

社会经济调查 方法与实务

■ 杜子芳 编著

国家统计局培训学院

社会经济调查方法与实务

杜子芳 著

国家统计局培训学院

[illegible]

作者在十余年教学与实践基础上,依据对学科的长期研究理解,对课程内容进行了重新界定,最重要的一点将统计分析的内容剔除在外,是调查与统计学区隔开来,避免课程内容重复,减轻学生的学习负担。同时对课程的章节结构也进行了较大调整,本书的结构安排主要遵循这样的思路:

(1) 社会经济调查只是调查的一个具体种类和领域, 因此对调查的基本概念、问题、思想和原理的介绍必须置于社会经济调查内容的展开之前。

(2) 调查的步骤之间存在某种固有的顺序, 先有知识方法, 后有实践应用; 有了方案和问卷才有可能去抽样并进行实地调查; 二手调查早于一手调查, 相关内容的章节安排应体现这一点。

(3) 并列的各种方式方法也有远近亲疏, 性质接近的理应安排在相邻近的章节。

(4) 方法与实务密不可分, 但方法总归更重要, 应视为设置课程的首要目标, 所以既坚持以方法为框架安排讲解的内容, 又务求使其水乳交融。

本书的读者对象设定为在校本科和大专学生, 调查机构的从业人员, 因此本书既可作为高等院校有关专业的教材, 同时可以作为调查机构的培训用书。本书在教学时并不要求以前修过统计学的课程。

在本书写作过程中, 许多人士都给予了作者热诚的帮助, 纪宏、赵彦云和贾俊平先生直接参与了大纲的讨论和制订, 王健女士对本书内容提出很多有益的意见, 并对全稿进行了校对。誊写、打字、帮助寻找资料等烦琐工作有赖我的妻子和家人的不辞辛苦。在此我要对他们谨表谢忱。至于一如既往给予多方支持帮助的国家统计局教育中心的石占前、丘京南、毛盛勇、霍丙森等诸位也是我要表示深切谢意的。

最后, 我要郑重说明的是, 作者虽然十分努力, 但毕竟水平有限, 诸多不妥乃至错误, 恐仍难避免, 敬请同行和读者谅解并不吝赐教。

杜子芳

2003年9月10日

“调查”一词有两重涵义：调查活动和调查方法。两者的联系为凡是调查活动必然使用调查方法，任何使用调查方法的行为都是调查活动。本书在使用“调查”一词时除非为了避免混淆，否则不予区分，读者只要留心当不致造成严重不便。

调查的应用范围十分广泛。事实证明,一方面政策制定、投资决策、科学研究、机构管理甚至司法实践中都贯穿着或多或少的调查活动。另一方面,调查内容也出现在许多学科中,如社会学、经济学、医学、农学、林学、管理学、生物学、统计学、教育学和工程学等。

调查的重要性是获得公认的。从孙子兵法“知己知彼，百战不殆”，毛泽东“没有调查研究，没有发言权”，到“实事求是”、“实践是检验真理的一切标准”等至理名言中，都直接或间接地肯定了调查的重要性。

但尽管如此，人们对调查的性质仍存在不少误解和争议。一些人认为调查既不是一门独立的科学，也不是什么系统的方法，最多只是从属于其他学科的一种工具性技术，例如在统计学的著作里，大多明说或暗示调查仅仅是整个统计过程的一个环节之一。另一些人则将调查看成“调查研究”的简称、缩写或“调研”的别称，因此，社会调查就是社会调研，市场调查就是市场调研，民意调查就是民意研究。前者是对调查的贬低，后者是对调查的拔高，其后果将

引起人们的困惑——调查究竟是什么？

第一节 调查的涵义与特点

要直接回答这一问题并不困难，但说服力可能不够。为此，使用排除法即通过说明“调查不是什么”来间接得到答案也许是明智的选择。首先，调查不是数学、统计学或类似于数学、统计学的数或数据的分析理论和方法，调查只能得到结果，不能得到结论。事实上调查只是科学研究或日常工作中社会分工体系里的一个环节，别外的学科或方法则构成与调查形成有机联系的另一些环节，只有它们的巧妙配合，才能终成正果。其次，调查并不涉及规范性分析的内容，即不存在任何价值观的判断。在传统的“立场、观点、方法”论中，调查本身不涉及立场，排斥观点，谨守方法，象其他数理学学科一样，是纯粹的方法论科学。换言之，调查本身只关心真相，不关心对错。

