

## 青年研究方法

### 青年研究的方法论

青年研究需要有科学的理论做指导。所谓科学的理论，主要包括两个层次：一是科学的世界观和方法论，二是与青年研究具体内容相关的具体的社会科学理论。马克思主义的辩证唯物主义和历史唯物主义是科学的世界观和方法论，具有普遍的指导意义，因此它应该作为青年研究的根本指导思想和方法论基础。

一、历史唯物主义的基本精神即把客观存在的社会事实当作第一性的东西，要客观地、老老实实地去认识它，反映它。要尊重事实、实事求是。这就要求研究者对所研究的事物采取客观的态度。深入实际，调查研究，尽量全面地获取第一手材料。坚持实事求是的观点，就必须克服形形色色的主观主义的思维方法，不能为了迎合上级领导机关或某些权威人士的意图，去任意剪裁或歪曲客观事实，不能为书本上已有的结论或框子所禁锢，而不尊重活生生的现实，不能为多数人不符客观规律的看法所左右。因此，青年研究中第一条重要的原则即尊重事实、客观、全面地收集资料，实事求是地引出应有的结论，这是必须牢固树立的指导思想。

二、社会不仅是客观存在的，而且是普遍联系的，不断发展的、充满矛盾的。用马克思主义的辩证唯物主义指导青年研究，也就是用动态的、发展的、联系的和矛盾运动的观点来观察和分析研究中所涉及的事实和材料，而不是用孤立的静止的和否定矛盾运动的观点看事物。

社会生活中的各个领域、各种现象都是密切联系在一起的。要正确认识社会现象，就必须坚持联系的观点，克服孤立的观点。研究青年和青年问题，任何时候都不能脱离他们所处的社会历史条件和环境，必须把他们置于一个大的社会系统中，着重揭示他们与社会其它现象和其它因素之间的内在、本质、必然的联系。

社会生活中的一切事物和现象都是不断运动、变化、发展的，都是做为有生有灭的过程而存在的，因此，在青年研究中，必须坚持运动、变化、发展的观点，克服静止、停滞、僵化的观点。这就是说，研究青年问题，不仅要研究现状，且要研究历史，研究发展趋势，研究新问题的产生、形成和

发展。

社会生活中充满着各种各样的矛盾，要正确认识社会现象，必须适用矛盾分析方法，即在青年研究中，对任何事物都要一分为二。即要看到肯定的一面，又要看到否定的一面，即要看到其对立，又要看到它们的统一，不可片面地肯定、否定一切。

联系的观点、发展的观点、矛盾的观点是唯物辩证法最基本的观点。又是指导我们进行青年研究的科学方法论。

### 青年研究方法的特点

一、综合性。青年研究离不开大量的具体材料，这些材料的收集与分析需要综合性的研究方法，由于研究课题的广泛，要求青年研究的方法多种多样，因此它不仅使用文献法、观察法、问卷法、访问法等不同的调查方法，而且要采用典型调查、重点调查、抽样调查、个案调查等多种调查类型。同时它还要采取统计分析等现代科学技术方法及电子计算机技术等新的方法。

二、社会调查性。青年研究中搜集资料的主要方法即社会调查法，要想研究问题、解决问题，必须向现实了解情况，必须做深入、广泛的社会调查，没有社会调查，也就无所谓青年研究。

三、技术性。社会调查是青年研究的主要方法，社会调查的主要问题不是一些抽象的理论知识而是一系列具体方法和操作技术，例如，各种调查方法和研究方法，以及抽样的技术、设计问卷和表格的技术、观察和访问的技术，整理资料的技术，使用电子计算机的技术等，因此技术性成了青年研究方法的一大特点。

### 青年研究的一般程序

- 一、选择课题与提出研究假设。
- 二、设计研究方案。
- 三、搜集资料。
- 四、整理分析资料。
- 五、解释结果及至验证假设。
- 六、撰写报告。

### 青年研究的应用研究和纯理论研究

青年研究可大致上分为应用的或纯理论的两类。应用研究即指直接应用于解决现实中的青年问

题的研究,它涉及青年研究的广阔领域,诸如青年的升学、就业、婚姻、住房、消费等现实问题,应用研究的主要目的在于提出对策解决问题,所以研究课题的选择更多地取决于现实的需要而非研究人员的学术兴趣。应用研究往往需要采用社会调查法,深入实际、搜集大量的资料,一般规模较大、耗资较高。

纯理论研究(有时又称“基础”研究)即指在青年研究的某一领域进行纯粹的理论性研究,探讨那些在学术上已经向研究人员提出了挑战,但在实际应用上不论现在或将来都可能没有、也可能没有什么用处的问题,目的在于发展理论和检验理论及假设,进行纯理论研究,研究者必须充分研究这一领域中已知的概念和理论,一般情况下,许多纯理论研究必须在以前研究的基础上,在统一的概念体系中进行,否则研究是不可能的,在这个意义上说,从事纯理论的研究实际上是选择了理论发展链条上的某个环节。

