

全国高等院校财经管理类规划教材

统 计 学

——以 Excel 为分析工具
(第 2 版)

宋廷山 葛金田 王光玲 主编



北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书是一部以 Excel 为分析工具的实用性很强的统计学教材。按照数据的搜集、整理、综合、推断分析的顺序,介绍了统计数据的收集、整理与显示、概括性度量,抽样及参数估计,统计假设检验,方差分析,时间序列分析,指数与因素分析,相关与回归分析等基本内容。本书侧重于统计思想的介绍,避开了深奥的数学证明,对于复杂的统计计算通过 Excel 软件来实现,着重培养学生的统计意识和统计思想,使统计学更加实用。第二版在第一版的基础上增加了新的案例,并对其使用方式做了详细介绍。

本书既适宜于作为高等院校经济管理类专业统计学课程的教材,也可供从事经济统计工作以及学习使用 Excel 的读者参阅。

图书在版编目(CIP)数据

统计学:以 Excel 为分析工具/宋廷山,葛金田,王光玲主编. —2 版. —北京:北京大学出版社,2012. 6

(全国高等院校财经管理类规划教材)

ISBN 978-7-301-20676-8

I. ①统… II. ①宋…②葛…③王… III. ①统计学—高等学校—教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2012)第 102053 号

书 名: 统计学——以 Excel 为分析工具(第 2 版)

著作责任者: 宋廷山 葛金田 王光玲 主编

责任编辑: 邱 懿

标准书号: ISBN 978-7-301-20676-8/F·3189

出版发行: 北京大学出版社

地 址: 北京市海淀区成府路 205 号 100871

网 址: <http://www.pup.cn>

电子邮箱: zyjy@pup.cn

电 话: 邮购部 62752015 发行部 62750672 编辑部 62754934 出版部 62754962

印 刷 者: 河北滦县鑫华书刊印刷厂

经 销 者: 新华书店

787 毫米×1092 毫米 16 开本 21.75 印张 529 千字

2009 年 1 月第 1 版

2012 年 6 月第 2 版 2012 年 6 月第 1 次印刷

定 价: 45.00 元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话: (010)62752024 电子邮箱: fd@pup.pku.edu.cn

计(第2、3、4、9、10、11章);为满足本科、研究生等较高层次读者的需要,适当增加了推断统计学的内容(第5、6、7、8、12章),尤其是增设了难度较大、应用水平较高、实用性较强的多元统计内容,兼顾了普及和提高两个方面。

本书由齐鲁师范学院宋廷山教授、济南大学葛金田教授担任主编;青岛农业大学王坚副教授、山东科技大学姜爱萍副教授、济南大学解志刚副教授担任副主编;青岛农业大学王宝海教授主审。编写提纲由宋廷山教授提出初稿,经主编、副主编讨论形成。参加编写的有:宋廷山(第1章、第12章);山东省贸易职工大学袁洪英教授(第2章);王宝海(第3章);王坚(第4章);济南大学王光玲教授(第5章);山东科技大学马媛讲师(第6章);解志刚(第7章);山东经济学院田金方副教授(第8章);姜爱萍(第9章);济宁职业技术学院温洪芝教授(第10章);葛金田(第11章)。初稿形成后,由主编进行了修改总纂,由王宝海教授进行了审校。对于本教材可能存在的章节安排等结构性的问题,由主编、副主编负责;对各章具体内容可能存在的问题,由编写者本人负责。

本书的编写者大多有二十年左右在高等学校从事本学科教学研究工作的经验,在统计学的某些方面都有较深刻的研究,都有自己的见解。本书作为教科书,所列内容均为目前统计学界大多认可的观点,并力求使本书的观点前后保持一致,但并不影响本书作者在其他场合各自阐述自己的学术观点。

本书参考了国内外出版的大量本学科教材和专著,敬列于参考文献中。对编写这些著作的学界前辈、专家和同行们,我们表示崇高的敬意和衷心的感谢!

由于水平所限,教材中可能还存在许多我们还没有发现的问题,衷心希望使用本教材的老师、同学和其他读者批评指正,有问题或建议可发电子邮件至 sdeusts@163.com,对于特殊性的问题我们将给予个别答复,对于具有普遍性的问题,将在再版时进行更正和说明。在此我们一并表示感谢!

主 编

2008年8月24日于泉城济南

目 录

第一章 统计学总论	(1)
第一节 统计与统计学	(2)
第二节 统计研究	(6)
第三节 统计学的基本概念	(8)
第四节 统计设计	(13)
第五节 统计应用软件简介	(18)
第二章 数据的搜集	(24)
第一节 统计数据的来源	(25)
第二节 调查方案设计	(28)
第三节 调查问卷设计	(30)
第四节 调查方案设计案例——某校风调查方案及问卷设计	(36)
第三章 数据的整理	(43)
第一节 统计数据的预处理	(45)
第二节 统计分组与频数分布	(52)
第三节 统计图	(62)
第四节 统计表	(67)
第四章 数据的概括性度量	(72)
第一节 总规模度量	(73)
第二节 比较度量	(76)
第三节 集中趋势的度量	(82)
第四节 离散程度的度量	(93)
第五节 偏态与峰度的度量	(102)
第六节 描述统计工具的使用	(103)
第五章 时序分析	(113)
第一节 时间序列分析的基本问题	(115)
第二节 时间序列的水平分析	(117)
第三节 时间序列的速度分析	(121)
第四节 时间序列的趋势分析	(127)
第五节 时间序列的季节变动分析	(142)
第六章 指数分析	(151)
第一节 统计指数概述	(153)
第二节 综合指数和平均指数	(155)

