

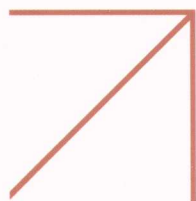


高等学校应用型本科管理学
“十二五”规划教材



统 计 学

Statistics



主 审 姚 旭
主 编 曹尔黎
副主编 刘美艳 李翠亭

C8
201413

P1

高等学校应用型本科管理学
“十二五”规划教材

统 计 学

主 审 姚 旭
主 编 曹尔黎
副主编 刘美艳 李翠亭



中国金融出版社

责任编辑：丁 芊

责任校对：张志文

责任印制：丁淮宾

图书在版编目 (CIP) 数据

统计学 (Tongjixue) / 曹尔黎主编. —北京: 中国金融出版社, 2013. 7

高等学校应用型本科管理学“十二五”规划教材

ISBN 978 - 7 - 5049 - 7030 - 5

I. ①统… II. ①曹… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2013) 第 137515 号

出版
发行 **中国金融出版社**

社址 北京市丰台区益泽路 2 号

市场开发部 (010) 63266347, 63805472, 63439533 (传真)

网上书店 <http://www.chinafph.com>

(010) 63286832, 63365686 (传真)

读者服务部 (010) 66070833, 62568380

邮编 100071

经销 新华书店

印刷 保利达印务有限公司

尺寸 185 毫米 × 260 毫米

印张 16

字数 352 千

版次 2013 年 7 月第 1 版

印次 2013 年 7 月第 1 次印刷

定价 34.00 元

ISBN 978 - 7 - 5049 - 7030 - 5/F. 6590

如出现印装错误本社负责调换 联系电话 (010) 63263947

前 言

随着社会、经济和科学技术的发展,统计在现代化国家管理、企业管理中的地位和社会生活中的地位越来越重要,可以说人们的日常生活和一切社会生活都离不开统计。本教材为人们学习统计学理论和方法而编写,可以作为高校经济管理类专业教材,也可以作为统计学爱好者的学习参考书。

我们赞同一些统计专家学者对统计学概念的定义,认为统计学是一门关于数据的科学,是搜索、整理、显示、分析数据的手段,其研究目的是探索数据内在的数量规律,预测现象未来的数量特征,教材内容按顺序依次介绍了搜集数据、整理数据、分析数据理论和方法。我们也赞同社会统计学与数理统计学之间是互补的、不可替代的关系的观点,认为描述统计是统计工作的基础,是统计研究的第一步;推断统计是统计学的核心和统计研究的关键环节。教材在重点介绍描述统计基本内容的同时,也较为详细地、系统地介绍了推断统计的内容。

计算机技术和一系列新技术、新方法在统计领域的开发和应用,使统计数据的搜集、处理、分析过程日益现代化,提高了统计工作的效能,计算机科学已经成为统计科学不可分割的组成部分。教材结合一些重点内容介绍了 Excel 软件在统计分析中的应用,便于读者掌握利用计算机软件解决统计问题的方法。我们也设计了数值较小、易于手工操作的例题,读者能够通过简单的计算掌握统计测度的含义和计算过程。在思考和练习部分设计了填空、判断、选择和计算等多种题型,帮助读者深刻地理解统计理论和方法。

教材的第1章、第2章由哈尔滨金融学院管理系曹尔黎编写,第3章由黑龙江工程学院乔冠华和黑龙江大学吴艳玲编写,第5章、第6章、第9章由哈尔滨金融学院管理系李翠亭编写,第4章、第7章、第8章由哈尔滨金融学院管理系刘美艳编写。由于编者水平有限,不当之处,恳请读者批评指正。

编 者

2013年2月27日

目 录

1 总论	1
1.1 数据与统计学	1
1.2 统计学的产生和发展	1
1.3 统计学的分科	2
1.3.1 描述统计和推断统计	3
1.3.2 理论统计和应用统计	3
1.4 数据的来源	4
1.5 统计数据的质量	4
1.6 统计学中的几个基本概念	5
1.6.1 统计总体和样本	5
1.6.2 标志和变量	6
1.6.3 统计指标和指标体系	7
本章小结	8
思考与练习	8
2 数据的整理与测度	13
2.1 数据的整理	13
2.1.1 数据的分组	13
2.1.2 分布数列	16
2.1.3 次数分布直方图	17
2.1.4 洛伦兹曲线与基尼系数	20
2.2 分布集中趋势的测度	21
2.2.1 众数	21
2.2.2 中位数	22
2.2.3 分位数	23
2.2.4 均值	24
2.2.5 几何平均数	26
2.2.6 众数、中位数和均值之间的关系	26
2.3 分布离散程度的测度	27
2.3.1 极差	27
2.3.2 内距	27
2.3.3 方差和标准差	28

