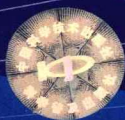




高职高专经济管理专业“十一五”规划教材

基础统计

高凯萍 张伟 主编



中国科学技术出版社
CHINA SCIENCE AND TECHNOLOGY PRESS

高职高专经济管理专业“十一五”规划教材

基 础 统 计

高凯萍 张伟 主编

中国科学技术出版社

CHINA SCIENCE AND TECHNOLOGY PRESS

· 北 京 ·

BEIJING

图书在版编目(CIP)数据

基础统计/高凯萍,张伟主编. —北京:中国科学技术出版社,2008.2

高职高专经济管理专业“十一五”规划教材

ISBN 978 - 7 - 5046 - 5087 - 0

I. 基… II. ①高… ②张… III. 统计学 - 高等学校:技术学校 - 教材 IV. C8

中国版本图书馆 CIP 数据核字(2007)第 199369 号

自 2006 年 4 月起本社图书封面均贴有防伪标志,未贴防伪标志的为盗版图书。

内 容 提 要

本教材以培养学生敏锐的定量思维能力、分析能力和实际应用能力为编写原则,用通俗易懂的语言,直观明了的统计图表和最新的统计资料为实例来讲述统计学的基本理论和方法。以帮助读者掌握统计的基础知识,为进一步学习专业打好基础。全书内容包括描述统计和推断统计两部分。内容涉及统计基础知识、统计数据的整理和显示、统计描述的综合指标、指数的编制及分析、统计推断的基础、参数估计与假设检验、相关与回归分析、时间数列的分析等。每章开始是“内容向导”,以帮助学生在学之前了解本章要学的主要内容;学完一节后有“本节知识要点网络图”和拓展训练,每章后有“内容回顾”以帮助学生总结。本书可供高等职业学校、高等专科学校、成人高等学校及本科院校的二级学院及各种经管类成人培训的学生使用。

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码:100081

策划编辑	林 培 孙卫华	责任校对	林 华
责任编辑	孙卫华 王 强	责任印制	安利平

发行部电话:010 - 62103210 编辑部电话:010 - 62103181

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

科学普及出版社发行部发行

北京蓝空印刷厂印刷

*

开本:787 毫米×1092 毫米 1/16 印张:13 字数:314 千字

2008 年 2 月第 1 版 2008 年 2 月第 1 次印刷

ISBN 978 - 7 - 5046 - 5087 - 0/C · 134

印数:1—4000 定价:22.60 元

(凡购买本社的图书,如有缺页、倒页、
脱页者,本社发行部负责调换)

高职高专经济管理专业“十一五”规划教材

编委会名单

主 任 邓泽民

副 主 任 陶学忠

专家编委 (以姓氏笔画为序)

丁增稳	马成旭	兰培英	刘晓玉	刘继伟	李学智
周延波	段兴民	袁放建	寇宝明		

委 员 (以姓氏笔画为序)

马西牛	马雨娟	马 惠	王 瑶	王永峰	王康美
刘芳霞	刘金波	刘振华	任晓丽	李从质	李建峰
张 伟	张 拓	张大田	张淑琴	张金城	杨 洁
杨 桦	杨淑娥	林敏捷	郑文昭	郑光顺	武妙仙
赵 轶	高凯萍	高建侠	康 丽	阎益国	温国兴
雷 剑	缙宇英	景海萍			

秘 书 长 劳 粟

《基础统计》编委会

主 编 高凯萍 张 伟

副 主 编 张献奇 马雨娟 郑光顺 阎益国

编 委 冯 蕊 杨 琴

前 言

《基础统计》是高职高专院校财经管理类各专业必修的公共基础课。根据近年来高职高专教育培养适应社会需求的应用型人才的目标,本书针对学生的实际文化水平与理解能力,在原有教学内容的基础上进行了修订,将“写给学生看、指导学生干”作为该教材编写目标。本着简单、实用、易于理解的原则,确立了“问题导入、任务驱动”的内容体系,将每一章节作为一个模块编写。通过对教材的学习和实践,使学生略通理论,侧重实用,具有极强的思考与应用能力。

