



普通高等教育“十三五”应用型本科规划教材

应用统计学 基础

秦春蓉 主 编
范保珠 陈 娇 胡佳婷 副主编
叶 璐 陈 郑 李晓燕 参 编



清华大学出版社



普通高等教育“十三五”应用型本科规划教材

应用统计学 基础

秦春蓉 主 编

范保珠 陈 娇 胡佳婷 副主编

叶 璐 陈 郑 李晓燕 参 编

清华大学出版社
北 京

内 容 简 介

本书系统阐述了统计学的基本理论与方法。全书共分8章,内容包括绪论、统计数据的搜集与整理、统计数据的描述性分析、时间数列、统计指数、抽样分布与参数估计、相关与回归分析等。为了突出统计学的实用性,本书特增加“Excel的统计数据分析应用”一章。

本书既可作为应用型本科院校非统计专业的经济管理类学生的必修教材,也可作为人口学、社会学、医药学、法学、新闻学等非经管类专业学生的选修教材。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

应用统计学基础/秦春蓉主编. —北京:清华大学出版社,2017

(普通高等教育“十三五”应用型本科规划教材)

ISBN 978-7-302-49534-5

I. ①应… II. ①秦… III. ①应用统计学—高等学校—教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2018)第 026629 号

责任编辑:陈 明

封面设计:傅瑞学

责任校对:王淑云

责任印制:杨 艳

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市国英印务有限公司

经 销:全国新华书店

开 本:170mm×230mm 印 张:17

版 次:2017 年 12 月第 1 版

印 数:1~2500

定 价:42.00 元

字 数:343 千字

印 次:2017 年 12 月第 1 次印刷

产品编号:076483-01

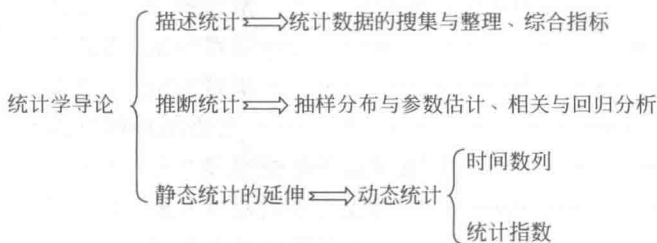
前言

随着现代信息技术的不断进步,数据化已经成为社会发展的一种趋势。在大数据时代,社会的运行、经济的活动、人类的行为都变得可数据化了。做好数据的搜集、整理与分析,许多问题才可迎刃而解。统计学作为进行数据搜集、整理与分析的重要工具性学科,它与人们的生活、工作密不可分,没有任何学科或领域能够真正离开统计学。

社会主义市场经济深入发展的当下,社会、经济、科技、教育彼此依赖的关系更紧密,进而对实用技能型和创新型经济人才的需求也更加迫切,这必然使得经济管理类专业的人才培养目标发生相应的变化,并逐渐趋于应用化和市场化。要使一本统计学教材能真正成为经济管理类专业的基础性、通用性的教材,其内容要从传统的“宏观统计”或“政府统计”转向“商务统计”,这样的统计学教材才更具生命力和适用性。

统计学也是高等院校经济管理类非统计专业的一门核心课程,是一门应用性很强的方法论课程,它能够为学生学习其他专业课程提供数量分析的方法。近几年来,应用型本科院校十分重视统计课程的教学工作,经济管理类专业普遍将统计学作为专业基础课程开设,这对于培养学生基本的数据分析能力十分重要。

本书共8章,在编写上采取由浅入深、循序渐进的思路,尽量避免繁杂的公式推导。各章内容结构设计为:“案例/问题导入——章节导言——统计理论——本章小结——练习题”。整个体系强调应用型本科院校经济管理本科专业的统计学课程教学中所必须的基本原理、基本方法、基本技能及其应用。本书的结构框架如下:



本书的特色是：覆盖统计学经典理论，取舍得当；融入现代教学方法，体系新颖，增加了趣味性；理论教学与实践教学有机结合。与其他应用型教材相比，本书最大的不同在于：在本书的第8章，以“中国旅游市场发展现状”为案例导向，数据源自《中国统计年鉴2016》，通过表格处理软件Excel 2013，从描述统计到推断统计，专门介绍数据的分析应用。包括数据整理与汇总、时间数列长期趋势分析、抽样与区间估计、线性回归分析等，着重探讨影响中国旅游市场发展的主要因素，从而更好地规划中国未来旅游产业的发展。从分析的结果可以清楚地了解未来中国国内旅游收入的发展趋势、城镇人均旅游消费水平趋势、国内旅游人数的变化等。

本书基于编者多年来对经济管理类专业本科教学的实践经验及体会，集思广益编写而成。全书由秦春蓉主编，各章的编写分工如下：李晓燕（第1章）、叶璐（第2章）、陈郑（第3章）、陈娇（第4章）、胡佳婷（第5章）、秦春蓉（第6章）、范保珠（第7章、第8章）。

本书的出版，得到了清华大学出版社的支持与帮助，在此对为本书出版付出心血的编辑们及有关同事、朋友的关心，致以衷心的感谢。还要特别感谢本书所参考和引用的相关资料、案例的作者。

由于编者水平有限，书中不当、疏漏之处在所难免，恳请广大同行及读者批评指正，以期有机会再版时予以修正与完善。

编 者

2017年8月

目 录

第 1 章 绪论	1
1.1 统计的含义	2
1.2 统计学的研究对象	3
1.3 统计学的学科分类	5
1.3.1 描述统计学和推断统计学	5
1.3.2 理论统计学和应用统计学	6
1.4 统计学的几个基本概念	6
1.4.1 总体与总体单位	7
1.4.2 标志与标志值	8
1.4.3 变异与变量	9
1.4.4 指标与指标体系	10
1.5 统计软件简介	12
1.5.1 Excel 简介	12
1.5.2 SPSS 简介	12
1.5.3 SAS 简介	13
1.6 本章小结	13
1.7 练习题	14
第 2 章 统计数据的搜集与整理	16
2.1 统计数据的类型	16
2.1.1 定类数据	17
2.1.2 定序数据	17
2.1.3 定距数据	17
2.1.4 定比数据	18
2.2 统计数据的搜集	18
2.2.1 统计数据搜集的组织形式	19
2.2.2 统计数据搜集的方法	24
2.2.3 统计数据搜集的工具	25

2.3	统计数据的整理	29
2.3.1	数据的预处理	29
2.3.2	定性数据的整理	31
2.3.3	定量数据的整理	38
2.3.4	统计表	50
2.4	本章小结	53
2.5	练习题	54
第3章	统计数据的描述性分析	60
3.1	总量指标	61
3.1.1	总量指标的概念	61
3.1.2	总量指标的作用	61
3.1.3	总量指标的种类	61
3.1.4	总量指标的计量单位	64
3.2	相对指标	65
3.2.1	相对指标的概念	65
3.2.2	相对指标的作用	65
3.2.3	相对指标的数值表现形式	65
3.2.4	相对指标的种类和计算方法	66
3.2.5	相对指标的注意事项	72
3.3	集中趋势指标	73
3.3.1	集中趋势指标概述	73
3.3.2	位置平均数	75
3.3.3	数值平均数	83
3.3.4	各种集中趋势指标的比较	93
3.3.5	集中趋势指标的计算和应用注意事项	94
3.4	离散程度指标	95
3.4.1	离散程度指标的概念	95
3.4.2	离散程度指标的作用	96
3.4.3	离散程度指标的种类和计算方法	96
3.5	偏态和峰度	101
3.5.1	偏态	101
3.5.2	峰度	102
3.6	本章小结	104
3.7	练习题	106

