

中国社会学函授大学教材之七

统计学原理

北京大学社会学系

郭崇德 主编

中国社会学函授大学编印

前 言

本教材是为中国社会学函授大学的同学们编写的。我们考虑到函大同学来自不同的岗位,文化程度不同,社会经历和背景不同,而且主要是自学,因此,我们尽可能做到深入浅出、通俗易懂。考虑到大部分同学没有接触过数理统计,课程中涉及较深的数学推导过程我们都予以省略。我们希望学有余力的同学,在使用本教材的同时参考一些有关数理统计的书籍,这样会对本课程的学习有较大的帮助。

在编写中,我们广泛参考了国内外出版的有关社会经济统计方面的书籍,尽可能将社会统计学的内容选入本书,以使同学们能对社会统计学有一定的了解。

本书主编是北京大学社会学系讲师郭崇德,研究生阎焱、钱江洪、张杰、傅康园也参加了编写。张杰编写第四章、第八章;阎焱编写第一章(部分)、第九章;傅康园编写第十章、第十二章(部分);钱江洪编写第十一章、第十二章(部分);其余各章由郭崇德执笔。全书各章由郭崇德和阎焱校阅。

本书不仅可作为函大教材,也适宜其他有志于社会学的同志自学,并可作为有关专业的教学参考书。

由于编者水平有限,加之时间、篇幅的限制,本书难免有不少缺点和错误,恳请广大读者指正。

编 者

1985年5月20日

目 录

第一章 导言	(1)
第一节 统计学的性质	(1)
第二节 统计学的研究方法	(2)
第三节 统计的作用	(4)
第四节 统计学与其他科学的关系	(4)
第五节 统计学中的几个基本概念	(5)
第二章 统计调查	(8)
第一节 统计调查的意义、作用和种类	(8)
第二节 统计调查的核心——统计调查方案的设计	(11)
第三节 几种统计调查方式	(16)
第三章 统计整理	(18)
第一节 统计整理的意义、原则和步骤	(18)
第二节 统计分组	(19)
第三节 次数分配	(22)
第四节 频数和频率	(34)
第五节 统计表	(39)
第四章 统计图示法	(45)
第一节 统计图示法的意义和作用	(45)
第二节 条形图	(46)
第三节 平面图、立体图和象形图	(51)
第四节 统计地图	(55)
第五节 曲线图	(58)
第六节 人口金字塔	(64)
第五章 统计指标	(67)

第一节	统计指标的概念和作用	(67)
第二节	统计指标的种类	(69)
第三节	相对指标的种类	(72)
第六章	现象集中趋势的计量和分析	(78)
第一节	现象集中趋势的意义和作用	(78)
第二节	各种平均指标	(81)
第三节	应用平均指标应注意的问题	(111)
第七章	现象离中趋势的计量和分析	(115)
第一节	现象离中趋势的意义	(115)
第二节	测定离中趋势的各类指标	(117)
第三节	离散系数	(131)
第八章	指数法	(133)
第一节	统计指数的意义及编制原则	(133)
第二节	指数的种类	(134)
第三节	统合指数体系	(136)
第四节	平均数指数体系	(146)
第九章	抽样调查方法	(151)
第一节	抽样法的作用和特点	(151)
第二节	抽样误差的概念	(153)
第三节	抽样误差的计算	(159)
第四节	抽样推断	(165)
第五节	必要抽样数目的确定	(168)
第十章	相关分析法	(170)
第一节	相关关系的概念和种类	(170)
第二节	对相关关系的定性判断和分析	(175)
第三节	相关表和相关图	(176)
第四节	相关系数	(181)
第五节	回归直线	(186)
第六节	估计标准误差	(191)

第十一章	动态数列分析	(194)
第一节	动态数列的意义、种类及编制原则	(194)
第二节	动态数列的绝对数分析指标	(197)
第三节	动态数列的相对数分析指标	(200)
第四节	动态数列的平均数分析指标	(203)
第五节	现象发展长期趋势的分析	(212)
第十二章	统计预测	(226)
第一节	统计预测概述	(226)
第二节	趋势预测	(227)
第三节	因果预测	(233)

