

第一章 统计学基础

第一节 总 论

一、统计的含义及其特征

统计一词有三种含义：统计工作、统计资料、统计学。统计工作是人们为了认识、研究客观现象，对其数量特征进行搜集、整理、分析的实践活动过程；统计资料是指通过统计实践活动过程所取得的数字资料 and 与之相关的其他实际资料，是统计工作的成果；统计学是一门以大量客观现象的数量方面为其研究对象的认识方法论科学，它既是统计工作的经验总结和理论概括，又是指导统计活动的原理和方法。

统计活动及其所形成的统计数据都具有客观性、数量性、总体性特征。

客观性是指统计是对客观存在的事物数量特征和数量关系的反映，它所遵循的指导思想是唯物论的反映论。

数量性是指统计的研究对象是客观现象的数量方面。在统计学中，无论是各种现象的规模、速度、结构、质量、水平，还是现象之间的数量关系以及质与量互变的数量界限，都是通过具体的数量来表现的。它是统计最重要的特征。

总体性是指统计研究的不是个体现象的数量方面，而是由大量个体现象构成的总体的数量方面。只有以总体为研究对象，才能消除偶然因素的影响 正确地揭示出事物的本质和规律性。

二、统计的作用和职能

统计的作用可归纳为认识和服务两个方面，其中认识作用是基本作用。人们要改造世界，首先要认识世界。统计在我国经济

建设中发挥着了解国情国力，指导国民经济和社会发展的重要作用。统计的认识和服务作用具体体现为它同时具有信息、咨询和监督三种职能。

统计的信息职能是指系统地搜集、整理和提供大量的以数量描述为基本特征的信息；统计的咨询职能，是指根据掌握的丰富的统计信息资源，经过统计分析，为科学决策和管理提供咨询意见和对策建议；统计的监督职能，是指根据统计调查和分析，从总体上对国民经济和社会运行状况进行全面、系统的定量检查、监测和预警，及时揭示决策运行中的问题和偏差，促使社会经济按照客观规律的要求发展。

统计的三种职能中信息职能是最基本的职能 是统计咨询和监督职能得以发挥的基础。统计的职能以法律的形式在《中华人民共和国统计法》中得到明确的规定：“统计的基本任务是对国民经济和社会发展情况进行统计调查、统计分析，提供统计资料和咨询意见 实行统计监督。”

三、统计的认识过程与工作过程

统计既然是一种认识活动，就必然有一个对客观事物的认识过程。统计的认识过程是从定性认识开始，经过定量认识，再到定量认识与定性认识相结合。

统计的认识过程具体体现为统计设计、统计调查、统计整理和统计分析四个工作阶段。

统计设计是指根据统计研究对象的性质和研究目的，对整个统计工作做出全面的计划和安排。其主要内容有：统计指标和指标体系的设计、统计分类和分组的设计、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等。

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、有计划地搜集原始资料的过程。

统计整理是对搜集来的原始资料进行加工整理，将反映个别

事物的个体特征资料，转化为反映总体数量特征的综合资料的阶段。统计整理既是统计调查的继续，又是统计分析的基础和前提，它在整个统计工作中起着承上启下的作用。

统计分析是运用科学方法对统计资料进行研究，对客观事物得出定量与定性相结合的深刻认识，进而提出建议和进行预测的阶段。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。

第二节 统计指标和统计分组

统计指标和统计分组是统计工作各个阶段都离不开的两个基本要素。所以，在统计设计阶段就应该全面考虑。

一、统计指标的含义

（一）统计指标的概念及其构成要素

统计指标有两种解释：一是指反映总体数量特征的概念或名词，如国内生产总值、人口数等等，它们都是可以量化的概念，我们称之为统计指标的设计形态；二是指反映总体数量特征的概念在一定时间、空间条件下的具体数值，如 2000 年中国国内生产总值 89404 亿元，年末全国总人口 126583 万人，等等，我们称之为统计指标的完成形态。一般教材中提到的统计指标都是后一种解释。