2.3.4 离散系数	30
2.4 分布偏态与峰度的测度	30
2.4.1 偏态及其测度	31
2.4.2 峰度及其测度	31
2.5 统计表和统计图	33
2.5.1 统计表	33
2.5.2 统计图	34
本章小结	37
思考与练习	37
3 概率、概率分布与抽样分布	45
3.1 事件和概率	45
3.1.1 随机试验、事件和样本空间	45
3.1.2 事件的概率	46
3.2 随机变量及其概率分布	46
3.2.1 随机变量	47
3.2.2 离散型随机变量的概率分布	47
3.2.3 离散型随机变量的数学期望和方差	48
3.2.4 常用的离散型随机变量的概率分布	49
3.2.5 连续型随机变量的分布函数	51
3.2.6 常见的连续型随机变量——正态分布	52
3.3 抽样调查的组织方式	56
3.3.1 简单随机抽样	56
3.3.2 分层抽样	57
3.3.3 系统抽样	57
3.3.4 整群抽样	57
3.4 抽样分布	58
3.4.1 抽样分布的概念	58
3.4.2 $\bar{x}$ 抽样分布的形式和特征	59
3.4.3 样本比率抽样分布的形式和特征	62
3.4.4 两个样本统计量的抽样分布	63
本章小结	64
思考与练习	64
4 参数估计	70
4.1 参数估计的基本原理	70
4.1.1 估计量与估计值	70

4.1.2	点估计与区间估计	70
4.1.3	评价估计量的标准	72
4.2	一个总体参数的区间估计	73
4.2.1	总体均值的区间估计	73
4.2.2	总体比率的区间估计	76
4.3	两个总体参数的区间估计	76
4.3.1	两个总体均值之差的区间估计 (大样本)	76
4.3.2	两个总体比率之差的区间估计	77
4.4	样本量的确定	78
4.4.1	估计总体均值时样本量的确定	78
4.4.2	估计总体比率时样本量的确定	79
	本章小结	79
	思考与练习	80
5	假设检验	85
5.1	假设检验概述	85
5.1.1	假设检验的基本概念	85
5.1.2	两种类型的错误	88
5.1.3	检验功效	89
5.1.4	单侧检验与双侧检验	90
5.1.5	利用 P 值进行决策	90
5.2	一个总体参数的检验	93
5.2.1	总体均值的检验	93
5.2.2	总体比率的检验	102
5.2.3	总体方差的检验	103
5.3	两个总体参数的检验	105
5.3.1	两个总体均值之差的检验	106
5.3.2	两个总体比率之差的检验	114
5.3.3	两个总体方差比的检验	115
	本章小结	118
	思考与练习	118
6	相关与回归分析	124
6.1	相关分析	124
6.1.1	相关关系的概念	124
6.1.2	相关系数	127
6.1.3	斯皮尔曼等级相关系数	130

6.2 一元线性回归分析	132
6.2.1 相关分析与回归分析的联系	132
6.2.2 总体回归函数与样本回归函数	133
6.2.3 回归系数的普通最小二乘估计	135
6.2.4 拟合优度的度量	140
6.3 线性回归的显著性检验与回归预测	142
6.3.1 回归系数显著性的 t 检验	142
6.3.2 一元线性回归模型的预测	143
本章小结	146
思考与练习	147
7 时间序列分析与预测	152
7.1 时间序列的描述性分析	152
7.1.1 时间序列的含义	152
7.1.2 时间序列的图形描述	153
7.1.3 时间序列的速度分析	154
7.2 时间序列及其构成因素	157
7.2.1 时间序列的构成因素	157
7.2.2 时间序列构成因素的组合同模型	159
7.3 长期趋势的变动分析	160
7.3.1 测定长期趋势的移动平均法	160
7.3.2 测定长期趋势的指数平滑法	163
7.3.3 测定长期趋势的模型法	166
7.4 季节变动分析	172
7.4.1 季节变动分析的原始资料平均法	172
7.4.2 季节变动分析的趋势——循环剔除法	174
7.5 循环变动分析	176
7.5.1 循环变动及其测定目的	176
7.5.2 循环变动的测定方法	176
本章小结	180
思考与练习	181
8 统计指数	186
8.1 统计指数的概念和分类	186
8.1.1 统计指数的概念	186
8.1.2 指数的作用	187
8.1.3 统计指数的分类	187

