

21世纪
经济管理类教材

全国统计教材编审委员会

推荐使用教材

国家精品课程

统计学

【第二版】

Statistics

陈珍珍 主 编
罗乐勤 副主编
黄良文 主 审



厦门大学出版社

XIAMEN UNIVERSITY PRESS

中国统计出版社出版

国家统计局编

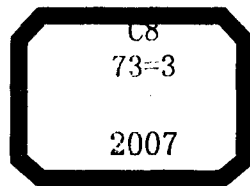


统计学

Statistics

中国统计出版社
北京

中国统计出版社
北京



21 世纪经济管理类教材

全国统计教材编审委员会推荐使用教材

国 家 精 品 课 程

统 计 学 (第二版)

Statistics

陈珍珍 主 编
罗乐勤 副主编
黄良文 主 审

厦门大学出版社

图书在版编目(CIP)数据

统计学/陈珍珍主编,罗乐勤副主编.—2版.—厦门:厦门大学出版社,
2007.2

21世纪经济管理类教材(全国统计教材编审委员会推荐使用教材)

ISBN 978-7-5615-1839-7

I.统… II.①陈…②罗… III.统计学-高等学校-教材 IV.C8

中国版本图书馆 CIP 数据核字(2001)第 094517 号

厦门大学出版社出版发行

(地址:厦门大学 邮编:361005)

<http://www.xmupress.com>

xmup@public.xm.fj.cn

厦门昕嘉莹印刷有限公司印刷

(厦门市前埔东路 555 号 邮编:361009)

2007年2月第2版 2007年2月第1次印刷

开本:787×960 1/16 印张:20.5

字数:358千字 印数:1~11000册

定价:26.00元

如有印装质量问题请与承印厂调换

第一版前言

随着我国社会主义市场经济的日臻完善,无论在宏观的经济调控领域还是在微观的企业管理领域中,人们要进行高效的监控和科学的管理,就必须准确及时地获得经济运行中的各类信息。统计学是一门提供数据信息的搜集、处理、归纳和分析的理论与方法的科学。随着社会的发展,统计的运用领域越来越广泛,不管是在经济管理领域,还是在军事、医学、生物、物理、化学等领域的研究中,人们对数量分析与统计分析都提出更高的要求。尤其在知识经济时代,人们对新知识的需求与日俱增,为适应市场经济对人才的需求,把握统计学为认识社会规律服务的方向,统计教育必须与知识经济时代相适应,统计教材的改革是关键。本书正是为适应新形势发展的需要而编著的,它可作为高等院校经济类、管理类专业统计学基础课的教材,也可作为从事统计工作人员的参考书。

统计学是经济类、管理类的核心课程,本书的编写是根据统计教学大纲的要求,在总结多年来的教学经验、并参阅大量的国内外相关资料的基础上完成的。参与本教材编著的全体成员,长期致力于统计教学与统计研究工作,探索出对经济类、管理类的学生进行统计基础教育的比较成功的道路。在内容安排上,贯彻“大统计”学科建设的思想,针对客观实际中存在的大量不确定现象,在书中对随机现象的统计处理引进数理统计方法,力求使社会经济统计与数理统计融为一体。在体系设计上,本书是根据人们认识客观现象的顺序进行安排的。具体章节安排如下:第一章总论,介绍统计学的基本轮廓;第二章统计调查、整理和数据显示,第三章统计分布的数值特征,研究的是如何搜集资料,进行数据处理并描述现象的数量特征,这属于一般统计研究理论与方法;第四章参数估计,第五章假设检验,研究的统计推断问题,这是属于随机现象的统计处理理论与方法;第六章相关分析与回归分析,第七章时间序列分析与预测,第八章统计指数与因素分析,第九章统计综合分析与评价,研究的是在各具体领域中统计分析方法的运用。该体系的设计强调了统计为认识社会现象的特征与规律服务,使统计的目标更加明确。为了能深入浅出地阐明统计原理,本书搜集大量的实例,并结合 EXCEL 软件的运用,从实例分析入手,阐明数理分析的方法,这有利于教学手段的现代化,增强学生学习的兴趣和提

STATISTICS

统计学

高学习效率。对书中打星号的部分,任课教师可根据具体情况酌情选择授课内容。

全书编写的具体分工如下:第一章,陈珍珍;第二章,陈珍珍;第三章,罗乐勤;第四章,朱平辉;第五章,朱平辉;第六章,皇甫秀颜;第七章,罗乐勤;第八章,罗乐勤;第九章,陈珍珍;附录和 EXCEL 的运用,袁加军、朱平辉。本书由陈珍珍教授担任主编,罗乐勤教授担任副主编,负责全书的设计、审订、修改、总纂和定稿工作。