纯理论的研究与应用研究并不相互排斥,纯理论的研究并不一定不能应用,反之亦然。青年研究的最终目的是解决青年问题,同时对青年研究的理论发展做出贡献,因而两者互相影响,互相促进。

### 描述性研究与解释性研究

描述性研究意在描述有关事件的存在状态或发展过程,它只告诉人们发生了什么,如在此基础上对研究的问题进行概括或抽象,以理论的形式来说明现象产生的原因,这就是解释性研究。

描述性研究,其目的在于详细描述现象(描述发生了什么情况),研究人员往往没有正式假设,比如,对城市青少年犯罪率的研究,青年人婚姻家庭状况的研究等,只要详细说明情况怎么样即可,而解释性研究一般是力图通过详细评述情况为何或如何发生来解释某一社会现象,解释的一个基本形式即回答为何和如何的问题,比如,为什么城市的青少年犯罪率高,为什么年轻家庭较中老年家庭的稳定性差等等。除解释外,解释性研究的一个有关目的是预测,不但对某一事件发生之后能够对其进行解释,我们也希望能够在事件即将发生的时候做出预言,进而可对事件的发生加以控制。

### 横剖研究与纵剖研究

横剖研究即在某一特定时间内做人口横断面的研究。所谓横断面是指一个具有不同职业,不同教

育和收入水平,不同种族与宗教等各种不同背景的人的宽广的样本,而纵剖研究是指在一段长时间内,数周、数月、数年内所做的资料搜集。纵剖研究的基本优点在于,资料能从大量的人们那里搜集来,而且由于资料不受时间变化的影响,因而是可供比较用的。采用调查法的研究,比如邮寄问卷、访问并从理论上说都是横剖式的,即使实际上调查过程可能要好几周或好几个月才能完成。纵剖研究的优点在于整个时期的变化均可加以研究,它侧重于观察一段时间内的社会现象,至少要有两点、三点,在这一点时间怎样、到第二点时间怎样、或者用不同点的3个时间来比较看有何发展变化,所以横剖研究也是一种动态研究和发展研究。由于纵剖研究时间长、费用大,通常不可能象横剖研究那样大面积地进行,而只能确定较少的单位进行研究。另外,凡希望考察趋势的人,必将进行一次纵剖研究。例如,假如你希望研究1950—1988年青年人婚恋观的变化,你的研究就会是纵剖的。

### 社会调查法

社会调查即研究人员根据一定的研究任务或课题,运用科学的手段和方法,深入社会实际,通过实地调查,从研究对象身上获取第一手材料,以了解掌握研究对象实际状况的一种研究方法。它是青年研究的基本方法。社会调查的类型很多,主要有普通调查、典型调查、重点调查、个案调查和抽样调查。社会调查的主要方法是:参与法、观察法、问卷法、访问法。

社会调查法侧重于从社会的现实出发,通过直接地调查来收集数据和资料。青年研究的一个很重要的任务即解决现实中存在的青年问题,而许多问题的解决是已有的统计资料和文献分析所力所不及的,所以必须借助于社会调查的方法。目前,社会科学方面重视实证研究的倾向已有显著发展,社会调查技术也取得惊人的进步,因此,社会调查受到普遍广泛的重视。

社会调查做为一个完整的过程,是有一定科学程序的,主要有以下步骤:

#### 一、准备阶段

1. 确定调查任务,包括明确调查课题,进行初步探索及提出研究假设等工作。

2. 设计调查方案,包括调查指标的设计和调查总体方案(调查目的、范围、对象、调查项目、调查方法等)的设计,并对设计的总体方案进行可



行性研究。

3. 组建调查队伍, 包括调查员的选择、培训和组织等工作。

### 二、调查实施阶段

这一阶段的主要任务是, 采取选定的调查方法, 按照调查设计的要求做好搜集资料的工作。这是社会调查的一个最关键的阶段, 此一阶段, 应及时发现并解决调查工作中出现的新情况、新问题, 同时要对调查资料进行严格的质量检查和初步整理工作, 以便就地更正。调查的组织者应亲临第一线, 加强领导, 力求以最少的人力、最短的时间、最好的质量完成搜集材料的任务。

### 三、研究阶段

研究阶段的主要任务是: 整理分析资料, 开展理论研究, 撰写调查报告。

调查资料的整理即对调查来的资料进行初步加工, 使之条理化、系统化, 并以集中、简明的方式反映调查对象的总体情况。在此基础上对资料进行分析推断, 并进一步对其加以解释, 便能将调查结果与建立的研究假设相互验证, 从而说明事物的前因后果, 预测事物的发展趋势, 作出自己的理论说明。

研究阶段是社会调查的保证、提高阶段。社会调查能否出成果, 以及成果作用的大小, 在很大程度上取决于这个阶段的工作。

调查报告是调查研究成果的集中体现。

## 文献法

文献法, 亦作历史文献法, 是根据一定的目的和研究课题, 通过搜集文献, 摘取其中有关信息来获得所需资料的一种方法。

文献的种类很多, 按照不同标准, 可作许多不同的分类。最广泛的文献形式是文字文献, 包括报刊、杂志、书籍、政府出版物、会议文献、科研报告、学位论文、档案、日志、日记和其它油印或打印材料。另外还有数字文献, 如统计年鉴、统计报表及其它各种数字材料, 以及电影、电视、图片、绘画等图象文献, 磁带、唱片等有声文献。