第 4 章 时间数列	113
4.1 时间数列概述	114
4.1.1 时间数列的含义与构成要素	114
4.1.2 时间数列的类型	115
4.1.3 时间数列的编制原则	116
4.2 时间数列的水平指标分析	117
4.2.1 发展水平	117
4.2.2 平均发展水平	117
4.2.3 增长量	126
4.2.4 平均增长量	128
4.3 时间数列的速度指标分析	128
4.3.1 发展速度	128
4.3.2 增长速度	130
4.3.3 平均发展速度	131
4.3.4 平均增长速度	132
4.4 时间数列的趋势分析	133
4.4.1 时间数列的构成因素与分解	133
4.4.2 长期趋势分析	135
4.5 时间数列的季节变动分析	142
4.5.1 同期平均法	143
4.5.2 移动平均趋势剔除法	145
4.6 本章小结	147
4.7 练习题	149
第 5 章 统计指数	156
5.1 统计指数概述	157
5.1.1 统计指数的概念	157
5.1.2 统计指数的作用	157
5.1.3 统计指数的分类	158
5.2 统计指数的编制	159
5.2.1 个体指数的编制	159
5.2.2 总指数(综合指数、平均数指数)的编制	160
5.3 指数体系与因素分析	165
5.3.1 指数体系	165

5.3.2	因素分析	165
5.4	几种常用的经济指数	172
5.4.1	工业生产指数	172
5.4.2	居民消费价格指数	172
5.4.3	货币购买力指数	173
5.4.4	股票价格指数	173
5.5	本章小结	174
5.6	练习题	175
第 6 章	抽样分布与参数估计	181
6.1	抽样调查概述	182
6.1.1	抽样调查的相关概念	182
6.1.2	抽样调查的特点	184
6.1.3	抽样调查的使用范围	185
6.1.4	抽样调查的原则	185
6.2	抽样方法	187
6.2.1	简单随机抽样	187
6.2.2	等距抽样	187
6.2.3	分层抽样	187
6.2.4	整群抽样	187
6.3	抽样的一般步骤	188
6.4	抽样估计	189
6.4.1	总体特征	189
6.4.2	样本特征	190
6.4.3	重复抽样与非重复抽样	190
6.4.4	抽样调查的理论依据	191
6.5	样本均值、样本成数的抽样平均误差	193
6.5.1	抽样误差的概念	193
6.5.2	样本均值、样本成数的抽样平均误差计算	194
6.5.3	样本均值、样本成数的抽样平均误差影响因素	198
6.6	总体均值、总体成数的参数估计	198
6.6.1	参数估计的概念	198
6.6.2	点估计	199
6.6.3	区间估计	200
6.7	简单随机抽样下必要样本容量的确定	203
6.7.1	样本容量的影响因素	203



6.7.2 必要样本容量的确定	203
6.8 本章小结	204
6.9 练习题	205
第7章 相关与回归分析	210
7.1 相关分析	211
7.1.1 相关关系的概念	211
7.1.2 相关关系的种类	211
7.1.3 相关关系的测量	212
7.1.4 相关系数	214
7.2 回归分析	216
7.2.1 回归分析概述	217
7.2.2 一元线性回归分析	218
7.2.3 多元线性回归分析	220
7.3 本章小结	220
7.4 练习题	222
第8章 Excel 的统计数据分析应用	228
8.1 Excel 在数据整理与汇总中的应用	231
8.1.1 统计数据分组、编制变量分配数列与绘制统计图	231
8.1.2 集中趋势与离散程度	238
8.2 Excel 在时间数列长期趋势分析中的应用	240
8.2.1 移动平均法	240
8.2.2 最小平方法	242
8.3 Excel 在抽样与区间估计中的应用	244
8.3.1 抽样分析	244
8.3.2 区间估计	244
8.4 Excel 在线性回归分析中的应用	246
8.4.1 一元线性回归分析	246
8.4.2 多元线性回归分析	250
8.5 本章小结	252
8.6 练习题	252
附表 标准正态分布概率表	255
部分练习题参考答案	257
参考文献	262