本书借鉴了黄良文教授的研究成果,在编纂过程中得到厦门大学计统系领导、厦门大学出版社的大力支持,黄良文教授对本书进行了认真细致的审阅,袁加军、林飞同志在本书的排版中做了大量的工作,在此,我们表示衷心的感谢。

由于编著者水平有限,加之时间仓促,书中难免有疏漏或错误之处,恳请同行专家和读者不吝赐教,以便我们修改与完善。

编著者

2001 年 12 月

第二版前言

本书是首届福建省精品课程“统计学”立项教材,其内容体系是根据培养“宽口径、厚基础、重应用、高素质”的经济类、管理类本科生的教学目标而设计的。本书自 2002 年出版以来,受到教师和学生的普遍欢迎,已经过 6 次印刷,总印数达到 3 万多册。

统计是信息时代人们迅速、准确地获得各类信息,处理各种数据,认识事物数量特征及其变化规律的工具。随着社会主义市场经济的发展,我国经济形势日新月异,因而社会各个方面对统计提出了更新、更高的要求,统计教材改革需要进一步深化。我们认为,如何更有效地提高统计的认识能力,关键就是要善于运用统计理论与方法去把握不确定性的现象的数量规律。在各个领域里,推断统计的运用越来越广泛。根据以上思路和近年在教材使用中的积累经验,我们推出了本书第二版。第二版传承了原版本“少而精”和“学以致用”的优点,对原有版本的章节进行了适当的调整,并增加了一些推断统计的内容。在第二版的修订中,选择了社会关心的热点问题和最新统计数据来设计案例和习题,使课堂教学与社会实际的联系更加紧密,以利于提高学生学以致用用的能力。

参加本书修订工作的有:陈珍珍教授(第一章、第二章、第十章)、罗乐勤教授(第三章、第八章、第九章)、朱平辉副教授(第四章、第五章、第六章)、皇甫秀颜博士(第七章)、袁加军讲师(附录和附表)。陈珍珍教授和罗乐勤教授负责全书的审订、总纂、修改和定稿工作。

多年来,本书能够顺利地出版和再版,一直得到厦门大学出版社和厦门大学计划统计系有关领导的大力支持。在本书使用过程中,不少兄弟院校对该书的修订方案提出宝贵的意见。这次修订,王建军博士也付出了辛勤的劳动。在此,我们表示衷心的感谢。

由于编著者水平有限,书中难免有不完善乃至错误之处,恳望同行专家和读者不吝赐教。

陈珍珍

2006 年 12 月于厦门珍珠湾花园

STATISTICS



第一版前言

第二版前言

第一章 总论	1
第一节 统计、统计学和统计数据	1
统计的含义	
统计学的研究对象及其特点	
统计数据的主要类型	
变量及其分类	
第二节 统计学的基本概念	5
统计总体与样本	
总体单位与标志	
统计描述和统计推断	
第三节 统计学与其他学科的关系	9
统计学与会计学的关系	
统计学与计量经济学的关系	
统计学与数学的关系	
第二章 统计数据的搜集、整理和显示	14
第一节 统计数据的搜集	14
统计数据搜集的基本理论与方法	
统计调查的组织形式	
统计调查体系	
统计调查误差	
第二节 数据整理	23
数据整理的内容与程序	
统计分组	

STATISTICS

统计学

* * * * *	统计指标	
* * * * *	第三节 频数分布	33
* * * * *	频数分布的基本概念	
* * * * *	变量数列的编制	
* * * * *	累计频数与累计频率	
* * * * *	频数分布的类型	
* * * * *	第四节 数据显示	38
* * * * *	统计表	
* * * * *	统计图	
* * * * *	第三章 统计分布的数值特征	52
* * * * *	第一节 分布的平均水平、集中趋势和位置的度量	52
* * * * *	统计平均数的含义与作用	
* * * * *	数值平均数	
* * * * *	众数、中位数和其他分位数	
* * * * *	第二节 分布的离散程度	70
* * * * *	变异指标的含义与作用	
* * * * *	极差与四分位差	
* * * * *	平均差	
* * * * *	方差与标准差	
* * * * *	变异系数	
* * * * *	箱线图在统计描述中的运用	
* * * * *	第三节 分布的偏度和峰度	80
* * * * *	统计动差	
* * * * *	偏度	
* * * * *	峰度	
* * * * *	第四章 概率、概率分布和抽样分布	87
* * * * *	第一节 随机事件与概率	87
* * * * *	随机事件	
* * * * *	概率	
* * * * *	第二节 随机变量的概率分布	90
* * * * *	随机变量的概念	