该教材的编写人员是从全国高职高专院校具有丰富教学经验的一线教师中严格筛选出来的,确保了教材的编写质量,使其具有实用性、工具性、动态性和前瞻性。

本教材的主要特点是:

1. 体现了教育部对高职高专教学改革的总体精神要求。教材按“以培养学生职业能力为本位、以学生就业为导向的新理念”要求而编写。

2. 把握统计学科发展方向。基础统计是一门关于数据的收集、整理、分析的学科。作为一门方法论的科学,其量化分析方法已渗透到社会各领域。本教材从应用的角度,分为描述统计和推断统计两部分介绍,将统计工作的过程作为隐线贯穿其中。

3. 采用“问题导入、任务驱动”的编写方式。基础统计是理论与实践综合性较强的课程,不经过系统的学习与训练,很难掌握这门课程的主要知识点。因此,本书采用“提出问题——分析问题——解决问题——归纳总结”的方式,将“问题导入、任务驱动”的思想体现在每一章节中,巧妙地将知识体系和基本技能体系融入其中。

4. 利于教师教学操作,便于学生课后自学。本教材遵循“通俗性、趣味性和易接受性”的原则,每章课前列有“章节问题”与“内容向导”,每节课后列有“本节知识要点网络图”和“拓展训练”,中间穿插“思考一下”,章后又有“内容回顾”,使学生展开思考、即刻动手、容易理解、便于记忆。考虑到各教材使用院校的具体情况,将目录中标有“*”的章节作为选学内容,可以适应不同层次的学生使用。

本书由高凯萍、张伟担任主编。各章执笔人是:高凯萍编写统计学入门、第六章、第七章和附录;张伟编写第一章;马雨娟编写第二章;郑光顺编写第三章;冯蕊编写第四章;阎益国编写第五章;杨琴编写第八章。本书由高凯萍设计编写方案并拟定编写大纲,最后对全书进行了修改与总纂。

本书在编写过程中,得到了相关院校的大力支持,在此我们表示衷心的感谢。

由于编者水平有限,书中难免有疏漏和不足之处,恳请广大读者及专家批评指正。

编 者

2007年11月

目 录

统计学入门	1
统计的含义	1
统计工作的过程	2
统计的主要研究方法	3
描述统计与推断统计	5
第一章 描述统计的基础	7
第一节 描述统计的几个基本概念	7
一、总体与总体单位	7
二、标志与标志表现	9
三、统计指标与指标体系	9
四、变异与变量	10
第二节 统计数据的来源	12
一、统计数据的种类	12
二、统计调查方案	13
三、数据的收集方法	17
四、统计调查的组织形式	18
五、统计误差	22
第二章 统计数据的整理与显示	25
第一节 统计数据整理的内容	25
一、统计数据的预处理	26
二、统计数据的分组	27
三、统计数据的汇总	29
四、统计数据的显示	30
第二节 品质数据的整理与显示	33
一、分类数据的整理与显示	33
二、顺序数据的整理与显示	35
第三节 数量数据的整理与显示	37
一、静态数值型数据的整理与显示	37
二、时间型数据的整理与显示	44
第三章 统计描述的综合指标	50
第一节 总量指标	50
一、总量指标的概念与作用	50
二、总量指标的种类	51

第二节 相对指标	54
一、相对指标的概念与作用	54
二、相对指标的种类及计算	56
第三节 平均指标	65
一、平均指标的概念与作用	65
二、平均指标的计算方法	67
第四节 变异指标	84
一、变异指标的概念与作用	84
二、变异指标的种类及计算	85
三、标准化分位值	89
第四章 指数的编制及分析	94
第一节 指数的概述	94
一、指数的概念与作用	94
二、指数的种类	95
第二节 总指数的编制与应用	97
一、综合指数	97
二、平均指数	102
三、指数的应用与解释	106
第三节 指数体系与因素分析	109
一、指数体系	109
二、因素分析	110
第五章 统计推断的基础	119
第一节 统计推断概述	119
一、统计推断的意义	119
二、统计推断的基本概念	121
三、随机抽样的设计	125
第二节 抽样分布	128
一、抽样分布的概念	128
二、中心极限定理与大数定律	130
三、样本均值与样本成数的抽样分布	132
第六章 参数估计与假设检验	138
第一节 参数估计	138
一、参数估计的方法	138
二、总体参数的区间估计	143
第二节 假设检验 *	148
一、假设检验的基本思想	148
二、假设检验规则与两类错误	149
三、假设检验的步骤与方法	151
四、假设检验与区间估计的关系	152

