

# 应用统计学

主 编 卫海英  
副主编 王 军 郭海华  
编 著 卫海英 王 军  
郭海华 尹居良  
王斌会

暨南大学出版社

图书在版编目 ( CIP ) 数据

应用统计学 魏海英主编 广州 : 暨南大学出版社 , 2004  
ISBN 7-307-04811-0

I ①应... II ①魏... III ①应用统计学 IV ①C8

中国版本图书馆 CIP 数据核字 ( 2004 ) 第 100000 号

责任编辑 : 翟云玉

责任校对 : 沈红凌

排版 : 暨南大学出版社照排中心

出版发行 : 暨南大学出版社 ( 广州·石牌 )

印刷 :

开本 : 787 毫米 1/16 印张 : 15.5 字数 : 300 千字

版次 : 2004 年 1 月第 1 版

印次 : 2004 年 1 月第 1 次

印数 : 1-3000 册

定价 : 25.00 元

# 前 言

当前，我国市场竞争日趋激烈，企业须作更为深入的市场研究来进行科学的预测决策，有更多的数据资料需要处理。每一位在校的经管专业大学生、在职学员和从事工商管理及市场研究的工作人员等都希望能掌握更多的应用性较强的统计数据分析方法以适应企业发展的需要，同时也非常渴望能获得一本既有理论阐述又有实际应用指导即理论与实际操作较好结合的工具书。

但是，纵观我国目前的统计学教材，普遍存在过多的理论阐述的不足，缺乏从实用的角度来阐述统计学理论。因此学与用严重脱节是目前我国高校统计教育中存在的普遍现象。而本书则恰好弥补了这一缺陷。

首先，作者们集十余年的教学科研经验，从实用的角度构造了一个在工商管理活动中较常用的统计方法体系，并结合大量经济活动中的实例，系统、全面地阐述了统计学。其次，作者收集了大量资料在阐明理论的同时，进行了生动独特的案例分析，给读者提供了清晰明了的统计方法选择与运用的思路。

书中所有的例题和案例均运用当今世界上最流行的统计软件包来~~分析~~来分析，这在我国目前的统计教材中尚未出现，而这也正是当今统计学学习者所急需和所必须的，从而本书解决了目前普遍存在的统计学教材与统计分析软件操作不配套不衔接的严重问题，并顺应了我国正提倡和开展的多媒体教学的潮流。

本书由卫海英构思并拟出全书大纲。本书共分十三章，各章的撰写者是：

卫海英：第一章、第十一章、第十二章部分；王军：第三章、第六章、第八章；郭海华：第二章、第四章、第十二章部分；尹居良：第五章、第九章；王斌会：第七章、第十章。最后由卫海英、王军、郭海华共同修改总纂定稿。该书与卫海英主编的《~~统计学~~统计学在经济管理中的运用》一书配套使用。

对该书的不足或错误之处，恳请读者和专家批评指正。

本书作者在编写过程中，参考了许多国内外专家学者的论著，受益匪浅，在此深表谢意，同时还非常感谢暨南大学出版社的王华、苏彩桃以及所有关心和帮助我们的人。

编著者

2002年 8月 愿日

# 目 录

前言 .....	( 员 )
第一章 绪论	
员员 统计学的产生和发展 .....	( 员 )
员圆 统计学的分类 .....	( 员 )
员猿 统计学的应用与误用 .....	( 圆 )
员源 统计学的若干基本概念 .....	( 猿 )
员缘 统计软件的应用 .....	( 远 )
第二章 统计数据的收集与整理	
圆员 数据资料的收集 .....	( 苑 )
圆圆 数据资料的整理 .....	( 员园 )
圆猿 分布数列 .....	( 圆园 )
圆源 数据资料的显示 .....	( 圆源 )
案例 圆员: 上海市远程护理服务市场调研方案设计 .....	( 圆缘 )
第三章 单变量描述统计分析	
猿员 集中趋势分析 .....	( 猿圆 )
猿圆 离中趋势分析 .....	( 猿圆 )
猿猿 偏态与峰度 .....	( 猿源 )
案例 猿员 振兴大学 .....	( 猿缘 )
第四章 统计指数	
源员 指数的概念 .....	( 缘圆 )
源圆 指数的种类 .....	( 缘圆 )
源猿 指数的计算 .....	( 缘圆 )
源源 几种现行指数的编制 .....	( 缘缘 )
第五章 概率论的基本概念和方法	
缘员 随机事件与概率 .....	( 苑圆 )
缘圆 随机变量及其分布函数 .....	( 苑圆 )
缘猿 随机变量的数字特征和极限定理 .....	( 怨缘 )
缘源 大数定律和中心极限定理 .....	( 员园员 )