第1章 绪 论



学习目标

1. 了解统计的含义及统计学的研究对象；
2. 认识统计学的学科分类；
3. 掌握统计学的几个基本概念。

基本概念

统计 总体 总体单位 标志 标志值 变异 变量 指标 指标体系

案例导入

1981年,首届国际《红楼梦》研讨会在美国召开。威斯康星大学讲师陈炳藻独树一帜,宣读了题为《从词汇上的统计论〈红楼梦〉作者的问题》的论文。他从字、词出现频率入手,通过计算机进行统计、处理、分析,对《红楼梦》后40回系高鹗所作这一流行看法提出异议,认为120回均系曹雪芹所作。



章节导言

在诺贝尔经济学获奖者中,2/3以上的研究成果与统计和定量分析有关。因此,著名经济学家萨缪尔森在其经典的教科书《经济学》第12版中特别提到:“在许多与经济学有关的学科中,统计学是特别重要的。”

通过数据来研究规律、发现规律,贯穿了人类社会发展的始终。人类科学发展史上的不少进步都和数据采集分析直接相关,例如现代医学流行病学的开端。伦敦1854年发生了大规模的霍乱,很长时间没有办法控制。一位医师用标点地图的方法研究了当地水井分布和霍乱患者分布之间的关系,发现有一口水井周围的霍乱患病率明显较高,借此找到了霍乱暴发的原因:一口被污染的水井。关闭这口水井之后,霍乱的发病率明显下降。这种方法,充分展示了数据分析的力量。

1.1 统计的含义

数据,已经渗透到当今每一个行业和业务职能领域,成为重要的生产因素。人们对于海量数据的挖掘和运用,预示着新一波生产率增长和消费者盈余浪潮的到来。2012年,大数据首次被麦肯锡咨询公司提出,随后哈佛大学社会学教授加里·金(Gary King)说:“这是一场革命,庞大的数据资源使得各个领域开始了量化进程,无论是学术界、商界还是政府,所有领域都将开始这种进程。”

“满城尽谈大数据”,很多人其实并不理解大数据的真正价值是什么。2017年初,哈佛大学政治学系金教授在上海交通大学举办了一场名为《大数据,重要的不是数据》(Big Data is Not About the Data)的讲座。讲座里,金教授用3个大数据研究案例说明:大数据的真正价值在于数据分析。有数据固然好,如果没有分析,数据的价值就无法体现。

数据足够大了之后,我们突然发现一切社会现象到最后都有统计规律。它不像物理学那样可以准确地去描述因果的关系,其本质就是一个统计规律。统计作为进行数据搜集、整理与分析的重要工具,与人们的生活、工作密不可分,没有任何学科或领域能够真正离开统计。只有掌握统计的基本知识和方法技能,才能及时获取各种统计信息,然后根据获取的信息认识社会经济发展的趋势和规律,并做出正确的决策和调控。

发展到现在,“统计”一词有三种含义:一是指统计工作,即统计是一种数据搜集、加工和提供的过程,是一种使用明确的概念、方法和程序,以有组织、有条理的方式,从一个总体的部分或所有单元中搜集感兴趣的指标信息的调查过程,并包括将这些信息综合编辑成有用的简要形式的所有活动;二是指统计资料,即统计活动的成果,包括在统计调查中所取得的原始资料和经过加工整理汇总的综合统计资料;三是指统计学,即分析统计数据的方法和技术,它是一门阐述如何去搜集、整理、显示、描述、分析数据和由数据得出结论的一系列概念、原理、原则、方法和技术的科学,是一门独立的、实用性很强的通用方法论科学。具体内容如表1-1所示。

表 1-1 统计的含义

类别	内 涵
统计工作	也称统计活动,是指为了研究客观事物,以有组织、有形式的方式,搜集、整理、分析等工作的总称
统计资料	是指在统计工作中所获得的所有资料的总称,这些资料形式可以是数据、文字、报表、分析报告等
统计学	是指通过科学地概括、总结统计实际经验,形成一门系统地阐述统计理论与方法的学科