那么，调查到底是什么？遗憾的是迄今为止，许多调查学者对此仍然莫衷一是。从学科分类标准和实际应用背景看，将调查定义为“一门关于搜集资料的方法论科学”是合适的。但即便在此意义上，调查的概念仍有广义与狭义之分。广义的调查是指任何有意地搜集信息的活动。而狭义的调查是指任何有意且系统地搜集信息的活动。“有意且系统”是指，调查乃是根据特定目标、按照科学方法、规范程序对有关总体的信息进行系统的搜集、整理活动。本书要研究的正是调查活动需要使用的科学方法和必须遵守的规范程序。

与广义调查中的非狭义调查相比，狭义调查具有三个特点：

（一）总体性

人们所最常见的一类广义调查是罪案调查和以社会热点或专业热点为报道对象的新闻调查。这些调查当然是有意的，说这些调查不系统似乎也与人们的日常印象不符，须知在法学领域中的刑侦学、法医学可是很早就拥有科学殿堂的正式席位；至于一个新闻调查节目往往也是各路人马巧思妙想协同作战的结果，与狭义的调查区别何在？问题在于，狭义调查的最终目的常常是直接指向“总体”的属性或特征，一家数字电视厂商委托一家调查公司对北京 1000 户居民进行的购买意向调查；一开始目的就是据此推算整个北京可能购买数字电视的人数或偏爱某个功能的人数比例；而上述那类广义调查的最终目的并不指向总体及其属性特征。例如南京建业区侦破一个抢劫杀人案件的直接目的仅仅是破案罢了，如果说其目的是为了增加全国的破案率则未免牵强，尽管客观

结果确实会提高全国的破案率。

狭义调查的总体性特点还有助于解释人们对案例调查认识上的变化，早期人们是将案例调查与所谓统计调查对照并列的，而今随着标准化的普遍采用，许多学者已经将案例调查也归于统计调查的范畴。实际上，因为统计学也以总体属性和特征为研究对象，故狭义调查在很多场合被称为统计调查。有关调查与统计的关系我们后面会详细讨论，在此应指出的是，调查的主要方法与程序皆因总体性而起。

（二）科学性

狭义调查的科学性表现在两个方面：一是对调查的代表性误差能做出相当可靠的估计，或者能够事先给予相当可靠的控制；二是许多方法凝结着许多先哲的知识与经验。因此调查结果的可信性极高，这些都是普通调查所不及的。

（三）规范性

狭义调查与非狭义调查的另一区别在于狭义调查存在着相当固定的程序，这套程序在科学界已经得到公认。对于正规的、严肃的调查来说，这种程序就象某些产品的质量认证和现今应聘者的学历证明一样，也许是不完善的，却非常的有效。规范性保证了调查结果的可检验性，从而为调查结果的获得承认和实际应用奠定基础。

不仅调查本身的含义有这样那样的理解，调查的同义词也很多，例如使用广泛的民意调查也被广泛地称为民意测验，难道智商测验不是关于智力的一种典型的调查行为吗？而在许多理工学科例如工程专业里，工程测量代替了工程调查的称谓；此外测试、度量、实验等等在相当多的场合里也是调查一词的代用品，然而这些代名词常常表示更精致、更准确的调查。

其实就调查本身只关心真相与本色的特点看，测量也许比调查本身更能清楚地反映出调查的本质，更能准确地刻划调查客体的特征。换言之，调查本质上是一种测量活动。

什么是测量？测量乃是主体依据某种法则用数字或符号表示客体的属性或特征。测量具有 6 个要素：测量主体、测量客体、测量对象、测量法则、测量工具、测量结果（数字/符号）。测量活动就是测量主体针对测量对象，依据测量法则使用测量工具向测量客体发射信号，获得数字或符号表达信息（即测量结果）的过程。

测量主体，即测量活动的发起者或从事者。

测量客体，是调查主体要用数字或符号来进行表达、解释和说明的属性或特征的载体。比如测量一个人的身高时，这个人就是测量客体。

测量对象，又称测量内容，即测量客体的某种属性或特征。实际上，在任
何一种测量中，我们所针对的虽然是某一客体，但所测量的对象却并非客体本
身，而是这一客体的属性或特征。比如桌子是我们的测量客体，而桌子本身我
们却无法测量，只有桌子的各种特征，比如它的高度、宽度、重量、颜色等等，
才能构成我们测量的内容。

