

高职高专经济管理专业教材
经济信息管理培训教材

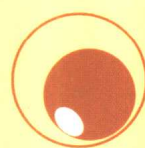
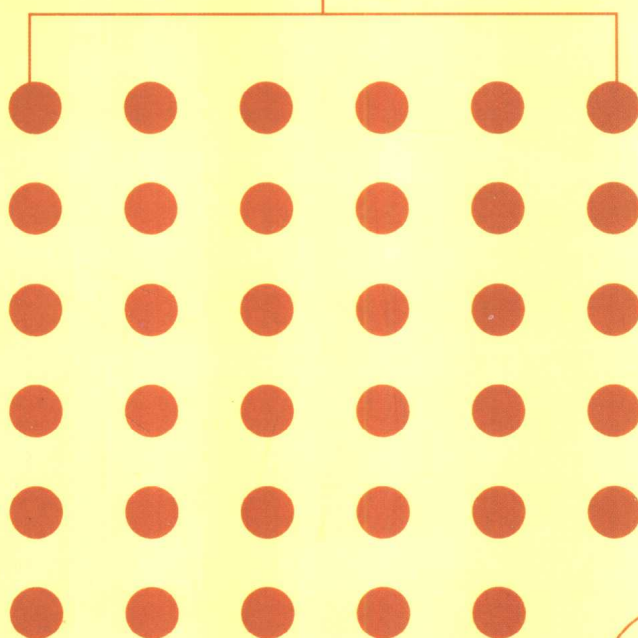
统计计算



分析

米小琴 主编

延静 李洁 副主编



清华大学出版社

高职高专经济管理专业教材 经济信息管理培训教材

统计计算与分析

米小琴 主 编
延 静 李 洁 副主编

清华大学出版社
北 京

内 容 简 介

《统计计算与分析》是一本统计入门读物,内容包括:总论、统计数据的收集、统计数据的处理、数
据分布的集中趋势和离散趋势、抽样技术和分析、相关和回归分析、时间数列分析、统计指数等。

本教材作为高职高专经济类专业的特色教材,采用一体化格式设计,包括:章首案例、正文、小结、
关键概念、思考题和实例分析题。全书定位明确、理论适中、贴近实际、操作性强,既避开繁杂的数学
公式,又系统完整地阐明统计学的科学思想和方法,表达通俗易懂,在描述原理时配合图表解释,便于
教学和自学。

版权所有,翻印必究。举报电话:010-62782989 13901104297 13801310933

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

统计计算与分析/米小琴主编. —北京:清华大学出版社, 2004.9

(高职高专经济管理专业 经济信息管理培训教材)

ISBN 7-302-06434-2

I. 统… II. 米… III. 经济统计—统计分析—高等学校:技术学校—教材 IV. F222

中国版本图书馆 CIP 数据核字(2004)第 084133 号

出 版 者:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客户服务:010-62776969

责任编辑:徐培忠

封面设计:付剑飞

印 刷 者:清华园胶印厂

装 订 者:北京鑫海金澳胶印有限公司

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:12 字数:295千字

版 次:2004年9月第1版 2004年9月第1次印刷

书 号:ISBN 7-302-06434-2/F·877

印 数:1~5000

定 价:17.00元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联
系调换。联系电话:(010)62770175-3103 或(010)62795704

编 委 会

主 任：牟惟仲

副主任：王纪平 吴江江 徐培忠 张建国 孙桂华 孟庆生
宁雪娟 李大军

编 委：宋承敏 张昌连 冀俊杰 仲万生 王 新 孟震彪
刘长鑫 林 亚 李贵保 徐 军 白文祥 冯仁华
栾茂茹 周 平 卫停战 孟乃奇 马爱杰 付绪昌
张国英 王伟光 李书胜 孙淑凤 阚晓芒 高光敏
盛定宇 孟繁昌 韩凤莲 周 宏 米小琴 李 洁
延 静 王 凯 孟建华 宋常桐 智 瑾 周 晓
周 兵 董 铁 王桂霞 邵海峡