统计指标就其完成形态而言，由以下要素构成：

1. 定性范围——包括指标名称和指标含义。
2. 定量方法——包括计量单位和计算方法。
3. 指标数值——即一定时间、空间条件下的具体数值。

（二）与统计指标有关的几个基本概念

为了正确地理解统计指标，首先必须搞清和统计指标有关的几个基本概念及其相互关系。

1. 统计总体和总体单位

统计指标反映的是统计总体的数量特征。统计总体是由客观

存在的某种性质相同的大量个体所组成的整体，简称总体。总体具有同质性、大量性、差异性三个基本特征。构成总体的个体单位称为总体单位。例如，要研究北京市工业企业的发展情况，那么北京市所有的工业企业是总体，每一个工业企业是总体单位。

2. 标志、变异和变量

标志是说明总体单位特征或属性的名称。通常分为品质标志和数量标志两种，品质标志表明事物属性方面的特征，只能用文字表示，例如人的性别、文化程度；数量标志表明总体单位数量方面的特征，例如人的年龄、工资。数量标志的具体表现值称为标志值。统计指标就是根据总体单位某一数量标志的标志值汇总计算得到的。在统计研究中，品质标志主要是作为统计分组的依据，进而计算出各组的总体单位数。数量标志除作为分组依据计算总体单位数外，还可以进行许多其他计算，如计算单位产品成本、人均劳动生产率等。

在数量标志中，不变的数量标志称为常量或参数；可变的数量标志称为变量。由于变量的函数仍为变量，所以由可变数量标志构成的各种指标也称为变量。变量的具体表现值称为变量值或标志值。变量按其取值的连续性又分为连续变量和离散变量两种。离散变量的取值表现为整数，通常以计数的方法取得，如人口数、企业数等。连续变量的取值是连续不断的，相邻两值之间可以作无限分割，其数值必须用度量或测量的方法取得，如人的身高、体重等。

各概念之间的关系如图 1—1 所示。从中我们可以看到，指标和标志分别对应于总体和总体单位，它们之间的关系是：标志是计算统计指标的基础，统计指标是标志值的综合表现；随着研究目的和任务的变化，指标和数量标志之间存在着变换关系。

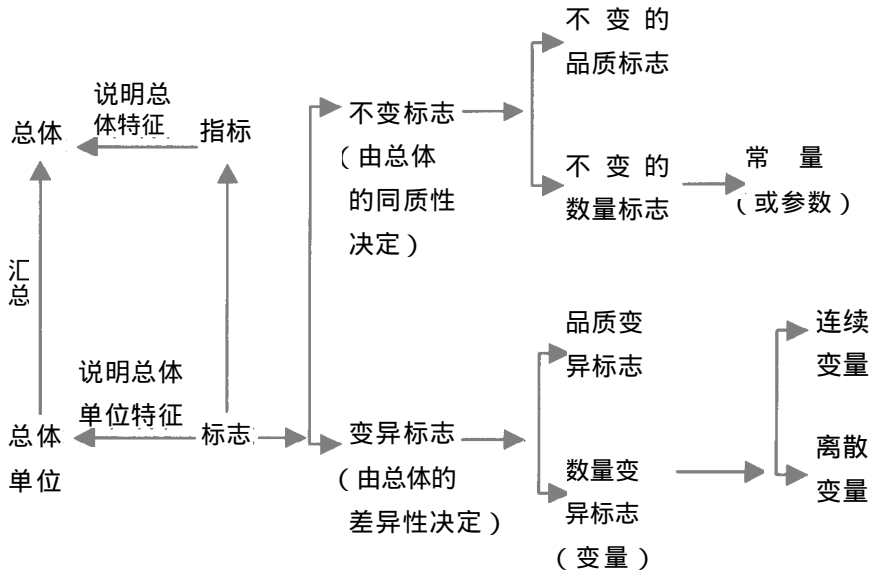


图 1—1 各概念之间的关系

二、统计指标的种类

(一) 统计指标按其反映总体现象的内容特征不同，可分为数量指标和质量指标。数量指标又称外延指标，是反映总体现象规模大小和数量多少的指标，一般表现为绝对数。质量指标又称内涵指标，是说明总体性质、内部结构、比例关系的指标，其数值表现为相对数或平均数。