8.2 总指数的编制方法	188
8.2.1 总指数编制的基本问题	188
8.2.2 加权总指数的编制	190
8.2.3 加权综合指数的主要形式	194
8.3 指数体系与因素分析	196
8.3.1 指数体系及其作用	196
8.3.2 总量变动的因素分析	197
8.4 几种常用的经济指数	198
8.4.1 消费者价格指数	198
8.4.2 生产者价格指数	199
8.4.3 股票价格指数	200
8.4.4 农副产品收购价格指数	201
8.5 综合评价指数	201
8.5.1 综合评价及其方法	201
8.5.2 构建综合评价指数的基本问题	202
8.5.3 综合评价指数的编制方法	203
本章小结	207
思考与练习	208
9 国民经济统计基础知识	213
9.1 国民经济统计的基本内容和基本原则	213
9.1.1 国民经济运行与国民经济核算	213
9.1.2 我国国民经济核算体系的基本框架	214
9.1.3 国民经济核算的基本原则	215
9.2 国民经济统计的主要分类	216
9.2.1 国民经济分类的作用	216
9.2.2 机构部门分类	217
9.2.3 活动部门分类	217
9.3 国民经济统计的基本指标	219
9.3.1 国民经济统计指标体系	219
9.3.2 国民经济生产指标	219
9.3.3 国民收入分配指标	223
9.3.4 对外经济往来指标	225
9.3.5 资产负债和国民财富指标	226
9.4 国民经济统计的常用分析指标	228
9.4.1 国民生产分析指标	228
9.4.2 收入分配分析指标	230

9.4.3 最终消费和投资分析指标	233
9.4.4 其他重要的分析指标	234
本章小结	236
思考与练习	237
参考文献	241
附录 常用统计表	242
附表1 标准正态分布表	242
附表2 t 分布表	243
附表3 χ^2 分布表	244

总 论

内容提要

统计学是关于数据的科学。本章围绕数据与统计学的关系展开,阐述了数据的来源、统计学的性质和研究目的,同时还介绍了统计研究中的一些最基本的概念。

1.1 数据与统计学

统计学是收集、整理、显示和分析数据的科学,其目的是探索现象的数量规律。一般来说,收集是指通过测量、调查等方法取得数据;整理是对数据分组、观察其分布的情况;显示就是用图、表的形式呈现数据的特征和现象的数量规律;分析是用统计方法研究数据,探索现象的数量规律。

关于数据。首先,由于个别现象的数据有特殊性、偶然性,而对大量同类个体数据的综合才具有相对的普遍性和稳定性,便于我们探索和发现现象的数量规律,因此,“数据”是大量同类个体的特征。例如,某市进行城市住户调查,对该市的200户居民家庭的劳动就业状况、收入和现金支出的数据逐一登记,目的是反映该城市全部居民家庭劳动就业、家庭收入和现金支出的数量特征和数量变化规律。其次,“数据”与数学中抽象的数量不同,它是个体在具体时间、地点、条件下的特征。另外,“数据”是对个体特征测量或登记的结果,因此数据不但有数字型的,也有非数字型的。例如,人对某一事物的态度是“赞成”和“反对”。在统计研究工作中通常会对非数字型的数据作数据化处理,例如用“1”表示赞成,“0”表示反对。

探索现象数量规律是统计研究的最终目的。数量规律是应用统计方法从偶然性中探索到的现象内在的、本质的数量规律。例如,出生婴儿的性别比为105-107;重复投掷均匀的硬币出现正面和反面的比率接近 $1/2$;某城市随机抽出300户居民家庭,其平均月生活费支出为4000元,用300户家庭的平均月生活支出,推断该城市全部家庭的平均月生活费的范围时的数量规律;研究企业某产品的广告费投入与产品销售额之间依存关系的数量规律等。

统计学是关于数据的科学,它提供了如何收集、整理、显示和分析数据,探索现象数量规律的方法。统计学来源于对数据的研究,目的也在于统计数据的研究。没有对数据的研究,统计学就失去了存在的意义。