测量法则，即进行测量时数字和符号与客体各种属性或特征内容之间的对
应规则，是用来区分客体不同特征或属性的操作程序和操作规则。比如“将桌
子放置在水平的地面，然后用直尺从地面垂直地靠近桌面边缘，桌面所对应的
直尺的刻度是桌子高度”，这句话所陈述的就是测量桌子高度的一种规则。

测量工具，测量主体根据测量法则用来对测量对象进行分配数字或符号的
工具，如衡量身高所使用的公制米尺，了解个人情况的问卷，工程测量的仪器、
仪表等等。

测量结果（数字或符号），即使用测量工具根据测量法则对测量对象所分
配的数字或符号的记录。比如 120cm、110cm 等就是使用直尺按照“将桌子放
置在水平的地面，然后用直尺从地面垂直地靠近桌面边缘，桌面所对应的直尺
的刻度是桌子高度”的测量法则测量桌子高度所得的结果。

根据取得测量结果的方法不同，可把测量分为：

（一）直接测量

把测量对象与作为测量标准的量直接进行比对，或用预先按标准校准的测
量仪器和预先确定的对应法则对测量对象进行测量，并能直接得到测量对象数
量大小的测量结果。

直接测量可用一般公式表示：

$$y = x$$

式中 y ---测量对象的数量大小或受测量；

x ---测量值

对诸如物体重量、长度，企业职工数目、工资总额，社区家庭数量、植树
数量以及工程所进行的测量，对时间、压力、温度、湿度、质量使用仪表进行
的测量都是直接测量。

（二）间接测量

受测量不能通过直接测量的方法得到，必须通过一个或多个另外的直接测
量值，利用一定的函数关系运算才能得到，此种测量称为间接测量。间接测量
可以用下面的一般公式来表示：

$$y = f(x_1, x_2, \dots, x_n)$$

式中 y --- 测量对象的数量大小或受测量;

$x_1, x_2, \dots, x_n$ ---- 各直接测量值

间接测量在科学研究中应用最为广泛, 因为大多数情况下, 或是直接测量方法不能奏效; 或是直接测量的效果不如间接测量精确, 过程不如间接测量方便。

(三) 综合测量

受测量不能通过直接测量或间接测量得到, 而必须通过一系列直接或间接测量值建立联立方程组, 然后通过解联立方程组的办法, 才能得到最后的测量结果。

第二节 调查的作用

调查服务于研究, 研究的内容无外乎提出理论假说或对这种理论假说进行实证检验, 经过实证检验成立的理论假说上升为确定理论, 成为决策依据或行动指南; 不成立的理论假说则被推翻。

提出或形成理论假说的途径主要有:

(一) 归纳推理法

将局部的观察结果和事实材料经过归纳, 形成具有普遍意义的假设。这是形成假设的最主要的方法。如根据自己在生活中的观察以及从报刊等新闻媒介上得到的信息, 对于“影响人们生育意愿的因素”这一研究课题就可以形成诸如文化水平越高, 生育意愿越低; 传统观念越浓厚, 生育意愿越高等假设。

(二) 逻辑推理法

根据以往的理论或事物发展的必然规律对事物的发展变化趋势做出某种推测的方法。如独生子女家庭的普遍化会导致若干年后小学生源的不足; 社会风气不好必然导致青少年犯罪率上升等等。

(三) 类比推理法

面对陌生的社会现象, 拿所熟悉的事物、过程、规律做出假设。如过度城市化会导致犯罪率的上升; 随着工业化的发展, 人口将逐渐向城市集中等等。

理论假说作为对某种现象的推测性说明,总是通过两个或者两个以上的概念之间的关系来表达的。其类型主要有:

(一)一元假设(x)。指某一个事实或现象的单义判断。如大部分独生子女受到家长的娇惯。

(二)二元假设(x, y)。指某一事实或现象与另一事实或现象之间关系的判断。如“城市人口密度越高,城市中的犯罪率也就会越高”,“生活水平的提高与生育率的下降有密切关系”。

(三)多元假设(x, y, z)。指某一事实或现象与另外两个或更多的事实或现象之间关系的判断。如,缺乏社会整合性(x)会导致精神沮丧(z),而精神沮丧的人会有行为偏差(y);又如,青少年犯罪原因(x),主要有家庭教育不当(y_1),社区环境不良(y_2),结交了不良伙伴(y_3)等等。

假说的表达方式,既可用简明的文字语言表达,也可用数学语言、图表来表达。通常的方式有:

(一)函数式。“ y 是 x 的函数”,即 y 与 x 成正向(或反向)变化关系。如:个人的理想子女数目会随着其教育程度的提高而递减。

(二)条件式。“如果有 x ,则有 y ”。如:不良家庭的子女具有较高的犯罪率。

(三)差异式。“如果 x 不同,则 y 不同”。如:不同职业的人,其时间的分配结构亦不同。

调查的作用在提出理论假说与实证检验两个环节上都有体现,但在第一个环节里起作用的只是一部分调查结果---提供了必要的然而未必完整可靠的信息;调查在实证检验这个环节里的作用最为人关注,这种作用往往被归纳为三个方面:概念化、操作化与实施化。我们在下面将要对概念化与操作化进行详细的讨论。

构成调查或测量对象的是测量客体的某些属性或特征,这些属性或特征有可见与不可见之分,可见属性或特征具有直觉、直观、简明,可以直接测量的

特点。不可见属性或特征具有抽象、复杂，只能间接测量——通过对另外的量进行直接测量实现目标的特点。也就是说欲对不可见特征进行测量，必须首先经过某种途径，得到不可见特征的某个或某些可见的转化态，然后对这些与可见特征具有类似显明直观特点的转化态进行直接测量，并把测量结果当作不可见特征的反映与代表。例如，儿童的智力发育水平是全社会关心的，但智力是一个不可见的特征，没有办法直接进行测量。通常的做法是，设置某种大多数专家都认可的智力测定标准即所谓智商测验表，将因此得到的智商分数当作智力水平的转化态，从而对智力水平的测量就转化为对智商分数测定。某个儿童的智商分数就被视为其智力水平的反映，有兴趣的读者不妨查阅一份实际用过的智商测定表，并认真体会其中的奥妙。

所谓概念化是指，对每一个测量对象（在这里系专指测量客体的不可见特征，因为可见特征已经满足概念化的标准了，同时又必须是理论假说中某些更高层次概念所涉及的不可见特征），从概念上进行具体的定义。众所周知，我们在日常的语言交流中，交流双方对同一概念的含义一般只有大致相同的理解，其界限往往含糊不清。一般来说，这并不妨碍人们的日常交流，但从科学的角度看问题，就不能允许对某个概念抱“差不多”的无所谓态度，而必须清晰地划定其外延，准确地阐明其内涵。

兹举例，假设测量对象包括“居民生活水平”这一项内容。在人们的日常生活中，提到生活水平，人们自然马上就想到日子过得怎么样，吃的、穿的、用的、住的如何，而对于究竟其确切含义是什么并不深究，当然也无此必要。但作为一个具体的测量对象，必须阐明表征该测量对象的概念在此到底指的是什么。如“居民家庭生活水平”作为某种调查中的测量对象，可以这样定义：“测量居民家庭中人均实际消费和占有的社会产品与劳务的数量与质量”。

任何调查要达到其目的，首先要依靠精确描述总体、概念、定义及其分类的合理系统方法，因此要圆满完成调查的重要前提是有一个严谨简明的概念体系。

经过概念化的步骤之后，不可见特征取得了更接近转化态的中间态，更具体更明确了，但中间态还不是直观、简明的转化态，还不能作为直接测量的对象。仍以居民生活水平问题为例，尽管已经像上面那样进行了明确的定义，但在实际测量中，如家计调查时，我们不能问“你家平均每人实际消费和占有的社会产品和劳务是多少，质量又如何？”不用说这样的问题是任何一个居民都难于回答的。这充分说明，概念性的定义依然不能成为直接的测量对象，在调查时直接用于调查询问，因为它还不是一个可测定的量，不具操作性。

概念化后还要有一个操作化步骤。

所谓操作化，是使不可见特征从已经取得概念性定义的中间态，通过赋予操作性定义进一步形成在实践中具有经验层次上之可测量性的转化态，完成从不可到可见的转化，最终成为直接测量对象的过程。

操作化是概念化的延伸与发展，它将概念性定义具体化为可测量的变量，如某个简单指标、调查表或问卷中的指标或项目，比如对于“居民家庭户生活水平”在上面的定义之后，操作化就是进一步规定采用家庭户内的人均收入、人均生活费用支出、耐用消费品拥有量、人均住房面积等一系列指标来衡量，并对这些指标层层深入分解，直至化为实际调查时可直接测量的指标。如规定

$$\text{户内人均居住面积} = \text{户内总居住面积} / \text{户内人口数}$$

调查的实践性主要表现在操作化与实施化上，因此调查的主要内容也可概括为两个方面：一是选择合适的指标反映特定的科学概念；二是通过合适的方法获得这些相关指标的数值。

实施化与操作化联系紧密，但任务迥然不同。操作化仅仅解决了实施的可能性，将从理论假设分解而来的一组概念化作一系列简单指标或可直接测量调查的问题或项目，但是怎样获得这些问题的答案、项目的内容，例如如何抽样、如何组织实地测量或调查等等，却是实施化的任务。实施化乃是本书的重点所在，从第三章到第八章的内容都是实施化的展开，在此不赘。