(二) 统计指标按其具体内容和实际作用不同，可分为总量指标、相对指标和平均指标。

1. 总量指标

总量指标是反映客观现象总体在一定时间、地点条件下的总规模、总水平的综合指标，用绝对数表示，也称绝对指标。例如，一个国家或地区一定时期的人口数、耕地面积、国民生产总值等，都是总量指标。

总量指标是认识总体现象的基本指标，也是计算相对指标和平均指标的基础。相对指标和平均指标一般是由两个有联系的总量指标对比计算出来的，是总量指标的派生指标。

总量指标主要有三种分类：

(1) 总量指标按反映总体内容不同，分为总体单位总量和总体标志总量。总体单位总量反映的是总体单位数的多少；总体标志总量反映的是总体单位某一数量标志值的多少。例如，将北京市统计局作为一个整体，统计局的全体职工总数就是总体单位总量；全部职工的工资总额就是总体标志总量。

(2) 总量指标按反映总体的时间状况不同，分为时期指标和时点指标。时期指标反映统计总体在一段时间内发展的总量，例如产品产量、商品销售额。时点指标反映统计总体在某一时刻状态上的总量，例如人口数、商品库存总额、固定资产原值。时期指标的数值可以连续计算，各时期的数值可以直接相加，数值大小与时期长短成正比。时点指标的数值不能连续计算，其数值大小与时间间隔长短没有直接关系。

(3) 总量指标按采用的计量单位不同，分为实物指标、价值指标和劳动指标。

实物指标是根据事物的自然属性和特点采用实物单位计量的指标。包括自然单位、度量衡单位、复合单位和标准实物单位。例如，人口以“人”为单位、粮食以“吨”为单位、货物周转量以“吨公里”为单位，等等。实物指标的主要特点是能够直接反映现象的具体内容和使用价值，从而具体地表明事物的规模 and 水平。但是该指标在使用上也存在着一定的局限性，即缺乏对不同类产品或商品的综合性能。因此不能用来反映现象的总规模、总水平。

价值指标是以货币单位计算的总量指标，如增加值、销售收入等。价值指标具有广泛的综合性和概括性，可以综合反映客观现象的总规模、总水平。

劳动指标是以劳动时间为单位计算的总量指标，如出勤工日、定额工时。该指标主要在企业范围内使用。

2. 相对指标

相对指标又称统计相对数，它是两个有联系的总量指标的比率，用以反映现象之间的数量对比关系。它以相对数的形式表现

为百分数、千分数、系数、倍数、成数及翻番数等，也可以用复合单位表示，如人口密度用“人/平方公里”表示。

相对指标在实际工作中应用范围广泛，它可以反映社会现象的相对水平、普遍程度、内部结构、比例关系和动态变化程度；可以使一些不能直接对比的现象找到共同的比较基础，从而判断事物之间的差别程度。

相对指标按照研究目的不同，主要分为六种：

(1) 计划完成程度相对指标：用于同一总体中的实际数与计划数对比，表明预定目标的完成情况，通常用百分数表示。

(2) 结构相对指标：用于同一总体中某一部分数值与总体全部数值的对比，反映总体内部的结构状况，一般用百分数表示。

(3) 比例相对指标：用于同一总体中不同部分之间数值的对比，表明总体内部各部分之间的比例关系。

(4) 动态相对指标：用于同一现象不同时期的数值之比，说明同类事物在不同时间上的发展方向和变化程度，一般用百分数表示。

(5) 比较相对指标：用于不同总体中两个性质相同的指标之比，借以说明同类事物在不同空间条件下的数量对比关系。一般用倍数表示。

(6) 强度相对指标：用于不同总体中两个性质不同但有一定联系的总量指标之比，用来说明现象的强度、密度和普遍程度。一般用复合单位表示。

3. 平均指标

平均指标是反映总体各单位某一数量标志一般水平的综合指标，如职工的平均工资、平均年龄。平均指标的特点是把同质总体内某一数量标志的差异抽象化了，是总体各单位标志值的代表值。利用平均指标可以反映总体分布的集中趋势，同时由于该指标消除了总体范围大小的影响，因此，可以对同类现象在不同单位、不同地区间进行比较研究，在实际工作中应用广泛。

平均指标按计算方法不同分为算术平均数、调和平均数、几

何平均数、中位数和众数等几种主要形式。

(1) 算术平均数 (\bar{x})