第一章 导 言

第一节 统计学的性质

统计是人类社会发展的产物。在原始社会它的最初表现是以石子、结绳为标记来计量对象物，以后随着社会生产力的发展，它逐步成为治理国家、管理社会经济活动的重要手段，主要表现为对国土、人口、自然资源的调查，统计方法也因此有了很大的发展。但统计上升为理论最后成为一门独立的学科，还是在近代。统计一词，英文为 Statistics，它源于拉丁文 Status，原意是指国家。随着统计应用范围的不断扩大，统计一词的涵义不再局限于某一现象的数量方面资料的收集和计算，而是泛指统计工作（实践），统计资料和统计学（理论）。统计资料是统计工作实践的成果，统计学是对统计实践经验的理论概括。

统计学的研究对象是大量的社会现象的数量方面，通过对社会现象的数量和数量关系的研究从而认识事物的本质和发展规律。它的研究对象具有以下特点：

统计学研究的数量是通过对事物大量观察得来的大量的数字资料来综合说明事物的特征，而不是研究个别事物的数量、单独个体的观察或少量的观察。没有大量数据的综合研究，不是统计研究。例如：研究一个国家的人口发展状况，它必须是全面掌握反映人口的许多指标数值说明人口发展的特征，而不是对某一个别人的特征进行研究。因此，统计学的范围必须是大量的，能代表总体的一般性综合资料。

统计学研究社会现象是以数字记录为基础，通过大量观察方法取得反映事实真象的数据。统计学研究问题不能离开统计数

据，因为，数字记载比其他任何形容词对现象的描述都具体和确切。例如：反映一个国家物质生活水平的高低，不能简单的说好、比较好或差的概念。好、比较好和差的概念基准就是用统计数量来加以区分。

统计学研究的大量社会现象的数量，必须是同质的总体的量。所谓同质总体是指研究对象根据研究目的在时间，空间和属性方面具体相同的性质。时间方面的同质是指被研究的静止的社会现象必须以同一时间标准为基础。例如：对比分析各地区的人口数，各地区必须以相同时间的人口数进行比较。空间方面的同质是指被研究的动态社会现象必须是同一地区在不同时间上的对比分析。例如：观察某地某种商品物价变动情况，只能用该地区该种商品不同月份或年份的物价水平进行对比分析才是有意义的。属性方面的同质是指事物之间进行对比分析时应把相同的事物放在一起对比分析才有实际的意义，或谓有“可比性”。

统计学是研究大量社会现象的数量表现和数量规律，它所显示的每一个数量，都是具体事物的数量表现，都有具体的社会经济内容。例如：仅表现为7,490的一个数，只能说明它是含有千位数值的一个数，若表明我国1981年全年完成工农业总值是7,490亿元，这个包含有实质性内容的统计数字才是统计学研究的具体数量。

第二节 统计学的研究方法

统计学的研究方法是遵循大数法则，其基本方法是大量观察法。它必须是在观察大量的现象之后，综合地反映大量现象的数量特征，而不是去研究个别事物的个别特征。在大量中去探求共性的、一般的规律。观察的总体单位越多，越能反映总体的特征，只有通过大量观察，才能消除偶然因素对事物的影响，显现出恒常因素的作用。这就是数理统计中所讲的“大数法则”作用

的结果。

统计学的研究方法还应遵循辩证唯物主义的基本观点去观察和研究现象的数量表现和数量规律。即必须从事物的相互联系和相互制约中去研究数量表现和规律。事物之间的联系和制约，必然从数量上反映出来，比如：一个国家的物质生产领域是由各物质生产部门构成，如果农业收成不好，势必影响工业、交通运输业等其它部门的生产量下降。因此，在研究现象的数量关系时，必须同现象之间的本质联系起来才能得到正确结论。统计学在具体研究现象过程中，必须注意到从事物的量变到质变中去研究事物的数量关系，考察事物的质变，必须通过观察事物的量变达到一定的界限才可能说明质的变化。比如：一个国家的工业化程度高低，是依据工业生产总值在国民生产总值所占的比重大小的数量指标来衡量。人口老、中、青的划分是由一定的年龄大小界限来判断的。统计学还应从事物的不断发展中去研究事物的数量变化及其规律，应用统计数量之间的对比来说明事物发展的快与慢，好与坏，优与劣。统计学在研究事物中还必须注意从事物的矛盾中去研究事物的数量关系，采用对比分析的方法从数量中发现事物的矛盾现象，为人们研究矛盾，解决矛盾，推动事物前进提供数量依据。