第七章 相关与回归分析	155
第一节 相关关系概述	155
一、相关关系的含义	155
二、相关关系的类型	157
三、相关关系分析的内容	158
第二节 直线相关分析	159
一、相关表与相关图	159
二、相关系数	161
第三节 直线回归分析	165
一、回归分析的概念	165
二、回归方程的建立	166
三、回归标准误差	167
四、总体 Y 值的推断 *	168
第八章 时间数列的分析	172
第一节 时间数列分析的前提	172
一、时间数列的特征	172
二、时间数列分析的基本假设	174
三、时间数列的速度分析	175
第二节 时间数列的测定与预测	176
一、时间数列的测定	176
二、时间数列的预测方法	186
附录	195
参考文献	199



统计学入门

一名 Sloan-Kettering 癌症研究中心的研究者有一次因伪造数据受到谴责。他的数据中所包含的数字来自 6 组老鼠，每一组有 20 只。这些数值表示每一组的存活率：53%、58%、63%、46%、46%、67%。这些数值中主要的缺陷是什么？^①

——章节问题

内容向导

- ◇ 统计的含义
- ◇ 统计工作的过程
- ◇ 统计的主要研究方法
- ◇ 描述统计与推断统计

统计的含义

在日常生活中，“统计”一词的使用有时是指工作过程，如统计一下本市的失业人数；有时是指统计资料，如我们经常谈到的“国民收入统计”或“人口统计”；有时是指学科本身，如本书是统计教材。由此可见，在不同场合使用“统计”一词的含义不尽相同，概括地讲，“统计”一词具有三种含义，即统计工作、统计资料和统计学。

※ 统计工作是对客观存在的事物或现象的数量方面进行的一种计量活动。

统计的认识对象不是事物或现象的全部，而是事物或现象的数量方面，包括数量多少、现象之间的数量关系和质量互变的数量界限。统计的目的，就是要反映事物或现象数量方面的现状及其发展变化过程。

统计的认识对象是事物或现象总体的数量方面。总体是由具有某种相同性质的全体事物所组成。如，人口统计不是要了解和研究个别的人，而是要反映一个国家或一个地区人口多少，男女各多少，各种年龄、各种民族、各种职业的人口有多少，出生率、死亡率多大，如何变化，这种量的变化有没有一定的规律，等等。人口统计是把人口作为总体看待的。

统计工作是对事物或现象的一种认识活动。认识是主体对客体的反映，认识活动就是主体对客体的反映活动。经过反映活动，在人们的头脑中留下的成果就叫做认识。主体对客体的反映有两种情况：另一种是不自觉地、被动地反映；另一种是自觉地、能动地反映。人们根据一定的目的，采用一定的方法，自觉地、能动地反映客观事物的活动，就是

^① 资料摘自《初级统计学》第 8 版，[美] Mario. Triola 著，清华大学出版社，第 13 页。

调查研究活动。因而,统计工作就是对客观存在的事物或现象的数量方面进行的一种调查研究活动,包含了一个工作过程。

※ **统计资料是统计工作的结果,也称统计数据。**

统计资料是通过统计工作所获得的能够说明现象特征的各项数字资料以及与之相关的其他实际资料的总称。其表现形式有统计表、统计图、统计年鉴、统计公报、统计报告及其他有关统计数字信息的载体等。如国家统计局每年发布的统计公报中的有关农业、工业和建筑业、固定资产投资、国内贸易、对外经济、交通、邮电和旅游、金融、证券和保险、教育和科学技术、文化、卫生和体育、环境保护、人民生活和社会保障等方面的各种具体数字资料,都是反映我国国民经济和社会发展情况的统计资料。