算术平均数是统计中最基本、最常用的一种平均数。它的基本计算形式是用总体标志总量除以总体单位总数。

在实际工作中,根据未分组资料计算平均指标一般采用简单算术平均数的方法,即将总体各单位标志值简单相加除以总体单位总数。用公式表示:

$$\bar{x} = \frac{\sum x_i}{n}$$

式中 x_i 为总体各单位标志值 ($i=1, 2, 3, \dots, n$), n 为总体单位数。

例如:某车间有 20 个工人,某月工资总额为 40000 元,则该月的平均工资是:

$$\bar{x} = \frac{40000}{20} = 2000(\text{元})$$

根据分组资料计算平均指标就需要用加权算术平均数的方法,即首先计算出各组的总量指标,然后加总计算出总体总量指标,再除以总体单位总数。用公式表示:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f}$$

式中, x_i 代表各组的标志值 ($i=1, 2, 3, \dots, n$), f_i 代表各组的次数,即权数。

例如上述 20 个工人的工资分组资料如表 1—1 所示:

表 1—1 某车间工人工资分组资料

工资(元) x	工人人数(人) f	工资总额(元) xf	各组人数所占比重(%) $f_i/\sum f$
合计	20	37200	100
1000	2	2000	10
1200	6	7200	30
2000	8	16000	40
3000	4	12000	20

$$\bar{x} = \frac{\sum x_i f_i}{\sum f} = \frac{37200}{20} = 1860 (\text{元})$$

计算结果表明该车间工人的平均工资为 1860 元，工人人数是权数，它在这里起权衡轻重的作用，从表 1—1 可以看出：平均工资最接近权数最大的一组。权数也可以用各组单位数占总体单位数的比重来表示，即：

$$\bar{x} = \sum x_i \cdot \frac{f_i}{\sum f} = 1000 \times 10\% + 1200 \times 30\% + 2000 \times 40\% + 3000 \times 20\% = 1860 (\text{元})$$

(2) 调和平均数 H

调和平均数是利用标志值倒数计算的平均数，所以也称倒数平均数。在实际工作中，经常会遇到只有各组标志总量和各单位标志值，缺少总体单位数的情况，这时就必须采用调和平均数的方法计算平均指标。

例如 某人在早市上用 1 元钱买了 1 千克西红柿，即每千克 1 元；又用 1 元钱买了 2 千克黄瓜，即每千克 0.5 元；再用 1 元钱买了 1.5 千克茄子，即每千克 0.67 元，要求计算三种蔬菜的平均价格。

平均价格是用总金额除以总数量，即：

$$H = \frac{1+1+1}{\frac{1}{1} + \frac{1}{0.5} + \frac{1}{0.67}} = \frac{3}{1+2+1.5} = \frac{3}{4.5} = 0.67 (\text{元})$$

归纳得到简单调和平均数的计算公式：

$$H = \frac{n}{\sum \frac{1}{x_i}}$$

加权调和平均数实际上是加权算术平均数的变形。当知道各组的标志总量 ($\sum xf$)，而不知道各组的次数 (f) 时，可以设 $m = \sum xf$ ，则 $f = m/x$ ，代入加权算术平均数的计算公式，即：

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{\sum x \frac{m}{x}}{\sum \frac{m}{x}} = \frac{\sum m}{\sum \frac{m}{x}}$$

由此得到加权调和平均数的计算公式：

$$H = \frac{\sum m}{\sum \frac{m_i}{x_i}}$$

其中， m_i 表示各组的标志总量，在这里作为权数。在具体应用时，如果掌握各单位标志值和相应的次数资料，就采用加权算术平均数公式；如果掌握各单位标志值和各组的标志总量，就采用加权调和平均数公式。

(3) 几何平均数 (G)

几何平均数是用 n 个变量值的连乘积开 n 次方计算的平均数，它适合于计算平均比率和平均速度。其计算公式为：

$$G = \sqrt[n]{x_1 \cdot x_2 \cdots x_n} = \sqrt[n]{\prod x_i}$$

式中： x_i ($i=1, 2, 3, \dots, n$) 代表各组的变量值， n 为变量值的个数； Π 为连乘符号。

(4) 中位数和众数

将总体中各单位标志值按大小顺序排列，居中间位置的那个标志值就是中位数。如果总体单位数是偶数，则居中间的两个标志值的算术平均数是中位数。由于中位数不受数列中极端值的影响，因此在一些情况下能更好地反映现象的一般水平。例如居民收入中位数就不受个别极高收入户的影响，比平均收入更具有代表性。