第三节	指数体系与因素分析	(160)
第四节	常用的经济指数	(167)
第七章	抽样调查及参数估计	(176)
第一节	抽样推断	(177)
第二节	抽样分布	(184)
第三节	参数估计	(197)
第八章	假设检验与方差分析	(219)
第一节	假设检验的基本问题	(219)
第二节	一个总体参数的检验	(224)
第三节	两个总体参数的检验	(229)
第四节	方差分析的基本问题	(239)
第五节	单因素方差分析	(242)
第六节	双因素方差分析	(247)
第九章	相关与回归分析	(262)
第一节	相关关系	(262)
第二节	相关分析	(264)
第三节	回归分析概述	(271)
第四节	一元线性回归分析	(272)
第五节	多元线性回归分析	(278)
第六节	利用回归方程进行预测	(284)
第七节	非线性回归和特殊变量的回归模型	(286)
第八节	运用 Excel 进行回归分析	(293)
第十章	多元统计分析	(306)
第一节	聚类分析	(306)
第二节	主成分分析	(314)
第三节	因子分析	(319)
第四节	判别分析	(325)
第五节	关联分析	(328)
主要参考文献	(338)

统计学总论

【引导案例】

新发展 新跨越 新篇章

我国“十一五”时期的五年,是积极应对来自国内外的各种风险和挑战,经济保持平稳较快增长、综合国力大幅提升的五年。

经济平稳较快增长。“十一五”前期,我国经济快速增长,2006年增长12.7%,2007年加速到14.2%,增速仅次于改革开放后最高的1984年。2008年,受百年不遇的国际金融危机的巨大冲击和影响,经济增速陡然回落到9.6%。面对严峻的国内外形势,党中央、国务院果断决策,迅速出台并不断丰富完善应对国际金融危机的一揽子计划,我国经济增速在世界各国中实现率先回升,2009年经济增长9.2%,与世界经济下降0.6%形成鲜明对照。2010年经济增长进一步回升到10.3%,明显快于世界主要国家的平均增速。2006—2010年,我国国内生产总值年均实际增长11.2%,不仅远高于同期世界经济年均增速,而且比“十五”时期年平均增速快1.4个百分点,是改革开放以来最快的时期之一。

经济总量不断迈上新台阶。2010年,我国国内生产总值达到397 983亿元,扣除价格因素,比2005年增长69.9%。经济总量居世界位次稳步提升。2008年,我国国内生产总值超过德国,位居世界第三位。2010年,我国国内生产总值按平均汇率折算达到58 791亿美元,超过日本,成为仅次于美国的世界第二大经济体。我国经济增长对世界经济的贡献不断提高。特别是2008年第三季度金融危机爆发后,在世界主要经济体均面临负增长或停滞困境时,中国经济依然保持了相当高的增速并率先回升,为世界经济复苏做出了重大贡献。

人均国内生产总值快速增加。在经济总量稳步增长的同时,人均创造价值水平也在不断提高。初步统计,2010年我国人均国内生产总值达到29 748元,扣除价格因素,比2005年增长65.7%,年均实际增长10.6%,比“十五”时期年平均增速快1.5个百分点。

国家财政实力明显增强。经济快速增长带来了国家财政收入的稳定增长。我国财政收入2007年超过5万亿,达到51 322亿元;2008年超过6万亿,达到61 330亿元;2010年超过8万亿,达到83 080亿元,比2005年增长1.6倍,年均增长21.3%。“十一五”时期,我国财政收入的快速增长,为加大教育、医疗、社保等民生领域投入,增强政府调节收入分配能力等提供了有力的资金保障。

国家外汇储备大幅增加。我国外汇储备 2006 年突破 1 万亿美元,达到 10 663 亿美元; 2009 年突破 2 万亿美元,达到 23 992 亿美元。2010 年年末,我国外汇储备已达到 28 473 亿美元,比 2005 年增长 2.5 倍,年均增长 28.3%。我国外汇储备规模自 2006 年超过日本,连续五年稳居世界第一位。

(摘自:国家统计局“十一五”经济社会发展成就系列报告之一:新发展 新跨越 新篇章
http://www.stats.gov.cn/tjfx/ztfx/sywcj/t20110301_402706119.htm)

第一节 统计与统计学

一、统计的含义

“统计”一词在各种实践活动和科学研究领域中经常出现。然而,不同的人或在不同的场合,对其理解是有差异的。比较公认的看法是,统计有三种含义,即统计活动、统计数据和统计学。