缘缘	数理统计中常用的几种分布 .....	( 缘缘)
第六章 抽样分布与参数估计		
缘缘	抽样分布 .....	( 缘缘)
缘缘	参数估计 .....	( 缘缘)
缘缘	样本容量的确定 .....	( 缘缘)
案例 缘缘	粤悦公司 .....	( 缘缘)
案例 缘缘	爱华公司 .....	( 缘缘)
第七章 假设检验		
缘缘	假设检验的基本原理与步骤 .....	( 缘缘)
缘缘	单个正态总体的假设检验 .....	( 缘缘)
缘缘	两个正态总体的假设检验 .....	( 缘缘)
缘缘	假设检验时应注意的问题 .....	( 缘缘)
案例 缘缘	葡萄糖生产厂家 .....	( 缘缘)
案例 缘缘	选何种材料制造产品为佳 .....	( 缘缘)
第八章 非参数统计		
缘缘	符号检验 .....	( 缘缘)
缘缘	威尔科克森符号等级检验 .....	( 缘缘)
缘缘	曼-惠特尼 哉检验法 .....	( 缘缘)
缘缘	柯尔莫哥洛夫-斯米尔诺夫检验 .....	( 缘缘)
缘缘	$\chi^2$ 检验 .....	( 缘缘)
缘缘	游程检验 .....	( 缘缘)
缘缘	等级相关检验 .....	( 缘缘)
案例 缘缘	晕葬杂缘公司 .....	( 缘缘)
案例 缘缘	广告公司 .....	( 缘缘)
第九章 方差分析		
缘缘	单因素试验的方差分析 .....	( 缘缘)
缘缘	双因素试验的方差分析 .....	( 缘缘)
缘缘	方差秩分析 .....	( 缘缘)
案例 缘缘	温度影响电池的寿命吗 .....	( 缘缘)
第十章 相关与回归		
缘缘	简单线性相关分析 .....	( 缘缘)
缘缘	一元线性回归分析 .....	( 缘缘)
缘缘	多重回归分析 .....	( 缘缘)
案例 缘缘	利税与工业总产值关系分析 .....	( 缘缘)

案例 11.1 财政收入多因素分析.....	( 11.1 )
------------------------	----------

## 第十一章 时间数列

11.1 时间数列的动态分析 .....	( 11.1 )
11.2 趋势变动分析 .....	( 11.2 )
11.3 几种常用的预测方法 .....	( 11.3 )
案例 11.4 根据资料作决策，囤积货物可以创造高价吗.....	( 11.4 )

## 第十二章 产品质量统计

12.1 质量、质量特性与质量职能 .....	( 12.1 )
12.2 排列图 .....	( 12.2 )
12.3 因果图 .....	( 12.3 )
12.4 控制图 .....	( 12.4 )
案例 12.5 电话订单处理过程是稳定的吗.....	( 12.5 )

附录一：习题答案 .....	( 附录 )
----------------	--------

附录二：常用统计表 .....	( 附录 )
-----------------	--------

参考书目 .....	( 附录 )
------------	--------

# 第一章 绪 论

## 1.1 统计学的产生与发展

统计一词具有双重的基本意义：第一，它代表由资料导出的实际数字，例如学生某课程考试的及格率、居民生活费物价指数、观看足球赛人数、公司的销售额和利润等；第二，它是一门学科，并称之为统计学。

通常所说的统计可以视为统计工作、统计资料和统计学。统计工作是人们对客观事物数量方面进行调查研究的认识活动，包括数据资料的收集、整理和分析等；统计资料是统计工作的直接结果，包括调查得到的经过整理具有信息价值的各种统计数据和分析报告；统计学则是一门收集、整理、描述、显示和分析统计数据的科学，是统计工作实践经验的理论概括，其目的是探索数据内在的数量规律性，是数据的科学。

统计学产生于 17 世纪欧洲资本主义生产方式形成和发展的初期，是从几个不同的领域开始的。1759 年英国的威廉·配弟在其代表作《政治算术》一书中用大量的数据资料对英、法、荷三国的经济实力进行比较，以“数字、重量、尺度”为语言表达他的思想和观点，马克思称他为统计学的创始人。英国一位事业有成的商店老板，名叫格朗特，因好奇心的驱使，花了相当长的时间收集钻研每周教堂的公告，其中“死亡表”记载有出生、受洗礼与死亡人数及其死亡原因。1700 年格朗特发表了《关于伦敦死亡表的观察》一文，在文中指出了不少规律性的现象，如男女人数占总人口数的比例大致相等；新生儿中男婴的比例稍高；男婴的死亡率比女婴死亡率略高等等。这篇文章可谓首次根据大量原始资料来对社会与生物现象作实证研究，因而获得后人尊崇，并称之为人口统计学的起源。法国的帕斯卡尔和费马特在数学家们对机会游戏研究的基础上将赌博中出现的各种具体问题，归纳为一般的概率原理，成为古典概率论的创始人。上述几位科学家从不同的角度研究，经过 100 余年的演变与发展，统计学趋于成熟，成为一门研究客观事物数量方面的独立的方法论科学，在研究对象、方法体系和学科体系方面都具有自己的特点。

