

.....● 普通高等教育规划教材

S tatistics

统计学

◎ 段淑梅 张纯荣 主编

 机械工业出版社
CHINA MACHINE PRESS



013059280

C8
238

普通高等教育规划教材

统 计 学

主 编 段淑梅 张纯荣
副主编 杨朝丹 肖 琦
参 编 王雪秋 韩 梅 杨 晶



机械工业出版社



北航

C1665607

C8
238

本书以培养学生的实际应用能力为目的,本着理论够用、强化技能的原则组织编写,略去抽象、复杂的数学计算与公式推导,强调处理实际问题的能力。本书架构体系从统计数据出发,以统计数据的收集、整理、分析和解释为主线,由描述统计到推断统计组织编写。全书共九章,包括总论、统计调查、统计整理、综合指标、抽样推断、假设检验、时间序列分析、统计指数以及相关分析与回归分析等内容。在写作方法上,力求简明扼要,深入浅出,理论联系实际。为便于学生学习和掌握书中内容,本书每章章前有本章要点,每章章后有本章小结、思考题和练习题。

本书可作为高等院校经济管理类专业教材,也可为企业事业单位管理人员以及对统计学感兴趣的其他人士提供参考。

图书在版编目(CIP)数据

统计学/段淑梅,张纯荣主编. —北京:机械工业出版社,2013.7

普通高等教育规划教材

ISBN 978-7-111-43042-1

I. ①统… II. ①段…②张… III. ①统计学 - 高等学校 - 教材
IV. ①C8

中国版本图书馆 CIP 数据核字 (2013) 第 136525 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑:曹俊玲 责任编辑:曹俊玲 马碧娟 何 洋

封面设计:张 静 责任印制:乔 宇

北京机工印刷厂印刷 (三河市南杨庄国丰装订厂装订)

2013 年 8 月第 1 版第 1 次印刷

184mm×260mm·13.25 印张·326 千字

标准书号:ISBN 978-7-111-43042-1

定价:26.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

社服务中心:(010)88361066 教材网:<http://www.cmpedu.com>

销售一部:(010)68326294 机工官网:<http://www.cmpbook.com>

销售二部:(010)88379649 机工官博:<http://weibo.com/cmp1952>

读者购书热线:(010)88379203 封面无防伪标均为盗版

前 言

统计学是一门收集、整理、分析和解释统计数据的方法论的学科,其目的是探索统计数据内在的数量规律性。现代社会,人们每时每刻都会接触到大量的、各式各样的数据,离开数据,人们将寸步难行。那么,这些数据从何而来,说明了什么问题,隐含着什么关系和规律?还缺少哪些数据,该如何去获得所需数据以得出所需结论?这些正是统计学要解决的问题。学习统计学,掌握基本的统计理论与方法,具备一定的收集、整理和分析统计数据的能力,是对经济管理类专业学生的基本要求,正因为如此,统计学一直被教育部列为经济类、管理类专业的核心课程之一。

本书从应用型人才培养的视角出发,本着理论够用、强化技能的原则组织编写。与以往的统计学教材相比,本书略去抽象、复杂的数学计算与公式推导,强调实际应用,让读者掌握常用的统计学原理、方法,提高依据实际问题进行统计设计和调查的能力,以及依据统计结果处理实际问题的能力。在写作思路,本书从统计数据出发,以统计数据的收集、整理、分析和解释为主线,由描述统计到推断统计组织全书结构。在写作方法上,力求简明扼要,深入浅出,理论联系实际。为便于学生学习和掌握书中内容,本书每章章前有本章要点,每章章后有本章小结、思考题和练习题。

本书由多所高校长期担任统计学教学工作的教师共同编写完成。具体分工如下:段淑梅编写第一章,并对全书进行审稿、总纂和定稿;张纯荣编写第三章、第四章、第八章,并负责文稿修改和总体架构设计;杨朝丹编写第二章、第九章;肖琦编写第五章、第六章和第七章;王雪秋、韩梅、杨晶对文稿进行了细致的校对。

在本书的编写过程中,我们参阅和借鉴了国内外同行的大量文献,针对大学学生应该掌握的知识和技能进行编写,并且特请长春工业大学许春燕教授给予审稿。在此一并表示衷心的感谢!尽力之余,书中不足之处在所难免,恳请读者多提宝贵意见。