11.1.1	离散型随机变量的概率分布	
11.1.2	连续型随机变量的概率分布	
11.2	抽样分布	96
11.2.1	抽样推断的基本概念	
11.2.2	重复抽样条件下样本均值和成数的抽样分布	
11.2.3	不重置抽样条件下样本均值和成数的抽样分布	
11.3	大数定理与中心极限定理	103
11.3.1	大数定理	
11.3.2	中心极限定理	
第五章	参数估计	107
15.1	第一节 点估计和区间估计	107
15.1.1	总体参数估计概述	
15.1.2	点估计	
15.1.3	区间估计	
15.2	第二节 参数估计中样本量的确定	114
15.2.1	估计总体均值时样本量的确定	
15.2.2	估计总体成数时样本量的确定	
15.2.3	在确定样本量时应注意的问题	
15.3	第三节 不同抽样组织形式下的参数估计	117
15.3.1	类型抽样形式下的参数估计	
15.3.2	整群抽样形式下的参数估计	
15.3.3	等距抽样形式下的参数估计	
15.3.4	不同抽样组织设计的比较	
第六章	假设检验	130
16.1	第一节 假设检验概述	130
16.1.1	假设检验的基本概念	
16.1.2	两种类型的错误	
16.1.3	检验功效	
16.2	第二节 总体参数检验	135
16.2.1	单侧检验与双侧检验	
16.2.2	总体平均数的检验	

* * * * *	总体成数的检验	
* * * * *	p -值检验	
* * * * *	第三节 非参数检验.....	142
* * * * *	自由分布检验概述	
* * * * *	符号检验	
* * * * *	秩和检验	
* * * * *	游程检验	
* * * * *	第七章 相关分析与回归分析.....	153
* * * * *	第一节 相关分析.....	153
* * * * *	函数关系与相关关系	
* * * * *	相关关系的种类	
* * * * *	相关图和相关表	
* * * * *	相关系数及其计算方法	
* * * * *	相关系数的检验	
* * * * *	第二节 简单线性回归分析.....	159
* * * * *	相关分析与回归分析的关系	
* * * * *	一元线性回归模型	
* * * * *	模型参数的估计	
* * * * *	一元线性回归模型的检验	
* * * * *	一元线性回归模型预测	
* * * * *	第三节 多元线性相关与回归分析.....	175
* * * * *	标准的多元线性回归模型	
* * * * *	多元线性回归模型的参数估计	
* * * * *	多元线性回归模型的检验和预测	
* * * * *	复相关系数和偏相关系数	
* * * * *	第八章 时间序列分析与预测.....	187
* * * * *	第一节 时间序列的水平分析.....	187
* * * * *	时间序列的概念与种类	
* * * * *	发展水平与平均发展水平	
* * * * *	增长水平与平均增长水平	

第二节	时间序列的速度分析·····	196
2.1.1	发展速度与增长速度	
2.1.2	平均发展速度与平均增长速度	
2.1.3	水平分析与速度分析的结合与应用	
第三节	时间序列的分解分析·····	203
3.1	时间序列的构成因素	
3.2	长期趋势的测定	
3.3	季节变动的测定	
3.4	循环变动的测定	
第四节	统计预测·····	219
4.1	加权最小二乘法	
4.2	指数平滑法	
4.3	三点法	
第九章 统计指数与因素分析·····		230
第一节	统计指数概述·····	230
1.1.1	统计指数的概念	
1.1.2	统计指数的种类	
1.1.3	统计指数的作用	
第二节	综合指数的编制与应用·····	232
2.1	综合指数的编制原理	
2.2	综合指数的编制方法	
2.3	综合指数的主要应用	
第三节	平均指数的编制与应用·····	241
3.1	平均指数的编制原理	
3.2	平均指数的编制方法	
3.3	平均指数的主要应用	
第四节	指数体系与因素分析·····	247
4.1	指数体系的概念	
4.2	连锁替代法	
4.3	平均指标变动的因素分析法	

STATISTICS

统计学

第十章 统计综合评价	259
第一节 统计综合评价概述	259
统计综合评价的概念	
统计综合评价的一般程序	
综合评价的局限性	
第二节 综合评价的常规处理方法	261
评价指标体系的建立	
同度量处理方法	
指标权重的确定方法	
评价项目的综合汇总	
第三节 模糊综合评价法	279
模糊综合评价的基本理论问题	
模糊综合评价的数学模型和步骤	
模糊综合评价的实际应用	
附录:Excel 在统计中的应用实例	291
附表 1 二项分布临界值表	304
附表 2 正态分布概率表	305
附表 3 t 分布临界值表	307
附表 4 X^2 分布临界值表	308
附表 5 F 分布临界值表($\alpha=0.05$)	309
附表 6 秩和检验表	311
附表 7 游程检验 R 临界值表	312
参考书目	313