第三节 调查结果好坏的衡量标准

调查既然是一种测量行为，其误差存在的必然性已被一切从事科学实验者所公认，并被表述为所谓误差公理：

任何测量结果都可能具有误差，而且误差始终存在于一切科学试验和测量过程中。

误差来源于许多方面和环节。

测量主体误差：测量者本身的能力（如表达力、分辨力）和态度造成的测量误差。

测量客体误差：选择出来作为测量客体的个体不能很好代表总体的情况，受访问的人不配合等。

测量工具误差：测量工具本身存在缺陷，如在问卷中一些项目设计不合理；

测量身高的尺子刻度不清等。

测量法则误差：程序不够完善，工具选择不当。

测量环境误差：测量时测量法则要求的条件受周围因素的影响无法得到严格保证。

上述来源的误差具有不同的性质和特点，例如有的可以通过人为努力得以避免或消除，有的则无法避免或消除。根据测量误差的性质和特点，误差可分为系统误差、随机误差和粗大误差。

（一）系统误差：在偏离规定的严格条件下多次测量同一属性或特征量时，误差的绝对值和符号保持恒定；或在测量条件改变时，误差按某一确定的规律变化的。系统误差主要来源测量法则和工具。

所谓确定的规律意思是，这种误差可以归结为某一个因素或某几个因素的函数，这种函数一般可以用解析公式、曲线或表格来表达。例如：钢尺长是温度的函数，某些电量是频率的函数等等。

（二）随机误差：在实际测量条件，多次测量同一属性或特征量时，误差的绝对值和符号的变化，时大时小，时正时负，没有确定的规律，但测量次数增多时，其平均值的误差趋于 0。随机误差是由多种对测量值影响微小，彼此互不相干的因素引起的，测量现场的各种环境条件如温度高低、燥声干扰如（访问时的电话铃声）、情绪波动等的无规律微小变化都可能产生测量误差，因此很难控制，也不能修正，但具有统计规律，可以估计。随机误差主要来源于环境因素及其对测量要素的间接影响。

由于随机误差在各项测量中的单个无规性，导致了众多随机误差之和有正负相消的机会，随着测量次数的增加，则随机误差的个数也增加，而随机误差平均值愈来愈小并以零为极限，因此，多次测量的平均值的随机误差比单个测量值的随机误差小，这种性质通常称之为抵偿性。抵偿性只发生在同一测量过程中产生的许多随机误差中，也称为本次随机误差。

（三）粗大误差：超出在规定条件下预期的误差，如测量时标记错了，将 3 记录为 4，将 7 记录为 1。粗大误差主要来源于测量主体，但在以人为客体的测量里，测量客体也是可能的来源。理想的测量结果不应该包含粗大误差，即所有的异常只有系统误差和随机误差两类，尽管这在非工程性测量时极难达到，但应尽量加以避免。

调查结果的好坏一般用所谓调查精度来表示。调查 A 的精度定义为误差 δ 的相反数：

$$A = -\delta$$

那么误差 δ 是什么呢？

广泛采用和承认的误差形式为

$$\delta = x - \mu$$

式中 δ ——绝对误差；

x ——测量值；

μ ——真实值。

即 误差=测量值-真实值

表示误差大小的指标一般有两类：以误差本身的量纲作为计量尺度的称绝对误差，如上式所示；以指定的真实值或其他量纲相同的值作为计量尺度的称相对误差。

绝对误差指标的不足之处是不能确切地反映测量的精确程度。为了弥补绝对误差的不足，提出了相对误差的概念。相对误差又叫相对真误差，它是绝对误差与真值的比值，通常用百分数来表示。若用 r 表示相对误差，则

$$r = \frac{\delta}{\mu}。$$

由于真实值一般无法确知（如果知道了，测量或调查本身就没有意义了），故不论绝对误差还是相对误差都只有理论上的意义，可以帮助我们把握正确的思考方向，并无多少实用价值。在实际场合，我们是根据误差来源，将精度分解为信度与效度两个指标加以度量。为了进一步对精度与误差展开讨论，对以下几个基本概念予以介绍是必要的。