编 者

目 录

前言

第一章 总论	1
本章要点	1
第一节 统计与统计学	1
第二节 统计的产生和发展	6
第三节 统计数据类型	10
第四节 统计学的基本概念	12
本章小结	16
思考题	16
练习题	17
第二章 统计调查	19
本章要点	19
第一节 统计调查概述	19
第二节 统计调查方案	22
第三节 统计报表	25
第四节 专门调查	29
本章小结	33
思考题	33
练习题	33
第三章 统计整理	36
本章要点	36
第一节 统计整理概述	36
第二节 统计分组	38
第三节 次数分布	43
第四节 统计表和统计图	48
本章小结	52
思考题	52
练习题	52
第四章 综合指标	55
本章要点	55
第一节 集中趋势指标	55

第二节 离散程度指标	63
第三节 变量分布状况指标	68
本章小结	70
思考题	71
练习题	71
第五章 抽样推断	74
本章要点	74
第一节 抽样推断的意义和作用	74
第二节 抽样平均误差	77
第三节 总体参数估计	81
第四节 抽样设计	85
第五节 样本容量的确定	91
本章小结	93
思考题	93
练习题	93
第六章 假设检验	95
本章要点	95
第一节 假设检验的一般问题	95
第二节 Z 检验	98
第三节 t 检验	103
第四节 非参数检验	107
本章小结	116
思考题	116
练习题	117
第七章 时间序列分析	119
本章要点	119
第一节 时间序列概述	119
第二节 长期趋势的测定	128
第三节 季节变动的测定	140
第四节 循环变动的测定	142
本章小结	145
思考题	145
练习题	145
第八章 统计指数	147
本章要点	147
第一节 统计指数概述	147



第二节 综合指数	149
第三节 平均数指数	151
第四节 指数体系和因素分析	153
第五节 几种常用的价格指数	156
本章小结	158
思考题	159
练习题	159
第九章 相关分析与回归分析	162
本章要点	162
第一节 相关分析概述	162
第二节 直线相关关系的测定	166
第三节 一元线性回归分析	173
本章小结	186
思考题	187
练习题	187
附录 常用统计表	191
参考文献	205

第一章 总 论

【本章要点】

1. 理解统计的含义、研究对象及其特点。
2. 明确统计工作过程及基本研究方法。
3. 掌握统计数据的类型。
4. 掌握统计学的有关基本概念。

在日常生活中,统计无时不有、无处不在。比如,在外出旅游时,需要关心一段时间内的详细天气预报;在投资股票时,需要了解股票市场价格的信息,了解某只特定股票的有关财务信息;在观看足球比赛时,除了关心进球的多少外,还要知道各支球队的技术统计,等等。要正确分析和理解这些问题,统计学知识必不可少。

第一节 统计与统计学

一、统计的含义

什么是统计?简言之,统计是研究事物数量特征的一种计量活动。然而,“统计”一词在不同的背景下,人们赋予它的含义是有差异的。一般从三个方面来理解统计的含义,即统计工作、统计数据和统计学。

(一) 统计工作

统计工作又称统计活动,是指收集、整理和分析统计数据,并探索数据的内在数量规律性的活动过程。社会经济统计的基本任务是对国民经济和社会发展情况进行统计调查和统计分析,提供统计资料,实行统计监督。

(二) 统计资料

统计资料又称统计数据,即统计活动过程所获得的各种数字资料和其他资料的总称。它具体表现为各种反映社会经济现象数量特征的原始记录、统计台账、统计表、统计图、统计分析报告、政府统计公报、统计年鉴等各种数字和文字资料。

(三) 统计学

统计学是指阐述统计工作基本理论和基本方法的科学,是对统计工作实践的理论概括和经验总结。它以社会经济现象总体的数量方面为研究对象,阐明统计设计、统计调查、统计整理和统计分析的理论与方法,是一门方法论科学。在本书中,对统计学作如下定义:统计学是收集、处理、分析和解释数据的科学,其目的是探索数据的内在数量规律性,以达到对客观事物的科学认识。

统计工作、统计资料和统计学之间有着密切联系。统计工作与统计资料之间是过程和成果的关系,统计资料是统计工作的直接成果。统计工作和统计学是实践与理论的关系,统计



学是统计工作实践经验的理论概括和科学总结,它来源于统计实践,又高于统计实践,反过来又指导统计实践,统计工作的现代化与统计科学的技术进步是分不开的。统计工作、统计资料和统计学相互依存、相互联系,共同构成完整的统一体。

二、统计的研究对象及特点

社会经济统计的研究对象是大量社会经济现象的综合数量特征和数量关系。它通过对社会经济现象数量关系的研究,揭示现象的规模、水平、结构、速度、趋势、各种比例关系和依存关系等,达到对社会经济现象的本质特征和规律性的认识。作为众多社会科学中的一门,统计学有其自身的特点,概括起来有以下五个方面:

(一) 数量性

统计学的研究对象是社会经济现象总体的数量方面,包括数量的多少、现象之间的数量关系、质量互变的数量界限。这一特点是统计学(定量分析学科)与其他定性分析学科的分界线。数量性是统计学研究对象的基本特点,因为数字是统计的语言,数据资料是统计的原料。一切客观事物都有质和量两个方面,事物的质和量总是密切联系、共同规定着事物的性质。没有无量的质,也没有无质的量;一定的质规定着一定的量,一定的量也表现为一定的质。但在认识的角度上,质和量是可以区分的,可以在一定的质的情况下,单独地研究数量方面,通过认识事物的量进而认识事物的质。因此,事物的数量是人们认识客观现实的重要方面,通过分析研究统计数据资料,研究和掌握统计规律性,就可以达到统计分析研究的目的。

(二) 总体性

统计的数量研究是对总体普遍存在着的事实进行大量观察和综合分析,得出反映现象总体的数量特征,说明现象变化的规律性。统计研究是从调查个别事实开始的,但观察、登记个别单位的标志表现,目的不在于研究个别事实的具体状况,而是要集合大量单位的事实,加以汇总和分析,来研究现象总体的规律性。统计也不是一概不研究个别事实。以大量观察为依据的综合数量特征形式来研究社会经济发展过程,会不可避免地趋于一般化、抽象化。因此,研究总体的统计数据资料不排除对个别事物的深入调查研究,但它是为了更好地分析研究现象总体的统计规律性。例如,工资统计分析,要反映、分析和研究一个地区的工资情况,先要从每个职工的工资开始统计,然后再综合汇总得到该地区的工资情况。只有从个体开始,才能对总体进行分析研究。

(三) 具体性

统计研究对象是社会经济现象在具体时间、地点、条件下所表现的数量方面,即它不是纯数量的研究,而是具有明确的现实含义的,这一特点是统计学与数学的分水岭。数学是研究事物的抽象空间和抽象数量的科学,而统计学研究的数量是客观存在的、具体实在的数量表现。统计研究对象的这一特点,也正是统计工作必须遵循的基本原则。正因为统计的数量是客观存在的、具体实在的数量表现,它才能独立于客观世界,不以人们的主观意志为转移。统计资料作为主观对客观的反映,必然是存在第一性,意识第二性,存在决定意识。只有如实地反映具体的已经发生的客观事实,才能为人们进行统计分析研究提供可靠的基础,从而分析、探索和掌握事物的统计规律性。

(四) 社会性

统计研究的对象是人类社会活动的过程和结果。人类社会活动是人们有意识、有目的的主观活动,各种活动都贯穿着人与人之间的关系,会受到一定的社会观和世界观的影响。而社会经济统计以社会经济现象作为研究对象,自然具有明显的社会性。例如,收入分配关系主要由所有制、产业布局、城乡差别等社会性因素决定。所以,统计在研究收入关系时,不是简单地按收入的高低进行分组,而是按所有制分组、按城乡分组等,这些方法就表现出了明显的社会性。同样,在计算平均收入水平时,也不能简单地把各种收入水平进行平均,而是按照同质性要求进行平均,把收入水平悬殊的各部分分开计算,这也体现了明显的社会性。所以,在研究社会经济现象时,必须充分认识到这些社会关系。

(五) 变异性

统计研究同类现象总体的数量特征,前提则是总体各单位特征表现存在着差异,而且这种差异并不是由某种固定的原因事先给定的。如果各单位不存在这些差异,也就不需要作统计;如果各单位之间的差异是可以按已知条件事先推定的,也就不需要统计方法了。统计上把总体各单位由于随机因素引起的标志表现的差异称为变异。由于社会经济现象具有广泛的、复杂的联系,各个单位所处的地位不同、条件各异,因而社会经济的个别现象有其特殊性和偶然性,所以对社会经济现象总体的数量研究,必须运用统计的方法,从各单位的变异中归纳概括出它们的共同特征,显示出现象的普遍性和必然性。

三、统计研究的基本方法

在对社会经济现象进行统计分析时,要应用各种专门的研究方法。统计研究的基本方法主要包括大量观察法、统计分组法、综合指标法、统计模型分析法和统计推断法。

(一) 大量观察法

大量观察法是统计学所特有的方法,也是统计调查阶段的基本方法。所谓大量观察法,是指在统计研究的过程中,对所研究的社会经济现象总体的全部或足够多的单位进行调查观察,借以掌握现象的综合特征,认识社会经济现象发展规律性的一种方法。统计工作之所以要运用大量观察法,是由统计研究对象的大量性和复杂性决定的。在现象总体中,个别单位往往受偶然因素的影响,如果任选其中之一进行观察,其结果不足以代表总体的一般特征;只有观察全部或足够多的单位并加以综合,影响个别单位的偶然因素才会相互抵消,现象的一般特征才能显示出来。大量观察的意义在于可使个体与总体之间在数量上的偏误相互抵消。