1.2 统计学的产生和发展

统计学产生于17世纪中叶,是从几个不同的领域开始的。统计学的一个源头是

1676年英国的经济学家威廉·配第（William Petty）的《政治算术》。在这部书中，他利用实际资料，运用数字、重量和尺度等统计方法对英国、法国和荷兰三国的国情国力，作了系统的数量对比分析，从而为统计学的形成和发展奠定了基础。因此，在《马克思恩格斯全集》中马克思称：“威廉·配第是政治经济学之父，在某种程度上也是统计学的创始人。”

统计学的另一个源头是英国的约翰·格朗特（John Graunt）。格朗特以1604年伦敦教会每周一次发表的死亡公报为研究资料，在1662年发表了《关于死亡公报的自然和政治观察》的论著。书中分析了60年来伦敦居民死亡的原因及与人口变动的关系，首次提出通过大量观察，可以发现新生儿性别比例具有稳定性和人口不同死因的比例具有稳定性等规律；并且第一次编制了“生命表”对人口的死亡率和寿命作了分析，从而引起了普遍的关注。他的研究表明了统计学作为国家管理工具的重要作用。

统计学的第三个源头是古典概率论，奠基人包括法国的布莱斯·帕斯卡（B. Pascal）和皮埃尔·德·费马（Pierre de Fermat）。帕斯卡和费马通过通信的方式，将赌博中出现的各种问题，例如掷三颗骰子可能出现的点数，为什么掷三颗骰子10点出现次数总是多于9点出现次数等问题，归纳为一般的概率原理，为后来概率论和统计学的发展奠定了重要的基础。

自17世纪中叶上述几位科学家从不同角度开始了统计学研究后，经过几代统计学家的努力，到19世纪末建成了古典统计学（主要是描述统计学）的基本框架。

20世纪初许多重要的统计方法和观点开始建立并逐步完善，如著名的统计学家费希尔（R. A. Fisher）建立的相关和回归分析、多元分析、以最大似然估计为中心的点估计理论、方差估计法、费希尔的F分布；1908年，英国的威廉·希利·高塞特（W. S. Gosset）导出的 t 分布，利用 t 分布可以采用小样本对全部产品的质量状况进行检验和推断；J. 奈曼（Jerzy Neyman）和E. S. 皮尔逊（E. S. Pearson）的置信区间估计和假设检验理论等。到20世纪中叶构筑了现代统计学的基本框架。

从20世纪50年代以来，统计理论、方法和应用进入了全面发展的新阶段。一方面，统计学受计算机科学、信息论等现代科学技术的影响，新的研究领域层出不穷，如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计、线性统计模型、探索性数据分析等。另一方面，统计方法的应用领域不断扩展，几乎所有的科学研究都离不开统计方法。因为不论是自然科学、工程技术、农学、医学、军事科学，还是社会科学都离不开数据，要对数据进行研究和分析就必然用到统计方法。可以说，统计方法与数学、哲学一样成为所有学科的基础。

1.3 统计学的分科

统计学的内容十分丰富，研究与应用的领域非常广泛。从统计教育的角度，统计学大致有以下两类：

1.3.1 描述统计和推断统计

描述统计是通过图形、表格和概括性的数字,对数据资料进行整理、分析的统计方法。描述统计又分为集中趋势分析和离散程度分析等部分。

推断统计是根据样本信息对总体进行估计、假设检验、预测或其他推断的统计方法。

统计学分为描述统计和推断统计,一方面反映了统计发展的前后两个阶段;另一方面也反映了统计研究、探索客观事物内在数量规律性的先后两个过程。我们知道,大量同类个体数据的综合,才能反映现象的数量规律和数量特征。而人力、物力、时间、破坏性试验和被研究个体现象的无限性等原因,又使得我们不可能对所有的个体都进行观察和登记,统计研究自然考虑采用由部分推算全部的方法,即研究总体的数量特征要经过两个阶段:首先利用描述统计的方法研究样本的数量特征,然后采用推断统计的方法来推算总体的数量特征和规律性。显然,描述统计是基础,是统计研究工作的第一步,没有描述统计收集可靠的数据并提供有效的信息,即使高明的统计学家和科学的推断方法也难以得出准确的结论。推断统计是现代统计学的核心和统计研究工作的关键环节,因为统计最终能否科学准确地探索到现象总体内在的数量规律与选用何种统计量,选用什么推断方法,如何进行推断有着直接的联系。一个出色的统计工作者的能力和技巧在推断统计中将得到充分的体现和检验。