## 1.2 统计学的分类

从统计方法的角度，统计学分为描述统计学和推断统计学，即反映了统计发展的前后

此为试读, 需要完整PDF请访问: [www.ertongbook.com](http://www.ertongbook.com)

两个阶段，又反映了统计方法研究和探索客观事物内在数量规律性的先后两个过程。描述统计学是整个统计学的基础和统计研究工作的第一步，它包括对客观现象的度量，调查方案的设计，数据的收集与整理，用图表方法和数量方法综合分析统计资料等。我们常见的报纸、杂志、报告或其他出版物上的数据表格、图形或数值都属于描述统计范畴。推断统计学是现代统计学的核心和统计研究工作的关键环节，它是根据概率论揭示随机变量的一般规律，利用样本信息对总体的某些性质或数量特征进行推断和检验的方法。在现实中一般需通过抽样推断方法才能认识事物内在的数量规律性，但若没有可靠的数据资料，再科学的推断方法也难以得出准确的结论，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发展的规律。此外，从其他的角度，统计学还可分为经济统计学和数理统计学、理论统计学和应用统计学等。

## 1.1 统计学的应用与误用

早期，统计应用在描述一地区或国家各种情势资料上，随着经济的发展，统计愈来愈多地应用于商务活动中。在 20 年代，美国两位著名的统计学家休哈特（Walter A. Shewhart）和戴明（W. Edwards Deming）对推动统计学在商务中的应用有卓越的贡献。休哈特先生是纽约贝尔实验室的一位统计专家，他曾运用抽样调查方法发现军人的军衣、军鞋尺寸的分布近似于正态分布，根据该分布规律设计出的军衣和军鞋符合大多数军人的需要，从而解决了急用军需品的规格和尺寸的设计问题。戴明先生则是美国普查局的一位统计专家，为改进最终产品的一致性和品质，两位专家运用统计质量管理技术（Statistical Quality Control）于企业实务中，监控生产过程，这些技术在二次世界大战期间对美国产业界的发展发挥了很大的作用。而战后，美国产业界大量放弃了这些统计技术，部分是因为美国是当时世界上少数几个能自给自足且有余力援助他国的国家，在无竞争压力情况下，导致美国产业忽略了统计技术的改进。日本战后废墟重建，但一开始因其产品缺乏品质管理，声誉不佳，直至 1954 年戴明先生访问日本，并忠告日本企业老板要将统计技术和管理哲学引进日本制造业后，日本在学习美国统计质量管理的基础上，结合本国的国情，进一步发展完善为全面质量管理（Total Quality Management），日本的产品才开始脱胎换骨，生产出高品质产品畅销世界市场，不但逐渐取代品质不佳的美国产品，亦因而赚取了大量利润。相反，美国则因其产品品质较差而市场占有率逐渐降低。目前，美国产业界已感受到统计观念在现今全球性市场的生存竞争中相当重要，统计思维就是竞争的根本，对了解统计学且有能力运用统计技术的人才需求若渴。现在大至国际政治经济形势的分析，国家的决策施政和管理，小至企业的业务经营和班组核算，从宏观到微观，统计已是无处不在，已成为国家和企业管理的有力工具。

统计方法的正确应用，有助于我们认清事物的真相、发现事物变化的数量界限、揭示事物发展的内在规律。相反，统计方法的错误使用，将造成事实的扭曲、读者的误解，甚至决策的失败。有些统计的误用者是无意的，但有些人却是为达到某种目的故意地隐藏不利资料而突出有利部分，所以统计可以变为制谎者的“工具”和政治家的“手段”。统计“谎言”的一种常见现象较多出现在小样本的观察结果上。如仅调查了 10 位牙医对某牌牙膏的看法就作出如下定论：每 10 位牙医中就有 9 位特别推荐该牌牙膏；将某药品给 10 位

病人服下有 怨位见效则宣称该药品有效率达 怨怨等等。再如某汽车公司的广告“近十年来本公司在贵国卖出的车子，有 怨怨还在路上跑”，消费者看到这一广告自会产生这样的印象：该公司的汽车性能一定不错才能开这么久。而实际上上述车子的 怨怨是近三年内卖出的，显然这是汽车制造商巧妙的广告说辞，且误导了消费者。有时数字本身似乎是令人迷惘的，如平均月工资为 愿愿愿元，给人的感觉似乎很精确值得相信，但若平均月工资为 愿元，则似没有上一数字精确而难以令人信服，事实上统计数字是否精确与其所含小数位多寡无关。显然，只有透彻地理解统计理论及方法，才能正确地应用而避免误用。

## 愿原 统计学的若干基本概念

### 愿愿愿 总体与总体单位

总体是在一特定研究中所有感兴趣的个体的集合，是由客观存在的具有某种共同性的许多个别事物组成的整体。构成总体的个体则是总体单位。例如，在商业企业职工情况调查中，所有商业企业职工就是一个总体，每一位商业企业职工就是一个总体单位。确定总体是为了确定调查研究的对象和调查研究的范围，确定总体单位则是为了确定调查登记项目的承担者。总体具有三方面的特征：

