

高等院校统计学系列教材

统计学

——理论·实务·案例（第三版）

主编 赵振伦 李洪静



立信会计出版社

LIXIN ACCOUNTING PUBLISHING HOUSE

高等院校统计学系列教材

统计学

——理论·实务·案例（第三版）

主编 赵振伦 李洪静



立信会计出版社

LIXIN ACCOUNTING PUBLISHING HOUSE

图书在版编目(CIP)数据

统计学:理论·实务·案例/赵振伦,李洪静主编.

—3 版. —上海:立信会计出版社,2017. 7

高等院校统计学系列教材

ISBN 978-7-5429-5531-9

I. ①统… II. ①赵… ②李… 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(2017)第 191762 号

策划编辑 赵志梅
责任编辑 赵志梅
封面设计 南房间

统计学——理论·实务·案例 Tongjixue

出版发行	立信会计出版社	
地 址	上海市中山西路 2230 号	邮政编码 200235
电 话	(021)64411389	传 真 (021)64411325
网 址	www.lixinaph.com	电子邮箱 lxaph@sh163.net
网上书店	www.shlx.net	电 话 (021)64411071
经 销	各地新华书店	

印 刷	上海肖华印务有限公司	
开 本	787 毫米×960 毫米	1/16
印 张	30	
字 数	560 千字	
版 次	2017 年 7 月第 3 版	
印 次	2017 年 7 月第 1 次	
印 数	1—3 100	
书 号	ISBN 978-7-5429-5531-9/C	
定 价	48.00 元	

如有印订差错,请与本社联系调换

第三版前言

统计学是一门比较难学的学科,记住公式,不等于理解统计;学会计算,不等于会用统计。统计的真谛在于它所体现的思想,在于它所提供的思维方式。

随着经济的快速增长、社会的不断发展以及经济指标体系的不断完善和更新,原有的统计学教材个别章节已不适应形式发展的需要,有些指标、理论计算手段需要更新,这就是改版的初衷。

第三版的出版仍然感谢立信会计出版社,感谢编辑和有关同志的辛勤劳动。感谢书后所列参考文献的作者,本书的部分素材选自所列参考文献。

书中疏漏之处,恳请读者指正。

编 者

前 言

随着经济改革的不断深入发展,统计的科学理论和分析方法在企业中得到了广泛应用。本人编写此书的目的,是为了给大、中专院校经济类专业学生以及企、事业单位统计人员、经济管理人员在学习和工作中多提供一些参考。

本书内容通俗易懂,案例较为生动、直观,理论与实际紧密结合。它的出版将有利于案例教学的发展,有利于统计学科教学质量的提高,使读者进一步明确在企业经济活动中该怎样应用科学的统计分析方法来进行统计分析。

本书第二编所征集的企业经济统计分析案例,本身就是统计工作的成果。从案例中可以看出,企业统计人员充分发挥了统计提供信息、实行监督、提供咨询、参与决策的功能,为企业创造更好的经济效益起到了良好的作用。

在编写过程中,由于本人水平所限,疏漏之处在所难免,恳请专家、读者批评指正。在此,也向案例提供者表示感谢!

编 者

目 录

第一编 统计理论和方法

第一章 绪论	3
第一节 统计及其应用领域	3
第二节 统计数据类型	5
第三节 统计学中的几个基本概念	6
本章要点	14
本章习题	14
第二章 统计数据的搜集	16
第一节 统计调查的概念、作用和种类	16
第二节 统计调查方案的设计	20
第三节 统计调查的基本方法	24
本章要点	27
本章习题	27
第三章 统计数据的整理	28
第一节 资料整理的一般问题	28
第二节 统计分组	31
第三节 分配数列	35
第四节 统计数据的展示	39
本章要点	45
本章习题	45
第四章 静态分析指标——数据描述	47
第一节 绝对指标	47
第二节 相对指标	50
第三节 分布的集中趋势	56