※ **统计学是一门关于数据的收集、整理、分析的学科。**

数据收集也就是取得统计数据。数据整理是将数据用图表等形式显示出来。数据分析则是通过统计方法研究数据,并对分析的结果进行解释与说明。

统计学是在大量的统计工作实践的基础上逐渐形成和发展起来的一门阐述如何收集、整理、分析和解释统计资料的理论和方法的一门科学。统计学的发展经历了 300 多年的历史,现已形成比较完整的学科体系。

统计一词的三种含义是紧密联系的,统计资料是统计工作的成果,统计工作与统计学之间是实践与理论的关系。

统计工作的过程

统计工作是对客观存在的事物或现象的一种认识活动。作为一种认识活动,就必然有一个对客观事物的认识过程。统计认识过程是:从定性认识到定量认识,再到定量与定性相结合。例如,要了解某地区居民家庭收入情况,必须首先搞清居民家庭收入的概念,明确计算范围和计算方法,才能进行调查,这属于定性认识。然后,从调查该地区每一户居民的收入情况开始,经过计算整理得出对该地区居民家庭收入水平总体情况的认识。这是从对个体数量表现的认识过渡到对总体数量特征的认识的过程,属于定量认识。然而,这种定量认识还不是统计认识的终结,还必须与新的定性认识相结合,即调查结果所形成的数据说明这一地区居民家庭收入究竟达到一个什么水平,是贫困、温饱,还是小康?一般统计认识的全过程分为四个环节或阶段,依次为统计设计、统计调查、统计整理和统计分析。

※ **统计设计即对整个统计工作做出全面安排和规划的阶段。**

统计设计的主要内容有:统计指标和指标体系的设计,统计分类和分组的设计,统计表的设计,统计资料收集方法的设计,统计调查、整理和分析研究方案的设计,统计工作各部门和各阶段之间相互协调与联系的设计,统计力量的组织与安排,等等。

统计设计是统计工作顺利进行的前提。因为这一阶段是对现象进行初始定性认识,为定量认识作准备的。只有事先进行设计,才能统一认识、统一步骤,使整个统计工作有秩序地、协调地进行,以保证统计工作的质量。

※ **统计调查即统计数据的收集阶段。**

统计数据从使用者的角度看一般来源于两种渠道:一种是来源于直接的调查或科学实

验，这是统计数据的直接来源；另一种是来源于别人调查或实验的数据，这是统计数据的间接来源。直接来源的统计数据的收集是指直接向调查单位（个体）搜集未经加工整理、需要由个体过渡到总体的统计数据的过程；间接来源的统计数据的收集是指根据研究目的，搜集已经经过加工整理、由个体过渡到总体、能够在一定程度上说明总体现象的统计数据的过程。但是，从统计数据本身的来源看，统计数据最初都是来源于直接的调查或实验。从这个角度来讲，统计数据的收集主要是指根据统计设计的要求，向调查总体中的各个单位搜集统计数据的工作阶段。

统计调查是整个统计工作的基础环节。因为这一阶段是统计认识活动由初始定性认识过渡到定量认识的阶段，这个阶段所收集的资料是否客观、周密，直接关系到统计整理的好坏，关系到统计分析结论的正确与否，决定着统计工作的质量。

※ **统计整理是将收集到的统计数据进行分组汇总的过程。**

统计数据的整理在统计数据采集和统计分析之间起着承前启后的衔接作用，它是统计分析的前提。通过统计数据的收集，我们得到的主要是表明个别单位情况的原始资料，其次还有次级资料，这些资料是分散的、不系统的。仅仅根据这些资料不可能认识现象的总体，也不可能从数量方面反映出现象总体特征和规律。为此，必须按照统计研究的任务和需要对统计数据进行整理。

由于统计调查收集到的数据来源不同，因而统计数据整理的具体内容有两种：一是对原始资料的整理，其目的是将大量的说明总体实际情况的总体单位的标志表现经过审核、分组、汇总，最终得到说明总体数量特征的统计指标的数值。二是对次级资料的整理，其主要目的是为了保证统计资料的动态可比性。但应该注意到，有时由于时间、客观条件的变化或统计分析研究目的发生变化，也会导致所需要的统计指标的计算口径、计算范围等发生变化，这就必须对次级资料按照新的统计方法和要求进行重新调整和整理。