愿愿愿 同质性，即构成总体的各个单位必须具有某一方面的共性，这个共性是我们确定总体范围的标准。

愿愿愿 大量性，即总体是由许多单位所组成的，而不是只有个别单位。

愿愿愿 差异性，即总体各单位之间，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。

总体按其包括范围的大小可分为无限总体和有限总体。总体单位很多且不可计数的总体叫无限总体，而总体单位可以计数的总体叫有限总体。对无限总体只能抽取一部分单位进行非全面调查，据以推断总体；对有限总体则可以根据具体情况选择全面调查或非全面调查。在社会经济统计中，大多数总体属于有限总体。

在推断统计中，总体又可分为全及总体和抽样总体。全及总体即为上述的研究整体，抽样总体则是从该整体中抽取的部分单位，也称为样本，样本是总体的一个子集。

总体和总体单位是根据统计研究目的确定的，随着研究目的的改变，总体和总体单位也将随之变化。

### 愿愿愿 标志和指标

标志是总体单位所具有的属性和特征。每个总体单位从不同的角度考察可以有多种属性和特征，例如作为总体单位，每位商业企业职工的 职业、性别、民族、年龄、文化程度、工资水平等属性和特征就是标志。标志可分为品质标志和数量标志，只能用文字表示的标志，如性别、民族等，称为品质标志；可以用数字表示的标志，如年龄、工资水平等，称为数量标志。标志还可分为不变标志和可变标志，在总体各单位之间具体表现完全相同的标志，如职业等，称为不变标志；在总体各单位之间具体表现不完全相同的标志，

如性别、年龄等，称为可变标志。即有不变的品质标志和数量标志，也有可变的品质标志和数量标志，其中可变的数量标志是统计研究的主要标志。

统计指标是综合反映总体数量特征的概念和数值，由指标名称和指标数值组成。例如2005年某商业企业职工总数为1000人、工资总额为1000万元等都是从不同的方面反映总体数量特征的指标。统计指标具有两个重要特点：①具体性。指标是总体某一方面质和量的具体统一，说明总体在具体时间、地点、条件下的数量特征；②综合性。指标是总体单位数量标志汇总整理的结果，说明的不是个别单位或部分单位的数量特征，而是构成总体的所有单位的综合结果，故也称为综合指标。

统计指标按不同的特点可分成不同的类别，若按反映数量特点的不同，可分为数量指标和质量指标。数量指标又称为总量指标，是指反映社会经济现象的总规模、总水平或工作总量的指标，如人口数、总产量、总产值等；质量指标是指反映社会经济现象的相对水平或工作质量的指标，如平均工资、劳动生产率等。

若按指标数值形式的不同可分为绝对数指标、相对数指标和平均数指标。绝对指标是反映社会经济现象总体规模和水平的总量指标，若按反映时间的不同，绝对指标有时期指标和时点指标。时期指标反映社会经济现象在一定时期内所达到的总数量，其值可直接累加，如工业总产值、商品销售总额等；时点指标反映社会经济现象在某一时点内所达到的规模或水平，其值不可直接累加，如年初人口数、储蓄存款余额等。

相对指标是两个有联系的绝对指标数值的比值，常用的相对指标有计划完成相对数、结构相对数、比例相对数、比较相对数、强度相对数和动态相对数等六种，它们的计算与特点可见表 1-1 所示。

表 1-1 几种相对数对比表

相对数名称	计算公式	分子与分母相比					举例
		时间	空间范围	指标含义	计算方法	能否颠倒	
计划完成程度	$\frac{\text{实际完成数}}{\text{计划完成数}}$	相同	相同	相同	相同	不能	空调年销量计划完成程度
结构相对数	$\frac{\text{总体某部分数值}}{\text{总体数值}}$	相同	不同	相同	相同	不能	我国人口占世界人口比重
比例相对数	$\frac{\text{总体某部分数值}}{\text{总体另一部分数值}}$	相同	不同	相同	相同	能	重工业与轻工业产值之比
比较相对数	$\frac{\text{甲国(单位)某现象数值}}{\text{乙国(单位)同现象数值}}$	相同	不同	相同	相同	能	北京市与广州市消费物价指数之比

相对数名称	计算公式	分子与分母相比					举例
		时间	空间范围	指标含义	计算方法	能否颠倒	
动态相对数	$\frac{\text{报告期数值}}{\text{基期数值}}$	不同	相同	相同	相同	一般情况下不能	广州市国内生产总值 $\frac{1999\text{年}}{1998\text{年}}$ 之比
强度相对数	$\frac{\text{某一总体指标数值}}{\text{另一有联系总体指标数值}}$	相同	不同	不同	不同	能	资金利润率