第四节 分布的离中趋势	75
本章要点	88
本章习题	89
 第五章 概率与概率分布	91
第一节 随机事件与概率	91
第二节 随机变量及其分布	97
第三节 大数定律与中心极限定理	106
本章要点	107
本章习题	107
 第六章 抽样推断	109
第一节 抽样推断概述	109
第二节 抽样调查中的几个基本概念	111
第三节 抽样调查的理论基础、组织形式及调查方法	113
第四节 抽样误差和抽样估计	117
第五节 抽样调查的组织方式及其误差的计算	130
本章要点	141
本章习题	142
 第七章 假设检验与方差分析	146
第一节 假设检验的基本思想和步骤	146
第二节 正态总体参数的检验	149
第三节 总体成数的假设检验	154
第四节 单因素试验的方差分析	155
本章要点	159
本章习题	159
 第八章 时间数列分析	161
第一节 时间序列的对比分析	161
第二节 长期趋势的测定	175
第三节 季节变动的测定	188
第四节 循环变动和不规则变动的测定	196
本章要点	199

本章习题	199
第九章 统计指数	202
第一节 指数的概念和种类	202
第二节 综合指数	204
第三节 平均指数	217
第四节 指数体系与因素分析	229
第五节 指数数列	236
第六节 减缩指数	241
第七节 常用经济指数	244
本章要点	250
本章习题	251
第十章 相关分析与回归分析	256
第一节 相关关系的概念和种类	256
第二节 一元线性回归方程与相关系数	260
第三节 回归方程的显著性检验与区间估计	270
第四节 一元非线性回归方程与相关系数	277
第五节 自相关与自回归分析	283
第六节 多元线性回归与相关分析	293
本章要点	303
本章习题	304
第十一章 Excel 在统计学中的应用	307
第一节 Excel 概述	307
第二节 Excel 在描述统计中的应用	309
第三节 Excel 在相关回归分析中的应用	316
本章要点	320
本章习题	321

第二编 企业统计分析案例

第十二章 市场统计分析案例	325
---------------	-----

第十三章 营销统计分析案例·····	332
第十四章 企业生产要素统计分析案例·····	345
第十五章 企业产品质量统计分析案例·····	359
第十六章 企业技术进步统计分析案例·····	366
第十七章 企业资金、成本、利润统计分析案例·····	383
第十八章 企业扭亏增盈统计分析案例·····	398
第十九章 企业综合经济效益统计分析案例·····	416
第二十章 企业实现经济增长方式转变统计分析案例·····	428
附录一 术语表·····	446
附录二 用 Excel 生成概率分布表·····	454
主要参考文献·····	468

第一编

统计理论和方法

第一章 绪 论

学习目标 通过本章的学习,了解统计学的含义、统计学的产生与发展;正确地理解统计学的对象与特点;掌握统计学的几个基本概念;把握统计指标设计的一般问题,并能根据研究目的进行简单的统计指标体系的框架设计。

关键概念 统计(statistics) 标志(symbol) 指标(index) 指标体系(index system) 总体(population) 变量(variable) 总体单位(population unit) 样本(sample) 参数(parameter) 统计量(statistic)

本章将介绍统计学的一些基本问题,包括统计学的含义、统计数据及其分类、统计中常用的一些基本概念等。

第一节 统计及其应用领域

一、什么是统计学

统计是处理数据的一门科学。人们给统计学下的定义很多。比如,“统计学是收集、分析、表述和解释数据的科学”;“统计是一组方法,用来设计实验、获得数据,然后在这些数据的基础上组织、概括、演示、分析、解释和得出结论”。综合地说,统计学(statistics)是收集、处理、分析、解释数据并从数据中得出结论的科学。

统计学是关于数据的科学,它提供的是一套有关数据收集、处理、分析、解释并从数据中得出结论的方法,统计研究的是来自各领域的的数据。数据收集也就是取得统计数据;数据处理是将数据用图表等形式展示出来;数据分析则是选择适当的统计方法研究数据,并从数据中提取有用信息进而得出结论。

数据分析所用的方法可分为描述统计方法和推断统计方法。描述统计(descriptive statistics)研究的是数据收集、处理、汇总、图表描述、概括与分析等统计方法。推断统计(inferential statistics)是研究如何利用样本数据来推断总体特征的统计方法。比如,要了解一个地区的人口特征,不可能对每个人的特征一一进行测量;对产品的质量进行检验,往往是破坏性的,也不可能对每个产品进行测量。这就需要抽取部分个体即样本进行测量,然后根据获得的样本数据对所研究的总体特征进行推断,这就是推断统计要解决的

