

统计学

S t a t i s t i c s

盛承懋 楼 江 余晓红 编著



东南大学出版社

统 计 学

盛承懋 楼 江 余晓红 编著

东南大学出版社

内 容 提 要

本书系统地介绍了工程管理中常用的统计分析方法。全书共分10章,内容有总论、统计调查与整理、统计量与抽样分布、参数估计、假设检验、方差分析、回归分析、统计指数、时间序列分析和多元统计分析。本书的特点是:侧重基本概念、基本理论、基本方法的介绍;强调实际应用,采用了较多的工程管理应用实例;结合统计软件(Excel, SPSS)的介绍,力求简明易懂,便于操作。每章配有小结和练习题,帮助读者总结与复习。

本书可供高等院校工程管理专业和其他经济管理类专业的本科生作统计学课程的教材或参考书之用。

图书在版编目(CIP)数据

统计学/盛承懋,楼江,余晓红编著. —南京:
东南大学出版社,2002.9

工程管理系列教材

ISBN 7-81089-012-3

I. 统... II. ①盛...②楼...③余...
III. 统计学-高等学校-教材 IV. C8

中国版本图书馆 CIP 数据核字(2002)第 059114 号

东南大学出版社出版发行

(南京四牌楼2号 邮编210096)

出版人:宋增民

江苏省新华书店经销 南京京新印刷厂印刷

开本:B₅ 印张:15 字数:294千字

2002年10月第1版 2002年10月第1次印刷

印数:1-4 000 定价:18.00元

(凡因印装质量问题,可直接向发行科调换。电话:025-3795802)

编写委员会名单

主任委员 成 虎 盛承懋

副主任委员 (以姓氏笔画为序)

王元钢 王卓甫 刘碧云

李启明 宋学锋 陆军令

陆惠民 杨鼎久

委 员 (以姓氏笔画为序)

刘钟莹 许 敏 连永安

周 云 黄月华 黄安永

黄有亮 温作民

出版说明

“工程管理”专业自 1999 年列入教育部本科专业目录,并在全国招生以来,其教材的问题一直为人们所关注。通常一个新兴专业的培养方案、课程教学大纲和教材体系的建立和完善必须经历很长时间,但由于现代社会对“工程管理”专业人才的需求量大,近几年设置“工程管理”专业的院校越来越多,仅江苏省就有十几所,这在全国各省市中是少有的。“工程管理”专业的招生量也在不断扩大。为此,我们只能先做起来,尽快地拿出这套教材,在以后的教学实践中再不断地完善。

按照高等学校工程管理专业指导委员会制定的“工程管理”本科专业的培养方案及课程教学大纲的要求,我们组织了江苏省有“工程管理”本科专业的十几所院校编写了这套“工程管理”专业的教材。

这套教材包括了“工程管理”专业的专业课程和方向课程。在本套教材的策划和编写过程中遇到了很多矛盾和问题:

1. 工程管理专业的学生需要综合的广博的知识,工程管理专业有两个一级基础学科——土木工程和管理学。这使得“工程管理”专业的课程数目多,而且这些课程横跨土木工程、管理学、经济学、法律等学科。组织这个课程体系的难度很大。
2. 教材的内容必须新颖。我国已经加入 WTO,我国的工程管理必须与国际接轨,在教材中必须既符合我国的国情,又要反映国际上最新的内容。正像许多工程实践结果显示的一样:这两者要有机地结合在一起是十分困难的。
3. 按照高等学校工程管理专业指导委员会制定的培养方案及课程教学大纲的要求,以及课时分配,每门课的课时都比较少。而每门课有自己完整的知识体系,我们又要求尽可能多地介绍一些新的内容,以扩大学生的知识面。
4. 虽然“工程管理”是一个新兴的专业,但是它的课程在名称上并不是新的。这些课程在过去的“土木工程”专业(或建筑工程专业)、“建筑管理工程”专业,以及相关的管理专业、房地产专业都曾开设过。但在“工程管理”专业中,这些课程的教学大纲和课时都已变化,它们的内容必须更新。
5. 在过去相关专业的教学体系中,这些课程内容有明显的重叠和内容不全,