1.3.2 理论统计和应用统计

理论统计是指统计学的数学原理。由于现代统计科学几乎用到了所有的数学知识,要成为优秀的统计工作者就必须经过严格的数学训练,特别是从事统计理论和方法研究的人,必须有坚实的数学基础。从广义角度来讲,统计学应该包括概率论,因为概率论是统计推断的数学基础,而概率论是数学的一个分支,所以理论统计应该是包括概率论在内的统计方法数学原理的研究。

在统计工作者中,从事理论统计研究的人只是很少的一部分,大部分是应用统计方法去解决实际问题的应用统计工作者。统计学是一门数据科学,由于在自然科学、社会科学的所有研究和实际工作中都要通过数据来分析和解决问题,统计方法的应用就自然而然地扩展到几乎所有的研究领域。例如,统计方法在物理研究中的应用形成了统计物理,统计方法在生物学中的应用形成了生物统计,统计方法在医学中的应用形成了医疗卫生统计,统计方法在风险管理与保险中的应用形成了保险精算学,统计方法在微观企业管理中的应用形成了管理统计等。以上这些应用统计学的不同分支所应用的基本方法都是一样的,即都是描述统计和推断统计的主要方法。但由于各应用领域都有其特殊性,统计方法在应用中就具有了不同的特点。作为一名优秀的应用统计工作者,不但要能熟练地掌握和应用各种统计方法,而且必须具备所研究和应用领域的专业知识。

1.4 数据的来源

数据主要来源于统计调查。统计调查就是根据统计研究的目的,收集、登记个体特征的工作过程。来源于直接组织的调查、观察和科学试验的数据,称为第一手数据或直接的数据;来源于其他机构研究成果的数据,称为第二手数据或间接的数据。研究现象的数量特征和数量规律,如果没有相关个体的数据可以利用,就要专门组织调查或试验来获取直接数据。例如,CCTV 经济生活大调查、互联网购物消费调查、家庭收支情况调查等,都是为了获取直接的数据。借助于已有的研究成果来分析现象的数量特征,收集的数据就是间接的数据。例如,在国家统计局网站收集我国 2000—2010 年的 GDP 总量以研究其变化规律,收集的数据就是间接数据;在报纸、杂志、统计年鉴、网络上收集数据,从调查公司、数据库公司购买的数据都是间接数据。

统计调查的主要方法有普查和抽样调查两种。

1. 普查。

普查是为某一特定目的,专门组织的一次性全面调查。世界各国都定期地进行人口普查、农业普查。例如,我国在 1982 年进行了第三次全国人口普查,1990 年、2000 年和 2010 年分别进行了第四次、第五次和第六次全国人口普查。普查收集的直接数据内容丰富,综合的结果能准确反映一国国情、国力。但是由于普查涉及面广、调查单位多,需要耗费大量的人力、物力和财力,通常进行一次普查要间隔较长的时间。2003 年,国家统计局、国家发展和改革委员会和财政部联合发出通知,对我国的普查项目和周期作了调整 and 安排,经济普查每 5 年进行一次,在逢 3、8 的年份实施;人口普查、农业普查每 10 年进行一次,在逢 0、6 的年份实施。

2. 抽样调查。

抽样调查是统计调查中应用最广、最重要的调查方法,它是通过随机样本对总体数量规律性进行推断的调查研究方法。虽然抽样调查不可避免地存在着由样本推断总体产生的抽样误差,但统计方法不仅可以估计出误差的大小,而且可以进一步控制这些误差。由于以上这些特点,加之其节省人力、财力、物力,又能保证实效性的特点,抽样调查已经成为科学研究及管理决策最重要的方法之一。

1.5 统计数据的质量

统计研究的整个工作过程就是对数据的加工过程。从原始数据的收集开始,经过整理、显示、样本信息的提取到总体数量规律的科学推断,都存在一个减少误差、提高数据质量的问题。但不同的统计工作阶段统计数据误差的原因不同,严重程度也不同。

统计调查阶段是统计研究的第一步,是直接收集数据的阶段。因而这一阶段统计数据的质量如何,直接影响到整个统计工作。在这一阶段中,从不同的角度分类,可以分为非抽样误差与抽样误差。