1. 统计活动

统计活动又称统计工作,是指收集、整理和分析统计数据,并探索数据的内在数量规律性的活动过程。

2. 统计数据

统计数据或称统计资料,即统计活动过程所获得的各种数字资料和其他资料的总称。表现为各种反映社会经济现象数量特征的原始记录、统计台账、统计表、统计图、统计分析报告、政府统计公报、统计年鉴等各种数字和文字资料。

3. 统计学

统计学是指阐述统计工作基本理论和基本方法的科学,是对统计工作实践的理论概括和经验总结。它以现象总体的数量方面为研究对象,阐明统计设计、统计调查、统计整理和统计分析的理论与方法,是一门方法论科学。

《不列颠百科全书》的定义:统计学是收集、分析、表述和解释数据的科学。

《中国百科全书·数学卷》的定义:统计学是一门科学,它研究怎样以有效的方式收集、整理、分析带随机性的数据,并在此基础上对所研究的问题作出统计性推断,直至对可作出的决策提供依据或建议。

统计学的英文是“statistics”。以单数形式出现时,表示统计数据或统计资料;以复数形式出现时,表示一门科学即统计学。

统计工作、统计资料和统计学之间有着密切联系。统计工作同统计资料之间是过程同成果之间的关系,统计资料是统计工作的直接成果。就统计工作和统计学的关系来说,统计工作属于实践的范畴,统计学属于理论的范畴,统计学是统计工作实践的理论概括和科学总结,它来源于统计实践,又高于统计实践,反过来又指导统计实践,统计工作的现代化同统计科学研究的支持是分不开的。

统计工作、统计资料和统计学相互依存、相互联系,共同构成了一个整体,这就是通常所说的统计。

推断统计学;从方法研究的重点来看,统计学可分为理论统计学和应用统计学。

(一) 描述统计学和推断统计学

描述统计学研究如何取得反映客观现象的数据,并通过图表形式对所搜集的数据进行加工处理和显示,进而通过综合、概括与分析得出反映客观现象的规律性数量特征。描述统计学的内容包括统计数据的搜集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与分析方法等。

推断统计学研究如何根据样本数据去推断总体数量特征,它是在对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。

描述统计学与推断统计学的划分,还反映了统计方法发展的前后两个阶段和使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),那么运用描述统计就可以达到认识总体数量规律性的目的;如果获得的只是研究总体的一部分数据(样本数据),那么要找到总体的数量规律性,就要运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。推断统计在现代统计学中的地位越来越重要,已成为统计学的核心内容,这是因为在对现实问题的研究中,所获得的数据主要是样本数据。但这并不等于说描述统计不重要,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,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,是统计学发展的巨大成就,也是统计学成熟的重要标志。

(二) 理论统计学和应用统计学

理论统计学即数理统计学,主要探讨统计学的数学原理和统计公式的来源。由于现代统计学用到了很多方面的数学知识,从事统计理论和方法研究的人员需要有坚实的数学基础。而且,概率论是统计推断的数学和理论基础,所以广义的统计学亦应包括概率论在内。理论统计学是统计方法的理论基础,没有理论统计学的发展,统计学也不可能发展成为今天这样一个完善的科学知识体系。理论统计学的主要内容有概率理论、抽样理论、实验设计、估计理论、假设检验理论、决策理论、非参数统计、序列分析、随机过程等。

应用统计学探讨如何运用统计方法去解决实际问题。其实,将理论统计学的原理应用于各个学科领域,就形成了各种各样的应用统计学。例如,统计方法在生物学中的应用形成了生物统计学,在医学中的应用形成了医疗卫生统计学,在农业试验、育种等方面的应用形成了农业统计学。统计方法在社会科学领域的应用也形成了若干分支学科,在经济领域的应用形成了经济统计学及其若干分支,在管理领域的应用形成了管理统计学,在社会学研究和社管理中的应用形成了社会统计学,在人口学中的应用形成了人口统计学,等等。应用统计学除了包括各领域通用的方法,如参数估计、假设检验、方差分析等之外,还包括某领域所特有的方法,如经济统计学中的指数法、现代管理决策法等。应用统计学着重阐明这些方法的统计思想和具体应用,而不是统计方法数学原理的推导和证明。

四、统计的职能与任务

统计是适应国家管理的客观需要而逐步产生和发展起来的。现代化国家管理系统包括决策系统、执行系统、信息系统、咨询系统、监督系统五个组成部分。国家统计兼有信息、咨询、监督三种系统的职能。

(2) 试验设计法,是通过设计实验取得所要研究的数据的方法。试验设计要遵循的原则:重复性原则,在相同条件下重复多次试验;随机化原则,在实验中对实验对象的分配和试验次序是随机安排的;区组化原则,组内差异大,组间差异小。