必须既尽量减少课程之间的内容的重叠,又作一些补缺。

在本系列教材的策划和编写过程中,除了十分注重上述这些矛盾和问题的解决之外,编委会与各本教材的主编还充分考虑到工程管理专业与其他管理专业、土木工程专业之间的相关性,因此,在教材的编写体系与内容上兼顾了这些专业的需要,也可为这些专业所采用。尽管本系列教材已经过多次的讨论和修改,但书中必然有许多不足、缺陷,希望本专业的同行们、学生们在应用中对本套教材中的问题提出意见和建议,以使我们能够不断地改进,将它做得越来越好。

这些参编的老师都长期在“工程管理”及其相关专业从事专业课程的科研、教学,具有丰富的研究成果和教学经验,曾编写过许多教材。有的老师还曾参加过国家级、省部级规划教材的编写。

从1999年5月份我们第一次讨论本系列教材至今已有3年。经过几年的努力我们终于将它奉献给大家,奉献给我们的专业。教材体系和内容的成熟也是一个专业成熟的重要标志之一。

本系列教材的出版,得到江苏省各有关高校领导的关心和支持,得到国内有关同仁的热情指导,得到东南大学出版社的鼎力相助,在此谨向各位表示衷心的感谢!

编 委 会

序

统计方法是现代社会各行各业发展不可缺少的重要工具,统计方法的应用几乎渗透到自然科学和社会科学的各个领域,为适应教育面向现代化的需要,为工程管理和工商管理类专业学生提供一本针对性强、易读易懂、便于应用的统计教材是我们编写此书的初衷。

本书系统地介绍了工程管理中常用的统计分析方法。其特点是:侧重基本概念、基本理论、基本方法的介绍;在介绍各种统计方法理论背景的同时配合了较多的工程管理应用实例;为帮助学生在将来的实际工作中更加得心应手地用好统计方法,我们增加了统计软件(Excel, SPSS)应用的内容。每章配有小结和练习题,帮助读者总结与复习。

全书共分 10 章,内容包括总论、统计调查与整理、统计量与抽样分布、参数估计、假设检验、方差分析、回归分析、统计指数、时间序列分析和多元统计分析。所需学时为 36~48 学时,要求学生具备概率论的基础。多元统计分析这部分内容可以作为课程设计的教学内容。

本书第 3,4,5,10 章由盛承懋编写,第 1,2,7 章由楼江编写,第 6,8,9 章由余晓红编写,书中关于统计软件的应用章节全部由余晓红编写,全书由盛承懋任主编,负责全书的修改、总纂和定稿工作。

本书在编写过程中得到了苏州科技学院党政领导的关心与支持,东南大学出版社的编辑为本书尽快与读者见面付出了辛勤的劳动,在编写过程中参阅了有关专家、学者正式出版的统计学教材,书中相当部分内容的打印得到了姬丽华女士的帮助,在此一并向各位表示衷心感谢。

由于作者水平有限,不妥之处敬请读者批评指正。

编 者

2002 年 5 月

目 录

1 总论	1
1.1 统计学的产生与发展	1
1.2 统计学中的几个基本概念	3
小 结	6
练习题	7
2 统计调查与整理	8
2.1 统计调查的方案设计及组织形式	8
2.2 统计整理与频数分布	16
2.3 统计表与统计图	19
小 结	26
练习题	27
3 统计量与抽样分布	28
3.1 总体、样本与统计量	28
3.2 描述分布集中趋势的统计量	29
3.3 描述分布离散趋势的统计量	35
3.4 抽样分布	38
小 结	42
练习题	43
4 参数估计	45
4.1 参数估计的基本问题	45
4.2 参数的点估计	46
4.3 参数的区间估计	53
小 结	64
练习题	64
5 统计假设检验	66
5.1 统计假设检验概说	66

5.2	正态总体均值的检验	72
5.3	两个正态总体均值相等的检验	76
5.4	正态总体方差的检验	80
	小 结	85
	练习题	86
6	方差分析	88
6.1	单因素方差分析	88
6.2	双因素方差分析	92
6.3	使用 Excel 进行方差分析	99
	小 结	101
	练习题	101
7	回归分析	104
7.1	概述	104
7.2	一元线性回归	105
7.3	多元线性回归	114
7.4	使用哑变量的回归	117
7.5	非线性回归	120
7.6	使用 Excel 进行回归分析	123
	小 结	125
	练习题	125
8.	统计指数	128
8.1	总指数的编制方法	128
8.2	拉氏指数与帕氏指数的比较	131
8.3	总指数的平均形式	132
8.4	固定权数与基期变换	135
8.5	指数体系和因素分析	136
8.6	我国零售价格指数的编制方法	139
	小 结	141
	练习题	141
9	时间序列分析	144
9.1	时间序列概述	144