※ **统计分析是指运用统计分析方法，对整理过的统计资料进行分析研究的过程。**

统计分析是统计工作的最终环节。因为这一阶段是对现象得出定量与定性相结合的深刻认识的阶段，也是发挥统计职能的关键阶段。

统计工作的四个阶段，虽然先后有序，内容不同，但它们之间是紧密联系的。任何一个阶段的工作做得好坏直接影响整个工作质量的好坏。由于统计设计的内容渗透于各个工作阶段，因而将其分别归并到统计调查、统计整理和统计分析三个阶段中讲述。这也正是我们将统计学归纳为一门关于数据的收集、整理、分析的学科的原因所在。

统计的主要研究方法

统计工作各阶段有着不同的工作内容和要求，与之相适应就需要使用各种专门的研究方法，即大量观察法、统计分组法、综合分析法与归纳推断法。

※ **大量观察法是指对总体中的全部或足够多的个体进行观察的方法。**

由于各种现象都是在诸多因素的作用下形成的，其中个体的特征表现之间存在很大的差异，如果孤立地对其中少数个体进行观察，其结果往往不足以反映现象总体的一般特征。必须对总体中的全部或足够多的个体单位进行观察，经过综合概括，抽象各单位之间的差异，才能显示出所研究的客观事物的本来面目，使我们能正确地 from 总体上把握客观事

物的本质特征和必然规律。例如,我们想了解我国人口的性别比例,如果只调查个别家庭或少数家庭,情况可能大不相同,有的男多女少,有的女多男少,但是,经过大量观察,男女人数则基本趋向均衡。据第五次人口普查资料显示,祖国内地 31 个省、自治区、直辖市和现役军人的人口中,男性为 65 355 万人,占总人口的 51.63%,女性为 61 228 万人,占总人口的 48.37%,性别比(以女性为 100,男性对女性的比例)为 106.74。这就表明,尽管个别现象受偶然性因素的影响出现偏差,但观察数量达到一定程度,就呈现出一定的规律性。统计研究就是从全部的或足够多的个体观察入手最后达到对总体数量的认识。

※ **统计分组法是指根据统计研究任务和被研究现象总体的特点,按照一定的标志,将研究总体划分为若干个不同性质或不同类型的组成部分的一种统计研究方法。**

统计在研究复杂现象总体时,各单位不仅有量的差别,而且还存在质的差别。从数量方面认识事物不能离开事物质的方面,将所研究的现象总体区分为不同性质的组成部分是统计进行加工整理和深入分析的前提。如对国民经济按经济类型、行业分组,可以使人们清楚地观察不同经济类型和不同行业的发展规模、发展变化规律以及相互之间的依存关系。再如,我们将某一班学生的某门课考试成绩收集来,若只从单个学生的情况来看,是观察不出规律的,但如果按照“成绩”分组并加以整理和汇总后,便可以看出该班学生本门课考试的基本情况。

※ **综合分析法是指运用各种综合指标,并对其进行分解和对比分析,以研究总体的数量差异和数量关系。**

对大量原始数据进行整理汇总,通过计算各种综合指标显现出现象在具体时间、地点以及各种因素共同作用下所表现的规模、水平、集中趋势和差异程度等,概括地描述总体的综合特征和变动趋势,这是综合分析法的主要内容。因而,综合分析法包括对比分析、平均分析、变异分析、指数分析等,既可以从静态上,也可以从动态上综合反映现象的数量表现及数量关系。

例如,我们可以运用某一时间的国土面积、人口总数、国内生产总值、人均国民收入等指标来说明我国特定时间的基本国情,也可以将不同时间的人口总数对比分析全国人口的动态变化情况。

※ **归纳推断法指的是由样本指标对全及总体的数量特征作出推断。**

归纳推断法是归纳法在统计推理中的应用。归纳法是指由个别到一般,从事实到概括的逻辑推理方法,它可以使我们从具体的事实中得出一般结论。在统计实践中,我们所观察的往往只是部分或者有限的单位,而所需要判断的总体所包括的单位却是大量的,甚至是无限的。这就需要我们根据局部的样本资料对整个总体数量特征做出统计推理,并且这种推理要有一定的置信标准要求。以一定的置信标准要求,根据样本数据来判断总体数量特征的归纳推理方法称为归纳推断法。它可以用于总体的参数估计和假设检验,现已成为我国统计分析的一种重要方法。例如,我们想了解某电子元件厂产品的平均使用寿命,就只能采取科学的抽样方法从该批电子元件产品中抽取很少一部分进行检验,以此来推断该批电子元件的平均使用寿命。