大量观察法的数学依据是大数定律。大数定律是随机现象的基本规律。大数定律的一般概念是:在观察的过程中,每次取得的结果不同,这是由偶然性所致的,但大量、重复的观察结果的平均值却几乎接近确定的数值。狭义的大数定律是指概率论中反映上述规律性的一些定理,表述平均数的规律性与随机现象的概率关系。

(二) 统计分组法

统计分组法就是根据一定的研究目的和现象的总体特征,将总体各单位按一定的标志,把社会经济现象划分为不同性质或类型的组别。统计分组法是统计研究的基本方法,主要用于统计整理阶段。

统计分组法是研究总体内部差异的重要方法。通过分组,可以研究总体中不同类型的性质以及它们的分布情况,如产业的经济类型及其行业分布情况;可以研究总体中的构成和比例关系,如三次产业的构成、生产要素的比例等;也可以研究总体中现象之间的相互依存关



系,如企业经营规模和利润率之间的关系等。

由于将总体单位分组是在取得资料后整理资料时进行的,因此,分组容易被认为只是整理方法,这是一种误解。其实,无论是从实际工作过程来讲,还是从作用上来讲,它都是贯穿统计活动全过程的一种重要的方法,它的作用在统计设计和分析研究中都十分显著。

(三) 综合指标法

综合指标法是指运用各种统计指标来反映和研究客观总体现象的一般数量特征和数量关系的方法。通过综合指标的计算,可以显示出现象在具体时间、地点条件下的总量规模、相对水平、集中趋势、变异程度,并进一步从动态上研究现象的发展趋势和变化规律。

指标和分组是密切联系、相互依存的,共同反映社会经济现象的质和量。统计分组如果没有一定的统计指标来反映现象的规模水平,就不能显示出现象的数量特征;而统计指标如果没有科学的分组,就往往容易掩盖矛盾,成为笼统的指标,甚至成为虚构的指标。所以在研究社会经济现象的数量关系时,必须科学地进行分组,合理地设置指标。

(四) 统计模型分析法

统计模型分析法是指根据一定的经济理论和假设条件,用数学方程去模拟客观经济现象相互关系的一种研究方法,如相关分析法、回归分析法和统计预测法。利用这种方法,可以对客观现象和过程中存在的数量关系进行比较完整和全面的描述,凸显所研究的综合指标之间的关系,从而简化了客观存在的复杂的其他关系,以便利用模型对所关心的现象变化进行评估和预测。

(五) 统计推断法

统计推断法是指以一定的置信标准,根据样本数据来判断总体数量特征的归纳推理方法。统计在研究现象的总体数量关系时,需要了解的总体对象的范围往往是很大的,有时甚至是无限的,而由于经费、时间和精力等各种原因,以致有时在客观上只能从中观察部分单位或有限单位进行计算和分析,根据局部观察结果来推断总体。例如,要说明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,推断出这批灯泡的平均使用寿命,并给出这种推断的置信程度。这种在一定置信程度下,根据样本资料的特征,对总体的特征作出估计和预测的方法称为统计推断法。统计推断法是现代统计学的基本方法,在统计研究中得到了极为广泛的应用,既可以用于对总体参数的估计,也可以用于对总体某些分布特征的假设检验。从这种意义上来说,统计学是在不确定条件下作出决策或推断的一种方法。

四、统计工作过程

统计工作的全过程可以分为四个阶段:统计设计、统计调查、统计整理和统计分析。

(一) 统计设计

统计设计是指根据统计研究对象的性质和研究目的,对统计工作的各个方面和各个环节所做的全面部署和安排。统计设计的最终结果表现为各种标准、规定、制度、方案 and 办法,如统计分类标准、目录、统计指标体系、统计报表制度、统计调查方案、普查办法、统计整理或汇总方案等。统计工作是一项高度集中统一和科学性很强的工作,无论是统计总体范围、统计指标的口径和计算方法,还是统计分类和分组的标准,都必须统一,不能各行其是。只有科学地进行统计设计,才能做到统一认识、统一步骤、统一行动,使整个统计工作有秩序地、协调地进行,从而从根本上保证统计工作和统计资料的质量。因此,统计设计是

统计工作的先导。统计设计的主要内容有：统计指标和指标体系的设计、统计分类和统计分组的设计、统计表的设计、统计资料搜集方法的设计、统计工作各个部门和各个阶段的协调与联系、统计力量的组织与安排等。