9.2	长期趋势分析	150
9.3	季节变动和循环变动分析	158
9.4	时间序列的预测方法	164
	小 结	171
	练习题	171
10	实用多元统计分析	174
10.1	聚类分析	174
10.2	判别分析	180
10.3	主成分分析	191
10.4	因子分析	197
10.5	使用 SPSS 进行多元统计分析	202
	小 结	209
	附 录	211
	参考文献	228

1 总 论

1.1 统计学的产生与发展

统计是一门关于数量资料的调查、整理、分析和解释的科学。

所谓数量资料调查,就是对数量做出观测或计数的过程,是取得统计数据的一步。正确的结论,只能来自适当搜集的、具有代表性的资料。本项工作在整个统计程序中,是很重要的一步。数量资料调查的一些代表性做法,将在第2章中介绍。

所谓整理资料,就是以适当形式展示搜集的统计数据,以便导出合乎逻辑的结论。这一步工作也有一些代表性的做法,也将在第2章中介绍。

所谓分析资料,就是从收集的资料中抽取有关信息的过程,从而得到一个概括的、全面的数量描述。分析资料常用的统计量有平均数、中位数、极差和标准差等。

所谓解释资料,就是从搜集的“小批量”对象的资料中所导出的结论,对同类的大量对象作出推断。这是本书的重要内容之一。

所以,统计是一门科学。一定程度上讲,统计所讨论的问题是能用数量信息,也就是由计数或观测得来的信息加以回答的问题。至于究竟是为生物或医学研究而对细胞或细菌计数,还是为工业生产而调查工厂的工人数或工时数,这都关系不大。我们要考虑的首先是确定信息的种类,其次是考虑如何尽量做到适当而有效地搜集和加工该项信息,最后是解释结果。

“统计”一词有三种不同的含义。一种含义是指一种科学程序,比如统计学一词中的“统计”,是指用来研究和评价数量资料的一种科学程序。统计又可以是指一种实践活动,比如“统计工作”一词中的统计,它是指搜集、整理、分析和编制统计资料等业务活动。统计还可以是指数量资料,例如“某某年鉴”、“某某报表”中有统计,这表明其中有数量资料。

最初,统计是为奴隶主、封建主管理国家的需要而搜集资料,比如为征召兵役、收纳贡赋、派使徭役的需要而开始的人口、土地、财富的记录、计数工作。由于当时社会经济发展缓慢,统计发展也很缓慢。应当说,当时的统计工作还处在原始的调查登记和简单记数这样的初级阶段。

到封建社会解体,由资本主义取而代之以后,社会经济有了迅速发展,社会的分工,使统计工作扩展到了工业、商业、税收、航运和外贸等领域,开始形成社会经济统计。17、18世纪以后,经济学、人口学、生物学等学科先后建立,统计计数和统

计方法在这些学科中也得到了应用。另外,当时在哲学和数学方面的发展又给统计方法的形成带来了深远的影响。所以那时候的统计无论是作为管理工具,还是作为科研方法,都为统计学的建立提供了条件,也对统计学的发展提出了新的要求。

在资本主义的统计活动发展到一定阶段时,人们开始总结统计实践活动的经验,逐渐形成了比较系统的统计理论知识,这就是统计的起源。目前,公认最早的有关统计理论的著作,是 17 世纪 60 年代,在英国开始的政治算术学派的论述,迄今已有 300 多年的历史。与此同时的还有德国的记述学派。政治算术学派的创始人是英国的威廉·配第(1623—1687),他在自己撰写的《政治算术》一书中,用数量比较的方法进行政治经济分析。使用方法的独创性,使他成了统计学的奠基人。记述学派又称国情学派或国势学派,“统计学”这个名称便是这个学派启用的,但这个学派的著作有统计学之名,而无统计学之实,他们对统计学研究领域的确是明确的,但缺少统计学中最重要的数量分析。