大量观察法是统计调查阶段的基本方法,统计分组法是统计整理阶段的基本方法,而综合分析法和归纳推断法是统计分析阶段的两个基本方法。之所以在统计分析阶段存在两种基本方法,是由于在统计调查阶段存在全面调查与非全面调查两种情况。如果是全面调

查, 我们可以通过统计数据的整理与计算得到综合指标对现象总体直接进行综合分析; 如果是非全面调查, 通过统计数据整理与计算得到的首先是部分个体 (即样本) 的指标数值, 因而必须据此推断现象总体的指标数值, 也就需要运用归纳推断法。所以, 从统计分析的角度来讲, 统计按其所用的方法又可以分为描述统计与推断统计。

描述统计与推断统计

※ 描述统计就是指如何从已知的观察资料, 收集、整理、分析研究并提供统计资料的理论与方法, 用以说明研究现象的情况和特征。

描述统计包括各种数据处理, 这些数据的处理是用来总括或描述数据的重要特征的, 而不必深入一层地去试图推论数据本身以外的任何情况。因此, 描述统计的主要作用是通过对现象进行调查或观察, 将所得到的大量数据加以整理、简缩、制成统计图表, 并就这些数据的分布特征 (如集中趋势、离散趋势等等) 计算出一些概括性的数字 (如平均数、标准差等)。借助于这些概括性数字, 从杂乱无章的资料中取得有意义的信息, 不仅可以描述该总体的特征, 而且便于对不同的总体进行比较, 从而作出结论。

※ 推断统计是指只凭样本资料以推断总体的数值的技术和方法。

推断统计可以利用样本资料来代替总体资料, 在观察资料的基础上深入一步地发掘进行分析、研究和推断, 以推知资料本身以外的情况和数量关系, 从而对不肯定的事物作出决断, 为管理与决策提供数据依据。

从样本调查结果以推断总体, 不仅限于推算总体平均数或比率, 还包括假设检验, 这在科学研究中是一个非常重要的方法。例如, 在工业生产中可以用于检验两种不同的工艺方法所生产的产品在质量上是否有差别, 从而判断一种新的工艺方法是否优于原有的工艺方法。又如, 在药品生产试用阶段, 就可以通过大量观察检验新药与旧药之间是否存在显著性的差异, 从而判定新药较旧药疗效是否更好、更安全, 并确定其临床推广应用的价值。

由此可见, 描述统计的应用有助于节约为提供全部数据所必须花费的时间和篇幅, 使客观现象的统计数据可以一目了然, 条理清晰, 使用方便。而推断统计由于节省时间、人力和物力, 因而备受人们的重视和欢迎。但是, 描述统计是推断统计的前提, 而推断统计是描述统计的发展。

内容回顾

统计一词有统计工作、统计资料和统计学三种含义。统计工作是对客观存在的事物或现象的数量方面进行的一种计量活动; 统计资料是统计工作的结果, 也称统计数据; 统计学是一门关于数据的收集、整理、分析的学科。三种含义之间是密切联系的, 统计资料是统计工作的成果, 统计工作与统计学之间是实践与理论的关系。

统计工作包含统计设计、统计调查、统计整理与统计分析四个阶段。设计是对整个统计工作做出全面安排和规划; 调查是收集统计资料的过程; 整理是将收集到的统计数据进行分析分组汇总; 分析是指运用统计分析方法, 对整理过的统计资料进行分析研究的过程。调

查是统计工作的基础阶段，整理是统计工作的中间环节，分析是统计工作的目的所在。

大量观察法、统计分组法、综合分析法和归纳推断法分别是统计调查、统计整理和统计分析的主要研究方法，贯穿整个统计工作过程，也是统计学科的灵魂所在。大量观察法是指对总体中的全部或足够多的个体进行观察的方法；统计分组法是指根据统计研究任务和被研究现象总体的特点，按照一定的标志，将研究总体划分为若干个不同性质或不同类型的组成部分的一种统计研究方法；综合分析法是指运用各种综合指标，并对其进行分解和对比分析，以研究总体的数量差异和数量关系；归纳推断法指的是由样本指标对全及总体的数量特征作出推断。