上述的基本观点体现在统计研究的基本方法中，统计研究的基本方法是：

一、大量观察法：即先取得大量的原始资料，然后对现象进行全面、综合地分析的观察方法。

二、分类分组法：已经获取了大量的原始资料之后，把不同性质的事物分门别类的区别开来进行比较观察的方法。

三、综合指标法：即利用统计方法和技术，对统计资料进行加工、整理、并计算出许多说明总体特征的指标，反映现象和数量上的本质和规律的方法。

统计研究的具体方法是：平均法、动态法、指数法和相关

法。

平均法是研究和分析现象的一般水平。动态法是从事物的连续性来研究和分析事物的发展变化。指数法主要是用来研究和分析事物的内部构成因素的变动。相关法是研究和分析事物之间的相互关系。

第三节 统计的作用（任务）

统计服务的有效手段，就是向各方面提供所需要的真实准确的统计资料。统计的作用，主要有三：认识作用、监督作用和管理作用，具体说来有以下几项：

（一）准确、及时、全面和系统地提供有关社会经济发展情况和资料，并加以分析，进行必要的预测，以此为国家制订政策和计划、指导社会 and 经济发展提供所需要的数据和情况。

（二）对政策和计划的执行情况进行统计检查和监督，这也要通过提供统计资料来实现。

（三）为各部门、各地区和各单位管理各项经济事业和社会事业提供所需要的统计及其分析资料。

（四）为科学研究和宣传教育提供统计资料。

统计数字的真实性是统计工作的生命，坚持如实反映情况是做好统计工作的首要原则。在我们今后的统计工作和社会调查中，真实性和科学性是我们必须严格遵守的原则。

第四节 统计学和其它科学的关系

统计学成为认识社会的有力武器，必须有它正确的理论指导，统计学和很多有关科学有着密切的联系，其中，尤其和哲学、政治经济学、社会学和数理统计学的联系最为密切。

（一）由于统计学本身方法论的性质，它与哲学的关系非常

密切。哲学是人类认识世界最一般的方法论科学，统计学既然也是一门认识社会的方法论科学，它就必然要以哲学作为它的方法论基础，如统计研究特别强调一切从实际出发，强调从质与量的密切联系中来认识事物，这都是哲学上最基本的观点。

(二) 统计学既然主要以社会经济现象为其研究对象，这就规定了它必然以社会经济理论科学作为其理论指导。例如：总产值，净产值，基本建设投资额，总和生育率等指标，都反映了一定的社会，经济范畴，在确定它们的具体概念和范围时，必须以相应的政治经济学和社会学理论所阐述的有关原理为依据。

(三) 统计学研究事物的数量关系，离不开数学方法，尤其是数理统计方法，因为在社会经济现象中，存在着大量的、随机性的偶然现象。同时，有许多关系可用一定的数学模型来研究。我们后面将讲到的统计平均数和标志变动度，抽样推断，相关分析，以及统计预测等等，都是数学和数理统计方法在社会经济研究中的应用。随着科学技术的高度发展，电子计算机日益普及，数学已经深入到社会经济研究的每一个领域，马克思曾经说过：一门科学，只有当成功地应用了数学才称得上完善。因此，我们建议学习本课程时，若有余力，希望能看些有关数理统计方面的书，这对于本课程的学习是会有很大的帮助的。