应用文献法获取资料的一般步骤是:

一、根据研究课题, 确定所要搜集文献的内容, 类别及时间范围。例如, 假如研究课题为年轻人婚恋观的变化, 那么文献搜集的内容就应紧紧围绕这一题目, 搜集有关的观念、事例、数据。文献的类别可以是一些理论性文章、婚姻家庭方面的杂

志, 中国青年报等等。另外, 研究变化就要有一个确定的时间界线, 如八十年代前后。

### 二、搜集文献。

### 三、整理分析文献。

文献搜集的方法很多, 主要有两种, 一是检索工具查找法, 即利用已有的检索工具查找文献资料。我国目前一般图书情报工作都以手工检索工具为主。手工检索工具, 按著录的形式, 可分为目录、索引和文摘等。另一种查找文献的方法, 是参考文献查找法, 即利用著作者本人在文章、专著的末尾所开列的参考文献目录, 或者是文章、专著中所提到的文献名目, 追综查找有关文献资料。通常, 人们往往将两种方法结合起来, 交替使用。无论采用哪种方法, 一般都要到图书情报机构去通过各种目录查找文献。因此, 学会查阅目录, 是文献法的一项基本功。目前, 文献目录索引的编排方法主要有5种, 即分类编排法、主题编排法、编年编排法、地区编排法和字顺编排法。查阅文献时, 可按这5种目录索引编排方法查找。另外, 使用文献法搜集资料, 许多涉及的是厂矿企业、街道或区、乡、村基层单位的文献, 这些文献大都设有统一的编目方法。因此, 调查时应主动说明来意, 介绍自己的研究课题和目的以及所希望找到的文献资料, 尽量争取有关领导和管理人员的支持和配合。

文献搜集的方法按摘录信息的形式分, 有标记、眉批、抄录、提纲、札记等形式。

进行文献搜集时应遵循以下几条基本原则,

一、紧紧围绕研究课题, 有选择、有分析地搜集与之有关的资料, 不能毫无选择地盲目搜集。

二、在内容上要尽可能丰富, 要注意资料搜集的全面性和完整性。过去、现在的材料, 正面反面的材料, 典型性材料, 综合性材料都要有所涉及, 总之, 从内容上看, 搜集的材料越丰富越好。

三、时序上要有连续性, 即围绕调查课题搜集的文献, 在时序上不能中断, 否则, 所需资料可能残缺不全, 无法反映调查对象发展变化的状况。

四、要考察资料的真实性和正确性。因为真实性问题是资料有无价值的核心问题。其次, 资料的准确性也十分重要, 有些资料缺乏准确性, 特别是有些统计数据包含不少水分, 搜集资料时应加以判断。

文献法具有许多与其它方法不同的特点。首先, 文献研究的一个基本优点就是, 它们使研究人员对其不能亲自接近, 从而不能以其它方法进行研

究的研究对象做研究。它可以超越时空条件的限制，从时间跨度而言，它可以了解到几十年、几百年、几千年前的历史人物和历史事件，因此，文献研究尤其适于作一长时期的研究，研究的对象往往是一种趋势。从空间跨度上看，它可以跨越国界、洲界，了解到世界各国及国内各地区的社会情况，这一点，是其它任何调查方法所力所不及的。

其次，文献法具有很少反应性或没有反应性的优点，因为用文字记载下来的情报，既不会对调查者的主观好恶作出任何反应，也不会出现在直接调查中经常发生的由于被调查者心理因素影响而产生的种种反应性误差，所以一般来说它比口头情报更真实、更准确、更可靠。

再次，文献法比其它调查方法更方便、更自由，同时效率高、费用低、质量好。它可以用更少的人力、更少的经费、更少的时间，获得较其它调查方法更多的情报。而且，许多文献，是由老练的专业人员撰写的，可能比例如邮寄问卷表的粗劣回答有价值得多。

同其它任何调查方法一样，文献法也有其自己的局限性。

首先，任何文献，都是一定时代、一定社会条件下的产物，都是一定的人撰写的，因此要受时代及社会条件的局限，受撰写者个人素质及立场、观点的制约。所以，通过文献法获得的资料，与客观真实之间总会有一定差距。

其次，社会是不断运动、变化、发展的，但总会有一些最新的事物、现象、情况、问题没有记录上文献，所以文献法获得的情报往往落后于现实一大截。

再次，文献资料很难找齐、找全，常会发生资料不足的缺憾，资料不足时又很难找到补救的办法，研究的深浅常常受现有材料的限制。

此外，文献法都是书面情报、纸上的东西，缺乏具体性、生动性。

## 实验法

实验法，亦作试验法，是指根据一定的研究目的和研究假设，人为地控制某些现象，突出某些主要因素，在“强化”的条件下通过观察、记录、搜集资料以证实和研究客观现象间因果关系的方法。