从统计分析方法的角度来看，统计又可分为描述统计和推断统计。描述统计就是指如何从已知的观察资料，收集、整理、分析研究并提供统计资料的理论与方法，用以说明研究现象总体的情况和特征。推断统计是指只凭样本资料以推断总体的数值的技术和方法。描述统计是推断统计的前提，而推断统计是描述统计的发展。

第一章 描述统计的基础

我国地大物博，人口众多，是一个多民族的国家。这个结论是如何得到的呢？你能拿出有力的证据来说明吗？

——章节问题

内容向导

- ◇ 描述统计的几个基本概念
 - ✧ 总体与总体单位
 - ✧ 标志与标志表现
 - ✧ 统计指标与指标体系
 - ✧ 变异与变量
- ◇ 统计数据的来源
 - ✧ 统计数据的种类
 - ✧ 统计调查方案
 - ✧ 统计数据的搜集方法
 - ✧ 统计调查方式
 - ✧ 统计误差

第一节 描述统计的几个基本概念

一、总体与总体单位

统计是对客观存在的事物或现象整体的数量方面进行的一种计量活动。当我们要了解全国工业企业的生产经营情况时，生产经营活动是一种现象，而这种现象是企业的行为过程，企业就是作出这种行为的具体事物，事物的整体就是我国所有工业企业。要对全国工业企业的生产经营情况整体的数量特征进行计量，需从每一个工业企业着手调查。仅对其中一个工业企业的生产经营情况的调查不能得到全国的整体的数据，但如果所调查的每一个工业企业生产规模、职工人数、生产的产品品种与数量等等所有的特征表现完全一样，也就不需要了解事物或现象的整体情况，只需调查一个工业企业即可。然而，当你调查的既有工业企业，又有商业企业时，最终得出的结论就不只限于全国的工业企业的生产经营情况，也包括了商业企业的生产经营情况。由此可见，全国所有的工业企业是所要了解的整体，而每一个工业企业是个体。使用统计术语来表述，全国所有的工业企业为总体，每一个工业企业为总体单位。即总体是根据一定的目的和要求所确定的研究事物的全体。

※ 总体是由客观存在的、在某种共同性质基础上所组成的许多个别事物的整体。每一个个别事物即为总体单位，简称单位。

要构成统计总体必须具备三个基本特征。

1. 同质性

同质性是指总体各单位必须具有某种共同的性质，才能构成统计总体。如工业企业总体中每个工业企业具有相同的经济职能，都是进行工业生产经营活动、向社会提供工业产品的，这样才能组成工业企业总体。同质性是总体的根本特征，是构成统计总体的前提条件。

2. 大量性

大量性是指总体是由许多总体单位组成的，一个或少数单位不能形成总体。这是因为统计研究的目的是要揭示现象发展变化的一般规律，而事物的发展变化规律只能在大量事物的普遍联系中表现出来。如前面谈到的要研究我国工业企业的生产经营情况，就不能只用个别几个企业的情况来代替，因为个别企业有其各自的特殊性和偶然性，我们只有对大量的或是足够多的企业进行研究，才能使个别单位偶然因素的作用相互抵消，从而显示出总体的本质和规律性。大量性是统计研究的必要条件。

3. 变异性

变异性是指总体各单位除了必须在至少一方面性质相同外，在其他方面必须有不同的表现。例如工业企业总体中，每个企业除了具有相同的经济职能外，其他方面如经济类型、从业人数、产值、成本、利税额等等，就各不相同。变异是普遍存在的，统计研究就是要在个别事物的差异中寻找共性，以揭示其活动的规律性。变异性是进行统计的前提条件。