专家组：李大军 齐众希 蔺 涛 杨学军 林 立

秘书长：李大军

前 言

生产离不开统计,流通也离不开统计。随着国家经济建设的迅猛发展,无论是在制造业,还是在商务活动、经营领域,统计的作用日趋明显;特别是在网络经济促动下,世界经济高度融合、社会生活节奏日趋快速化,对信息的需求明显加快和增多,其处理技术也更加复杂和多样化,作为统计信息基础的统计计算和分析技术就显得越来越重要,同时也在发挥着越来越重要的作用。

在每一门社会科学以及相关学科中都会有一个基本工具,既用于描述的目的,也作为评估的基础。统计计算与分析既是一种工具,又作为一种手段而被广为接受,被频繁地用于各种各样的应用性研究中,如社会学、心理学、管理学、人口学、政治学、经济学以及公共卫生等领域。统计作为高职高专经济管理专业的专业基础课程,学生更需要掌握其扎实的统计理论和过硬的计算与分析技能,以便为其他后续课程的学习打好基础。

本书广泛吸收了国内外优秀教材的成果,注重对统计基本理论和技术的阐述,并运用实际数据和实例作说明。全书重点阐述了三个问题,即如何收集数据、如何处理数据以及如何分析数据。全书以设计新、观念新、案例新为特色。所谓设计新,表现在全书采用一体化格式设计,包括:章首案例、正文、小结、关键概念、思考题和实例分析题。所谓观念新,是指全书既避开繁杂的数学公式,又系统完整地阐明统计学的科学思想和方法,表达通俗易懂,举例生动实用,在描述原理时配合图表解释,便于教学和自学。所谓案例新,是指全书以培养学生的应用能力为主线,选材包括实践中经常会遇到的许多具体疑难问题,以指导学生进行实质性问题的研究,真正体现高职高专的教育特色。

本书共九章,内容包括:总论、统计数据的收集、统计数据的处理、数据分布的集中趋势和离中趋势、抽样技术和分析、相关和回归分析、时间数列分析、统计指数软件应用等。现代统计分析已离不开数据的查询,为了使学好并用好统计方法,本教材在附录中介绍了如何利用计算机网络进行宏观统计数据的查询。

本书在编写过程中,翻阅参考了大量国内外最新有关统计的书刊、文件及资料,收集了最新的、具有实用价值的典型案例,力图使读者在阅读时能够有所收获。希望本书的出版能够促进我国统计知识的普及和教育水平的提高。

本书作为高职高专经济管理专业的特色教材,定位明确、理论适中、知识面宽、贴近实际、操作性强、自成体系、适用范围宽泛。既适用于高职高专、大中专院校经济管理专业(会计、理财与税收、会计信息管理、保险营销、工商管理、金融、经济信息管理、电子商务、计算机网络管理、物流管理等)学生的学历教育,又适用于广大工商企业人员的短期培训,还适用于广大社会自学者。

本教材由李大军秘书长进行总体方案的策划，米小琴主编并统稿，延静、李洁为副主编，国家统计局的蔺涛博士、处长对全书进行了审定。参加编写的人员有：延静(第一章、第八章)，班丽冬(第二章、第三章)，李洁(第四章)，米小琴(第五章、第六章)，苏向明(第七章)，智瑾(第九章)，李大军(附录)。由于时间紧，在教材编写过程中难免存在不足之处，因此诚恳地请各位专家和读者指正。

