惯,对大脑资源的占用也是较为经济的。

4. 量化的方法由粗到精,并形成了一些常用的度量尺度

对现象进行数量化,根本的目的是体现、表示事物的差异及其顺序,而且量化的过程也是先易后难,逐步精细化。

(1)有些事物因为种属的不同,人们可以赋以概念或名称来定义,如动物与植物、树与草、柳树与松树,这种方法能够形成“列名尺度”(Nominal Scale),可以区别事物,但无法确定事物之间数量的差异,用数学语言说,就是只有相等和不相等的关系,而没有大于和小于的关系。比如,人口的民族特征、性别特征等,都只能用文字来表述,各民族、各性别之间只有性质的差异,而不存在数量上的比较关系。

(2)反映数量差异较粗陋的方法是用较明显的差异来反映顺序,如多和少,又如小孩、成人、老人,另有一类事例也可以说明这种方法的原始性:南部非洲布须曼人的语言,除了“一”、“二”和“多”外,再没有别的数字,英文的 thrice 和拉丁文的 ter,同样有双重意义——三倍和许多,拉丁文的 tres(三)和 trans(超过)之间有着可信的联系。这种方法可以形成“顺序尺度”(Ordinal Scale),可以按照某种数量对几类事物进行排序,有大于和小于的比较,但无

法准确描述数量,也就不能进行算术运算。比如学生成绩的优、良、中、及格、不及格,人口按照儿童、少年、青年、壮年、老年进行划分等。

(3)比较精确的方法是“间隔尺度”(Interval Scale),基本上有两种思路,第一种方法是把某一现象的最大差异范围或一个已知的差异范围等分为若干等份,由事物不同的状态体现不同的差异水平,如摄氏温度是把常压下液态水温度的区间定义为等分的100度,形成测量温度的标准;第二种方法主要针对不知其最大差异范围或难以划分界限的现象,人为规定某个尺度,使之通行以便进行数量信息的交流,如尺、千克、加仑、光年等度量衡单位。间隔尺度的计量结果可以对数量进行运算,从而能够进一步探讨数量规律。

统计学中,可以把以上三种尺度所形成的具体表现——如性别的男女、服装的型号、商品的销售量等——同等地视为“数据”,因为它们都是某些现象的特征表现,当然,在进行统计分析时,采用的方法会有所不同。

能够看到,在人们日常对数量现象和特征进行观察时,存在着一个习惯性的特点——模糊,虽然有其合理性,但也有很大的局限性——不够精确,不利于发掘深层的信息。固然,我们并不需要时时处处掌握精确而详细的数量信息,这与人使用大脑的效率有密切关系。但我们还应看到:仅仅具有模糊的数量认识是不够的,模糊的数量认识导致模糊的数量界限,不利于我们深入、严格地探讨事物的关系。

二、质与量规定性的关系

质与量是事物的两种基本规定性,二者既相互联系,又相区别,共同对事物的存在做出基本的规定。质的规定性具有惟一性,是只有该事物才特有的,而不是数种事物所共有,一物一质是事物

存在的根本格局。量是质的表达方式,量是质的等级、规模、范围和结构的表现,是事物可以由数和形来表示的规定性。比如,少年与青年的区别就是从年龄、生理和心理特征等内容上划分的。质的变化其实就是事物通过自身的数量关系和数量变化来重新表达质,使新的质产生了变化,不同于以前的质。黑格尔认为,质变是“渐进过程的中断”。从质与量的哲学关系可以看出,量变是时刻不停的,我们总是置身于无止无休的量变之中。

由于不同的质具有明显的区别,我们更多的是靠事物的质的差异来区分各种事物,也就是说,质变或质的差异更容易被观察到,也更容易受到重视,形成鲜明的印象,相应地,量的变化容易被忽视,或者习焉不察。在现实生活中,人们是一分钟一分钟地生活,而不是几年几年地生活。我们的认识过程经常是:每天面对的都是数量,多或少,正或负,总是在观察、应对着或大或小的数量变化,量变悄然积累许多之后,我们偶然回首,蓦然发觉已今非昔比。

三、归纳逻辑在数量方面的运用

既然数量的表现和变化充斥时时处处,人们就会通过对数量关系的长期观察而积累大量的经验,进而从中发现一定的规律性。这就是由特称判断推出全称判断的归纳逻辑,即从个别性的前提推出一般性的结论,是对经验事实的概括。我国有许多农谚就是在大量观察之后总结出来的,如北方地区有“瑞雪兆丰年”,“雪”与“丰”是两种不同的现象,长期的观察,使人们发现两者有比较紧密的联系——如果正月里下雪,当年丰收的可能性很大,对这一现象的总结逐渐传为谚语。其实这也符合客观的道理——降雪既可以使土壤保持良好的墒情,也可杀灭有害的细菌和虫卵,保证春夏季农作物的正常生长;渔民有“无风来长浪,不久狂风降”的谚语,它说的现象是:如果在海上看到“长浪”,很快会有大风吹来。它的道理是:海上风暴引起的巨浪,传到风力平静的海区,因受空气的阻

力,波高(振幅)减低,波长变长,形成“长浪”,这种波浪传播的速度,比风暴中心移动的速度要快两三倍,这对海上渔民可以起到预警的作用。统计方法对事物之间统计规律的探索机制与此相似。