上述三个特征缺一不可，只有同时具备这三个特征，才能形成统计总体，才能进行统计研究。

思考一下

要检验一批灯泡的使用寿命，总体与总体单位分别是什么？研究一个学校某班学生的学习情况，总体与总体单位又是什么？而要研究合理化建议的提出和采纳情况呢？

总体和总体单位之间的关系属于整体与个体的关系。两者的划分不是固定不变的，而是相对的，它们会随着研究目的的改变而变换。当研究目的和任务确定后，统计总体和相应的总体单位就产生和固定了下来。例如，研究全国工业企业的经营情况，全国所有工业企业是总体，每一个工业企业是总体单位。如果研究目的改变为研究某一个工业企业的经营情况，则这个企业就不再是总体单位，而变成总体了。

当我们对事物或现象的总体情况进行了解时，必须从每一个总体单位开始，如果总体中所包含的个别事物即总体单位是可以计数的，称之为有限总体。如上面提到的我国工业企业总体。如果总体中所包含的个别事物即总体单位是难以计数的，称之为无限总体。如，太空星球总体、世界上的植物总体、动物总体、大量连续生产的某种小件产品总体等等则属于无限总体。从理论上讲，总体都应是有限的，只是由于各种条件的限制，使得有

些总体的个别单位不可计数,才把这些总体假定为无限总体。对有限总体既可以进行全面调查,也可以进行非全面调查,而无限总体只能进行非全面调查。

二、标志与标志表现

要对全国工业企业的生产经营情况整体的数量特征进行计量,需从每一个工业企业着手调查。如调查得到某工业企业的从业人数是10 000人,年利税额200万元。每个工业企业都有从业人数、年利税额,企业不同其具体数值也存在不同,这是由总体的差异性所决定的。这里用来说明工业企业特征的从业人数、年利税额等就是标志,“10 000人”“200万元”则是该工业企业从业人数、年利税额的特征表现,也就是标志表现。

※ 标志是说明总体单位特征的名称。标志表现是指每一个总体单位所表现的具体属性或数量特征,一般用文字或数值来表现。

例如,某学生的性别是男、年龄19岁、文化程度是高中,这里,“性别”、“年龄”、“文化程度”是标志,而“男”、“19岁”、“高中”就是标志表现。

标志表现可能是文字,也可能是数值。按标志表现是文字还是数值将标志分为品质标志和数量标志。品质标志是说明总体单位属性特征的名称,一般用文字表示。某学生的性别是男、民族是汉族,这时“男”、“汉族”分别表明该学生在“性别”和“民族”方面的属性。这里“性别”、“民族”是品质标志,“男”、“汉族”则为该学生这两个品质标志的具体表现。数量标志是说明总体单位数量特征的名称,用数值(包括计量单位)表示。例如,某学生年龄为19岁。这里的“年龄”是数量标志,而“19岁”是“年龄”的标志表现,又称标志值。

总体不仅具有差异性,还具有同质性。因而在同一总体中至少会出现一个所有总体单位的其标志表现完全相同的情况,将这一标志称之为不变标志。例如,全国工业企业为总体,每个工业企业的经济职能是相同的,都是向社会提供工业产品的,“经济职能”就是一个不变标志。反之,当一个总体各单位在某一标志名称下的具体表现不完全相同时,这一标志就是可变标志。如前例中的,从业人数、年利税额等,其具体表现在总体各个单位上不尽相同,属于可变标志。

三、统计指标与指标体系

通过对客观存在的个别事物即总体单位的调查了解,目的是要得到反映的事物或现象总体的数量特征。如调查得到各个工业企业的从业人数,年利税额、企业类型等具体的标志表现,经整理汇总得到全国工业企业总的从业人数、年上缴利税总额和不同类型的工业企业所占的比重等一些数值。这些说明总体数量特征时所用到的从业人数、年上缴利税总额和国有企业所占比重即为统计指标。由此可见,标志与指标是相互对应的。

※ 统计指标是反映事物或现象总体数量特征的概念。

在不同的场合,其使用时有所不同,有时仅指反映事物或现象总体数量特征的名称。有时是指反映事物或现象总体数量特征的名称和具体数值。

一般认为,对统计指标的这两种理解都是合理的,只是适用于不同场合。在一般性统计设计时,只能设计统计指标的名称、计量单位和计算方法,这是不包括数值的统计指标。然后,经过搜集资料、汇总整理、加工计算可以得到统计指标的具体数值,用来说明