第五节 统计学中的几个基本概念

为了进一步研究统计学的基本原理和方法，有必要先介绍一下统计学中常用的几个概念：

一、统计总体与总体单位

统计总体是指被观察事物的全体，它是由客观存在的同一性质的许多个别事物结合而成。

统计总体单位是指构成统计总体的个别事物。

统计总体和总体单位是一个事物的两方面。总体是对事物的

整体而言，总体单位则是指个体而言。例如：观察一个工业企业的职工收入水平，总体就是工业企业的全部职工，而企业的每一个职工就是总体单位，总体与总体单位，由于观察的角度不同可以相互转换，例如：观察全国的工业企业职工收入水平，此时，全国所有工业企业构成总体，每一个工业企业则为总体单位。

二、标志

标志是说明事物的属性或特征的名称。例如，人口数按男、女性别分类统计，人口的“性别”一词，是说明人口的自然属性的名称，就是一个标志。又如：人口数按年龄分组统计，一岁年龄组的人口与五十岁年龄组的人口，在生理特征上有很大不同。因此，凡是用来说明事物的属性标志称为品质标志（如人口按性别分，按民族分，企业按所有制形式分等）。而用来说明事物的数量特征或数量界限的称为数量标志（如人口按年龄分组，职工人数按工资多少分组等）。

三、指标

指标是用来说明总体在一定时间、地点条件下的规模和水平的数量化的科学概念。例如：中国1981年完成的工农业总产值是7490亿元，工农业总产值是一个特有的经济概念，只有属于工业和农业的生产总值才能统计在内，不属于工业和农业的生产活动成果则不能统计在内。这个7490亿元就是具体说明1981年在中国这个具体地点、条件下工农业生产达到的规模和水平，所以，指标是一个分析社会经济现象的科学概念，未经实现的指标为计划指标，已经实现的指标为统计指标。

标志和指标是两个不同的概念，两者间既有联系又有区别，二者的主要区别有二：

（一）指标是说明总体特征的，而标志则是说明总体单位特征的。

（二）标志有不能用数值表示的品质标志与能用数值表示的数量标志两种，而统计指标都是用数值表示的，没有不用数值表

示的统计指标。

指标和标志的联系，主要也有两点：

(一) 计算上的联系，有许多统计指标的数值是从总体单位的数量标志汇总来的，如：一个县的粮食总产量是所属各乡、村粮食产量的总和。

(二) 指标和标志间存在着变换关系，由于研究目的不同，原来的统计总体如变成总体单位了，则相对应的统计指标也就变成标志了，反过来也是如此。

四、差异、变量和变量值

差异是指现象之间客观存在的数量差别，客观现象存在的数量差别是统计研究的基本内容，没有差别，也就不必进行统计和统计研究。变量是说明总体的差异标志。反映总体的差异标志很多，可以因研究问题的目的而定。例如：反映人的差异标志有身高、体重、年龄等等都可称为变量，变量值是指差异标志的各个具体的数值。

差异、变量和变量值是一个事物的统一体，例如：观察某班大学生的平均年龄，班内每一个大学生的年龄很可能是不相同的，甲的年龄是18岁，乙的年龄是19岁，丙的年龄是20岁……等等，18岁与19岁和20岁之间在具体数值上的不同，这个不同的差别就是差异，观察大学生的年龄，“年龄”就是一个变量，班内每个大学生的具体不同年龄就是变量值。

变量又可分为连续变量和非连续变量。

连续变量的数值在变动范围内是连接不断的，相邻两值间可作无限分割，例如：人的身高、体重均是连续变量，连续变量可以细算到许多位小数值。

非连续变量的数值在变动的范围内只能取有限数值，且在数值之间以整数断开，例如：统计一个地区的工厂数有多少，农村公社数有多少，人口数是多少等，只能以个数、人数表示，不能以500.5个厂或300000.05个人表示。

第二章 统计调查

第一节 统计调查的意义、作用和种类

统计是认识社会的有力武器。要认识社会，重要的方法就是向社会作调查。人的认识离开社会实践，离开对实际情况的调查，认识也就成为无源之水、无本之木。统计调查就是按照预定的目的，采用科学的调查方法，有组织地向客观实际搜集资料的过程，其目的是为了获得被研究的客观现象或事物在数量上的表现。这一阶段，属于感性认识阶段。