总体中出现次数最多的标志值叫众数。利用众数作为社会经济现象一般水平的代表在具体工作中有一定的实用性。例如要了解农贸市场某种商品的价格水平，只要观察该种商品多数成交价格即可。

以上介绍了总量指标、相对指标和平均指标，除此之外，还有一种反映总体中各单位标志值差异程度的指标，称为标志变异指标。它综合反映了总体各单位标志值的差异性，可以衡量平均指标代表性的大小。如果说平均指标说明的是分配数列中变量的集中趋势，那么标志变异指标反映的就是变量的离中趋势。将平均指标和标志变异指标结合运用，能更全面地认识总体的数量特

征。标志变异指标主要有：

全距：又称极差，是总体各单位标志值中最大值与最小值之差，用来表示标志值的变动范围；

平均差：总体中各单位标志值与其算术平均数的离差绝对值的算术平均数；

方差：总体中各单位标志值与算术平均数离差平方的算术平均数，用 σ^2 表示；

标准差：方差的平方根，用 σ 表示；

变异系数：用以上各项变异指标分别与该总体的算术平均数对比，形成全距系数、平均差系数、标准差系数。它是用相对数形式表示的标志变异指标。当两个总体平均水平不一致时，一般用变异系数来比较其平均指标的代表性。

变异指标反映的内容和具体用途有特殊性，所以在统计学中将其单独列为一类指标。

三、统计指标体系

一个指标只能表明统计总体某一特征或某一侧面的情况，而在实际工作中需要全面、系统地反映总体各方面的数量特征，这就要求我们必须设计一套科学的统计指标体系。

统计指标体系是由一系列相互联系的指标群所构成的整体，用以说明所研究总体各个方面相互依存、相互制约的关系。

统计指标体系大体上可以分为两大类，即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反映国民经济和社会发展各个组成部分基本情况的指标体系；专题统计指标体系是针对某一个经济或社会问题而制定的专项指标体系。如经济效益指标体系、残疾人状况指标体系等。

统计指标体系是以指标为基本要素构成的整体，因此具有成套性、结构性、适用性的特点。

四、统计分组

（一）统计分组的含义和作用

统计分组是根据统计研究任务的要求和统计总体内在的特点，按照一定的标志将总体所有单位划分为若干个性不同又有联系的组成部分的方法。统计分组是统计整理的主要方法，也是一切统计研究的基础。通过分组可以使同一组内各总体单位的性质相同，不同组的各总体单位性质相异，即统计分组的结果形成组内同质，组间异质。统计分组的作用主要表现在以下三个方面：

1. 划分总体现象的类型。现象之间的差异使之构成不同的类型，不同类型的现象具有不同的特征和变化规律。通过统计分组可以把现象之间的类型加以区别，从而揭示出总体的内在特征。

2. 分析现象内部的结构和比例。统计总体经过分组，被划分为若干个性不同的部分，计算各个组成部分的总量在总体总量中所占的比重以及各个部分之间的比例关系，可以更深入地反映总体内部的结构和数量关系。

3. 分析现象之间的相互依存关系。任何现象都不是孤立存在的，利用统计分组分析和研究现象之间相互联系、相互制约、相互依存的关系，有助于人们全面、深入地认识事物。

（二）统计分组的方法与步骤

掌握统计分组方法，关键在于正确选择分组标志和划分各组界限。

第一，确定统计分组的内容和分组体系。首先要根据研究的目的确定分组的内容。其次要确定一个分组体系：如果是对同一个总体进行多个简单分组，就形成平行分组体系；如果是对同一个总体进行多个复合分组，则形成复合分组体系。

第二，选择分组标志。分组标志的选择应该服务于统计研究的目的和总体的性质、特点。分组标志有品质标志和数量标志之分，前者反映总体单位属性的差异；后者反映总体单位某种数量特征的差异。