非抽样误差是由于调查过程中各有关环节工作失误造成的。它包括调查方案中有关规定或解释不明确所导致的填报错误、抄录错误、汇总错误,不完整的抽样框导致的误差,调查中由于被调查者不回答产生的误差等。非抽样误差在普查、抽样调查中都可能发生。显然,从理论上讲,这类误差是可以避免的。克服或降低非抽样误差,一方面要加强对统计调查人员的培训,增强他们工作责任心,提高数据质量意识,加强填报和汇总时的检查;另一方面要掌握获取完整抽样框的方法,以及科学抽样的方法与技术。在非抽样误差中还有一种人为干扰造成的误差,即有意瞒报或低报数据,这是需要给予特别注意的。例如,对地方 GDP 统计时,由于层层加水造成的误差;在调查市场物价时,某些负责人为表现自己的工作业绩,无视有关统计的法律法规,强行调低物价指数。这种虚报、低报等瞒报的行为都触犯了《统计法》,统计人员要坚决抵制并予以揭露。

抽样误差是利用样本推断总体时产生的误差。由于样本只是总体的一部分,用样本的信息去推断总体,或多或少总会存在误差,因而抽样误差对任何一个随机样本来讲都是不可避免的。但它又是可以计量的,并且是可以控制的。在坚持随机原则的条件下,一般来讲,样本量越大,抽样误差就越小。

概括地讲,非抽样误差特别是其中的系统偏差是可以避免的。但如果不注意,这类偏差造成的结果对调查质量来说又是致命的。美国统计学会于 1995 年专门编写了一本题为《调查误差的主要来源是什么?》的小册子,列出了十种容易犯的错误并给出了应采取的措施。加强统计数据质量的管理要体现在统计研究的全过程中,在描述统计和推断统计阶段都要时刻注意统计方法的科学、准确,注意统计方法的前提条件和假设,要根据统计数据的特点和研究的目的选择统计方法,在统计分析时要注意定性分析与定量分析的结合,等等。在后面的各章中将根据不同的方法,从不同的角度强调统计数据的质量问题。

1.6 统计学中的几个基本概念

统计学是关于数据的科学,收集个别事物的属性、特征,是统计工作的起点;综合、整理以及显示现象数量特征和数量规律是统计工作的最终目的。本节介绍的概念是一直贯穿于统计研究过程的一些基本概念,这些概念有总体、样本、标志、变量、统计指标等。

1.6.1 统计总体和样本

统计总体简称总体,它是由具有某一相同性质的许多个别事物构成的整体。构成总体的个别事物称为总体单位。例如,研究某高校在校学生的生活消费情况,该校所有的在校生组成统计总体,每一位在校生都是一个总体单位。“性质相同”的具体体现是:他们都是某校的在校生。研究某市工业企业的生产、经营情况,该市的所有工业企业构成统计总体,每一个工业企业都是一个总体单位,同属某市、经济职能相同,就是性质相同。

总体是根据研究目的确定的,因此总体和总体单位不是固定不变的。统计研究

目的和任务发生变化时,原来的总体可能成为总体单位,总体单位也可能变为总体,也可能成为与新的研究目的无关的事物。例如,要研究我国汽车生产企业的生产、经营情况,则全国所有的汽车生产企业就构成统计总体,每一个企业就是一个总体单位。如果要研究某一汽车生产企业职工的生活情况,则该企业就是总体(由所有的职工组成),每个职工是总体单位。

样本是从总体中抽取的一部分总体单位构成的整体。构成样本的个别事物称为样本单位。探索总体现象的数量特征和数量规律的起点是收集大量的数据,也就是要调查、登记一些总体单位的数据。但是由于总体中所含总体单位的数目是庞大的和无限的,因此没有必要也不可能对每一个总体单位的数量特征都逐一登记。此时采用推断统计的方法,登记每一个样本单位的数据,整理、分析研究出样本的数量特征,然后估计总体的数量特征。例如,从某城市的全部居民家庭中抽出 300 户家庭,300 户居民家庭构成的总体是样本,抽样的目的是用样本的平均月食品消费支出推算全市居民的平均月食品消费支出等。

1.6.2 标志和变量

标志是说明总体单位属性或特征的名称。标志表现是总体单位在标志上具体表现的属性或数量。例如,所有在校生构成总体,每个在校生是总体单位,年龄、家庭住址、统计成绩、性别、专业都是名称。以“年龄”为例,每个总体单位在这个名称上的表现是不同的,甲的年龄是 20 岁、乙的年龄是 21 岁、丙的年龄是 22 岁,20 岁、21 岁、22 岁都是标志表现。标志表现体现了每个总体单位的属性或特征,而“年龄”是说明总体单位属性或特征的名称,称为标志。总体单位是标志的承担者,没有标志和标志表现就不可能整理、综合出总体的数量特征,标志和标志表现是研究总体数量特征的基础和前提。