第一章 总 论

* * * * *

第一节 统计、统计学和统计数据

一、统计的含义

何谓统计？人们在许多场合都接触过统计，例如，在证券市场上，投资者要预测某只股票价格的走势；在生产作业线上，质检人员抽检某些产品以估计批产品的合格品率；在药物开发领域里，药剂师想了解新研制的药物对某种疾病的治愈率及其是否值得推广等。统计为人们提供了获取相关信息、认识客观现象的工具。人们也常常从报纸杂志、电视新闻中获悉我国的经济增长速度、消费者价格指数或固定资产投资规模等经济数据资料。统计正渗入经济管理、科学实验和日常生活等各个领域。

统计这一词，在不同的场合，人们赋予它不同的含义。一般认为，统计的含义有三种：一是统计实践活动，二是统计数据，三是统计学。如上所列举的预测股票价格的变动趋势、估计批产品的合格率和检验新药物是否值得推广等，就是统计实践活动；经济增长速度、价格指数等，就是统计资料；而我们现在要学习的，就是统计学。

统计作为一种社会实践活动已有悠久的历史。据历史记载，我国在西周朝代已建立了统计报告制度，到秦建立中央集权的国家时，从中央到地方形成了比较完善的“上计”报告制度。统计被认为是治国创业图强的重要手段，正如春秋战国时管子所说的“举事必成，不知计数不可”。

统计，英文为 statistics，与“国家”、“状况”同一词根，可以说，自从有了国家，就有统计实践活动。最初，统计只是为统治者管理国家搜集资料、提供数

量依据。随着社会经济的发展,统计的应用领域越来越广泛,不仅仅只局限于经济管理领域,在军事、医学、生物、物理、化学和司法等领域中也大量地运用统计方法。统计实践活动,就是人们为认识客观事物,通过实验或调查搜集有关数据,并加以整理、归纳和分析,而后对客观事物规律性的数量表现做出统计上的解释。统计实践活动的过程实质上也是人们认识客观世界的过程。

人们通过统计实践活动所得的成果即统计数据,统计实践活动与统计数据的关系是工作过程与工作成果的关系。工作过程的好坏关系到工作成果质量的高低。人们对统计数据的要求是:客观性,即它能反映客观事实而不受任何偏见的影响或任何势力的干扰;准确性,即统计数据的偏差不能超过根据统计研究目的而事先确定的允许误差范围;及时性,即统计数据应及时搜集、及时加工、及时公布。

统计学是关于数据搜集、整理、归纳、分析的方法论的科学。统计学研究的是如何进行数据的搜集、加工和整理,如何从复杂纷繁的数据中得出结论,并科学地解释这个结论,以达到正确、深刻地认识客观现象的目的。统计学与统计实践活动的关系是理论与实践的关系,理论源于实践又高于实践,反过来又指导实践。

二、统计学的研究对象及其特点

统计学的研究对象亦即统计要认识的客体,一般说来,是指客观事物的数量特征和数量关系。统计学的研究对象具有如下特点:

(一)数量性

数据是统计的原料,离开了数字,统计工作就成了无米之炊。数量性是统计学研究对象的基本特点。统计数据是客观事物量的反映,通过数据以测度事物的类型、量的顺序、量的大小和量的关系。

(二)总体性

统计研究虽然是从个别入手,对个别单位的具体事实进行观察研究,但其目的是认识总体数量特征。例如,进行居民家计调查,虽然是对具体的每个调查户进行观察,但其目的并非研究个别居民户的家庭基本状况,而是要反映一个部门、一个地区乃至一个国家居民的收入、消费和就业等状况。

(三)变异性

统计研究的是同质总体的数量特征,其前提是总体各单位的特征表现存在着差异,而这些差异不是由某些特定的原因事先给定的。统计上把总体各单位由于随机因素引起的某一标志表现的差异称为变异。例如,一个学校的