问题。

二、统计的应用领域

统计方法是适用于所有学科领域的通用数据分析方法,只要有数据的地方就会用到统计方法。随着人们对定量研究的日益重视,统计方法已被应用到自然科学和社会科学的众多领域,统计学也已发展成为由若干分支学科组成的学科体系。可以说,几乎所有的研究领域都要用到统计方法。比如,政府部门、学术研究领域、日常生活中、公司或企业的生产经营管理中都要用到统计。下面将给出统计在经济管理中的一些应用。

1. 企业发展战略

发展战略是一个企业的长远发展方向。制定发展战略一方面需要及时了解和把握整个宏观经济的状况及发展变化趋势,了解市场变化;另一方面还要对企业进行合理的市场定位,把握企业自身的优势和劣势。所有这些都离不开统计,需要统计提供可靠的数据,利用统计方法对数据进行科学的分析和预测等。

2. 产品质量管理

质量是企业生命,是企业持续发展的基础。质量管理中离不开统计的应用。在一些知名跨国公司,6 σ 准则已成为一种重要的管理理念。质量控制已成为统计学在生产领域的一项重要应用。各种统计质量控制图被广泛用于监测生产过程。

3. 市场研究

企业要在激烈的市场竞争中取得优势,首先必须了解市场。要了解市场,则需要作广泛的市场调查,取得所需的信息,并对信息进行科学的分析,以便作为生产和营销的依据。这些都需要统计的支持。

4. 财务分析

上市公司的财务数据是股民投资选择的重要参考依据。一些投资咨询公司主要是根据上市公司提供的财务和统计数据进行分析,为股民提供投资参考。企业自身的投资,也离不开对财务数据的分析,其中要用到大量的统计方法。

5. 经济预测

企业要对未来的市场状况进行预测,经济学家也常常要对宏观经济或某一方面进行预测。在进行预测时要使用各种统计信息和统计方法。比如,企业要对产品的市场潜力作出预测,以便及时调整生产计划,这就需要利用市场调查取得数据,并对数据进行统计分析。经济学家在预测通货膨胀时,要利用有关生产价格指数、失业率、生产能力利用等统计数据,然后通过统计模型进行预测。

6. 人力资源管理

利用统计方法对企业员工的年龄、性别、受教育程度、工资等进行分析,并作为企业制定工资计划、奖惩制度的依据。

当然,统计并不是仅仅为了管理才有用,它是为自然科学、社会科学的多个领域而发展起来的,为多个学科提供了一种通用的数据分析方法。从某种意义上说,统计仅仅是一种数据分析的方法,与数学一样,统计是一种工具,它是一种数据分析的工具。

利用统计方法可以简化繁杂的数据。比如,用图表展示数据,建立数据模型,有人认为统计的全部目的就是让人看懂数据,其实这仅仅是统计的一个方面,统计更重要的功能是对数据进行分析,它提供了一套分析数据的方法和工具。不同的人对数据分析的理解也会不一样,曲解数据分析是一种常见的现象。在有些人的心目中,数据分析就是寻找支持,他们的心目中可能有了某种“结论”性的东西,或者说他们希望看到一种符合他们需要的某种结论,而后去找些统计数据来支持他们的结论。这恰恰歪曲了数据分析的本质:数据分析的真正目的是从数据中找出规律,从数据中寻找启发,而不是寻找支持。真正的数据分析事先是没有结论的,通过对数据的分析才能得出结论。统计不是万能的,它不能解决你所面临的所有问题。统计可以帮助分析数据,并从分析中得出某种结论,但对统计结论的进一步解释,则需要你的专业知识。比如,吸烟会使患肺癌的几率增大,这是一个统计结论,但要解释吸烟为什么能引起肺癌,这就不是统计学家所能做到的,需要有更多的医学知识才行。

第二节 统计数据类型

统计数据是对现象进行测量的结果。比如,对经济活动总量的测量可以得到国内生产总值(GDP)数据;对股票价格变动的测量可以得到股票价格指数的数据;对人口性别的测量可以得到男或女这样的数据。下面从不同角度说明统计数据的分类。