（一）真实值 是指测量对象在一定条件下客观存在的量值。根据误差公理，除非该真实值本身是主观给定的（例如家庭人口数）或碰巧的场合，否则要想测量到真实值几乎是绝对不可能的。实际应用场合，往往把下列三种量值定义为真实值。

1. 理论真值：也称理论值或定义值，即根据一定的理论，在严格确定的条件下，按定义确定的数值。例如根据几何学三角形的内角和为 180° ；根据热力学，温度的最低点为绝对零度；绝对平等表现为洛伦兹曲线化为 45° 角的直线。但在目前科学技术水平下，这些值的准确测量仍是难题，不能消除误差，

然而这些都是客观存在的。

2. 约定真值：国际计量大会定义了长度（米）、质量（千克）、时间（秒）、电流强度（安培）、热力学温度（K）、物质的量（摩尔）、发光强度（坎德拉等七个基本单位）和平面角（弧度）、立体角（球面度）等作为全世界统一的约定真值。例如一秒是铯¹³³（ Cs^{133} ）原子基态的两个超精细能级之间跃迁的辐射在真空中的 9192631770 倍的持续时间。北京时间也是约定真值，我们平时的钟表所表示的时间实际上都不是真正的北京时间。

3. 相对真值 由于约定真值的获得十分困难，而在实际测量中对测量结果的精度要求也不需那么高，因此在满足实际需要的前提下，相对于实际测量所考虑的精度，其测量误差可以忽略的测量结果，称为相对真值或传递真值。例如我们平时使用的尺子、秒表等所反映的长度、时间就是一种相对真值。

（二）测量值 是指测量结果的数字或符号表示。由于测量误差的普遍存在，所有测量值都是测量对象真值的近似。测量值可用下式表示：

$$\text{测量值} = \text{真实值} + \text{误差}$$

因此为了得到真实值，关键在如何处理误差。一般的做法是，对于普通测量，直接用测量值作为测量结果；而对于高精度的测量，则按照误差理论等知识对测量值进行加工处理作为测量结果。

1. 单次测量值（ x ）

若对测量结果的精度要求不高或有足够的把握，经过一次测量所得测量值，能够满足对测量结果的精度要求时，就用单次测量值 x 来近似地表示受测量的真值。

2. 算术平均值（ $\bar{x}$ ）

在单次测量不能满足实际需要的测量精度情况下，为了能达到尽可能高的测量精度，必须经过多次测量，在相同测量条件下，取有限 n 次测量值 $x_1, x_2, \dots, x_n$ ，用这些测量值来估计受测量的真实值 μ 。在对测量数据进行处理中，得到普遍应用的是取所有测量值的算术平均值 $\bar{x}$ ，来代替不能测得的受测量真值 μ 。

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

数理统计理论已经证明，算术平均值 $\bar{x}$ 是受测量的真实值 μ 的最佳估计。

可以证明,对测量的算术平均值 $\bar{x}$ 的数学期望就是受测量的真实值 μ , 即 $\bar{x}$ 是 μ 的一个无偏估计值。

$$E(\bar{x}) = E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} E\left(\sum_{i=1}^n x_i\right)$$

比较严格地讲,只有当测量次数 n 为无穷大时, $\bar{x}$ 才会依概率收敛于数学期望 μ 。当 n 为有限时, $\bar{x}$ 本身也是一个随机变量。即 $\bar{x} \neq \mu$, $\bar{x}$ 在 μ 的左右摆动,只有当 n 相当大时 $\bar{x}$ 才趋于稳定,随着测量次数 n 的增大, $\bar{x}$ 越接近真值 μ 。

用算术平均值 $\bar{x}$ 作为测量结果有如下优缺点:

优点:

- (1) 计算简便,容易理解;
- (2) 容易用代数的办法对测量值进行处理;
- (3) 测量值 x_i 对真值 μ 所产生的不同方向的偏离,经过取算术平均值,有一定的相互抵消作用;
- (4) 随着测量次数 n 的增加, $\bar{x}$ 则趋于稳定,概率论中大数定理可证明此点。

缺点:

- (1) 算术平均值 $\bar{x}$ 受测量值的特大值或特小值的影响较大;
- (2) 有时算术平均值不是表示受测量的实际表示值。例如,必须用整数表示的量,求平均值后可能得到小数;
- (3) 依品次,程度或等级分类的系列,不能用求算术平均值的办法表示测量结果;
- (4) 对所研究的对象已明确数列的两端为不定时,无法求到算术平均值。