在政治学派和记述学派的争论还没有完全解决时,在 18 世纪以后便出现了新的学派即数理学派。数理学派的代表人物是比利时的阿道夫·凯特勒(1796—1874),凯特勒把概率论的基本原理应用到统计研究中,从而开辟了统计学的新领域,产生了数理统计学,从此统计学获得了迅猛发展。至今,国内外的统计学,都包括这样两个门类:社会经济统计学和数理统计学。

凯特勒是指明生物学与人类学的数量遵循正态曲线的第一人,也是认识大数守恒的第一人,还是表明在一个研究领域中发展起来的统计技术可以应用于其他大部分领域的创始人之一。到 19 世纪后期,卡尔·皮尔森(1857—1936)使用遗传学中的问题,提出了回归与相关的概念,他广泛地研究了抽样误差的影响,提出了卡方检验法。在上个世纪初,一个以笔名“学生氏”的作者,发表了关于解释抽样得来资料的文章,第一个认识从小样本抽取可靠资料的方法,有其特殊重要性。引进“原假设”这一名词和发展方差分析的统计技术的就是费喧。在整个 20 世纪,杰出的统计学者不胜枚举,他们在发展新理论和应用方面都做出了巨大贡献。电子计算机的应用,更是大大地促进了这些发展。当今的科研人员,无不认为统计是最为有用的工具之一。

根据数字信息做出的判断,对日常生活的影响正与日俱进。假设实验与假设检验已成为处理许多问题的常用方法。今天,建立在概率论基础上的现代统计学,在物理和生物科学中、在经济学和社会学中、在生理学和教育学中、在药剂学和农业学中、在政府和实业界中,都证明是不可缺少的重要工具。天文学家根据统计方法预言天体未来的位置;遗传分离律是由统计方法确定下来的;人寿保险费与年金赔偿额是以统计记录为基础的死亡表来核定的;电力公司若没有负荷需要的记录,就不能有效地供应电力;农业研究人员从统计上的考虑来决定农业田间实验的显

著性;生产人员发现抽样调查理论在控制产品质量方面的价值是无法估计的;企业领导人和政府首脑们使用统计方法作出决察。统计应用的范围尽管如此不同,但所用的统计方法,则大部分是相同的,只是各方面的侧重点有所不同。一般说,各个领域所使用的都是同样一些统计方法。

在使用统计方法时,有一点要注意,原始资料必须正确,方法应用必须得当,结果要由不仅懂得方法本身、还需懂得方法应用领域的人来解释。否则,统计方法像一把双刃剑,如果原始资料残缺不全、差错很多,或使用方法不得当,都会导致工作失误或决策错误,并带来不应有的损失。

1.2 统计学中的几个基本概念

1.2.1 总体和总体单位

所谓总体,是指客观存在的、在某种同质性的基础上结合成的、由许多个别事物构成的整体。总体既可以由人、物、组织单位等实体构成,也可以由现象、活动过程等非实体构成。例如,对市场价格变动的研究,除了商品价格,还包括各种劳务价格,“劳务”是活动的过程。

构成总体的个别事物,称为总体单位。要了解总体的数量特征,就必须要从总体单位一个个的登记开始。例如,我们要研究我国建筑施工企业的效益情况,那么全国的建筑施工企业就是一个总体,各个建筑施工企业就是总体单位;要了解全国建筑施工企业的经济效益情况,就得从这些单位取得资料,加以汇总整理,得到我们所要了解的情况。各个建筑施工企业所以能够构成总体,是因为它们都有建筑施工生产经营活动的职能,所以各个事物具有同质性是形成总体的基础。

总体有无限总体与有限总体之分。无限总体就是总体含有无限多的总体单位,多得不可计数。例如,我们要研究海洋中的鱼类,或是研究某地区的空气污染程度,则海洋中的鱼类、某地区的空气就是无限总体或可以看成无限总体。有限总体就是总体所包含的总体单位是有限的,可以计数的。例如,在某一时点上的人口或工业企业的数量所组成的总体,都是有限总体。