（二）统计调查

统计调查是指根据统计方案的要求，采用各种调查组织形式和调查方法，有组织、有计划地对所研究总体的各个单位进行观察、登记，准确、及时、系统、完整地搜集统计原始资料的过程。统计调查是统计认识活动由定性认识过渡到定量认识的阶段，这个阶段所搜集的资料是否客观、周密、系统、及时，直接影响到统计整理的好坏，关系到统计分析结论的正确性，决定着整个统计工作的质量。所以，统计调查是整个统计工作的基础。

（三）统计整理

统计整理是指根据统计研究的目的和任务，对统计调查阶段所取得的原始资料进行审核、分组和汇总，将分散的、零星的反映总体单位特征的资料转化为反映各组 and 总体数量特征的综合资料的过程。统计整理是将对总体单位特征的认识过渡到对总体数量特征的认识的桥梁和纽带。它既是统计调查的继续，又是统计分析的必要前提，在统计工作中处于中间环节，起着承上启下的作用，是沟通统计调查和统计分析的纽带和桥梁。

（四）统计分析

统计分析是指在统计调查和统计整理的基础上，用科学的分析方法，对所研究的现象总体进行全面、系统的数量分析，认识和揭示事物的本质和规律性，进而向有关单位和部门提出咨询建议以及进行必要的分析、预测的统计工作过程。统计分析是统计工作的最后阶段，也是统计发挥信息、咨询和监督职能的关键阶段。

一般来说，统计工作过程的四个阶段是依次进行的，各有自己的特定内容。同时，它们又相互联系、相互制约，任何一个阶段的工作失误，都会影响到整个统计工作的大局。在某些情况下，为了保证从整体上取得良好效果，各阶段也可以相互渗透、交叉进行。例如，根据实际工作需要，可以边设计、边调查、边整理、边分析；有时在调查、整理阶段进行一些必要的分析，或者对原设计方案进行适当的改进；有时在统计分析中因为现有资料不能满足需要而作一些必要的补充调查、加工整理和计算工作等。

五、统计的职能

统计具有信息、咨询和监督三大职能，它们共同构成了统计的整体职能。

（一）信息职能

信息职能是指统计部门根据科学的统计指标体系和统计调查方法，灵敏、系统地采集、处理、传递、存储和提供大量的以数量描述为基本特征的社会经济信息。因而要不断拓展统计信息的内容，保证统计信息的可靠性，完善统计信息的自动化建设，实现统计信息生产和使用的社会化程度。

（二）咨询职能

咨询职能是指利用已经掌握的丰富的统计信息资料，运用科学的分析方法和先进的技术手段，深入开展综合分析和专题研究，为科学决策和管理提供可供选择的咨询建议和对策方案。



(三) 监督职能

监督职能是指根据统计调查和分析,及时、准确地反映经济、社会和科技运行的实际状况,并对其实行全面、系统的定量检查、监测和预警,以促进国民经济持续、稳定、协调地发展。

统计的信息、咨询和监督职能是彼此依存、相互联系的,它们共同构成了一个完整的有机整体。在发挥各自职能的基础上,对三种职能进行优化和整合,形成合力,就能充分发挥统计在现代化管理过程中的作用。

第二节 统计的产生和发展

统计是随着人类社会政治、经济的发展与治理国家的需要而产生和发展的,其产生和发展包括两个不同层次的内容:一是统计作为一种社会实践活动的产生和发展;二是统计作为一门科学的产生和发展。

一、统计实践活动的产生和发展

统计实践活动是顺应人类社会政治、经济发展以及国家管理的需要而逐步产生并发展起来的,距今已有四五千年的历史。

早期的统计实践孕育于原始社会末期简单的计数活动,人们对猎物的清点与分配可看成是最早的统计行为。在我国古代有结绳记事、结绳计数的方法,郑玄的《周易正义》中提到“事大,大结其绳;事小,小结其绳;结之多少,随物众寡”,就体现了统计分组、汇总的基本思想。

到了奴隶社会,为了顺利征税、征兵、征劳役,统治阶级有必要详细了解社会基本情况,这些促成了有关人口、财富以及军事方面的统计实践的长足发展。在夏朝,有资料记载当时全国分九州,占地 2 430 万顷,总人口数为 1 355 万。到了西周,已建立了基本的统计报告制度,并依报送时间的长短不同分为“日成”“月要”和“岁会”。同处于奴隶社会的古埃及、古希腊、古罗马等国,也都有了关于人口、土地等方面的统计实践活动。