从马克思主义的认识论出发，只有感性的材料十分丰富和合乎实际（不是错觉）才能根据这样的材料引出正确的概念和理论来，因此，统计调查在统计研究的全部工作中占有极为重要的地位。因为，统计调查在统计工作的整个过程中，担负着提供基础资料的任务。所有统计计算和统计研究，都是在原始资料搜集的基础上建立起来的。如果统计调查得来的资料不真实、不及时、不系统、不完备，就会影响整个统计工作，使统计不能很好地发挥认识社会的作用，只有搞好统计调查，才能保证统计工作得以实现对于客观事物规律性的认识，为党和国家的政策制订、计划检查等提供准确、及时、可靠的依据。

统计调查根据统计研究的目的和被研究对象的不同而分为以下几种：

一、根据调查对象所包括的范围不同，可分为全面调查和非全面调查。

（一）、全面调查

全面调查是指在统计调查过程中，对调查对象的所有单位都一一加以调查，例如，普查和统计报表制度。全面调查的目的是

为了取得较为全面、系统的统计资料，譬如，要了解全国人口的基本状况，就必须对每个人的有关情况进行一一的调查，否则，就不可能得到有关全国人口状况的统计资料。全面调查虽能对某一现象获得全面、系统的统计资料，但在调查过程中要耗费大量的人力、物力和时间，且牵涉面很广，如果因准备和组织工作不完善将会直接影响统计资料的准确性和可靠性，所以，在进行全面调查之前应慎重考虑，充分准备。

（二）、非全面调查

非全面调查是指在统计调查过程中，对调查对象的一部份单位进行调查，例如，抽样调查，典型调查和重点调查等。非全面调查的目的是通过局部调查资料来说明或代表被研究现象的总体情况。譬如，为了解全国解放三十年来少数民族人民的知识结构改变状况，可以对全国各个少数民族人民进行全面统计调查，但耗费人力和物力大，而少数民族聚居的地区是较集中的，在此情况下，就可以采取非全面调查中的重点调查，只要抓住了重点地区少数民族的知识结构状况，也就说明一般情况了。在非全面调查过程中，若方法得当，符合科学原理，比之全面调查会起到事半功倍的效果。所以，非全面调查在统计调查中具有重要的作用，具体表现为：

1、适用性强。非全面调查可以不受客观条件的约束，有些社会经济现象不适于搞全面调查，如产品的破坏性（如炸弹）质量检验，只能搞非全面调查，在遵循科学原理基础上，非全面检验就能代替全面检验。

2、能提高统计资料的时效性。与全面调查比较，非全面调查可以节省人力、物力、缩短调查时间，在较短的时间内迅速完成调查工作，从而提高统计资料的及时性。

3、能提高统计资料的准确度。由于非全面调查涉及面比全面调查少，可以使调查人员集中力量，深入细致地搜集资料，发现差错及时得以纠正，从而保证资料的准确程度。

4、可用于对某些特殊或专门问题的研究。譬如，居民货币收支情况调查，疾病发病原因调查，个人对现代化态度调查等等均可以广泛采用非全面调查。

二、根据调查组织形式的不同可分为统计报表制度和专门调查。

(一)、统计报表制度

统计报表制度是统计部门按照国家或上级主管部门统一规定的表式、内容、计算方法、时间和报送程序，自下而上的搜集基本统计资料的方式。新中国建立以来，国家统一制定了国民经济基本统计报表制度，各业务部门根据各个行业的特点和统计工作的需要，相应地建立了各项专业统计报表制度。统计报表制度成为各级政府、领导机关和生产主管部门取得国民经济计划执行情况以及经营管理所必须的基本统计资料。及时地反映各个部门、基层单位的生产经营管理基本情况。它是编制各类计划进行经济核算、考核生产单位、个人的工作质量、检查生产成果的重要依据。

(二)、专门调查

专门调查是根据特定的目的和要求，而组织的搜集统计资料的一种调查方式。例如，普查、重点调查、典型调查等都可以称为专门调查。专门调查可以用来补充统计报表全面搜集资料之不足。将专门调查与统计报表制度结合运用，可以使统计调查工作做得深入细致，了解情况全面。譬如，人口的出生与死亡登记是通过统计报表来取得资料的，如果用普查人口的出生与死亡资料，可以检验统计报表中的出生和死亡的准确程度。