在实际运用中，要注意绝对指标和相对指标的结合运用。

综上所述，统计指标是统计表现的最基本方式，指标和标志既有区别又有联系。两者的区别主要体现在指标是说明总体的数量特征，所有的指标都用数字表示；而标志则说明总体单位的特征，既有用数字表示的数量标志，又有用文字表示的品质标志。两者的联系主要表现在许多指标的数值都是由总体各单位数量标志值汇总得到的，同时由于总体和总体单位随研究目的的变化而变换，所以指标和数量标志在一定条件下也可变换。

### §§§§ 变异和变量

变异是指可变标志在总体各单位之间所表现出来的差异，是统计研究的基础。变异是一种现象，不存在变异值。变量是可以取不同值的量。可变的数量标志和统计指标就是变量，如年龄、工资水平等。变量的具体数值称为变量值，如年龄 18 岁、月工资 1000 元。变量按其数值形式不同可分为离散变量和连续变量，离散变量只能取整数，如人口数等。连续变量可取任意小数，如产值、重量等。变量按其所受影响因素不同可分为确定性变量和随机变量，确定性变量是受确定性因素影响的变量，其变量值将发生多大变化及变化的方向均可确定。随机变量是受不确定性、偶然性因素影响的变量，其变量值变化的大小及方向均无法准确确定。

### §§§§ 统计量和参数

统计量是根据样本数据计算的指标，是样本的一个不依赖于任何未知参数的量，通常用小写字母表示，如样本均值  $\bar{x}$ 、样本标准差  $s$  和样本成数  $p$  等。参数是反映总体的指标，基本上是由样本统计量推断而得，通常用大写字母表示，如总体均值  $\mu$ 、总体标准差  $\sigma$  和总体成数  $P$  等。

# 缘 统计软件的应用

统计软件的应用始于 20 世纪 50 年代后期，目前常用的有 SAS (统计分析系统)、SPSS (社会科学统计软件)、SYSTAT (生物医学资料处理) 等。

SAS 系统为模块式设计，用户可选择不同功能的模块组合以满足不同的需要，它集数据访问、管理、分析和显示于一体，除统计分析外还有制图、矩阵运算、线性规划、质量控制和新药临床试验等功能，为经济管理、社会科学、生物医学、质量控制等领域的众多用户所采用，是公认的较完善全面的软件包，但其操作略为繁琐。

SPSS 具有操作简便、统计功能齐全、数据交换强大以及视窗组合等特点，在商务、政府部门、教学与科研单位的定量研究中发挥了巨大的作用。目前使用该软件的企业和各界人士正在急剧增加。本书的案例分折均附有该软件的简单操作步骤。

SYSTAT 是一套一般性用途的统计分析软件，它可以应用于统计资料的整理与分析，由于它结果显示易于理解、所占的记忆容量不大，所以可作为大型研究的先期探究和小型研究的主力。

SYSTAT 可称得上是统计软件包的元老，它有 100 多个相互独立的程序模块，较易增加新的模块，统计分析功能强大，在生物医学界和研究部门应用很广。

当前，统计分析的每一环节几乎均离不开统计软件的应用，由于软件功能的强大，已使稍具数学基础者都可应用统计方法，但对统计学原理的透彻领悟乃是正确解释电脑处理及分析结果的前提。

## 习题一

1. 统计学的概念及分类。

2. 统计总体与总体单位的关系。

3. 统计指标与标志的关系。

4. 相对数有几种，它们有何区别？

5. 统计软件 SAS 的基本特点。

## 第二章 统计数据的收集与整理

统计作为一种从数量方面认识事物的科学方法，它在对社会经济现象进行统计研究时，其研究的基础是大量的统计资料。因此首先要对所研究的对象进行调查，收集相关的数据资料，然后对数据资料进行整理使之系统化，为下一步的统计分析打下基础。本章将介绍统计数据收集和整理的常用方法。

### 第一节 数据资料的收集

数据资料的收集，一般称为统计调查。它是根据统计研究的目的和要求，采用科学的方法，有组织、有计划地收集统计数据的过程。

#### 一、数据资料收集的组织形式

##### （一）统计报表制度

统计报表制度是由政府统计机关或统计调查的组织机构按照统一的内容自上而下地统一布置，自下而上地逐级提供统计资料的调查组织形式，按统一规定填报报表所形成的制度，可以分为全面统计报表和非全面统计报表，按其报送周期，可以分为定期报表和年报，包括日报、旬报、月报、季报、半年报、年报等。统计报表按其填报单位，可以分为基层报表和综合报表。基层报表是由企业、事业等基层单位填报的统计报表，综合报表是由主管部门或统计部门根据基层报表逐级汇总填报的统计报表。

在我国，统计报表还按其内容和实施范围，分为基本统计报表和业务统计报表。基本统计报表的内容，主要是有关国情国力和国家宏观管理、决策需要的基本的综合性指标；业务统计报表的内容，主要是有关各业务主管部门生产管理需要的指标。