根据标志的性质不同,可以把标志划分为品质标志和数量标志两种。品质标志是表明总体单位质的特征的名称。例如,企业的经济类型、所属系统、姓名、性别、文化程度等都是品质标志。品质标志的标志表现一般是文字。例如,“企业经济类型”的标志表现为股份制企业、集体企业、国有及国有控股企业、港澳台及外商投资企业等。

数量标志是表明总体单位量的特征的名称。例如,企业的销售收入、累计利润总额、占地面积、产品产量、纳税额等都是数量标志。数量标志的标志表现为数值。例如,“纳税额”的标志表现为 50 万元、70 万元、130 万元等。数量标志的标志表现称为标志值。“50 万元”就是标志值。

根据标志表现是否相同,可以把标志分为不变标志和可变标志两种。

不变标志是指在总体单位上标志表现相同的标志。例如,全国所有国有企业构成总体,企业的“经济类型”就是不变标志;某校所有在校生组成总体,“所属院校”就是不变标志。总体单位至少存在一个不变标志,是个别事物组成总体的必备条件。

可变标志又称变动标志,是指在总体单位上标志表现不相同的标志。例如,全国所有国有企业构成总体,企业的“利润总额”、“占地面积”、“销售收入”等都

是可变标志；某校所有在校生组成总体，“择业城市”、“喜欢的运动项目”、“每天读课外书籍的时间”等是可变标志。

可变标志在总体单位上具体表现的属性或数值是不同的，这种不同称为变异。例如，可以说各工业企业在“利润总额”上存在变异。可变的品质标志和可变的数量标志都是统计分组和统计分析研究的基础。

变量是可变的数量标志。可变标志有数量标志也有品质标志。例如，在上例中，“利润总额”、“占地面积”、“销售收入”、“每天读课外书籍的时间”、“体重”、“每天看电视时间”都是可变的数量标志，可以称这些数量标志为变量。变量所表现的具体数值称为变量值。例如，“每天读课外书籍的时间”的具体表现为45分钟、60分钟、90分钟等都是变量值。

在统计学的概念中收集数据就是调查、登记总体单位或样本单位在标志上的标志表现，由于品质标志的标志表现可以作数据化处理，因此变量实际上就是可变的标志。

根据变量值是否可列，可以把变量分为离散型变量和连续型变量两种。

如果某一变量的取值是有限的或可列的，该变量称离散变量。“企业数”、“职工人数”、“设备台数”、“年龄”（实际工作中一般只取整数）是离散变量。

如果某一变量的取值是不可列的，该变量称连续变量。“利润总额”、“占地面积”、“销售收入”、“体重”等都是连续变量。

1.6.3 统计指标和指标体系

统计指标（统计测度）是反映现象总体数量特征的概念及具体数值。例如，2012年我国社会消费品零售总额比上年增长14.3%、全年国内生产总值519322亿元、人口出生率为12.10‰、出生人口性别比为117.70、城镇居民家庭恩格尔系数为36.2%等都是统计指标。一个名称（概念）加上具体的数值，用来反映总体在一定时间、地点、条件下的数量特征时就是统计指标，一个完整的指标一般包括六个要素：指标名称、指标数值、时间范围、空间范围、计算方法、计量单位。统计指标是反映现象总体数量特征的唯一方法。

在统计设计阶段统计指标还是一个概念，所以统计指标的概念也定义为：说明现象总体数量特征的概念。

统计指标具有可量性、综合性和具体性三个特点：

统计指标反映的是社会经济现象总体的数量特征，该数量特征用数值来表示，是可以度量或计数的。指标数值是统计指标的构成要素之一，所以统计指标具有可量性。

统计指标是对总体单位数量和属性特征的整理、综合和抽象。根据统计研究目的，利用科学的统计方法，采用标志对总体分组后汇总出各组的总体单位数或对总体各单位的变量值汇总、计算都可以形成统计指标。例如采用“性别”对学生总体分组，计算出各组的总体单位数和各组所含总体单位数在总体中占的比重，两者都是统计指标。统计指标是个别差异的综合和抽象，反映了总体的综合数量特征，所以统计指标具有综合性。