第三，划定组间界限。分组标志确定后，就要根据事物的特征严格划清组间界限，做到不重复、不矛盾、不遗漏。

第三节 统计调查

一、统计调查的概念

统计调查就是按照统计设计所确定的统计指标和统计分组，采用科学的方法，有计划地搜集反映总体各个单位有关标志的大量统计资料的过程。

统计调查的基本任务是向调查单位搜集原始资料。原始资料是指从调查单位搜集没有经过任何加工整理的个体资料。统计调查是整个统计工作过程的基础，因此要求必须做到准确、及时、全面、系统。统计调查的效果直接关系到整个统计认识过程的成败。

二、统计调查的种类

（一）按调查对象所包括的范围不同分为全面调查和非全面调查。全面调查就是对被调查对象的全部单位无一遗漏地进行调查，如人口普查。非全面调查则是对调查对象中的一部分单位进行调查，如抽样调查、重点调查。

（二）按调查登记的时间是否连续分为经常性调查和一次性调查。经常性调查要求随着调查对象的发展变化，连续不断地进行调查登记，因此又称连续调查。例如，要了解企业的产品生产、原材料消耗等情况，就必须在观察期内连续登记。一次性调查是间隔一个相当长的时间所作的调查，一般是为了研究总体现象在一定时点上的状态。

（三）按组织方式不同分为统计报表和专门调查。统计报表是按照统一规定的表式要求自上而下的布置自下而上地逐级提供统计资料的一种方式。专门调查是为了研究某些情况或问题专门组织的调查在我国目前的统计工作实践中应用广泛。其主要方式有：

1. 普查：即专门组织的一次性的全面调查，主要用来调查属

于一定时点上或时期内的现象总量。目前，经国务院批准的周期性的普查项目有人口普查、农业普查和经济普查。其中人口普查和农业普查每隔 10 年进行一次，分别在逢 0、6 的年份进行。经济普查每隔 5 年进行一次，在逢 3、8 的年份进行。与其它调查方式比较，普查可以得到更全面、更具体、更详细的资料，但是它需要的调查时间长，花费的人力、物力、财力巨大，因此一般适用于了解有关国情、国力的重要资料。

2. 抽样调查：是按照随机原则从调查总体中抽出一部分单位作为样本进行调查，并根据样本资料推算总体数值的一种调查方式。所谓随机原则是指从总体中抽取样本单位时完全不受主观因素的影响，使总体中的每个单位都有同等被抽中的机会，这样才能保证样本的结构和分布近似于总体，从而使样本对总体具有代表性。根据样本推算总体的数值虽然存在一定的抽样误差，但抽样误差的范围是可以计算并加以控制的，因此对总体的推断结果具有一定的可靠性。

3. 重点调查：是指在所要调查的总体中选择一部分重点单位进行的调查。所谓重点单位是指这些单位虽然数量较少，但其标志总量在总体标志总量中占有较大的比重，能够大致反映总体的基本情况。重点调查的投入少，见效快，既可用于一次性调查，又可用于经常性调查，实用性较强。

4. 典型调查：是根据调查目的和要求，在对所研究总体做全面分析的基础上，有意识地选择少数有代表性的典型单位进行深入、细致调查的一种方法。

目前，我国已基本形成以周期性的普查为基础，以经常性的抽样调查为主体，以必要的统计报表、重点调查、综合分析等为补充的多种调查方法综合运用的统计调查体系。

（四）按搜集资料的方法不同分为直接观察法、访问法、报告法、电话调查法等多种具体方式。

1. 直接观察法：是由调查人员到现场对调查对象直接进行观察或计量以取得资料的一种方式。如对农产品产量进行实割实测

调查。

2. 访问法：是由调查人员携带调查表向被调查者面对面逐项询问，并将答案填入表内的方法。

3. 报告法：是由被调查单位按照调查机构的统一表式和要求填报数据的一种调查方式。我国统计部门长期采用这种方法搜集企、事业单位的统计数据。近年来兴起的由被调查者自填式的问卷调查法也属于报告法。