采用自上而下布置统计报表、自下而上逐级报送统计资料的组织形式，在有健全的统计组织和报送渠道以及完善的基层企业原始记录和统计台账的条件下，能够取得比较完整的、准确的统计资料，便于连续地、定期地进行观察对比，但这种统计调查形式比较繁琐，所需人力、物力、财力较多，主要适宜于国有和集体企业、事业以及国家机关、团体，而对于私人企业、个体单位、家庭和个人，一般不宜于采用。

##### （二）普查

普查是一种专门组织的一次性全面调查，如人口普查、农业普查、工业普查、物资普查等。普查主要用来搜集某些不能够或不适宜于用经常调查而又要求准确的全面统计资

料，以了解重要的国情国力，普查一般用来调查一定时点上的现象，也可用来反映时期现象。普查有按一定周期进行的，也有根据需要临时进行的；有包括准备时间在较长时期才能完成的，也有在短期内快速完成的。随着现代电子计算和传输技术的迅速发展，多数普查趋向按照一定周期进行，如人口普查，世界上许多国家都规定为每五年进行一次。

### （猿）抽样调查

抽样调查是一种非全面调查的组织形式。在通常意义上，是指按照随机原则，从统计总体中抽取一定数量的单位作为样本，进行调查观察，用以推断总体的统计调查方式。抽样调查的特点是：它根据概率论和大数定律的原则，科学的计算，从总体中抽取少量单位进行调查观察，就能够用来推断总体；抽样调查调查单位少，可以用较少的人力、物力、财力和时间，轻便灵活，比较容易组织实施。常用的抽样调查组织形式有以下几种：

①简单随机抽样，简单随机抽样又称纯随机抽样。是指按随机原则直接从总体单位中逐个抽取，共抽  $n$  个单位组成样本，并使得各个不同的可能样本中的每一个样本被抽出的机会都相等的抽样组织形式。常用的方法有抽签和用随机号码表抽取样本。其理论上最符合随机原则，但在实际应用时受到一定的限制。如：当总体单位数很多时，难于编号和直接抽样。

②分层随机抽样，简称分层抽样，又称类型抽样。是指先按某一标志将总体分成若干组，然后再从各组中随机抽取若干个单位组成样本的抽样组织形式。例如：对城市职工进行家计调查，先把职工按所有制形式分为国有经济单位职工、集体经济单位职工和其他所有制单位职工，然后再从各组职工中随机抽样本单位。

分层抽样的分组标志，一般应选择与被研究现象有关的重要标志。通过分组，尽量缩小组内各单位标志值的差异，增大组间各单位标志值的差异，以便降低抽样误差。分层抽样又有等比例分层抽样与不等比例分层抽样之分。如果各组之间的差异大而各组内的差异程度差别不大时，可采用等比例分层抽样。如果各组之间的差异大，而且各组内的差异程度也显著不同，则可采用不等比例分层抽样。差异程度较大的组应该多抽几个，而差异程度比较均匀的组可少抽几个，以便降低抽样误差。

分层抽样适宜于由差异较大的单位所组成的总体。它将分组法与随机原则结合起来，减少了各组内标志值的差异程度，使各组都有抽取样本单位的机会，有利于提高样本的代表性，能得到比简单抽样更为准确的结果，因此在实际工作中应用较广泛。

③系统随机抽样，简称系统抽样，又称等距抽样、机械抽样。是指先将总体各单位按照一定标志顺序排列，然后依固定顺序和间隔，按样本容量  $n$  将所有总体单位  $N$  分为  $n$  个相等的部分（或叫确定抽样距离），每个部分包含  $\frac{N}{n}$  个单位。再用随机抽样方法确定在每个部分中的抽样序号  $k$ ，从每个部分的  $\frac{N}{n}$  个单位中抽取排序为  $k$  的那一个单位组成样本的一种抽样组织形式。

在系统抽样时，用来排队的标志有两种类型：一种是与被调查的内容无关，称为无关标志，在这种情况下，系统抽样与简单随机抽样相仿，因为被研究总体是随机总体。另一种是与被调查的内容有关，称为有关标志。在这种情形下，系统抽样与分层抽样相仿，因为它类似于按某个标志等距分组，利用了总体的有关信息。在实际工作中采用系统抽样确定抽样序号的方法通常有三种：第一种：随机起点，固定等距抽取单位。第二种：半距起点，固定等距抽取单位。第三种：随机起点，对称距离抽取单位。

系统抽样与分层抽样的区别主要在于：系统抽样在所分的每个相等部分中只抽取一个单位，且中选单位总处在每个部分中的同样位置或对称位置。分层抽样在所分的每个组中抽取若干个单位，每个组的单位可以相等或不等。即使每个组只抽一个单位，但每组的中选单位不一定处于相同位置或对称位置。系统抽样在实际工作中广为采用，它能使被抽选的各单位更均匀地分布在总体各个部分，提高了样本对总体的代表性。