与其它调查方法相比，实验法有如下特点：

一、可以使被研究的对象以较纯粹的状态出

现，即“纯化”实验对象。社会现象错综复杂，往往不易使人发现其中起作用的究竟是哪些因素，主导因素是什么。在实验中，人们可以借助于一定的手段和设备，根据研究目的的需要，排除各种偶然因素、次要因素和外界干扰，从中选择一些因素，改变、减少一些因素，把要研究的因素分离出来，单独进行观察、分析和研究。

二、强化并再现客观现象，在实验中，人们可以凭借自己的手段，创造出自然过程中难以出现或暂时不能出现的特殊条件，从而使客观事物的性质和规律能够强烈地表现出来。例如，要研究青年工人的生产积极性，可以对被实验者实行新的管理方法及奖励办法，以观察其生产积极性是否有所变化。

由于分类标准不同，实验法被分为许多不同的种类，依调查目的不同，可分为研究性实验和应用性实验。依调查环境不同，可分为实验室实验和现场实验。实验室实验研究者对实验的环境可进行严格的有效控制，现场实验，则是在自然的、现实的环境下进行的实验，实验者只能部分地控制实验环境的变化。按照调查的组织方式不同，实验调查可分为单一实验组实验，实验组对照组实验和多实验组实验。

一、单一实验组实验，即只选择一批实验对象做为实验组，通过实验活动前后实验对象的变化来做出实验结论。

二、实验组对照组实验，即选择一批实验对象作为实验组，同时选择一批与实验对象相同或相似的对象作为对照组，并且努力使实验组和对照组处于相类似的实验环境之中；然后，只对实验组给予实验刺激，而对对照组不给予实验刺激，最后，根据实验组和对照组前后检测的变化对比，作出实验结论。

三、多实验组实验。即选择若干批实验对象组成若干个实验组，在各自的实验环境下，通过实验活动前后实验对象的变化，作出各组的实验结论，然后再对各组实验结论进行对比研究，作出总的实验结论。

实验法的基本步骤如下：

1. 确定实验目的和实验课题，建立实验假设（或称理论假设），进行实验设计并完备实验条件。

实验假设离不开实验变量，实验变量就是指自变量和因变量。如果我们研究影响青年升学率的因



素, 就可以把升学率做为因变量, 把影响升学率的诸因素, 比如学校的师资力量, 教学设备情况, 教学管理等做为自变量。变量必须是可观察和可测量的。这就要求自变量和因变量都不能是大抽象的概念, 它们必须能被转变成测量指标。

实验设计即建立实验组和对照组。方法主要有两种: ①匹配法: 实验者将实验对象依照实验的目的, 使他们同时在所有有关特性上相配, 组成匹配对, 然后将每对中的一个分到实验组, 另一个分到对照组。②随机法: 实验者先制定一个抽样框架或者一张符合条件的实验对象清单, 然后用一些随机程序的方法, 例如随机数表, 从抽样框架中选取每组的成员。

2. 实验的实施, 实验实施工作主要是观察或测量。测量的目的就是要获得实验组和对照组中的某些指标值。步骤为:

实验组	对照组
①前检测	①前检测
②实验刺激	②不予实验刺激
③后检测	③后检测

实验效应 = 实验组(后测—前测) — 对照组(后测—前测)。

实验实施中还有很重要的一点即对实验过程的控制, 主要包括两个方面, 一是对变量的控制, 二是对非实验因素的控制。对变量的控制即指除那些选作因变量的特征外, 必须使研究对象的其它特性在实验前后保持相同状态。对非实验因素的控制就是要努力排除或减少非实验因素对实验过程的干扰。

3. 分析、研究实验前后的结果, 并写出实验报告。

## 普查

又称全面调查, 指对调查对象所包括的全部单位毫无遗漏地进行调查。它可以是全国性、地区性的, 也可以是行业性、部门性的。人口普查即一例。

普查具有以下特点:

一、全面调查。研究对象总体中的每一个单位都要被调查。

二、时间性强。一般要求在标准时点的几天内完成。

三、组织工作量大。

进行普查一般有这样几个步骤:

一、组织专门的普查机构。

二、设计普查方案。主要包括确定普查的总体和单位、普查的项目、时间、经费及预算等。

三、普查人员的培训。要有专门的培训机构和培训工作人员来进行。

四、普查试点, 以修订普查方案和细则。

五、普查登记。

六、资料汇总与分析。

七、公布普查结果。

进行普查, 应注意以下几个问题:

一、普查项目必须简明。普查时, 调查人员多, 调查对象广, 组织工作复杂, 调查项目过多, 不仅难于进行准确的调查, 而且难于进行统计分析。

二、普查时间必须同一, 统一规定调查资料所属的标准时点。

三、在尽可能短的时间内, 迅速完成调查任务。

四、尽可能按一定周期进行, 并尽可能保持普查内容的一致。这样才便于对普查资料进行对比分析, 掌握社会现象的发展趋势及其规律性。

与其它类型的调查比较起来, 普查所搜集的材料是最全面的, 而且误差小, 精确度高。但普查也有明显的局限性, 一是工作量大, 花费大, 组织工作异常复杂, 而且时效性差。二是调查的内容有限, 只能调查一些最基本、最一般的社会现象, 而很难对社会问题进行深入细致的研究。

## 典型调查

典型调查, 即指在对被调查对象进行全面分析的基础上, 依据调查目的的不同, 有计划地选择一个或几个典型单位所进行的调查。典型调查的特点即其个别性和代表性。另外, 典型调查一般比较深入、细致, 主要是定性的调查。

正确选择典型, 是保证典型调查科学性的关键步骤。典型即调查总体中一些具有代表的基本单位。如何选择典型? 首先, 应根据确定的调查目的和研究主题来选择典型。其次, 调查者对研究总体应有初步认识和总体印象, 这样选出的典型才真正具有代表性。再次, 选择典型, 要有实事求是的态度, 保证典型的客观性和真实性, 不能主观地确定典型, 更不能人为地树立典型。

典型调查能获得比较真实可靠的第一手材料, 可以调查到广泛、深入、系统的内容。但是, 典型

调查选择的只是个别或少数几个单位,其代表性总是不完全的,此外,典型调查主要是一种定性调查,它很难对调查对象的总体进行定量研究,不能不加限制地把调查结果直接推论到总体,因此,典型调查只有与其它方法,如抽样调查,普查等方法相结合,才能发挥更大的作用。

### 重点调查

即在全部调查单位中选择一部分能够反映研究总体基本情况的重点单位进行调查。比如,要了解大城市中青少年犯罪的情况,在北京、天津、上海等几个重点城市进行调查,便可获得调查研究必要的资料。

采用重点调查需要两个基本条件:一是调查的主要任务是掌握研究总体的基本情况,二是调查的某些项目或指标在少数重点单位能够得到集中反

虽然重点调查的调查对象也是一个或几个单位,但它与典型调查有明显区别:典型调查是选择同类事物中具有代表性的单位作为调查对象,重点调查则是选择同类社会现象中具有集中性的单位做为调查对象;典型调查的主要目的是认识同类事物的本质及其发展规律,主要是作定性调查;重点调查的主要目的是为了对某种社会现象总体的数量状况做出基本的估计,主要是作定量调查。相对说来,重点调查的选择比典型调查容易,只要知道被调查总体单位的所有相关指标,就可以确定调查对象。重点调查,被调查的单位一般比较少,相对说来,在人、财、物、时等方面比较节约。

### 个案调查

指对一个团体,一个组织(包括家庭、社区)或一个人以及一个事件进行详尽的调查研究的方法。比如,研究某个青少年犯罪团伙,往往要研究其各自的生活特征、成长过程及家庭背景,对每个人的过去和现在,对团伙的形成和发展都要进行全面、细致、深入的了解和认识。

个案调查的步骤为:

一、确定调查个案。

二、对案主进行访问,了解个案本身的材料及背景材料。

三、搜集有关资料,除与被调查者谈话而外,考察有关的文字记载也是个案调查的重要一环,比如有关案主出生、入学、犯罪等的情况。

四、诊断,不仅包括资料或证据的核实、修订、补充、整理分类和分析,而且包括通过分析研究后,针对存在的问题,提出解决的建议或方案。

在个案调查过程中,与被调查者谈话是重要的一环,应注意以下几点:

一、态度诚肯、热情,打消对方的种种顾虑,以取得谅解、信任与合作。

二、认真听取对方的谈话,仔细做好记录,不轻易打断对方,必要时可启发或引导对方把话题集中在调查的内容上。

三、注意调动被调查者的积极性,注意发现在谈话过程中出现的新线索,顺藤摸瓜。

四、态度从容不迫、提问得法。

另外,个案记录应注意:

一、既要写案主的情况,也要写个案调查的过程,如访问过程、谈话记录、访问背景等。

二、依时间先后顺序,注明年、月、日。

三、适时进行个案调查的整理、分类,及时发现问题,以便修改。

四、文字简练,条理清晰。

个案调查不仅应用于青年的劳改、劳教、工读学校,计划生育等部门工作中,而且还用于婚姻、恋爱、家庭、犯罪等研究。

### 抽样调查

抽样调查是指按照随机原则从总体中抽取一部分调查单位进行调查,然后根据样本值推论总体值的一种调查方法。抽样调查具有以下特点:一、以概率论和数理统计为基础,按随机原则选取调查样本。二、以样本值推论总体值,且推论是能计算出准确率大小的。三、抽样误差可以计算,并可加以控制。

一、抽样调查的程序一般包括以下几步:

1. 确定调查总体(范围、数量)。

2. 抽取调查样本。

3. 收集样本资料,即采取各种方法对样本单位进行实际调查。

4. 计算样本资料并推断调查总体。

二、抽样调查的几个基本概念:

1. 总体,指为一定研究目的和要求所确定的被调查对象的全体,如在某城市进行青年就业情况的调查,则该市的全体青年就构成一个总体。

2. 样本,即从总体中随机抽取出来进行调查的一部分单位。



3. 总体值, 即根据总体计算出来的某项指标值。

4. 样本值, 即根据样本计算出的指标值。

5. 抽样误差, 指样本值与总体值之差。是由于抽样的随机性而产生的不包括登记误差的代表性误差, 其数值主要受样本数目的多少, 以及被研究标志的变异程度的影响, 在一定的抽样方式下, 样本数目越多, 抽样误差越小, 总体中被研究标志的方差越小, 抽样误差也最小。一个总体可组成很多样本, 每个样本的样本指标都具有抽样误差, 这很多抽样误差的平均数叫做抽样平均误差, 计算公式如下:

$$\text{抽样误差(平均值)} = \sqrt{\frac{\text{总体方差}}{\text{样本单位数}}}$$

6. 重复抽样。指从总体中随机抽取一个样本后将其放回, 再从全部总体中抽取样本, 即每次抽样总体的单位都不变, 同一单位有多次被抽的可能。

7. 不重复抽样。指抽取一个样本后不再放回, 而在余下的总体中继续抽取, 每个总体单位只有一次被送的可能。

### 三、抽样的类型和方法;

抽样按是否随机, 可分为随机抽样和非随机抽样。

#### 随机抽样:

1. 简单随机抽样, 又称纯随机抽样, 即对总体单位不进行任何组合, 仅按随机原则直接抽取样本, 其中有3种不同的抽法:

①直接抽选法, 即从调查总体中直接随机抽取样本进行调查。

②抽签法, 即对总体单位统一编号, 号码标于卡片上, 将卡片拌匀后随机抽取, 直到抽满预定的样本数为止。

③随机数表法, 将总体中的所有单位统一编号; 根据编号位数决定随机数表中多少数字连成一组, 若总体单位数在10—99之间, 则取两个数字列成一组, 然后根据随机原则, 决定某行、某列的数字定为起点, 由左至右或由上至下查出号码, 凡总体单位编号与随机数表中数字相同者, 即为样本。

2. 类型抽样, 又称分层抽样, 研究者将总体单位按某种标准分成若干类(或叫层), 然后根据类型单位数与总体单位数的比率, 确定从各类型中抽取样本单位的数量, 最后, 按随机原则从各类型中抽取样本。比如进行某市青年价值观的研究, 就

可将青年分为学生、工人、农民、个体户等类型。

类型抽样最适合于总体单位数多, 且各单位间差异性大的研究对象, 分类将有明显差异性的单位分开, 使样本单位的分布更接近于总体分布, 有利于提高样本的代表性。另外, 分类后, 影响抽样误差的只是各类内的差异, 而各类间的差异已不属影响抽样误差。

类型抽样的关键是分类的标准要科学、要符合实际情况。

3. 等距抽样, 又称机械抽样或系统抽样, 即在研究总体中隔等距的单位抽取样本的方法。比如, 总体数为1000, 样本数为200, 那么抽样距离为

$$\frac{1000}{200} = 5, \text{如果要调查1000名青年, 就可以把其}$$

姓名有次序地排列起来, 每隔4人抽一个。

等距抽样, 样本在总体中的分布比较均匀, 具有较高的代表性, 抽样误差小于简单随机抽样, 且简单易行, 但缺点是, 调查总体的单位不能太多, 而且要有完整的登记册, 否则难于进行。此外, 使用等距抽样的方法, 还必须注意避免抽样间隔和总体排列的某种周期性重合。

4. 整群抽样, 又称聚类抽样, 即将调查总体按一定标准分成许多群, 然后按随机原则从这许多群中抽取若干群做样本, 对群内的每一单位逐个调查。比如, 若对某校学生的学习情况作调查, 可以随机抽取几个班, 对每个班的学生逐个进行调查。

整群抽样, 样本单位比较集中, 调查工作比较方便, 节省人力、物力和时间, 但它的样本分布不均匀, 代表性较差, 同时样本数无法精确控制, 因为在整群抽样中, 各群中的单位数是未知数。

#### 非随机抽样:

1. 依意抽样, 又叫方便抽样, 便利抽样等, 即按调查者的方便任意选取样本。如在街头路口、剧院、商店、车站等公共场所随便选择一些年轻人进行访问调查。这种方法简便易行而且节省, 但样本的代表性差, 不能推论一般。

2. 判断抽样, 即按调查者的主观判断来选择样本。为了提高判断抽样的价值, 提高根据这些样本推论一般的可靠性, 应尽量选取“多数型”或“平均型”的单位做为样本。判断抽样的准确性和代表性, 取决于调查者对调查对象的了解程度, 取决于调查者的判断能力。