4. 电话调查法：是由调查者通过电话按问卷内容进行提问，被调查者自愿回答的一种调查方式。

三、统计调查方案的设计

进行统计调查的首要问题是设计调查方案，一个完整的调查方案包括以下内容：

1. 确定调查的目的和任务。即明确在调查中要解决哪些问题，通过调查要取得什么样的资料及这些资料的用途。

2. 明确调查对象、调查单位和报告单位。调查对象是在确定了调查目的、任务之后所要调查的范围，即调查总体；调查单位就是组成调查总体的个体，它是调查中各个调查项目的承担者；报告单位也叫填报单位，是负责向统计调查机构提交调查资料的单位。

3. 制定调查项目和调查表。调查项目是指对调查单位所要调查的主要内容，即调查单位的各个标志的名称。调查项目一般采用调查表（或调查问卷）的形式。调查表是按一定的结构和顺序排列制成的表格。

4. 选择调查的方式和方法。根据调查任务的性质、要求期限、经费等具体情况，决定采用哪一种调查方式。

5. 规定调查的时间和组织计划。如果调查的是时期现象要规定出提取调查资料的起止时间；如果是时点现象要明确规定统一的标准调查时点。组织计划即确保调查实施的工作计划，包括调查机构的设置、人员的配备和培训、经费来源、工作步骤等内容。

第四节 统计整理

一、统计整理的概念和意义

统计整理是根据统计研究的目的，把统计调查所搜集到的原始资料进行科学的加工，使之系统化、条理化、科学化，成为能够反映事物总体特征的综合资料的过程。统计整理既是统计调查的继续，又是统计分析的基础和前提，在整个统计工作中起着承上启下的作用。统计整理的结果能否如实地反映客观情况，决定着统计资料的科学价值，也直接影响到统计分析的准确性和真实性。

随着信息技术的发展，计算机和互联网等现代化传输方式得到广泛应用，统计整理的内容也随之发生改变。具体包括两个方面：一是统计数据的管理，即分组、汇总、制表；二是统计数据的传输、贮存、更新、输出。

二、统计整理的步骤

（一）设计统计资料整理方案

首先，要根据统计研究的目的把已经确定的统计指标体系、统计分组体系设计到整理表或汇总表中，并且要规定出整理或汇总的具体方法。其次，根据搜集的原始资料和整理表的工作量，制定出具体的可行的工作计划。

（二）对原始资料进行审核

对原始资料进行审核的内容主要是资料的完整性和准确性。对资料完整性的审核主要是看资料是否系统周密、合乎逻辑，调查项目是否齐全，调查单位是否有重复和遗漏。对资料准确性的审核包括逻辑检查和计算检查两个方面，逻辑检查主要审核原始资料内容的合理性，有无矛盾和不符合实际的地方。计算检查是看各项指标的计算方法是否恰当，指标间的平衡关系是否成立，计算单位是否得当等。

（三）对原始资料进行统计分组和统计汇总

使用计算机处理需要先将原始资料的数据通过录入设备记载到磁介质上，然后进行分组和汇总，计算出各组的单位数和总体单位数及各综合指标的数值。

运用计算机对采集的数据汇总主要有以下方式：

1. 逐级汇总，即由区县（直报综合单位）、乡（镇、街道）级统计部门对采集的数据进行汇总上报给上一级统计部门；上一级统计部门只对上报的已汇总的综合数据进行再加工的数据处理方式。

2. 超级汇总，即由上一级统计部门直接对报告单位上报的统计数据进行审核、汇总加工的数据处理方式。

三、次数分布

（一）次数分布的概念

次数分布是将总体中的所有单位按某个标志分组后所形成的总体单位数在各组的分布。分布在各组的总体单位数叫做次数或频数。各组次数与总次数之比叫做比重或频率，各组的频率之和等于 1 或 100%。

次数分布是统计整理的一种重要形式，也是统计描述和统计分析的重要方法。它实际上是表明统计总体中的各总体单位在各组间的分布状态和分布特征的一个数列，因此也称为次数分布数列，简称分布数列。

根据分组标志特征的不同，分布数列可以分为品质分布数列和变量分布数列。

（二）变量数列的分类

变量分布数列又有单项变量数列和组距变量数列之分。

1. 单项变量数列是以一个变量值为一组所形成的次数分布数列（如表 1—2）。

单项变量数列一般在变量值不多且变动范围不大，变量呈离散型分布的情况下编制。