进入封建社会后,随着人类社会生产的发展,统计的范围也开始向社会经济生活的各方面扩展。秦时的《商君书》中提到“强国知十三数”,包括反映当时基本国情国力的人口、粮食、农业生产资料以及自然资源等指标,秦始皇统一中国后,更建立了从中央到地方的“上计”报告制度。长期的中央集权制国家也使得中国在户籍、田亩等领域的统计活动有了很大的发展。不论在统计的方法、制度还是组织上,中国都处于当时世界的领先水平。然而,自给自足的封建自然经济在阻碍社会生产力发展的同时,也阻碍了统计实践的发展,统计的功能仍局限在国家管理的需要上,其主要工作也多局限在原始调查登记以及简单的汇总分析工作上。

统计实践活动的广泛开展始于资本主义社会。在 17—18 世纪的资本主义上升时期,由于工、商、农、贸、交通业的发展,统计实践逐渐扩展到社会经济活动的众多领域,开始出现各种“国情普查”,并形成制度。18 世纪末,西欧各国开始设立专门的统计机构负责统计工作,收集各方面的统计信息,定期或不定期地举行人口、工业、农业、贸易、交通等专项调查,社会统计、科技统计、环境统计等从经济统计中分离出来,形成了比较完整的统计内

容体系,抽样方法的引入更将统计的认识能力推向了新的高峰。由于各国政府的重视与支持,1853年在比利时的布鲁塞尔召开了第一次国际统计会议,1885年成立了国际统计学会,1887年在罗马召开了第一届大会,此后每两年召开一次,1995年8月在北京召开了第50届会议。

20世纪30年代,抽样方法的普遍采用将统计从对现象的客观描述发展到抽样推断、预测分析方面;50年代后,国民经济账户体系和投入产出法的应用,各种新的数学方法的介入,更是极大地丰富了统计实践活动。随着计算机等现代科学技术的发展,统计实践活动进入了一个全面发展的新阶段,统计实践已经从早期的国家管理、社会经济管理方面扩展到人类的一般认知活动,大到一个国家或洲际的管理决策,小到一般家庭的投资决策;从研究社会经济现象的数量方面,深入到自然技术现象的数量方面,甚至连文科领域的法律、历史、语言等专业也都开始重视统计工作,统计已不仅仅是一个专业领域,还与数学、哲学一样成为众多学科的专业基础。相对地也要看到,我国现当代统计的发展滞后。我国于1906年设立了统计局,直到新中国成立后,统计工作才有了真正繁荣发展的制度环境,但受到前苏联理论界的影响,中国的统计工作长期偏向于描述统计,直到20世纪80年代,数理统计才得到人们的认可,在理论与实践发展上开始与世界同步。而于1984年1月1日颁布施行的《中华人民共和国统计法》经过1996年5月和2009年6月两次修订,明确指出“统计的基本任务是对经济社会发展情况进行统计调查、统计分析,提供统计资料和统计咨询意见,实行统计监督”,为我国的统计工作提供了有效的法律制度保障。

二、统计学的产生和发展

统计理论来源于统计实践,是统计实践活动的经验总结和理论概括。统计的起源很早,对统计发展的历史可追溯到远古的原始社会。但是,人类的统计实践上升到理论并予以总结和概括而成为一门系统的科学——统计学,却是近代的事情,距今只有300多年的历史。了解统计学的渊源及其发展过程,对于人们学习统计学的理论和方法,提高统计实践和理论水平,都是十分必要的。

在统计理论的形成过程中,由于各国的历史背景、经济水平与思想渊源不同,对统计的研究内容和表达方式也互异,因此,后来陆续形成了不同的统计学派。从统计学的产生和发展过程来看,大致可以划分为三个时期:古典统计学时期、近代统计学时期和现代统计学时期。

(一) 古典统计学时期

古典统计学时期,约自17世纪中叶至19世纪初叶。这段时期出现了三个学派:

1. 国势学派

国势学派也称记述学派,是伴随德意志的兴盛而产生的,代表人物是赫尔曼·康令(Hermann Conring, 1606—1681)和阿亨瓦尔(G. Achenwall, 1719—1772)。国势学派以叙述国家显著事项和国家政策关系为内容,并给出了统计学这一名词。但国势学派主要用文字表述,缺乏数字内容,和现代统计学比较起来,有一些实不副名,即没有太多的实际内容符合统计学的名称。

2. 政治算术学派

政治算术学派产生于17世纪的英国,代表人物是威廉·配第(William. Petty, 1623—



1687) 和约翰·格朗特 (John Graunt, 1620—1674)。威廉·配第著有《政治算术》一书, 在书中他用数字来表述, 用数字、重量和尺度来计算, 并配以直观的图表, 为现代统计学提供了一个完美的开端。约翰·格朗特对英国伦敦人口的出生率和死亡率分别进行统计, 并编制出世界上第一张“死亡率”统计表。他们的工作和使用的方法, 在现代统计学中都是至关重要的。他们虽未创立“统计学”之名, 但其所用于探索社会和经济现象数量规律性的方法却具有“统计学”之实。