一、分类数据、顺序数据、数值型数据

按照所采用的计量尺度不同,可以将统计数据分为分类数据(categorical data)、顺序数据(rank data)和数值型数据(metric data)。

(1) 分类数据是只能归于某一类别的非数字型数据,它是对事物进行分类的结果,数据表现为类别,是用文字来表述的。例如,人口按照性别分为男、女两类;企业按行业属性分为医药企业、家电企业、纺织品企业等,这些均属于分类数据。为便于统计处理,对于分类数据可以用数字代码来表示各个类别。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。

(2) 顺序数据是只能归于某一有序类别的非数字型数据。顺序数据虽然也是类别,但这些类别是有序的。比如,产品可以分为一等品、二等品、三等品、次等品;考试成绩可以分为优、良、中、及格、不及格等;一个人的受教育程度可以分为小学、初中、高中、大学及以上;一个人对某一事物的态度可以分为非常同意、同意、保持中立、不同意、非常不同意,

等等。同样,对顺序数据也可以用数字代码来表示。比如,用 1 表示“非常同意”,用 2 表示“同意”,用 3 表示“保持中立”,用 4 表示“不同意”,用 5 表示“非常不同意”。

(3) 数据型数据是按数字尺度测量的观察值,其结果表现为具体的数值。现实中所处理的大多数数据都是数值数据。

分类数据和顺序数据说明的是事物的品质特征,通常是用文字来表述的,其结果均表现为类别,因而也可统称为定性数据或品质数据(qualitative data);数值型数据说明的是现象的数量特征,通常是用数值来表现的,因此也可称为定量数据或数量数据(quantitative data)。

二、观测数据和实验数据

按照统计数据的收集方法,可以将其分为观测数据(observational data)和实验数据(experimental data)。

(1) 观测数据是通过调查或观测而收集到的数据,这类数据是在没有对事物人为控制的条件下得到的,有关社会经济现象的统计数据几乎都是观测数据。

(2) 实验数据则是在实验中控制实验对象而收集到的数据。比如,对一种新药疗效的实验数据,对一种新的农作物品种的实验数据。自然科学领域的大多数数据都为实验数据。

三、截面数据和时间序列数据

按照被描述的现象与时间的关系,可以将统计数据分为截面数据(cross-sectional data)和时间序列数据(time series data)。

(1) 截面数据是在相同或近似相同的时间点上收集的数据,这类数据通常是在不同的空间上获得的,用于描述现象在某一时刻的变化情况。比如,2011 年我国各地区的国内生产总值数据就是截面数据。

(2) 时间序列数据是在不同时间上收集到的数据,这类数据是按时间顺序收集到的,用于所描述现象随时间变化的情况。比如,2006 年至 2011 年我国的国内生产总值数据就是时间序列数据。

第三节 统计学中的几个基本概念

统计学和其他科学一样,也有它的研究范畴和概念。本节介绍的是统计学中常有的几个基本概念,其他概念将在后面各有关章节的内容中予以叙述。

一、统计总体和总体单位

统计总体和总体单位,简称总体和个体,是反映统计认识对象的基本概念。

统计总体是由客观存在的某些性质上相同的许多个体所组成的,组成统计总体的个体称为总体单位。例如,某大学所有学生组成一个总体,每个学生就是总体单位。

统计总体的形成必须具备一定条件,其客观条件主要有三条:第一,客观性,即总体和总体单位必须是客观存在的,可以实际观察的。任何主观臆想的东西,都构成不了统计总体。第二,同质性,即组成总体的所有个体必须在某些性质上是相同的。例如,我国的人口普查统计总体包括的每一个人,都具有中国国籍和居住在中国境内的共同性质;工业企业总体,必须是由工业生产经营的基层单位组成的。但是,总体的同质性并不是固定不变的,根据研究的目的不同,同质性也会随之变化。按某一种属性划分,总体可能是同质的,而按另一种属性划分,则可能是不同质的。例如,某地区全部国有工业企业,若按经济类型划分是同质的,但若按行业划分,就不能看成是同质的。第三,差异性,即构成总体的各单位除了同质性一面还必须有异性,否则就不需要进行统计调查研究了。例如,职工这个总体中的每个职工,除了有性质相同的一面,还有差异的一面,如工种,性别,年龄、文化程度、工资等,这样才构成社会统计调查研究的内容。

统计总体的形成还取决于主观条件,这主要有两个方面:

第一,统计研究的目的。统计研究的具体目的不同,作为认识对象的总体和个体便有所不同。例如,研究一个国家或地区的工业生产情况时,将这个国家和地区的每个企业作为总体单位,而在研究一个工业企业的生产情况时,则将该企业的每个车间或者每个班组作为总体单位。

第二,统计机构的状况。统计机构的立场、观点、工作条件、了解实际的深入程度等也决定了统计总体的形成。例如,要调查研究煤矿井下工人的健康状况,如果人力和经费充足,可以把全国煤矿的全体井下工人都列为总体单位形成统计总体;但是,若人力与经费有限,则可以选定大、中、小型煤矿各若干个,以其井下工人为总体单位形成统计总体。

总体与总体单位是既有联系又有区别的不同的概念,但是它们也会因研究对象不同而相互转化。同一个研究对象,在一种情况下是总体,在另一种情况下则变成了总体单位。例如,在研究某高等院校在校学生人数时,该省所有的高等院校是总体,省内每所高等院校是总体单位;而当研究的是某一院校内各系的在校学生人数时,则该院校被看成是总体,而该校的各系则是总体单位。

对于以实物形态存在的总体来说,可以分为有限总体与无限总体,大总体与小总体,可加总体与不可加总体等。如果总体中所包括的单位数是有限的,则称为有限总体;如果总体中包括的单位数是无限的,则称为无限总体。社会经济现象从理论上来说,都是有限总体,但倘若其总体单位数在某一有效的时间内不能一一列举出来,也可视其为无限总体,如电子元件产品、粮食产量等。有限总体与无限总体的划分,有助于选择合适的调查方法和计量方法。总体一经确定,则该总体可称为大总体,而大总体的各组成部分(即组)

则称为小总体,这样做有助于对指标和标志的理解。如果总体单位可以合计在一起,则该总体称为可加总体,如人数、播种面积等;如果总体单位不能合计在一起,则该总体称为不可加总体,如产品产量、固定资产原值等。不可加总体不能计算总体单位数,只能计算各种标志总量。

二、统计标志与变量

(一) 标志

在统计中,总体单位所具有的属性或特征称为标志。例如,每个工人都具有性别、工作、文化程度以及年龄、工龄、工资等属性特征,这些就是工人总体中各个单位的标志。就工业企业来说,企业的经济类型、主管系统、产品产量、固定资产原值等,也是工业企业总体中各个单位的标志。

由此可见,总体由总体单位构成,而总体单位是标志的载体和承担者,统计研究是从登记标志开始,并通过对标志的研究来综合反映总体的数量特征。因此,标志是统计研究的起点。

标志可分为品质标志与数量标志。工人的性别、工种、文化程度、企业的经济类型、主管系统等,是以品质方面的差别来表现的,这类标志称为品质标志;而工人的工龄、工资、企业的工人数、产品产量、固定资产原值等,则是以数量的多少来表现的,这类标志称为数量标志。

一般情况下,同一总体中的各个单位存在着许多标志,其中有的标志在各单位之间的表现是共同的,根据这样的标志将所有单位结合在一起,构成总体同质性的基础。例如,在国有工业企业总体中,在经济类型这一标志的表现上,各个企业都是相同的,这样的标志称为不变标志。除此之外,大多数的标志在各单位之间的表现是不同的,这样的标志称为变异标志。例如,上面所列举的生产工人与工业企业的各项标志就是变异标志。在统计中,变异标志是普遍的、大量的,这是由客观事物所具有的属性决定的。标志在各单位之间的变化称为变异,变异也是统计中的一个重要的概念。

(二) 变量

顾名思义,变量是指可变化的量。变量按反映的对象不同可以分为三种类型,即数值型变量、分类型变量和顺序型变量。

(1) 数值型变量就是可变化的数量标志。所有定量变量(数据)都属于数值型变量。

(2) 分类型变量是用于区分单元或个体属于不同类别组别的变量,如性别、种族、宗教、颜色都是典型的分类型变量。

(3) 顺序型变量是指能够反映单元或个体属于不同等级或顺序的变量。

需要说明的是,数值型变量也包括另外两种类型变量(数据)的数量变种。例如,0~1变换就是将一些分类型变量(数据)转化为数值型变量的常规工具。事实上,各种类型的