作者

2004年6月

目 录

第一章 总论	1
第一节 什么是统计	1
一、统计的含义	1
二、统计的研究对象及其特点	1
三、统计分析的常用方法	4
第二节 描述统计和推断统计	5
一、统计学的分类	5
二、描述统计学与推断统计学	6
第三节 统计中的几个基本概念	7
一、总体和样本	7
二、变量和变量值	8
第二章 统计数据的收集	11
第一节 统计数据的调查方式	11
一、常用的统计调查方式	12
二、第二手统计数据的收集	16
第二节 调查问卷的设计	16
一、问卷的含义和设计要求	16
二、问卷设计程序	17
三、问卷中问题的主要类型	18
四、与问卷设计有关的计量误差及对相关问题的处理	21
第三节 统计数据的收集方法	22
一、自填式	22
二、人员面访	23
三、电话访问	23
四、其他方法	25
第三章 统计数据的处理	28
第一节 统计数据处理的意义和内容	28
一、统计数据处理的意义	28
二、统计数据处理的主要内容	29
第二节 统计资料的预处理	29
一、编码	30
二、数据录入	30

三、审核	30
四、插补	31
第三节 统计分组	32
一、统计分组的概念	32
二、统计分组的作用	33
三、选择分组标志的原则	34
四、按品质标志分组和按数量标志分组	35
五、等距分组的步骤	38
第四节 汇总和计算	39
一、汇总	40
二、计算	41
三、分布数列	41
第五节 统计表和统计图	42
一、统计表	43
二、常用的统计图	44
第四章 数据分布的集中趋势和离中趋势	49
第一节 集中趋势指标	50
一、集中趋势指标概述	50
二、集中趋势指标的计算与应用	51
三、算术平均数与中位数、众数的比较	65
四、正确运用集中趋势指标的原则	66
第二节 离中趋势指标	67
一、离中趋势指标概述	67
二、离中趋势指标的计算与应用	68
第五章 抽样技术与分析	78
第一节 抽样方法	78
一、什么是抽样	78
二、非概率抽样	79
三、概率抽样	80
四、抽样技术的选择	87
第二节 估计	88
一、简单估计值的计算	88
二、误差限与置信区间	91
第三节 样本量的确定	94
一、影响样本量的因素	94
二、对均值估计计算样本量的步骤	96
三、对比例估计计算样本量的步骤	96

第六章 相关与回归分析	100
第一节 相关分析与相关系数	100
一、相关分析的概念	100
二、相关数据的收集和相关关系的种类	101
三、相关系数	104
第二节 回归分析	106
一、一元线性回归分析	106
二、回归方程的显著性检验	108
三、预测及应用	110
第七章 时间数列分析	113
第一节 时间数列的构成要素和分类	114
一、时间数列的概念、构成要素	114
二、时间数列的种类	114
三、时间数列的编制原则	115
第二节 时间数列分析的水平指标	116
一、发展水平	117
二、平均发展水平	117
三、增长量	122
四、平均增长量	122
第三节 时间数列分析的速度指标	123
一、发展速度	123
二、增长速度	124
三、平均发展速度与平均增长速度	125
第八章 统计指数	130
第一节 统计指数的概念和分类	130
一、概念	130
二、统计指数的作用	131
三、统计指数的分类	131
第二节 综合指数	133
一、综合指数的特点	133
二、数量指标指数的编制	134
三、质量指标指数的编制	137
第三节 平均数指数	139
一、加权算术平均数指数	139
二、加权调和平均数指数	141
第四节 指数体系和因素分析	142
一、总量指数	142

二、指数体系和因素分析·····	142
第五节 几种常用的价格指数·····	144
一、商品零售价格指数·····	144
二、消费价格指数·····	145
三、股票价格指数·····	146
第九章 Excel 软件应用 ·····	149
第一节 制作统计图表·····	149
第二节 集中趋势值的计算·····	153
第三节 离中趋势值的计算·····	158
第四节 回归分析·····	162
一、绘制离散图·····	162
二、用 Excel 软件进行回归分析的步骤 ·····	166
附录 1 利用网络进行统计数据的查询 ·····	174
附录 2 正态分布表 ·····	180

第一章 总 论

今年“五一”黄金周期间，全国共接待旅游者 1.04 亿人次，比 2002 年“五一”黄金周增长 19.6%；实现旅游收入 390 亿元，比 2002 年同期增长 17.8%；旅游者人均花费支出 375 元。（因 2003 年遭遇“非典”，故用 2002 年数据做比较——作者注）