3. 概率论学派

概率论学派以法国学者为主, 这一学派最早起源于对赌博中投骰子输赢问题的研究。其创始人是帕斯卡 (B. Pascal, 1623—1662) 和费马 (P. de Fermat, 1601—1665), 他们以通信的方式讨论赌博的概率问题。惠更斯 (C. Huygens, 1629—1695) 著有《骰子赌博理论》。棣莫弗尔 (A. de Moivre, 1667—1754) 发现正态方程式。同一时期, 瑞士数学家贝努里 (I. Bernoulli, 1654—1705) 提出了二项分布理论。此后, 拉普拉斯 (P. S. Laplace, 1749—1827) 在 1814 年发表《概率分析论》一书, 构筑了古典概率理论的完整体系, 并用于自然和社会现象的研究。泊松 (S. D. Poisson, 1781—1840) 提出泊松分布。高斯 (K. F. Gauss, 1777—1855) 提出最小平方法, 他还从观察天象中发现误差正态曲线。

(二) 近代统计学时期

近代统计学时期, 约自 19 世纪初叶至 20 世纪初叶。这一时期以比利时学者凯特莱 (A. J. Quetelet, 1796—1874) 为起点。他著有《社会物理》一书, 提出了偶然误差的概念, 并指出某一学科的统计方法可以用到其他学科。在统计研究方面, 他先研究天文、气象方面的统计资料, 后又用统计数字研究植物界和人类社会。他以概率论作为理论基础, 用大量观察和综合平均的方法进行研究, 从而把概率论、国势学和政治算术观察群体现象进行数量分析的方法, 融合为一门学科, 奠定了近代统计学的基础。凯特莱于 1851 年在比利时首都布鲁塞尔召开第一届国际统计学会会议, 该会又于 1855 年在伦敦召开, 改称为“国际统计学会” (International Statistical Institute), 这一名称沿用至今。

此外, 克尼斯 (K. G. A. Knies, 1812—1898) 和恩格尔 (C. L. Engel, 1821—1896) 以大量观察法寻求社会现象规律, 称为社会统计学派。高尔顿 (F. Galton, 1822—1911) 发现百分位数, 他还从研究遗传学和优生学中创立了回归分析的概念。皮尔逊 (K. Pearson, 1857—1936) 提出经验分布函数、相关分析、卡方检验和大样本的抽样理论, 并且完成了描述统计学的体系, 有人由此认为他是近代统计学的创始人。

(三) 现代统计学时期

现代统计学时期, 自 20 世纪初至今。在这一时期, 戈塞特 (W. S. Gosset, 1876—1937) 以笔名“学生” (Student) 发表 t -分布, 这是小样本抽样理论的基础。博雷尔 (E. Borel, 1871—1956) 奠定了现代概率理论的基础。费希尔 (R. A. Fisher, 1890—1962) 提出 Z -分布、显著性水平、假设检验、自由度、实验设计和方差分析等方法与概念。在美国, 斯内德克 (G. W. Snedcor, 1881—1974) 将 Z -分布转换成 F -分布, 瓦尔德 (A. Wald, 1902—1950) 提出决策理论和序贯抽样法, 冯·诺伊曼 (J. V. Neumann, 1903—1957) 和摩根斯特恩 (O. Morgenstern, 1902—1977) 提出博弈论, 使决策理论更加系统化。奈曼 (J. Neyman, 1894—1981) 和戴明 (W. E. Deming, 1900—1993) 提出抽样调查法, 对质量控制及生产管理贡献很大。维纳 (N. Wiener, 1894—1964) 的控制论和香农 (C. E. Shannon, 1916—

2001)的信息论,使推断统计学的理论更加健全。美国的大学自1950年把统计学设为独立的学系,1955年开始颁授统计学的高级学位。从20世纪50年代起,统计学受计算机、信息论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、随机过程、非参数统计、时间序列分析等。据美国学者估计,现代统计学是以指数式加速度发展的,新的研究分支不断增加,统计应用领域不断扩展。统计方法在各学科领域的应用又进一步促进了统计方法研究的深入和发展。