统计工作、统计资料和统计学三者相互联系,是理论与实践的辩证统一关系。统计资料的质量好坏、数量的多少取决于统计工作,统计资料是统计工作的成果,统计工作是搜集、整理、分析统计资料的活动过程。统计工作的发展和深化需要统计理论的指导。统计学来源于统计工作,是统计工作经验的理论概括,同时反过来又指导统计工作,推动统计工作的不断发展和提高。三者之间的关系如图 1-1 所示。

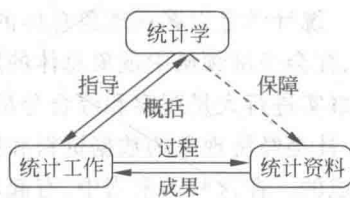


图 1-1 统计工作、统计资料、统计学的关系

总体而言,统计是人们正确运用理论和方法搜集数据、整理数据、分析数据和由数据得出结论的实际操作过程,是人们从数据层面对客观世界的一种认识活动和结果。



统计工作、统计资料、统计学的定义是什么?三者的关系是什么?

1.2 统计学的研究对象

在不同的场合,“统计”的含义或语义不同。但不论何时何地,只要提到统计,人们总是要把它与总计、差异比较、大量的数据分析联系起来。这说明数量性、总体性及相关性是统计学的基本特性。世界上各类现象的发展变化规律,都表现为质与量的辩证统一。要认识其客观存在的规律性,就必须认识其质与量的辩证关系,认识其数量关系的特征及度的界限。因此,统计的研究对象是指统计研究所要认识的客体。一般来说,统计的研究对象是客观现象总体的数量特征和数量关系,以及通过这些数量方面反映出来的客观现象发展变化的规律性。它往往具有如下特点:

1. 数量性

数量性是统计学研究对象的基本特点。因为,数字是统计的语言,数据资料是统计的原料。一切客观事物都有质和量两个方面,事物的质与量总是密切联系、共同规定着事物的性质。没有无量的质,也没有无质的量。一定的质规定着一定的量,一定的量也表现为一定的质。但在认识的角度上,质和量是可以区分的:可以在一定的质的情况下,单独地研究数量方面,通过认识事物的量进而认识事物的质。因此,事物的数量是认识客观现象的重要方面,通过分析研究统计数据资料,研究和掌握统计规律性,就可以达到统计分析研究的目的。例如,要分析和研究国民生产总值,就要对其数量、构成及数量变化趋势等进行认识,进而正确地分析和研究国民生产总值的规律性。

2. 总体性

统计学是以客观现象总体的数量作为自己的研究对象。统计学的研究对象是自然、社会经济领域中现象总体的数量方面,即统计学的数量研究是对总体普遍存在着的事实进行大量观察和综合分析,得出反映现象总体的数量特征和规律性。自然科学、社会经济现象的数据资料和数量对比关系等一般是在一系列复杂因素的影响下形成的。在这些因素当中,有起着决定和普遍作用的主要因素,也有起着偶然和局部作用的次要因素。由于种种原因,在不同的总体单位中,它们相结合的方式和实际发生的作用都不可能完全相同。所以,对于每个总体单位来说,就具有一定的随机性质。而对于有足够多总体单位的总体来说又具有相对稳定的共同趋势,显示出一定的规律性。统计学研究对象的总体性,是从对总体单位的实际表现的研究过渡到对总体的数量表现的研究。在研究总体的统计数据资料过程中,为了更好地分析研究现象总体的统计规律性,不排除对个别事物的深入调查研究。