三、根据对被研究对象搜集资料的时期分，可以分为经常性调查和一时性调查。

(一)、经常性调查

经常性调查是指对被研究对象进行不间断地连续性搜集统计资料的一种调查方式。经常性调查的目的是为了获得反映现象随时间的变化、发展在数量上的变动情况。例如，人口的经常性变

动是通过各级公安机关户口的经常登记取得统计资料。在社会经济现象中随着时日的变动，需要经常掌握其变动发展状况，相应地就得建立经常性统计报表制度作为搜集资料不可缺少的一种调查方式。从这个意义上讲，统计报表制度就是经常调查的具体体现。通过这种方式搜集的资料称为时期数（时间数列）。

（二）、一时性调查

一时性调查是指对被研究对象选择其某一时点上的状态进行搜集统计资料的一种调查方式。其目的是为了获得被研究对象在某一时刻上的静态统计资料。譬如，一个国家的人口数无时无刻不在变动，要能得到一个精确的人口数据，必须将人口的活动固定不变之后，才可以进行清点。这个一刹那时刻的人口数就称为时点数。我国一九八二年的人口普查就是采用的该年的七月一日零时为统计标准时点。诸如企业职工人数、生产设备调查、产品库存额等都应采用一时性调查。

在一时性调查中，又可分为定期与不定期调查。

定期调查是指每隔一定时期进行一次调查，其时间间隔大体相等，如美国人口普查，每隔十年进行一次。不定期调查是根据某时期的实际需要所组织的一次性调查，例如，我国不定期的人口普查。

第二节 统计调查的核心——统计调查方案的设计

统计调查方案是保证统计调查顺利进行的指南，如果没有周密、详尽的统计调查方案，统计调查的科学性便无从谈起。因此，统计调查方案是统计调查的核心问题。

统计调查方案就是在进行调查之前制定的实施调查计划，尤其以组织较大型的或普查性质的统计调查，必须要拟定（设计）调查方案，调查方案一般应包括以下几方面的内容：

一、确定调查任务和目的

制订调查方案的首要问题是要明确调查的任务和目的。如果任务不清楚，目的不明确，就无法确定向谁调查，调查什么，以及用什么方式调查等问题，或者调查目的和任务含混不清，由此调查得来的资料，要么不能用，要么不全面，不能达到预期的目的。在确定调查任务和目的时，一方面从研究问题的需要出发；另一方面也应考虑到实际可能性。作一次调查只能解决其主要问题，不可能主次不分，面面俱到。

二、确定调查范围、调查对象和调查单位

当调查任务和目的明确之后，随即要明确调查范围、调查对象和调查单位，即是解决向谁调查，由谁来具体提供资料。

所谓调查范围就是需要调查的那些社会现象的总体，它是由性质上相同的许多调查对象所组成的。调查单位就是组成社会现象总体的那些个体所属单位。例如，全国人口普查的调查范围就是在一国内的全部人口，每一个具体的人就是调查对象。必须回答调查员问题的每个家庭，就是调查单位。又如，某主管工业局需要了解全局所属企业的生产设备使用状况，该局的全部生产设备是调查范围，每一台生产设备就是调查对象。这些生产设备所在的车间或班组，就是调查单位。只有正确地、科学地确定调查范围、调查对象，才能划清所要研究的总体界限，这对于保证调查资料准确地反映实际情况，是很重要的（不重复，不遗漏）。

三、确定调查项目（调查内容）

调查项目是调查任务和目的具体体现。调查项目的选择直接关系到调查资料的价值，必须加以慎重考虑。确定调查项目既要便于调查得到准确、可靠的资料，又要使被调查者不发生顾虑和猜疑，每一个调查项目的含义，内容，以及如何回答都必须详加说明。调查项目可以用问答卷的形式或调查表的形式表现出来。

问答卷简称问卷。即调查者将所要调查了解的项目采用提问的方式让被调查者逐项回答，这种调查要求问卷回答人按规范照答，问题明确，若采用机器汇总资料的方式简便易行。兹举例说明。