(3) 统计描述法,是用综合指标、统计表、统计图等形式描述研究总体现象的数量特征的方法。包括统计分组法和综合指标法。

(4) 统计推断法,指在一定的置信标准要求下,由样本信息推断总体数量特征的归纳推理方法。包括参数估计、假设检验、方差分析等方法。

(5) 统计模型法,是根据统计资料,运用统计方法,对研究现象的结构或过程建立一种统计表达式,进行有关分析的方法。统计模型一般包括四个基本因素:变量、关系式、模型参数、随机项。

三、统计研究过程

和人类其他认识活动一样,统计研究也要经过由现象到本质、由矛盾的特殊性到矛盾的普遍性、由感性认识到理性认识的不断深化的过程。从具体统计认识活动来看,统计工作由统计设计、统计调查、统计整理、统计分析和统计资料的积累开发与应用等环节组成。

(一) 统计设计

统计设计是指根据统计研究对象的性质和研究目的,对统计工作的各个方面和各个环节所作的全面部署和安排。统计设计的最终结果表现为各种标准、规定、制度、方案和办法,如统计分类标准、目录、统计指标体系、统计报表制度、统计调查方案、普查办法、统计整理或汇总方案等。统计工作是高度集中统一、科学性很强的工作,无论是统计总体范围、统计指标的口径和计算方法,还是统计分类和分组的标准,都必须统一,不能各行其是。只有科学地进行统计设计,才能使整个统计工作有秩序地、协调地进行,从根本上保证统计工作和统计资料的质量。因此,统计设计是统计工作的先导。

统计设计的主要内容有:统计指标和指标体系的设计、统计分类和统计分组的设计、统计表的设计、统计资料搜集方法的设计、统计工作各个部门和各个阶段的协调与联系、统计力量的组织与安排等。

(二) 统计调查

统计调查是根据统计方案的要求,采用各种调查组织形式和调查方法,有组织、有计划地对所研究总体的各个单位进行观察、登记,准确、及时、系统、完整地搜集统计原始资料的过程。

统计调查是统计认识活动由定性认识过渡到定量认识的阶段,这个阶段所搜集的资料是否客观、周密、系统、及时,直接影响到统计整理的好坏,关系到统计分析结论的正确性,决定着整个统计工作的质量。所以,统计调查是整个统计工作的基础。

(三) 统计整理

统计整理是根据统计研究的目的和任务,对统计调查阶段所取得的原始资料进行审核、分组和汇总,将分散的、零星的反映总体单位特征的资料转化为反映各组 and 总体数量特征的综合资料的过程。

统计整理是将对总体单位特征的认识过渡到对总体数量特征的认识的桥梁和纽带,它既是统计调查的继续,又是统计分析的必要前提,在统计工作中,处于中间环节,在沟通统计调查和统计分析中起着承上启下的作用。

（四）统计分析

统计分析是指在统计调查和统计整理的基础上,用科学的分析方法,对所研究的现象总体进行全面、系统的数量分析,认识和揭示事物的本质和规律性,进而向有关单位和部门提出咨询建议以及进行必要的分析、预测的统计工作过程。统计分析是统计工作的最后阶段,也是统计发挥信息、咨询和监督职能的关键阶段。

从认识论的角度来说,统计设计属于对社会经济现象进行的定性认识;统计调查和统计整理,是实现对事物个体特征认识过渡到对总体数量特征认识的关键环节,属于定量认识的范畴;统计分析则是运用统计方法对资料进行比较、判断、推理和评价,揭示社会经济现象的本质和规律性的重要阶段。统计设计、统计调查、统计整理和统计分析的有机统一,体现了统计要在质与量的辩证统一中研究社会经济现象总体数量特征的原则要求。

第三节 统计学的基本概念

在论述统计学的理论与方法的过程中,要运用一些专门的概念,熟悉这些概念是掌握统计学的基础。

一、总体与样本

（一）总体

凡是客观存在、在某一共同性质基础上结合起来的许多个别事物的整体,叫做统计总体(简称总体)。例如,要研究某地区非公有制工业企业的生产经营情况,那么该地区全部非公有制工业企业就构成了一个总体,统计设计、统计调查、统计整理和统计分析都要围绕这一对象来进行;再如,要研究我国的人口状况,则全国人口就构成了一个统计总体,从设计普查方案、普查登记、资料汇总到最后公布普查数据等,也都要围绕这一对象来进行。

需要注意的是,在统计研究过程中,统计研究的目的和任务居于支配和主导地位,是考虑问题的出发点。统计总体取决于统计研究的目的和任务,有什么样的研究目的就要求有什么样的统计总体与之相适应;统计研究方法、步骤等也要体现统计研究的目的要求。

统计总体分为有限总体和无限总体两种类型。有限总体是指总体中的总体单位数可以计数或穷尽的总体。例如一个企业的全体职工、一个国家的全部人口等都是有限总体。如果总体中的单位数是一个无穷大量,或准确地度量它的单位数是不经济或没有必要的,这样的总体称为无限总体。例如在连续生产的生产线上产出的全部零件数,一片树林中生长的林木数,江河湖海中生长的鱼的尾数等。