根据国家旅游局、国家统计局联合制定的“黄金周”旅游信息统计调查制度，全国 31 个省、自治区、直辖市分别对辖区内“五一”黄金周的接待规模和效益进行了统计调查。全国假日办 8 日发布的统计结果显示，“五一”黄金周期间，在全国接待的 1.04 亿人次旅游者中，过夜旅游者（仅限于住在宾馆饭店和旅馆招待所）为 2 612 万人次，比 2002 年同期增长 16.1%；一日游游客为 7 801 万人次，比 2002 年同期增长 20.8%。黄金周期间，民航客运收入 16 亿元，比 2002 年同期增长 18.5%；铁路客运收入 11.4 亿元，比 2002 年同期增长 13.8%；31 个重点旅游城市实现旅游收入 159 亿元；其他旅游城市和景区实现旅游收入 203.6 亿元。

“五一”黄金周期间，31 个重点旅游城市监测的 143 家商业企业，共实现销售收入 30.75 亿元，比 2002 年同期增长 23.1%；监测的 140 家餐饮企业，共实现营业收入 2.35 亿元，比 2002 年同期增长 32.2%。

摘自新华网，2004-05-08

第一节 什么是统计

一、统计的含义

看了上面的报道，你能说出什么是统计吗？是这些数字还是这项工作？今天，“统计”一词已被人们赋予多种含义，因此很难给出一个简单的定义。在不同场合，“统计”一词可以有不同的含义。它可以指统计数据的收集活动，即统计工作。如上例，根据国家旅游局、国家统计局联合制定的“黄金周”旅游信息统计调查制度，全国 31 个省、自治区、直辖市分别对辖区内“五一”黄金周的接待规模和效益进行了统计调查。它也可以指统计活动的结果，即统计数据。如上例，全国假日办 8 日发布的统计结果显示，“五一”黄金周期间，在全国接待的 1.04 亿人次旅游者中，过夜旅游者（仅限于住在宾馆饭店和旅馆招待所）为 2 612 万人次，比 2002 年同期增长 16.1%。它还可以指分析统计数据的方法和技术，即统计学。如上例，“五一”黄金周期间，31 个重点旅游城市监测的 143 家商业企业，共实现销售收入 30.75 亿元，比 2002 年同期增长 23.1%；监测的 140 家餐饮企业，共实

现营业收入 2.35 亿元，比 2002 年同期增长 32.2%。这里使用了重点调查的方法和动态对比分析方法。

随着社会管理的日趋复杂，仅仅用数字计量客观现象已不能满足社会的需要。把数字算出来了，但是这些数字为什么是这样而不是那样，有什么规律？人们试图对客观现象进行定量的统计分析。要进行统计分析就需要有科学的方法和技术。

统计一词在不同的场合有三种含义：统计、统计数据和统计学。统计是一种社会实践活动，它通过数据的收集、整理和分析，为人们提供定量化的认识。统计活动是过程，统计数据是结果。统计学则是一门收集、整理和分析统计数据的方法科学，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。

统计数据的收集是取得统计数据的过程，是进行统计分析的基础。如何取得准确、可靠的统计数据是统计学研究的内容之一。

统计数据的整理是统计数据的加工处理过程，目的是使统计数据系统化、条理化，符合统计分析的需要。数据整理是数据收集与数据分析之间的一个必要环节。

统计数据的分析是统计学的核心内容，是通过统计描述和统计推断的方法探索数据内在规律的过程。

可见，统计学是一门有关统计数据的科学，统计学与统计数据有着密不可分的关系。在英文中，“statistics”一词有两个含义：当它以单数名词出现时，表示作为一门科学的“统计学”；当它以复数名词出现时，表示“统计数据”或“统计资料”。从中可以看出，统计学与统计数据之间有着密不可分的关系。统计学由一套收集和处理统计数据的方法组成，这些方法来源于对统计数据的研究，目的也在于对统计数据进行研究。离开了统计数据，统计方法就失去了用武之地，统计学也就失去了它存在的意义。而统计数据如果不用统计方法加以分析也仅仅是一堆数据而已，得不出任何有益的结论。