总体和总体单位不是固定的,总体的范围、包含的总体单位也不是固定不变的,它们都会随着统计研究目的的变化而改变。例如,要调查全国的工业企业的情况,则每个工业企业便是总体单位;若要研究一个工业企业内部劳动力构成情况,则一个工业企业就成了总体,而企业内部的每个职工就是总体单位。

尽管总体和总体单位是相对而言,但是作为总体,其四个特征:客观性、同质性、大量性、变异性都必须具备。

1.2.2 标志与指标

标志是说明总体单位所具有的属性或特征。每个总体单位有许多属性和特征,如企业中每一个职工作为总体单位考察时,有性别、民族、文化程度、年龄、工资等属性和特征。这些都是每个职工的标志,而它们在总体单位之间各有一定的具体表现,既有相同的,也有不同的。

标志按其性质不同,可以分为品质标志和数量标志。凡只能用文字表示、说明事物的品质属性的标志,如职工的性别、文化程度、民族等称为品质标志;凡是用数值表示、说明事物的数量特征的标志,如职工的年龄、工龄、工资等称为数量标志。

标志还可以分为可变标志和不变标志。标志在总体各单位之间的具体表现不完全相同,如职工的年龄、性别、工资等称为可变标志,这是统计认识事物的内容和根据。若在总体各单位中的具体表现完全相同的标志,如企业中每个工人的社会成分都是工人,则称此类标志为不变标志。任何总体中的各个总体单位至少有一个不变标志,以体现总体的同质性。

统计指标(简称指标)是综合反映统计总体数量特征的概念和数值,表明某一客观事物在具体时间、地点条件下的规模、水平,如一个国家在某一特定时点的国民生产总值、土地面积、人口总数、职工人数等。统计指标具有三个特点:(1)数量性。即指标都是用数值来表现的。(2)综合性。统计指标是对总体单位某一特征进行调查、登记并加以汇总整理,而得到的数据构成总体单位的综合结果,不是说明个别总体的数量特征。(3)具体性。统计指标是说明总体某一特征或属性的质与量的统一,是在一定时间、地点、条件下的具体数量的表现。

统计指标由指标名称和指标数量组成,两者既有联系又有区别,指标名称是统计所研究现象某一特征的科学概念,是对现象的质的规定,反映其内容所属的范畴;指标数值是统计所研究对象某一特征具体数值的综合结果,对现象特征从数量上加以说明,是对统计指标的量的规定。统计指标名称和指标数值有机结合,辩证统一地反映客观现象的质与量。

指标与标志既有联系又有区别。两者的联系是:(1)总体单位的某一标志与总体的某一统计指标的名称往往是一致的。例如,工业总产值是工业企业的一个标志,同时也是全国、地区或部门工业总产值指标的名称。(2)指标与标志之间存在着变换关系,由于研究目的不同,原来的统计总体变成统计单位,或原来的总体单位变成统计总体,则相应的统计指标变成标志,或相应的统计标志就要变成统计指标。(3)统计指标的数值是由各总体单位的标志值汇总或计算得来的。例如,全国的工业职工平均工资这一统计指标的数值,是根据所有工业企业的“工资总额”和“职工人数”这两个标志的数值计算出来的。

指标和标志又有明显的区别:(1)统计指标是说明整体特征的,标志是说明总体单位特征的。(2)因为统计指标都有一定的数值,所以只有数量标志才能形成统计指标。品质标志不能形成指标,是由于它不能用数值表示。(3)实际工作中只用指标而不用标志这个名词,标志只是统计的一个概念。

统计指标按其反映的数量特征不同,可以分为数量指标和质量指标。数量指标是反映现象总体规模大小、数量多少的总量指标,一般用绝对数表示,如职工人数、国民生产总值、企业个数等。质量指标是表明现象总体质量的指标,反映现象的相对水平或工作质量,一般用相对数或平均数表示,如劳动生产率、资金利润率等。

统计指标按其数值形式不同,可分为总量指标、相对指标、平均指标。总量指标(或称绝对指标)是说明现象规模、水平或工作总量的指标,如工资总额、产品产量等。相对指标是两个有联系的总量指标相比的结果,说明现象总体的结构、发展程度的指标,如产品的合格率、发展速度等。平均指标是按总体某些数量标志值计算的说明总体一般水平的统计指标,如平均工资、劳动生产率等。