划分有限总体和无限总体对于统计工作的意义在于可以帮助人们进行统计调查方法设计。在统计调查方法体系中,全面调查是对调查对象(总体)的全部单位无一遗漏地进行的调查,而非全面调查则是对调查对象中的一部分单位进行调查,然后再估计总体指标的方法。很显然,对于有限总体,既可以使用全面调查,也可以使用非全面调查,但对于无限总体就只能使用非全面调查。

统计总体具有同质性、大量性和变异性三个特点。

同质性,指构成统计总体的各个单位必须在某些方面而且至少在一个方面具备某种共

同的性质。同质性是构成统计总体的前提。

大量性,指统计总体是由总体的全部单位组成,只有一个单位的统计总体是不存在的。当然,研究目的不同,统计总体就不一样,总体中所包含的总体单位的数量也就不同,一个统计总体究竟包含多少总体单位,最终取决于统计研究的目的。统计总体具有大量性的特点,这是由统计研究对象决定的。如前所述,社会经济统计学的研究对象是大量社会经济现象总体的数量特征,因此,只能而且必须将多个具有某种共同性质的单位组合成一个完整的整体,作为统计研究的具体对象。

从统计研究的角度来说,变异性是指构成统计总体的各个单位之间存在的差别。简言之,变异就是事物之间的差别或不同。例如,工人的性别具体表现为男、女,工人家庭人口数表现为1人、2人、3人、4人、5人,工人的月工资表现为600元、700元、780元、890元、970元、1050元、1130元,等等。

在此,有三个问题需要特别说明:首先,变异是客观的,没有变异的事物是不存在的;其次,变异对于统计非常重要,没有变异就没有统计,如果总体单位之间不存在变异,只要了解一个总体单位的资料就可以推断总体情况了;第三,变异性和同质性之间相互联系、相互补充,是辩证统一的关系,不能用同质性否定变异性或用变异性否定同质性。

(二) 样本

统计研究最终是要确定总体的数量特征,但是有时总体的单位数很多,甚至无限,不可能或无必要对每个总体单位都做调查。这时,就要借助样本来研究总体了。样本就是按照一定的概率从总体中抽取并作为总体代表的一部分总体单位的集合体。也有学者称总体为母体,样本为子样。

样本是统计学中非常重要的概念,对这一概念的理解要注意三方面问题:其一,构成某一样本的每一单位都必须取自某一特定的统计总体,不允许该总体之外的单位介入该总体的样本。其二,样本单位的抽取应是按一定的概率进行的,而具体样本的产生应是随机的,因此必须排除人的主观因素对样本单位抽取和样本生成的干扰。其三,样本是母体的代表,带有母体的信息,因而能够推断母体;然而,样本只是母体的一个子集,且具有随机性,故由样本去推断总体会产生代表性误差。其实,如何从母体中抽取子样,怎样控制样本对总体的代表性误差,是推断统计学研究的主要问题。

(三) 总体单位与标志

1. 总体单位

构成统计总体的个别单位称为统计总体单位,简称总体单位。如全部工业企业中的每个企业、全国人口中的每个人也是总体单位。

总体和总体单位的关系是整体同个体、集合同元素的关系,两者相互依存、相互联系,不存在没有总体的总体单位,也不存在没有总体单位的统计总体。

总体和总体单位的具体形式随着统计研究目的的不同而不同,可以是人,也可以是物,还可以是组织(企业或家庭)或时间、空间、行为等。

总体和总体单位的关系不是一成不变的,随着研究目的的变动,两者可以相互转化。在一定研究目的下,一个事物可以作为总体而存在,然而当研究目的发生变化后,这个事物可能就成为总体单位了。例如,研究我国电子工业的发展情况,那么电子工业行业的所有企业就是一个统计总体,每个企业就是一个总体单位;而要研究一个企业的生产经营情况,那么一个企业就构成了统计总体;如果要研究整个国民经济的发展情况,国民经济所有行业组成

统计总体，而其中的电子工业行业又变成了总体单位。

2. 标志

统计是从对个体的观察开始，逐步过渡到对总体数量特征的认识的。标志是指说明总体单位特征的名称，例如，一个企业作为总体单位，这个企业的“所有制类型”、“生产能力”、“年产量”、“销售收入”、“职工人数”、“工资总额”等都是标志。

标志分为品质标志和数量标志两种类型。品质标志是说明总体单位属性特征的名称，如工人的“性别”、“民族”、“工种”等，只能用文字而不能用数值表示。数量标志是说明总体单位数量特征的名称，如工人的“工资额”、“工龄”、“年龄”等，可以用数值表示。