此外，统计数据在英文中是以复数形式出现的，这表明统计数据不是指单个的数字，而是由多个数据构成的数据集。单个的数据显然用不着统计方法进行分析，仅凭一个数据点也不可能得出事物的规律，只有经过对同一事物进行多次观察或计量得到大量数据，才能利用统计方法探索出事物内在的必然规律性。

二、统计的研究对象及其特点

统计学和统计活动是理论与实践的关系，它们的研究对象是一致的。统计的研究对象是大量现象总体的数量方面，其特征是在质与量的辩证统一中研究大量现象总体的数量方面，反映现象发展变化的规律性在具体时间、地点和条件下的数量表现，揭示事物的本质、相互联系、变动规律性和发展趋势。统计学和统计活动的区别在于：统计学从理论角度进行研究阐述，统计活动则是从实践上进行具体研究。

辩证唯物主义告诉我们，不论是自然现象还是社会现象，都存在质与量两个方面，二者是辩证统一、密切联系的。事物的质是通过量表现出来的，没有数量也就没有质量，量的积累，达到一定界限，将引起质的变化。因此，要研究事物的存在和发展，并掌握其发展规律性，必须研究事物的量的方面，研究事物发展规律性在具体时间、地点、条件下的数量表现。这里所说的数量方面是指现象的规模、水平、结构、速度、比例、普遍程度等

等。事物的质和量是密切联系的，是辩证的统一，因此，统计要对社会经济现象的数量方面进行研究，必须和其质的方面结合起来。换句话说，首先要明确现象的质的特征，而后才能正确反映其量的表现。例如，要统计国内生产总值，如果不明确什么是国内生产总值及国内生产总值的核算范围，就无法统计。又如，要统计主要工业品产量，就要明确哪些产品属于主要工业品。统计如果以社会经济现象为其研究领域，其具有的特点归纳起来可概括成如下三个方面：数量性、总体性和具体性。

（一）数量性

统计的研究对象是社会经济现象的数量方面，包括社会经济现象的规模、水平、现象间的数量关系，以及决定现象质量的数量界限。统计研究对象的数量性，是统计区别于其他社会经济调查研究活动的根本特点。必须指出，统计对社会经济现象数量方面的认识是定量认识，但必须以定性认识为基础，要和定性认识结合起来，遵循定性一定量一定性的科学的认识规律。例如，要了解和研究国内生产总值的数量、构成及其变化，首先必须了解国内生产总值的本质属性，即国内生产的最终产品的市场价值，然后才能根据这种认识去确定国内生产总值的口径、范围和计算方法，进而才能据以处理许许多多复杂的、具体的实际统计问题。

（二）总体性

统计研究对象不是个体现象的数量方面，而是由许多个体现象构成的总体的数量方面。例如，劳动生产率统计，不是研究某个人具体的劳动效率，而是研究一个国家、地区、部门或一个企业总体的劳动生产率及其变动。统计研究对象的总体性这个特点，是由社会经济现象的特点和统计研究的目的决定的。由于社会经济现象错综复杂，各个个体现象所处条件不同，它们既受共同因素的影响，又受某些个别的、偶然的因素影响。因此，个体现象的数量特征和变动趋势是难以说明社会经济现象总体的本质和规律的。只有以社会经济现象的总体为研究对象，即以构成总体的全部或足够多数的个体现象为研究对象，才能消除偶然因素的影响，正确地显示出社会经济现象的本质和规律性。但是，总体是由个体所构成的，要认识社会经济现象总体，就必须从调查了解个体现象的情况开始，从个体到总体。如人口统计必须从了解每个人的情况开始，然后经过分组、汇总、计算工作，过渡到说明人口总体数量的特征。