统计指标按其计量单位不同,可分为实物指标和价值指标。实物指标用实物计量单位,直接反映产品使用价值的数量。价值指标是用货币计量的计算指标。

统计指标按其作用不同,可以分为描述指标、评价指标和预警指标。描述指标是用来反映现象基本情况的指标,如社会劳动力资源总数、外汇储备数、在校学生人数等。评价指标是用来对客观现象活动的结果进行评估和考核的指标,比如工业企业经营活动效益评价的产品销售率、流动资金周转速度、劳动生产率等指标。预警指标是对现象宏观运行进行监测,并据此对可能出现的总体失衡、结构性矛盾、突发异常情况做出预报的指标,如通货膨胀率、固定资产投资增长率、失业率、人口增长率等,这类指标涉及面广,敏感性强,对国民经济的发展和社会稳定具有重要作用。

客观现象是错综复杂的,各种现象之间存在着相互联系、相互制约的关系。单个统计指标只能反映总体某一方面的特征。为了全面地说明现象的发展过程和它的各个方面,就需要一整套统计指标。这种具有内在联系、相互制约的一系列统计指标的整体,称为统计指标体系。如为研究工业企业生产经营的全貌,需要设置产品产量、产值、品种、质量指标,还要设置固定资产、流动资产占用额、资金周转速度、创利水平、上缴税金等财务指标等。通过一套完整的指标体系,把企业的全貌和发展变化过程及其因果关系反映出来。

统计指标体系可以分成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反映国民经济和社会发展基本情况的指标体系,它包括社会指标体系、经济指标体系和科技指标体系。专题统计指标体系是针对社会或经济问题而制定的专题指标体系,如企业经济效益指标体系、物价指标体系等。

统计指标体系能全面地反映现象之间的有机联系和发展过程,用以研究客观

现象,并做出全面客观的分析判断,避免片面性。统计指标体系会随着各种客观现象的发展变化而变化,但指标体系一经制定,应力求保持相对稳定,以便积累历史资料,进行系统的比较分析。

1.2.3 变异与变量

总体单位之间标志的不同表现,称为变异。包括品质标志和数量标志的差别,如表现在不同人身上的男女性别、年龄大小、文化程度高低等,都称为变异。

可变的数量标志,则称为变量。其在总体各单位所表现的标志值,就是变量值。

变量按其数值形式不同,可以分为离散型变量和连续型变量。离散型变量一般取整数值,如职工人数、企业个数、设备台数等。连续型变量可以对变量值作无限分割,取任意小数,如人的身高、体重、产值、利润等。

变量按其所受因素的影响不同,可分为确定性变量和随机性变量。受确定性因素影响的变量称为确定性变量。这种影响变量值变化的因素是明显的,是可以解释的,是人为的或者受人控制的,其影响变值的大小、方向都可以确定。如产品总成本的变化,无非是受产品产量和单位成本两个因素的影响,而这两者都是人为可以控制的变量,并且对生产总成本影响的大小和方向也是确定的。受随机性因素影响的变量称为随机变量,所谓随机因素,是指各种不确定的、偶然性的因素,这种因素对变量影响的大小和方向都是不确定的,通常是微小的。如某些机器零配件的尺寸大小,即使在完全相同的情况下进行生产,也总会存在差异,究其原因,可能与原材料的质量,电压的高低,气温、湿度的变化,操作者的情绪,甚至太阳黑子的爆炸等多种不确定的、偶然因素变化的影响有关。在客观现象中,既有确定性变量,也有随机性变量;影响其变量变化的因素有确定性因素,也有随机性因素。统计研究要按其目的和要求,作不同处理。

小 结

本章主要介绍了在自然科学和社会科学的诸多领域中,通过统计调查,将搜集的资料进行整理和分析,以综合指标的形式反映客观现象在一定时间、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,为社会和各有关部门提供信息和咨询。不仅阐述了统计的产生和发展,还叙述了统计研究的特点和方法,提出了统计学中几个基本的概念。

练习题

1. “统计”一词有哪几种含义？它们之间有何关系？
2. 统计学科的性质是什么？它有哪些特点？
3. 试举例说明什么是总体、总体单位、变量、变量值、统计指标、标志、统计指标体系。
4. 试说明标志与指标之间的联系与区别。