统计中反映总体单位特征的标志很多，如果按总体单位在标志上的具体表现是否存在差异来看，可分为不变标志和可变标志。当各个总体单位在某一标志上的具体表现都相同时，则为不变标志。不变标志体现总体的同质性。组成一个总体的各个总体单位必须有一个或几个不变标志，不变标志是使许多个别单位组合成为总体的前提。例如，以全国国有大中型机械工业企业为总体，这里的各工业企业均有所有制、企业规模、工业部门这三个不变标志。可变标志是指具体表现在总体各个单位上不相同或不完全相同的那些标志。例如，把全部国有大型机械工业企业作为一个统计总体，那么厂址、隶属关系、职工人数、资金额、生产能力、工业增加值、工业总产值、劳动生产率、平均工资、利税额等就是可变标志。和标志相联系的另外一个概念是标志表现，指总体单位特征在某一标志上的具体体现。如一个人的性别是“男”、年龄“50岁”、民族“汉族”；某企业是“股份制公司”、年产值“10亿元”。同标志一样，标志表现也分为品质标志表现和数量标志表现两种类型。

在引入标志表现这一概念之后，变异也可以表述为：统计总体单位之间在某一标志上具有不同标志表现的现象。

二、指标与变量

（一）统计指标

1. 指标的概念

指标是说明现象总体数量特征的概念或范畴。它有三大特点：数量性，所有统计指标都能用数量来表示；综合性，统计指标都是用来说明总体的；质的规定性，统计指标都是具有一定经济内容的。

统计指标，一般情况下是指指标名称或名称加数值。例如，“2010年中国国内生产总值（GDP）或国内生产总值 397 983 亿元”，“全国财政收入达到 83 080 亿元，比上年增长 21.3%”都是统计指标。

2. 指标的构成要素

统计指标的构成因素有：指标的名称、指标的概念、指标的内容、指标的计算方法、指标的计量单位、指标的空间范围、指标的时间范围、指标的数值等八大因素。

3. 标志与指标的区别和联系

（1）两者的主要区别是：标志是说明总体单位特征的，而指标是说明总体特征的；标志中的数量标志可以用数值表示，而品质标志不能用数值表示。所有的统计指标都是用数值表示的，不存在不能用数值表示的统计指标。

（2）两者的联系表现在：有些统计指标的数值是在总体单位的数量标志值基础上直接汇总得到的，如一个县的粮食总产量是所属各乡村粮食产量的合计数；在一定条件下（研究

品、次品等；“考试成绩”也是一种顺序尺度，它可将成绩分为优、良、中、及格、不及格等；“对某一事物的态度”作为一种顺序尺度，可将人们的态度分为非常同意、同意、保持中立、不同意、非常不同意，等等。显然，顺序尺度对事物的计量要比名类尺度精确些，但它至多测度了类别之间的顺序，而未测量出类别之间的准确差值。因此，顺序尺度的计量结果只能比较大，不能进行加、减、乘、除等数学运算。

3. 数值数据(Metric data)

数值型数据有两种衡量尺度：一是定距尺度，这种尺度是能测度事物类别或次序之间间距的数量标志，更具体些说，区间尺度是可将事物区分为不同类别，对这些类别进行排序，并较准确地度量类别之间数量差距的一种计量尺度。该尺度通常使用自然或物理单位作为度量单位，如收入用人民币“元”度量，考试成绩用“百分制”度量，温度用摄氏或华氏的“度”来度量，重量用“克”度量，长度用“米”度量等。区间尺度的计量结果表现为数值。区间尺度的数值可做加、减法运算，例如，考试成绩80分与90分之间相差10分，一个地区的温度20℃与另一个地区的25℃相差5℃，等等。但不能做乘、除法运算。二是比尺度，这种尺度可以准确地计量两个数值之间的倍数。比尺度数据可以做加、减、乘、除法运算。例如，一个人的月工资收入为6000元，另一个人3000元，可以得出一个人的收入是另一个的两倍。

区别这两类数据的显著特征是：用定比尺度测定的数值型数据有一个绝对固定的“零点”，“0”表示的是数值，即“没有”如花费是“0”，则表示没有花费；而定距尺度测定的数据的“0”，表示的是一种水平，比如温度为“0”，不是说没有温度，而是温度在“0”这样的一个水平上。

(二) 按与时间关系分，统计数据分为：时间数据、截面数据和混合数据

1. 时间数据(Time series data)

时间数据是同一空间在不同时间上的数据，它所描述的是现象随时间而变化的情况。如1950—2010年我国历年国内生产总值数据就属于时间序列数据。

2. 截面数据(Cross-sectional data)

截面数据是在同一时间不同空间上的数据，它所描述的是现象在某一时刻或某一时间段不同空间的数量变化情况。如2011年我国国内生产总值数据就属于截面数据。

3. 面板数据(Panel data)

面板数据又叫混合数据，是不同空间在不同时间上的数据。如1950—2010年历年我国国内生产总值数据就属于面板数据。

(三) 按统计数据的表现形式不同，统计数据分为：绝对数、相对数和平均数

1. 绝对数(Absolute number)

其现象的总体规模一般都以绝对数形式表现，一个地区的总人口、国内生产总值、货物周转量等都是绝对数。

2. 相对数(Relative number)

相对数是由两个相互联系的绝对数对比得到，反映事物的相对数量。

3. 平均数(Average)