④整群抽样，又称成批抽样。是指抽样时，先把总体按自然形态分为若干群，每群内均包含若干个总体单位，然后以群为单元，从总体中随机抽取一部分群组成样本，并对已抽中群群内的每个单位进行全面调查的一种抽样组织形式。例如：对某月连续生产的某工业产品进行质量检验，以一个小时内所有生产的产品为一个群，然后每隔一定时间抽选小时产品全面调查便是整群抽样。

整群抽样与前三种抽样组织形式有明显不同。前三种形式都是一个一个地从总体中抽取单位组成样本，可称为个体抽样。而整群抽样则是一群一群地从总体中抽取单位组成样本。每一群中包含的单位数可以相等也可以不相等。整群抽样只需编制总体各群的名单，比较节省费用和时间，取得结果比较迅速，主要用于大量生产条件下工业产品质量的检验。特别是总体缺乏可靠的总体内各单位名单或者编制总体单位名单的费用太大时的大规模调查中采用。由于整群抽样抽选单位比较集中，显著影响了在整个总体中各单位分布的均匀性，因此整群抽样比其他抽样形式的抽样误差要大些，这可通过增加抽选群数，减少每群内总体单位数的方法来弥补。

必须指出：整群抽样的分群和分层抽样的分组，其要求是不同的。分组要求增大组间差异，缩小组内差异，以利于提高样本代表性；而分群抽样要求缩小群间差异，增大群内差异来提高样本代表性。因为对中选群是作全面调查，群间差异越小，中选群的代表性就越大。

⑤多阶段抽样。单阶段抽样是指从总体中通过一次抽样就产生一个完整的样本的一种抽样方法。如前述四种抽样组织形式都是单阶段抽样。多阶段抽样是指把抽样过程分为若干个阶段，通过对多个阶段的抽样后才产生完整的样本这样一种抽样方法。在前几个阶段以群体的形式抽取单元，但并不把中选单元中的单位作为样本进行具体项目的调查；而在最后一个阶段才以个体的形式从上一阶段中选单元中抽取调查单位组成样本进行调查，或者以群体形式从上一阶段中选单元中抽取群并以中选群中全部单位组成样本进行调查。例如，我国职工家计调查，把抽样过程分为两个阶段。第一阶段从全国各城市中抽选调查城市；第二阶段从第一阶段中选的城市中抽选调查户组成样本。

多阶段抽样在大规模抽样调查工作中应用较多，但组织工作复杂，计算推断也较麻烦。在实行多阶段抽样时，各阶段采用的抽样方式可以不同，也可以相同。总的目的是使样本尽可能具有充分代表性，同时也要考虑经费与人财物力的可能。

(源)典型调查。根据调查的目的和要求，在对调查对象进行初步分析的基础上，有意识地从挑选有代表性的单位进行研究的一种非全面调查，人们则常称之为判断抽样。它是非随机抽样中的一种。唯物辩证法中关于个性与共性、特殊性与普遍性的辩证关系的理论，是典型调查的科学基础。典型调查从研究个别到普遍，然后又以普遍指导个别，是符合人类认识运动的规律，是认识事物的科学方法。而且，它灵活方便，反映迅速，省时省力，深入具体，可以把数字和情况结合起来，有它特殊的作用。表现为：对个别或少数具

有代表性的典型单位进行深入的、周密的调查研究的一种非全面调查的组织形式。典型调查的特点就如“解剖麻雀”，对某些问题作深入细致的调查研究，藉以研究现象的特点和规律性，可以利用典型调查资料，验证、补充全面统计或其他非全面调查的不足，了解新情况、新问题，研究新生事物，推广新鲜经验。

搞好典型调查的关键是要正确地选择典型单位，使典型单位真正具有代表性。如果典型单位选择不当，就会降低典型调查的作用，甚至会从中导出错误的结论，造成危害和损失。因此，选择典型单位时，要注意以下几点：

①要根据统计调查的目的和调查对象的特点，去正确划分社会现象的类型和确定所选典型的类别。

②要注意典型的代表性，要从事物的联系中去挑选典型。

③典型单位数目的多少要依总体单位之间的差异程度而定。差异程度大时，典型单位应多些，或采取“划类选典”的方法，从各类中分别选择典型；差异程度小时，典型单位数则可少些。

以上调查组织形式还可以从不同的角度进行分类，从调查的时间分为经常性调查和一次性调查，经常性调查是对调查对象进行连续不断的经常性的登记，一次性调查是对调查对象在某一时刻的情况进行调查，同类的调查以后在某一时刻可以再次进行，也可以不再进行。按调查的范围分为全面调查和非全面调查，全面调查是对调查对象的所有单位毫无例外地逐个进行调查登记，非全面调查则是只对调查对象的一部分单位进行调查登记。按调查的组织形式还可分为统计报表制度和专门调查等。