3. 具体性

统计学研究的是一定时间、地点、条件下具体事物的量,不是抽象的量。这是统计学和数学的一个重要区别。但是,由于统计学是从量的方面研究总体现象,因此,在许多方面要使用数学方法进行统计分析。统计学的这一特点,要求统计工作者必须按照具体事物的本来面目进行调查,保证调查资料对具体事物的准确反映。即在研究客观事物的数量方面,统计学研究的是具体对象的具体数量,而不是抽象的量。而数学则是研究抽象的数量关系和空间几何形式,是舍弃了具体对象质的规定性,是抽象的量。因此,尽管统计学有许多数学公式和应用各种数学方法,但不等于数学,必须在质与量的紧密结合中研究现象量的方面。

4. 变异性

统计学研究的是同类现象总体的数量特征,其前提是总体各单位的特征表现存在着差异,否则就没有进行统计研究的必要。现实中的差异性,有的是由固定原因引起的,有的是由多种原因引起的。由固定原因引起的差异性可以按照已知条件事先推定,不需要运用统计方法进行研究。由多种原因引起的差异性无法事先确定,需要运用统计方法进行研究,进而探索出现象总体的特征和规律。

5. 社会性

统计学研究社会经济现象。这一点与自然技术统计学有所区别。自然技术统计学研究自然技术现象。自然现象的变化与发展有其固有的规律,在其变化进程中,通常表现为随机现象,即可能出现或可能不出现的现象。而统计学的研究对象是人类社会活动的过程和结果。人类的社会活动都是人们有意识、有目的的活动,各种活动都贯穿着人与人之间的关系,除了随机现象外,还存在着确定性的现象,即必然要出现的现象。所以,统计学在研究社会经济现象时,还必须注意正确处理好这些涉及人

与人之间关系的社会矛盾。



统计的含义是什么？其研究对象有什么特点？

1.3 统计学的学科分类

统计学作为一门研究客观事物数量特征和数量关系的方法论科学，其内容构成错综复杂，既有层次性，又有交叉性。截至目前，统计学大致有以下两种分类（如图 1-2 所示）。



图 1-2 统计的学科分类

1.3.1 描述统计学和推断统计学

描述统计学指以总体全面资料或非随机性局部资料为基础的统计理论与方法体系，包括统计总体论（有关总体、指标和分组等理论）、统计设计、统计调查、统计整理、统计指数、动态分析理论、统计平衡理论、统计数据库等。它不同于仅研究如何整理和概括大量数据的“描述统计学”。

例如，想知道对监狱情景进行哪些改造，可以降低囚徒的暴力倾向。这就需要将不同的囚室颜色基调、囚室绿化程度、囚室人口密度、放风时间、探视时间进行排列组合，而后让每个囚室进行实验处理，最后用因素分析法找出与囚徒暴力倾向相关系数最高的因素。

推断统计学是研究如何利用样本数据来推断总体特征的统计方法。它是在对搜集的样本数据进行描述的基础上，对统计总体的未知数量特征作出以概率形式表述的推断。其中，总体是指所要研究事物的全体，样本是指从总体中抽取的一部分单位的集合。

例如，要了解消费者对某种食品的口味偏好，不可能对每个消费者一一进行测试，而只能选取一部分消费者，通过这部分消费者的情况对全体消费者的情况进行推断。这就用到了推断统计学。

描述统计学和推断统计学的划分，一方面反映了统计方法发展的前后两个阶段，同时也反映了应用统计方法探索客观事物数量规律性的不同过程。

统计研究过程的起点是统计数据，终点是探索出客观现象内在的数量规律性。

在这一过程中,如果搜集到的是总体数据(如普查数据),则经过描述统计之后就可以达到认识总体数量规律性的目的了;如果所获得的只是研究总体的一部分数据(样本数据),要找到总体的数量规律性,则必须应用概率论的理论并根据样本信息对总体进行科学的推断。