平均数反映现象总体的一般水平，包括静态平均数和动态平均数。

三、统计指标设计

（一）统计指标的设计原则

设计一个统计指标，要遵循下列原则要求：

1. 指标结构(要素)要完整

其结构应包括指标名称、指标含义、指标内容、时间范围、空间范围、计算方法、计量单位等内容。如2010年中国国内生产总值(GDP)397 983亿元。当然在设计阶段，指标数值是没有的。提供指标时，就应该包括指标数值。

2. 指标的名称要有科学的理论依据，要做到名称和内容要统一。

应让人们看到名称就大体上知道指标所反映的内容。

如居民消费者物价指数(Consumer price index, CPI)反映居民生活中的产品和劳务价格所统计出来的物价变动指标，通常是作为观察通货膨胀水平的重要指标。

3. 定义指标的概念。

用确切的语言界定指标的内涵。

4. 确定指标的外延或指标的内容。

确定哪些现象数值计入哪些不计入。

5. 要有科学的计算方法和统计口径。

例如，工业总产值指标。工业总产值是以货币表现的工业企业在一定时期内生产的已出售或可供出售工业产品总量，它反映一定时间内工业生产的总规模和总水平。它包括：在本企业内不再进行加工，经检验、包装入库（规定了需包装的产品除外）的成品价值，工业性作业价值，自制半成品、在产品期末初差额价值（生产周期较长的企业计算）。工业总产值采用“工厂法”计算，即以工业企业作为一个整体，按企业工业生产活动的最终成果来计算，企业内部不允许重复计算，不能把企业内部各个车间（分厂）生产的成果相加。但在企业之间、行业之间、地区之间存在着重复计算。

（二）统计指标的分类

1. 总量指标、相对指标和平均指标

统计指标按其所反映的内容或其数值表现形式，可以分为总量指标、相对指标和平均指标三种。

（1）总量指标是反映总体规模的统计指标，通常以绝对数形式表现，故又称其为绝对数，如人口总数、国内生产总值等。

（2）相对指标是两个统计指标之比，称为相对数，如性别比例、第三产业所占比重、计划完成程度、经济增长率、股票价格指数等。

（3）平均指标又称为平均数或均值，它所反映的是现象在某一空间或时间上的平均数量状况，如人均消费水平、某种股票一周平均价格、某班级统计学的平均成绩、单价、单产等。

广义上的相对数包括平均数。统计指标从表现形式上又分绝对数和相对数两种。所以统计分析一般要从绝对数和相对数两个方面进行。

2. 数量指标和质量指标

统计指标按其所反映的数量特点或性质不同，分为数量指标和质量指标。

- (2) 目的性,即指标体系的设计要考虑管理的要求或研究目的。
- (3) 全面性,即指标体系的设计要全面,从不同侧面反映事物。
- (4) 统一性,即指标体系的设计要三大核算统一。
- (5) 可比性,即指标体系的设计要不同空间、不同时期可比。
- (6) 核心性,即指标体系的设计要确定核心指标。
- (7) 可行性,即指标体系的设计要保证每一指标都能取得。
- (8) 互斥性,即指标体系的设计要使指标之间相关程度弱,注重指标的代表性。

(三) 统计指标体系的形式

社会经济总体现象的构成往往是复杂多样的,其间的联系形式是多种多样的,因而构成统计指标体系的各项指标之间的联系也有多种不同的形式,归纳起来主要有两种:

一类是内容衔接结构式统计指标体系。是用一系列统计指标所构成的从不同侧面、不同角度、不同环节、不同层次来反映研究对象不同数量特征的指标体系。如反映一个企业从人、财、物、产、供、销等方面设计的指标体系。

另一类是恒等式数量关系指标体系。这是由三个及三个以上的统计指标所形成的具有一定数量对等关系的指标体系。这类指标是建立统计指数体系的基础,也是进行因素分析的基本依据。如原材料消耗额、产品产量、单位产品原材料消耗量、单位原材料价格。

(四) 统计指标体系的种类

1. 描述性指标体系、评价性指标体系、监测性指标体系

按统计指标体系的功能层次来分,有描述性统计指标体系、评价性统计指标体系、监测性统计指标体系三种。其中,描述性统计指标体系是基础,它的任务是尽量全面、详细描述研究现象总体各方面数量特征,因此指标数量多、指标形式不一、指标之间重叠交叉现象普遍。评价性统计指标体系则是在描述性统计指标体系中根据评价任务精选出来的若干指标,或称“指标样本”,其指标数量相对少而精,在不失全面的情况下突出重点,指标形式一般比较统一,指标之间重叠交叉现象一般尽量避免。监测性或决策性指标体系则更加精炼,一般是由少数几个精干的指标构成,甚至只由一些高度综合的指标构成。

2. 社会统计指标体系、经济统计指标体系、科技统计指标体系

按指标体系反映的内容不同可分为社会统计指标体系、经济统计指标体系和科技统计指标体系。

3. 宏观统计指标体系和微观统计指标体系

按指标体系反映现象的范围不同可分为宏观统计指标体系和微观统计指标体系。

4. 基本统计指标体系和专题统计指标体系

按指标体系反映现象的方面侧重点可不同分为基本统计指标体系和专题统计指标体系。基本统计指标体系是由反映社会经济发展基本情况的统计指标构成。专题统计指标体系是针对某一专门问题而设计的指标体系,如经济效益指标体系。