第二节 统计学的源流

史前人类鸿蒙初辟,计数活动便已产生,如结绳记事、刻符记事,为后世许多与数量有关的学科之滥觞。有史料记载的早期人类文明,大量记录了一些可视为统计方法萌芽的计数活动。如《史记·夏本纪》:“禹平水土,定九州,计民数”;约五千年前古埃及法老每两年派人清查一次全国人口;古巴比伦在四五千年前就有清点人数的活动。随着文明的进步,生产能力提高,社会结构日趋复杂,各种人际交往大量增加,政事日繁,大规模的、制度化的统计活动在各主要文明中持续发展,所涉范围多为人口、税赋、土地、经济生产等,但未尝对统计活动进行系统化的探讨。

作为一门学问的统计学大约产生于 17 世纪中叶。可称为统计学源头的概有三端:国势学、政治算术、概率论。其后各种流派互相砥砺激荡,到 19 世纪上半期,统计活动在许多方面出现新景象,形成统计发展史上的一次高潮,后人称为“统计时代”。20 世纪初叶,数理统计得到重大发展,统计学逐渐成为效验非常的通用方法论学科。

一、国势学

统计学之产生和发展与国家、政治统治是紧密相关的。想对一个系统施加控制或影响,必须对该系统的状况有必要的了解。从这个意义上讲,统计学的萌芽出现较早,如 1536 年德国有《世界志》,1562 年意大利有《国家制度志》,1589 年意大利有《公民、政治统计与军事科学》,1614 年意大利有《世界诸国志》,1656 年德国有

《德意志王国志》等。这些著作的主要内容是概括介绍各国的地理、人口、产业、财富、兵力、典章制度、历史等情况,皆为资治之作。

最早讲授国势学的代表人物是德国的大学教授海尔曼·康令(Herman Conring, 1606-1681),1660年在德国的大学里以“欧洲最近国势学”为题,用对比的方法讲授当时欧洲各国国势,内容包括社会秩序、立宪、行政、人口、土地、财政、国家组织等。该讲座把德国在国际上所处的落后地位凸显出来,迅速引起社会各阶层的巨大反响。同时,他把国情的一般叙述变成一种系统学问的研究,也引起不少学者的兴趣。在康令的研究中,几乎不用数量表达有关现象,只用一些笼统的形容词,如“人口稠密”、“土地广阔”等,这种缺陷直到18世纪才被逐渐克服。

对国势学贡献最大的当属德国的大学教授阿亨瓦尔(Gottfried Achenwall, 1719-1772)。他从1748年起在大学里讲授政治统计课程,最先提出了“统计学”的名称,一生著述甚丰。他把能够左右国家繁荣富强的事项称为“国家显著事项”——“组织、人口、军队、财产、地面和地下资源”,国家显著事项的总和就构成这个国家的内容。阿亨瓦尔认为,统计学的最终目的,在于通过对有关国家重大事项资料的搜集和分析,通过对各国状况的了解和对比,使国家增加财富,增殖人口,发展科学、工业和商业,改善国家组织上的缺陷,其思想与中国春秋时代管仲的“举事必成,不知计数不可”和战国时代商鞅“强国知十三数”的思想在本质上是相同的。

以后,随着社会的进步,在英国传入的政治算术思想的影响之下,为了更好地适应国家管理及经济发展的需要,国势学出现分化,出现了表式学派,用列表的方法研究社会经济现象的数量特征,开始体现出统计学的典型特点。

二、政治算术

政治算术产生于17世纪的英国,代表人物有威廉·配第

(William Petty, 1623 – 1687)和约翰·格朗特(John Graunt, 1620 – 1674)。

威廉·配第禀赋聪慧,做过大学教授、政府土地官员,开办过企业,并且是英国皇家学会的创始人之一。在其名著《政治算术》中,运用数量分析的方法,对比英格兰、法兰西、荷兰、爱尔兰等国的“财富和力量”,为英国称霸提供了富有说服力的实证分析资料,自称“我进行的这项工作所使用的方法,在目前还不是常见的。因为和使用比较级和最高级的词语以及单纯作思维的论证相反,我却采用了这样的方法(作为我很久以来就想建立的政治算术的一个范例),即用数字、重量和尺度的词汇来表达我自己想说的内容,只进行能诉诸人们感官的论证,和考察在性质上有可见的根据的原因”。这一思想是认识方法的重大进步,事物在本质上的差异体现得比较明显,但在等级的交界处容易产生模糊,不利于精细地认识事物。了解大范围(国家)的事物必须有足够的数量认识。配第还在统计的计量方法、图表方法、分组方法、推算方法等方面有卓著的贡献。

约翰·格朗特主要是个商人,1662年出版了《关于死亡表的自然和政治的观察》(Natural and Political Observation upon the Bills of Mortality),引起学术界的特别重视,并成为英国皇家学会会员。他是利用大量数据研究社会人口变动规律的创始人。该书分析了60年间居民死亡的原因及与人口变动的关系,探讨了人口数量的变动规律以及如何处理人口统计资料的方法,其崭新的观点和分析结果改变了许多关于人口现象的传统见解。从统计学的角度看,格朗特发现了生命运动中的大数法则,创制了人口分析的重要工具——生命表,提出了统计人口的推算公式,对后世开展统计推断也有一定的启发。