## 四、数据资料的收集方法

(一) 直接观察法。直接观察法是调查人员亲临现场，对调查对象进行观察、测定，以取得第一手资料的一种方法。这种方法的特点，主要是调查人员凭借自己的感觉器官或借助仪器设备对调查对象的活动和现场情况加以考察、测定和记录。直接观察法的优点是能够保证资料的准确性，但需要训练有素的调查人员和精良的仪器设备，且所需人力、物力、财力较多，只能用于非全面调查。

(二) 报告法。报告法是我国统计调查中最常用的一种方法。我国的报表制度和某些一次性调查，就是采用报告法。报告法注重调查单位有健全的原始记录和核算制度，统计基础工作较好，调查结果具有完善的行政渠道和报告系统的条件下能够有效地取得资料。

(三) 采访法。①面谈：调查员向被调查者逐一询问、逐一记载以取得资料的方法。它的优点是调查人员对调查项目有统一理解，能按统一的口径询问和取得资料，所获得的资料比较可靠。但需要花费较多的人力和时间。②小组座谈会：调查者与被调查者直接见面和交谈，主持人可以清楚地向被调查者说明调查目的、要求和具体内容，被调查者有什么疑虑和不明白的地方可以随时提出，由调查者解答，彼此间比较容易沟通，反馈性强，可以取得比较深入细致的、丰富的信息和较为准确可靠的资料。

(四) 邮寄法。是由调查者将调查提纲或调查表邮寄给被调查者，向被调查者搜集资料的方法。这种方法在国外进行的一些社会、经济调查，特别是市场调查和民意测验方面被广泛采用。

(五) 电话访问。电话访问是调查人员通过电话向被调查者询问、交谈以获得所需资

料。这种方法具有节约人力、费用和时间的优点，适用于对一些简明扼要、数量有限的问题进行调查。电话采访要求具有比较完善的电话系统和完整的电话簿可供利用。因而这在西方国家是一种普遍使用的调查方法，随着电子计算机系统的应用，电话访问系统将会日益完善，目前在外国流行的有中心控制电话访谈法和电脑辅助电话访谈法（悦群）。

（远）实验法。实验法是利用自然科学中的实验求证方法用于市场调查中，是对市场现象的实验。实验法中调查者有目的地控制一个或几个市场因素的变化，以研究某市场现象在这些因素影响下的变动情况和规律。

### （苑）网上调查

#### ① 网上调查的优点

网上调查在 20 世纪 90 年代开始热门起来。这种现象主要有几个原因：

由于省略了印制、邮寄和数据录入后过程，问卷的制作、发放及数据的回收速度均得以提高。可以短时间内完成问卷并统计结果及报表。

费用问题，印刷、邮寄、录入及调研员的费用都被节省下来，而调研费用的增加却很有限。因此，进行大规模的调研较其他如邮寄或电话调研方法省下可观的费用。

随着网上固定样本调研的出现，调研员能够通过跟踪受访者的态度、行为和时间进行纵向调研。复杂的跟踪软件能够做到根据上一次的回答情况进行本次问卷的筛选，而且还能填补落选项目。

打一个电话却只提两三个问题在费用上是不值得的。但在网上，调研内容可以很容易包含在市场、商贸或其他一般站点上。例如，如果一个人上了银行主页，击活“信用卡”连接，在进入正式网页之前，他可以被询问几个有关被认为是最重要的信用卡特性问题。

网上可以接触很多人。目前很难想象还有什么媒体可以提供那么大的调研群体。

问卷调研还有一个独一无二的优点，即它们在视觉效果上能够吸引人。互联网的图文及超文本特征可以用来展示产品或介绍服务内容。对于那些有较新版本（两个最为流行的网上浏览器）的用户，声音及播放功能还可以加入到问卷中。这是其他调研方式所无法比拟的。

#### ② 网上调查的缺点

网上调查在目前来说还有不少缺点。最大的一点恐怕就是上网的人不能代表所有人口。使用者多为男性，教育水平高、有相关技术，较年轻和较高收入的人。不过，这种情形正有所改变，越来越多的人开始接触互联网。

商业网（尼尔森互联网人口统计研究（浏览者曾访问过多少个网站））利用万维网进行调研，同时结合电话调研以评估互联网数据的偏差。随着这两种调研方式差异的减少，互联网已成为一种大众化的市场工具。虽然这还需要一些年，但在新技术快速变化、传播及采用的时代，这注定会成为现实。

计算机产品购买者或是互联网使用者，是使用互联网调研的理想对象。利用互联网的企事业单位使用者也是不错的可发展的调研对象。目前估计有 1/3 的企事业单位已上网。

另外一个问题是互联网的安全性。现在很多使用者为私人信息的安全性担忧，加上媒体的报道及针对使用者的各种欺骗性文章，更使人忧心忡忡。然而，考虑到对互联网的私人信息，诸如信用卡账号之类进行担保的商业目的，提高安全性仍是互联网有待解决的重要问题。