五、常用统计分类标准

(一) 经济成分的划分

经济成分分为:公有经济(包括国有经济和集体经济)和非公有经济(包括私有经济、港澳台经济、外商经济)。

(1) 国有经济单位,指生产资料归国家所有的各种企业、事业单位,以及各级国家机关、人民团体等单位。

【练习题】

一、名词解释

统计；统计总体和总体单位；统计指标和标志；指标体系；时期指标和时点指标；总量指标、相对指标、平均指标；数量指标、质量指标；参数、统计量。

二、简答题

1. 简述统计一词的三个含义及其相互关系。
2. 简述统计数据的基本分类。
3. 简述统计指标的基本分类。
4. 简述统计指标和标志之间的区别与联系。
5. 简述统计研究的工作过程。
6. 简述统计的职能与任务。
7. 简述统计设计及其内容。
8. 简述指标体系的设计原则。
9. 简述三次产业分类。

三、填空题

1. 研究某市居民的生活水平，则该市全部居民户是_____，每一居民是_____。
2. 标志是说明总体单位特征的名称，有_____和_____两种。
3. 研究一个行业的发展情况，总体是_____，总体单位是_____。
4. 一个企业作为总体单位，该企业产值、产量、生产能力、设备的原值等属于_____标志，而所有制、行业性质等属于_____标志。
5. 统计指标反映总体的数量特征，数量标志反映_____的数量特征。
6. 统计指标的构成要素包括_____、_____、_____、计量单位、时间、空间和计算方法等方面。
7. 统计工作和统计资料之间是_____的关系，统计学和统计工作之间是_____的关系。
8. 如果一个变量是以整数形式取值，该变量被称为_____变量，如果它在一个区间内可取无穷多个值，该变量称为_____变量。
9. “统计”一词有三种含义，即_____、_____、_____。
10. 一个完整的统计工作过程包括_____、_____、_____、_____。
11. 所有的_____指标都用绝对数表示，而_____指标都用相对数和平均数表示。
12. 总量指标按其反映的总体内容不同，分为_____和_____；
按时间状况不同可以分为_____和_____；
按计量单位不同分为_____、_____、_____和_____。
13. 某企业年末商品库存量属于_____指标，其数值_____相加；商品销售收入属于_____指标，其数值_____相加。

例如,2010年的11月1日零时为第六次全国人口普查登记的标准时间(调查资料所属时间)、普查登记时间为2010年11月1日开始到11月10日。

(五) 调查预算

一项调查经费是一定的,制订经费预算方案是提高调查质量同时节省费用的必要环节。一项完整的调查可能涉及的费用有:(1)调查方案的策划费与设计费;(2)抽样设计费;(3)问卷设计费;(4)问卷印刷、装订费;(5)调查实施费;(6)数据编码、录入费;(7)数据统计分析费;(8)调查报告撰写费;(9)办公费;(10)其他费用。

(六) 调查实施

为了保证整个统计调查工作顺利进行,在调查方案中还应该有一个周密的组织实施方案。主要内容包括调查机构成立、调查人员挑选和培训、调查资料报送办法、调查前的宣传发动、调查文件的准备以及试点工作等。

第三节 调查问卷设计

一、问卷及问卷设计标准

(一) 问卷格式

问卷(或调查表、统计表)是专门为从被调查者那里获得有关某个主题的信息而设计的一组或一系列问题。

不同的调查问卷在具体结构、题型、措辞、版式等设计上会有所不同,但在结构上一般都由开头部分、甄别部分、主体部分和背景部分组成。

1. 开头部分

开头部分一般包括问候语、填表说明和问卷编号等内容。

(1) 问候语。在自填式问卷中,写好问候语十分重要,它可以引起被调查者的重视,消除顾虑,激发参与意识,以争取他们的积极合作。问候语要语气亲切、诚恳礼貌,文字要简洁准确,并在结尾处对被调查者的参与和合作表示感谢。

(2) 填写说明。在自填式问卷中要有详细的填表说明,让被调查者知道如何填写问卷,如何将问卷返回到调查者手中。这部分内容可以集中放在问卷的前面,也可以分散到各有关问题的前面。

(3) 问卷的编号。问卷的编号主要用于识别问卷以便校对检查、更正错误。

2. 甄别部分

甄别也称为过滤,它是先对被调查者进行过滤,筛选掉不需要的部分,然后针对特定的被调查者进行调查。

通过甄别或过滤,一方面可以筛选掉与调查事项有直接关系的人,以达到避嫌的目的;另一方面,也可以确定哪些人是合格的被调查者,哪些人不是。甄别的目的是确保被调查者合格,能够作为该调查项目的代表,从而符合调查研究的需要。

3. 主体部分

该部分是调查问卷的核心内容,包括了所要调查的全部问题,主要由问题和答案组成。