显然,描述统计和推断统计是统计方法的两个组成部分。描述统计是整个统计学的基础,推断统计则是现代统计学的主要内容。由于在对现实问题的研究中,所获得的数据主要是样本数据,因此,推断统计在现代统计学中的地位 and 作用越来越重要,已成为统计学的核心内容。当然,这并不等于说描述统计不重要。如果没有描述统计搜集可靠的统计数据并提供有效的样本信息,即使再科学的统计推断方法也难以得出切合实际的结论。从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学发展成熟的重要标志。

1.3.2 理论统计学和应用统计学

理论统计学侧重于统计方法的数理基础,包括概率论、经典统计理论、贝叶斯理论、统计判决理论等。应用数理统计学(现代意义上的数理统计学)则侧重于统计方法的应用形式,包括抽样技术、试验设计、相关分析、方差分析、多重应答分析、多元统计分析、序贯分析、线性统计模型、时间数列分析、非参数统计等。

应用统计学只涉及某一特定现象、领域的统计研究,又可以分为核算统计学和实验统计学。核算统计学是通过核算手段研究社会现象及其过程的数量特征或统计规律性的理论与方法体系,包括经济统计学、社会统计学、科技统计学、环境统计学等。而实验统计学是运用实验手段研究自然现象自身及其过程的数量特征或统计规律性的理论与方法体系,包括统计物理学、生物统计学、天文统计学、气象统计学、心理统计学、农业试验统计学、工程技术统计学等。

从统计学的学科分类可以看出,统计学的内容是十分丰富的,其研究和应用的领域非常广泛。



什么是描述统计?什么是推断统计?两者的关系是什么?

1.4 统计学的几个基本概念

统计学的研究对象表现出总体性,而总体是由总体单位构成。通过对各总体单位的标志表现进行综合,形成反映总体本质特征和规律的统计指标。因此,需要掌握几个基本概念:总体与总体单位、标志与标志值、变异与变量、指标与指标体系。

1.4.1 总体与总体单位

1. 总体与总体单位的概念

统计学中,研究对象的全体称为**总体**(population),组成总体的每一个个体称为**总体单位**(population unit)。总体是所要认识的研究对象的全体,是具有某种共同性质或特征的许多总体单位的集合体。总体所包含的总体单位数目,称为**总体容量**(population size),通常用 N 来表示。例如,2017 年重庆市的生产总值,其总体由重庆市所有区县的生产总值构成。

2. 总体的分类

根据总体包含的单位数是否有限,总体可以分为**有限总体**和**无限总体**。

总体所包含的单位数如果是有限的,称为有限总体,如人口数、企业数、商店数等。

总体所包含的单位数如果是无限的或在一定条件下是无法确定的,则称为无限总体,如连续生产的某种产品的数量、大海里的鱼等。

3. 总体的特征

在确定总体单位时,需要首先确定总体。那么判断总体的依据是什么呢?作为一个总体,必须同时具备以下几个特征:

(1) 同质性

同质性是指构成总体的各个单位必须具有至少一个共同的特征和性质。例如,重庆市工业企业总体在性质上与每个工业企业的经济职能是相同的,都是从事工业活动的基本单位。同质性是总体的根本特征,只有总体单位是同质的,统计才能通过对总体单位特征的观察研究,归纳和揭示出总体的综合特征和规律性。

(2) 大量性

大量性是指总体中必须包含足够多的总体单位。仅个别或极少数总体单位不能构成总体。这是因为统计研究的目的是为了描述事物总体的规律性。由于个别总体单位都受偶然因素的影响,其特征不能说明总体的特征,只有相当多的个别总体单位结合起来形成总体,才能使偶然因素的作用相互抵消,从而显示事物的共同性质即总体的特征。

(3) 差异性

差异性或称变异性,是指构成总体的各单位只是在某一性质上相同,但在其他性质或特征上具有一定的差异。例如,重庆市全部工业企业虽然具有相同的经济职能,但是每个工业企业从事的活动内容不同、企业法人不同、员工人数不同。这就是总体内在的差异性。

同质性是构成统计总体的基础,大量性是统计研究的前提,差异性则是统计研究