（三）具体性

统计所研究对象的量是具体的数量，不是抽象的数量，这是统计和数学的重要区别。数学虽然以现实世界的空间形式和数量关系为研究对象，但它是非常抽象的。而统计所研究的量是具体事物在具体时间、地点和条件下的数量表现，它总是和现象的质密切结合在一起的。例如，2003年我国的钢产量为22 233.60万吨，原煤产量为16.67亿吨，原油产量为1.7亿吨等，显然不是抽象的量，而是我国在2003年这一具体条件下钢、原煤、原油生产的数量表现。如果抽掉具体的内容，不是在一定时间、地点和条件下进行研究，那就不能说明任何问题，也就不称其为统计，其数字也就不是统计数字。

三、统计分析的常用方法

统计学提供了探索数据内在规律的一套方法。那么，什么是统计数据的内在数量规律性？为什么统计方法能够通过对数据的分析找出其内在的数量规律性？我们用下面的几个例子来说明。

1. 投掷硬币

我们都知道投掷硬币和掷骰子的游戏，随机地投掷一次硬币或骰子，出现正面、反面或某个点数是不确定的，完全是偶然的。但我们进行多次的重复投掷，就会发现一枚均匀硬币出现正面和反面的次数大体相同，即比值接近于 $1/2$ 。投掷的次数越多，就越接近于 $1/2$ 这一稳定的数值。同样，在掷骰子时，出现 $1\sim 6$ 点的比例也逐渐接近于 $1/6$ 。这里的 $1/2$ 和 $1/6$ 就是掷硬币和掷骰子出现某一特定结果的概率，也就是投掷硬币或骰子时所呈现的数量规律性。

2. 男女性别比

就单独一个家庭来观察，新生婴儿的性别可能是男性，也可能是女性。如果不对人口生育进行任何限制，有的家庭的几个孩子可能都是男孩，而有的家庭的几个孩子可能都是女孩。从表面上看，新生婴儿的性别比例似乎没有什么规律可循。但如果对大量的家庭新生婴儿进行观察，就会发现新生婴儿中男孩略多于女孩，大致为每出生 100 个女孩，相应地就有 107 个男孩出生。这个性别比例 $107:100$ 就是新生婴儿性别比的数量规律，古今中外这一比例都大致相同，这是由人类自然发展的内在规律所决定的。人类社会要发展，就要保持男女人数上的大致相同。尽管从新生婴儿来看，男性婴儿略多于女性，似乎并不平衡，但由于男性婴儿的死亡率高于女性，到了中年时，男女人数就大体相当了。进入中老年后，男性的死亡率仍然高于女性，导致男性的平均预期寿命比女性短，老年男性反而少于女性。人口生育在性别上保持大体平衡，保证了人类社会的进化和发展。

3. 农作物产量和施肥量

在进行农作物生产过程中，我们发现某种粮食作物的产量会随着某种肥料施肥量的增加而增加。当最初增加施肥量时，产量增加较快，以后增加同样的施肥量，粮食产量的增长逐渐放慢。当施肥量增加到一定数值时，产量不再增加。这时如果再增加施肥量，产量反而会减少。粮食产量与施肥量之间的这种数量关系，就是我们所要探索的数量规律性。如果我们能从大量的试验数据中，用统计方法找出产量与施肥量之间的数量关系，就可以确定出最佳的施肥量，以求得最大的效益。

4. 商品销售额和商品流通过费率

在商业活动中商品销售额与商品流通费用之间存在着依存关系，经营者发现商品销售额增大，流通费用额也会增大。但是，当销售额达到一定规模时，商品流通过费率就开始降低了。所以，通过统计方法找到流通过费率降低的商品额的临界值可以更有目的地开展经营活动。

上述这些现象说明，就一次实验或一次活动而言，结果可能是随机的，但是通过多次观察或实验得到大量统计数据，利用统计方法是可以探索出其内在的数量规律性的。在统计分析中常用的基本方法有以下几种：