3. 定额抽样, 是相当于分层抽样的非随机抽样, 另外它还要求每一层次样本代表它在总体中所

占的比例。

能选择一些控制特性作为划分总体的标准，把总体分成若干层，然后按各层在总体中的比例决定各层中样本的大小，最后按这些比例数字在各层中随意抽取。比如调查某市青年人对政治的参与意识，若青年人中团员占30%，非团员占70%，则选取1000个样本，就要抽取团员300名，非团员700名。

### 参与法

参与法，即研究人员进入到被调查对象当中去，做为他们中的一员，亲自参与其活动，通过切身体验及观察，搜集大量深入、丰富的资料的方法。比如研究单身青年工人的生活状况，研究者即可深入到某工厂一段时间，与单身青年们住在一起，白天一起工作，闲暇时间参与其娱乐活动。

参与法是一种定性调查，由于是研究者深入其中，所以调查比较深入、细致、全面、真实，能获得许多生动的材料。但它不能了解到很大范围的情况，不能推论一般。进行参与调查应注意以下几点：

一、取得被调查者的充分信任，调查人员最好不要让被调查者知道自己的真实身分，因为被调查者如果知道调查人员在对其进行调查研究，其行为就会丧失自然性，他们会有意无意地显示或掩饰某些方面，从而影响调查的真实性。

二、调查人员由于做为被调查者中的一员，参与其一切活动，所以看法、观点很容易站在被调查者一边，从而影响判断问题的准确性。研究人员应切记看问题站在旁观者的角度上，采取中立态度，不带任何感情倾向。

三、参与调查一般适用于对某一问题或某一次活动进行较细致的调查，“蹲点儿”即是参与调查的一种，它只是一种定性调查。另外，许多问题不能应用参与法进行研究，比如犯罪，所以参与法的适用性不是很强。

### 观察法

观察法，即研究者根据一定的调查目的，运用自己的感官或辅助工具去直接观察被研究对象的方法。

观察的类型很多。

一、依观察者是否参与被观察对象的活动，可分为参与观察与非参与观察。参与观察即前文所提

的参与法。非参与观察是指观察者不直接参与被观察对象的活动，而是以局外人的身分，从侧面对其进行观察，其中又有对被观察者暴露自己身分和不暴露身份的观察两种。

二、依观察者是否直接观察被研究对象，可分为直接观察与间接观察。直接观察即观察者直接用自己的感觉器官去进行观察被研究对象。间接观察法即研究者借助仪器或委托他人进行观察，比如观察一青少年犯罪团伙的日常活动，研究者不能随时跟踪观察，有时可以把观察的任务委托给居委会或这伙青少年的邻居，这种观察就是间接观察的方法。

三、依观察内容是否有统一设计的、有一定结构的观察项目和要求，可分为结构性观察和非结构性观察，结构性观察对研究的问题有很严格的界定，观察的步骤和内容都有比较明确和严格的规定。非结构性观察则灵活性较大，它只要求观察者有一个总的观察目的和要求，或一个大致的观察内容和范围，观察所得的材料比较零散。

观察法的一般原则和要求：

一、准确无误地反映真实情况。

二、客观性，尽可能避免自己的偏见或以往经验的影响。

三、全面性，全面观察，如实反映客观事物的全貌。

四、具体性，观察不要限于表面现象，要具体、深入。

进行观察时应注意的几个问题：

一、努力减少观察活动对被观察者的影响。当观察对象是人或人的活动时，特别是在被观察者的人数很少而又比较稳定的情况下，素不相识的观察者的出现和观察活动的进行，会使被观察者产生某种紧张、不安、好奇或羞涩的心理，有意或无意地显示或掩饰自己的行为表现。观察者一定要注意这一点，努力减少观察活动对被观察者的影响。

二、观察的同时积极思考，如果只满足于观察事物，而不对有关的问题积极认真地思考，那么即使观察到了有价值的现象，有时也会视而不见、听而不闻，轻易地放了过去。因此，观察中要善于把观察与思考结合起来，这样才能搜集到有价值的观察材料。

三、做好观察记录，最好是在观察的同时进行记录，不宜记录的场合，应在观察后及时追记。记录的内容要具体、详实。



应用观察法进行调查,可获得直接、具体、生动的第一手资料,与间接的书面调查和直接的口头调查比较起来,可靠性要高得多。另外它有利于对不能够或不愿意进行语言交流的社会现象进行调查,而且简便、易行、灵活性较大,但是,观察法也有其缺点和局限,首先,观察法最大的缺点即其表面性和偶然性,观察中观察者看到的一般都是些表面的现象,而且是在特定条件下发生的,带有一定的偶然性。其次观察法受时空条件的限制,许多现象不能或不宜于观察。另外,采用观察法搜集的资料往往会有一定的观察误差,受观察者主观因素的影响。但观察仍然是社会调查的一种最基本的方法。