学生,其年龄的大小、学习成绩的高低存在着差异,这才有必要研究其年龄、成绩的平均水平及其各层次的结构等状况。如果各单位不存在差异,也就无须进行统计了。

三、统计数据类型

(一)按数据的计量尺度分类

统计数据是对客观现象进行计量的结果,根据对研究对象计量的不同精确程度,将计量尺度由低到高、由粗略到精确分为两大层次:定性数据和定量数据。

1. 定性数据。定性数据常用文字表述,其计量结果表现为类别。定性数据又可以分为定类数据和定序数据。

定类数据亦称为列名数据,它是按照客观现象的某种属性对其进行平行的分类,此时,若用数字表示,该数字仅作为各类的代码,度量各类之间的类别差,不反映各类的优劣、量的大小或顺序。例如,人口按性别分为男女,用“1”表示男性,用“0”表示女性。定类尺度的主要数学特征可以用“=”或“≠”表示。在统计处理中,虽然可以计算单位数,但它不能表明第一类的一个单位可以相当于第二类的几个单位。

定序数据亦称为顺序数据,它是对客观现象各类之间的等级差或顺序差测度的数据,是比定类数据的计量尺度更高一级的数据。定序数据不仅可以研究对象分成不同的类别,而且还可以反映各类的优劣、量的大小或顺序。例如,学生成绩可以分为优、良、中、及格和不及格等五类,在这里,定序数据虽然无法表明一个优等于几个良,但却能确切地表明优高于良,良又高于中……定序尺度的主要数学特征可以用“<”或“>”表示。在统计的变量数列中可以确定其中位数、分位数等指标的位置。

2. 定量数据。定量数据比定序数据的计量尺度又更高一级。定量数据是对客观现象进行计量的结果,表现为具体的数值。它不但可以用数表示客观现象各类别的不同和顺序大小的差异,而且可以用确切的数值反映现象之间在量方面的差异。对于定量数据,一般可以运用数学的运算方法处理。在统计实践中,我们所处理的大多数是定量数据,定量数据在统计中占据重要的地位,例如,各类总量指标、相对指标和平均指标都是定量数据。

(二)按对客观现象观察的时间状态分类

在统计研究中,根据对客观现象观察的时间状态不同,统计数据可分为横截面数据和时间序列数据。

1. 横截面数据。横截面数据又称为静态数据,它是指在同一时间对不同单位的数量表现进行观察而获得的数据,例如,2006年6月我国各省市的消费价格指数是横截面数据。

2. 时间序列数据。时间序列数据又称为动态数据,它是指在不同时间对同一单位、同一现象的数量表现进行观察而获得的数据,例如,1978—2006年我国的消费价格指数是时间序列数据。关于时间序列,将在本书第八章进行详细介绍。

(三)按数据的表现形式分类

统计数据通常表现为绝对数、相对数和平均数。

1. 绝对数。现象的规模、水平一般以绝对数形式表现,例如,国内生产总值、产品总产量、人口数、进出口额等。绝对数的计量单位一般为实物单位或价值单位,有时也采用复合单位。实物单位可以是自然计量单位,也可以是物理计量单位,如人口数用人计量,机器数用台计量,粮食产量用吨计量,耕地面积用公顷计量等,对于一些化工产品,常常折合成为标准实物单位。价值单位是以货币形式进行计量的,如国内生产总值、进出口总额是以价值单位为计量单位。复合计量单位是由两种或两种以上计量单位复合而成的,如以“吨公里”为货物周转量的计量单位,以“千瓦时”为用电量的计量单位。

绝对数按其反映的时间状态不同,分为时期数据和时点数据。时期数据是反映现象在一段期间内发展过程的总量,它具有连续统计和可加性的特点,其数值大小与所属的时间长短有直接关系,如国内生产总值、进出口总额。时点数据是反映现象在某一特定时点所处的状态,它是采用间断登记方式取得资料的,不具有可加性,其数值大小与时点间隔长短无直接关系,如期末人口数、期末在建工程投资额等。

绝对数按其反映对象的内容不同,分为总体单位数和总体标志总量。

2. 相对数。相对数是由两个绝对数对比而得的,常用的相对数有结构相对数、动态相对数、比较相对数、比例相对数、强度相对数、利用程度相对数、计划完成相对数等。相对数的计量单位大部分是无名数,但也有一些是采用有名数为计量单位。把对比基数抽象为100而计算的相对数为百分数,把对比基数抽象为1000而计算的相对数为千分数,这些都是无名数。如果把某地区的人口数与该地区的土地面积对比所计算的相对数是一种强度相对数,被称为人口密度相对数,其计量单位为“人/平方公里”,这就是有名数。

3. 平均数。统计平均数是用于反映现象总体的一般水平或分布的集中趋势,数值平均数是由总体标志总量与总体单位数进行对比而计算的。关于这