1. 大量观察法

大量观察法就是对现象的总体(足够多)进行大量观察和综合分析的统计方法。

2. 统计分类法

统计分类法就是根据现象的特点和统计研究的目的,将现象按不同类型或不同性质划分成若干个部分的统计方法。

3. 综合分析法

综合分析法就是运用各种统计分析方法对现象数量方面进行分析研究,表明大量现象的一般特征、探索现象规律性的统计方法。具体方法包括以下几种:

(1) 动态分析法

动态分析法就是运用时间数列进行对比分析,描述现象变化的过程、变化方向和速度,并预测现象未来的发展趋势的一种统计方法。

(2) 抽样推断法

从个别到一般、从具体事实到抽象概括的推理方法,在逻辑上称为归纳法。统计研究从观察各个单位的具体特征着手,经过综合概括,得出关于总体一般特征的信息,所以归纳法是统计研究的基本方法。由于我们通常所观察的只是总体的一个样本,所以还必须根据样本资料对总体的状况进行推断或估计。以一定的可信程度,根据样本数据来判断总体数量特征的方法称为抽样推断法。

(3) 相关与回归分析法

自然现象和社会现象虽然错综复杂,但是各个现象之间总是存在着一定的联系。如果找到它们之间的联系,且确定一个因素,一个确定的结果马上就会显现出来,并且还有确定的对应数值。我们把对这种现象的分析叫做相关与回归分析。

(4) 指数分析法

自然现象和社会现象中有许多是复杂现象,其变动要受许多因素的影响。通过编制指数,进行指数分析,可以分析现象中各因素的影响方向和影响程度。

第二节 描述统计和推断统计

一、统计学的分类

现代统计学已发展成为一种多门类、多分支、多层次的科学。对此从不同的观察角度往往有不同的分类。根据统计方法研究和应用,可将统计学分为理论统计学与应用统计学;根据统计方法的构成,可将统计学分为描述统计学与推断统计学。

(一) 理论统计学与应用统计学

1. 理论统计学

理论统计学是关于数据的收集、整理和分析的最基本的原理、原则和方法,一般既适用于社会经济现象数量特征的观察和分析,也适用于自然现象和科学实验数据的分析研究,它是应用统计学共同的理论基础。所谓统计学是一门通用方法论科学,就是对理论统

计学而言的。

2. 应用统计学

应用统计学是运用于某一特定领域的统计理论和方法。由于在各个领域的研究和实际工作中都要通过数据来分析和解决问题,统计理论和方法的应用就自然而然地扩展到几乎所有的研究领域,形成了各种应用统计学。如统计方法在生物领域的应用就形成了生物统计学;统计方法在医药领域的应用就形成了医药统计学;统计方法在金融领域的应用就形成了金融统计学;统计方法在经济学和企业管理中的应用就形成了经济统计学、经济计量学、管理统计学等。应用统计学的不同分支所应用的基本方法都是一样的,即描述统计学和推断统计学的主要方法。但由于各领域都有特殊性,统计方法在不同领域的应用中就具有了不同的特点。应用统计学除了包括各领域通用的方法,也包括特定领域特有的方法,如经济统计学中的指数法、管理决策法等。

二、描述统计学与推断统计学

1. 描述统计学

描述统计学研究如何对客观现象的数量特征进行计量、观察、概括和表述。具体地说,描述统计学的主要内容包括:确定所要研究的数量特征及其计量层次、设计用来说明现象的数量特征的统计指标、收集和整理数据、计算统计指标并用图表显示。客观世界中,有的现象的数量特征比较直观,很容易观察和描述,如学校的学生人数和学生的身高、体重、学习成绩等。有的现象特别是社会经济现象却十分复杂,对其数量特征的描述就不那么轻松容易了。如对通货膨胀的描述、对国民经济的运行状况的描述等,这类描述问题涉及很多方面,首先必须结合现象的实质,在定性分析的基础上,正确界定有关的范围、概念、分类和统计指标。