上述三个阶段的划分只是大致的,其实,统计学的发展是渐进的、错综的,并没有明确的时间界限。而且从发展趋势看,统计理论研究的分支仍会不断增加,统计学将越来越具有交叉学科的性质,应用的范围也将更加广泛。一些过去与数字毫无联系的学科,如政治学、历史学、法学等,也正在并将更多地应用统计方法进行研究和分析。目前,阻碍统计方法推广应用的主要是定性资料,如国民经济行业、部门等分类资料,不同的政治观点、思想准则等。如何对这些定性问题进行量化处理分析,是统计学家们正在研究的重要课题之一。探索性数据分析的应用前景也很广阔,这种方法重视数据的直观显示、原始数据中信息的提取以及对特殊数据的识别和具体分析。贝叶斯统计理论研究将在统计推断、预测和决策等领域继续显示其强健的生命力,成为统计实践中有力的工具。

三、统计学的分科

统计方法已被应用到自然科学和社会科学的众多领域,并逐渐发展成为由若干分支学科组成的学科体系。从统计方法的构成来看,统计学可以分为描述统计学和推断统计学;从统计方法研究和统计方法的应用角度来看,统计学可以分为理论统计学和应用统计学。

(一) 描述统计学和推断统计学

描述统计学(Descriptive Statistics)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,进而通过综合概括与分析得出反映客观现象的规律性数量特征。其内容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分析方法等。

推断统计学(Inferential Statistics)则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。

描述统计学和推断统计学的划分,一方面反映了统计方法发展的前后两个阶段,另一方面也反映了应用统计方法探索客观事物数量规律性的不同过程。

统计研究过程的起点是统计数据,终点是探索出客观现象内在的数量规律性。在这一过程中,如果搜集到的是总体数据(如普查数据),则经过描述统计之后就可以达到认识总体数量规律性的目的了;如果所获得的只是研究总体的一部分数据(样本数据),要找到总体的数量规律性,则必须应用概率论的理论并根据样本信息对总体进行科学的推断。

显然,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,推断统计学则是现代统计学的主要内容。由于在对现实问题的研究中,所获得的数据主要是样本数据,因此,推断统计学在现代统计学中的地位和作用越来越重要,已成为统计学的核心内容。当然,这并不等于说描述统计学不重要,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切合实际的结论。



从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学发展成熟的重要标志。

(二) 理论统计学和应用统计学

理论统计学(Theoretical Statistics)是统计学的数学原理,它主要研究统计学的一般理论和统计方法的数学理论。由于现代统计学用到了几乎所有方面的数学知识,因此从事统计理论和方法研究的人员需要有坚实的数学基础。此外,由于概率论是统计推断学的数学和理论基础,因而广义地讲,统计学也是应该包括概率论在内的。理论统计学是统计方法的理论基础,没有理论统计学的发展,统计学也不可能发展成为像今天这样一个完善的科学知识体系。

在统计学的研究领域中,从事理论统计学研究的人是相对很少的一部分,而大部分则是从事应用统计学(Applied Statistics)研究的。应用统计学是研究如何应用统计方法去解决实际问题的。统计学是一门收集和分析数据的科学。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。例如,统计方法在生物学中的应用形成了生物统计学,在医学中的应用形成了医疗卫生统计学,在农业试验、育种等方面的应用形成了农业统计学。统计方法在经济和社会科学研究领域的应用也形成了若干分支学科。例如,统计方法在经济领域的应用形成了经济统计学及其若干分支,在管理领域的应用形成了管理统计学,在社会学研究和社会管理中的应用形成了社会统计学,在人口学中的应用形成了人口统计学,等等。以上这些应用统计学的不同分支所应用的基本统计方法都是一样的,即都是描述统计学和推断统计学的主要方法。但由于各应用领域都有其特殊性,统计方法在应用中又形成了一些不同的特点。

第三节 统计数据 的类型

获得有用的信息对于人们从事任何活动来说都是至关重要的。数据包含了在一个特定情况下作出更明确决策所需要的信息。在很多情况下人们都需要数据,如需要用数据来评价某种产品的质量是否符合标准,需要用数据来帮助人们在决策过程中形成可供选择的行动方案,等等。熟悉统计数据的类别性质,对人们在统计研究中有针对性地确定合适的统计分析方法很有帮助。

一、分类数据、顺序数据和数值型数据

按照所采用的计量尺度不同,统计数据可以分为分类数据、顺序数据和数值型数据。

1. 分类数据

只能归于某一类别的非数字型数据,称为分类数据(Categorical Data)。

分类数据是对事物进行分类的结果,该数据表现为类别,是用文字来表述的。分类数据主要是由分类尺度计量形成的。例如,人口按性别可分为男、女两类,企业按行业属性可分为医药企业、家电企业、纺织品企业,等等。为了便于统计处理,对于分类数据可以用数字代码来表示各类别。比如,用“1”表示“男性”,“0”表示“女性”;用“1”表示“医药企业”,“2”表示“家电企业”,“3”表示“纺织品企业”,等等。