2. 推断统计学

推断统计学研究如何根据部分数据去推论总体的数量特征。例如,要检验一批电子元件的使用寿命,不可能将所有产品都一一测试,只能从中随机抽取一部分产品来测试,并以其检验结果推论整批产品的使用寿命。要研究全国居民生活状况,就没有必要对所有的居民家庭进行调查,只需随机抽取一部分居民家庭来调查并据以推断全国居民生活状况即可。从总体中抽取的一部分个体构成一个样本,样本来源于总体,因此样本信息包含总体信息。但由于所调查的样本只是总体的一部分,样本所包含的总体信息必然不完备,据以推断总体必然产生一定的误差。而且由于样本是随机抽取的,推断的结论是否可靠也不确定。为了保证这种推断的精确度和可靠度,推断统计学就需研究样本的抽取方式方法、样本的大小、样本的分布、样本估计量的选择、对总体特征进行估计或推断的方法、误差的计算和控制等问题。

统计活动的最终目的都是为了说明所研究现象总体的数量特征和规律性,为此首先必须收集反映客观事实的数据。如果所收集的数据是总体数据(即总体中全部个体的数据),那么采用描述统计方法即可直接达到研究的目的;如果所收集的数据是样本数据(即总体中一部分个体的数据),那么采用描述统计方法得出样本的数量特征后,还必须应用推断统计方法根据样本信息对总体数量特征进行科学的推断。显然,描述统计学是整个统计学

的基础,而推断统计学是现代统计学的核心。如果没有描述统计确定的应该研究现象什么数量特征及其怎样度量和描述,没有描述统计提供有效的样本信息,那么推断统计就只能成为无源之水,无本之木,再高明的推断统计方法也难以得到准确的结论。另外,在实际工作中,所研究的现象总体常常包含无限多或近似无限多的个体,调查收集到的数据大多是样本数据,如果没有推断统计,就不可能得到对所研究现象总体数量特征和规律性的认识,也不可能使现代统计学建立在科学的基础之上。可见,推断统计离不开描述统计,只靠描述统计也难以揭示事物的内在规律。

第三节 统计中的几个基本概念

统计中有许多概念,其中有几个概念是我们经常要用到的,我们在此单独作一介绍。这些概念包括总体和样本、变量和变量值等。

一、总体和样本

(一) 总体

在统计学中把所研究对象的全体称为总体,也可以说是全部个体(数据)的集合。通常是由具有某种共同性质的许多个体组成,如由多个超市构成的集合,多个居民户构成的集合,多个网吧构成的集合等。组成总体的每一个元素称为个体,在由多个超市构成的总体中,每一个超市就是一个个体;由多个居民户构成的总体中,每一个居民户就是一个个体;由多个网吧构成的总体中,每一个网吧就是一个个体。总体具有三方面的特征:

1. 同质性

即构成总体的各个个体必须具有某一方面的共性。

2. 大量性

即总体是由许多个体所组成的。

3. 差异性

即总体的各个个体之间,除了必须在某一方面有共性之外,在其他方面还需存在差异,这些差异是统计研究的基础。

总体范围的确定有时比较容易。比如,要检验一批彩电的使用寿命,这一批彩电构成的集合就是总体,每一台彩电就是一个个体,总体的范围很清楚。但有些场合总体范围的确定则比较困难,比如,某个企业在电视台作了广告,想要了解看电视的人中有多少人看了该广告,并留下了印象。这里要统计所有看电视的人就比较复杂,因为看电视的人时间有长有短,有的人经常看,有的人则偶尔看,年龄上的差别也很大,是否都应列入总体?事实上,总体范围的确定十分复杂。当总体的范围难以确定时,可根据研究的目的来定义总体。

总体根据其所包含的单位数目是否可数可以分为有限总体和无限总体。有限总体是指总体的范围能够确定,而且元素的数目是有限的、可数的。比如,由若干个超市构成的总体就是有限总体,一批待检验的彩电构成的总体也是有限总体。无限总体是指总体所包括