### 访问法

访问法,即由访问人员依据按一定研究目的编写的调查提纲或问卷,面对面地向被调查者访问,从而获得有关资料的一种方法。

#### 一、访问法的一般步骤如下:

1. 做好访问前的准备。首先,访问以前,应尽可能掌握一些有关被访对象的基本情况,诸如性别、年龄、婚姻状况、文化程度、性格、习惯等,这样有利于在谈话中对被访者的把握。其次,要准备好详细的访问提纲和问题,学习与调查内容有关的各种知识。与问卷相比,访问提纲是一种访问和回答相对自由的搜集资料的工具,纲目较粗,结构较松散。粗线条的访问提纲应具体化为一系列的访谈问题。另外,访问前要选择好访谈的具体时间、地点和场合。

2. 实地访问。访问的第一步是接近被访问者。访问员应先做自我介绍,直接说明调查目的、意义和内容,态度要诚恳,以利于取得对方的信任。访问开始时,提问的方式可以多种多样,究竟采取哪一种方式,需依问题本身的性质和特点、被访者的具体情况以及访问者与被访问者的关系而定,开门见山,单刀直入,或是循循善诱,谨慎迂回。访谈过程中不仅要提问,而且要引导和追问,当被访者对提出的问题理解不正确,有所顾忌,或对某些问题一时回忆不起来,或谈话离题太远时,访问者都应给予及时的引导。另外,当被访者的回答模棱两可,自相矛盾,或不肯吐露真情时,访问者应进行追问,以便获得真实、具体、准确、完整的回答。访问的同时,做好访问记录。

3. 整理分析所搜集的资料。

#### 二、访问应注意以下几个问题:

1. 主动向被访者介绍自己的访问目的、内容和意义,以求得对方对访问的理解和支持,努力消除对方可能有的种种怀疑和顾虑,使他们相信他们完全有能力回答问题。

2. 态度谦虚,以礼待人,要有甘当小学生的精神。在访谈过程中,访问者应以谦虚的态度提出问题,聚精会神地听取对方的回答,应不时以点头、肯定的目光和手势等非语言信息来鼓励对方继续谈下去。

3. 访问过程中,访问者应始终保持客观、中立的态度,不应有倾向性或诱导性的任何表示,以避免被访问者的迎合心理。

4. 注意所提问题的顺序,一般应遵循先易后难,先泛谈后具体的原则,先谈一些比较容易回答且轻松愉快的题目,以便访问者与被访者建立良好的人际关系和谈话气氛,而后再谈那些比较深入、敏感的问题。

5. 谈话语言应地方化、口语化、通俗化。

6. 重视访问中的非语言信息。非语言行为,人的外貌、行为、表情及周围环境等可从侧面提供一些很有价值的信息,访问者应多加观察。

7. 做好访问记录。记录内容上应抓住几点,一是记要点,即主要事实、主要过程。二是记特点,即有特色的事件和情节。三是记疑点,留待访问后期询问或访问后再去做调查。四是记易忘点,如人名、时间、地点及各种数据等。五是记自己片断的思考、联想和感受。

8. 对不合作者做耐心的说服工作。只要方法对头,就能赢得他们的信任。

#### 三、访问法的优点主要有:

1. 广泛、深入、细致。由于访问法能做到问、听、看三结合,所以它不仅可以获得广泛而丰富的材料,了解事物的各种现象,而且能对某些问题追根求源,加深了解。

2. 访问调查富有弹性,在调查中,访问者能及时发现新问题,了解新情况。

3. 访问法是面对面的口头调查,在访问过程中,访问者能及时引导、解释和追问,同时可获得许多非语言信息,这一切都有利于调查的可靠性,有利于对被调查者回答的可靠性做出正确的评价。

4. 访问调查可以使用较复杂的调查问卷,同时可以克服邮寄问卷调查中回收率低的缺点。

缺点有:

1. 由于访问的标准不一, 访问结果不易做定量分析。

2. 访问质量难以把握, 要由访问员的素质而定。

3. 有些问题不能当面询问, 由于被访问者不能匿名回答问题, 所以, 对于某些敏感问题、尖锐问题和隐秘问题, 或者是访问者不宜当面访问, 或者是被访问者不愿当面回答和不做真实回答。这都会对访问结果产生不利影响。

4. 由于访问调查是面对面进行的, 因而费人、费时、费钱, 且不易做大面积的调查。

## 问卷法

是调查者适用统一设计的问卷向被调查者了解情况的一种方法。问卷即调查表, 是按照一定的理论假设设计出来的, 由一系列变量指标所组成的一种搜集资料的工具。

按照问卷发放形式的不同, 问卷调查可分为邮寄问卷、送发问卷和访问问卷。邮寄问卷即调查者通过邮局向被选定的调查对象寄发问卷, 并要求被调查者按照规定的要求和时间填答问卷, 然后通过邮局将问卷寄还调查者。送发问卷即调查员派专人将问卷送给选定的调查对象, 待其填完后, 再派专人收回。访问问卷即调查员面对面地访问被调查者, 按其回答填写问卷。

问卷调查的一般程序如下:

一、设计调查问卷。先要经过选择调查课题, 进行初步探索, 提出研究假设等几个先行步骤, 然后进入具体的设计阶段, 问卷设计十分复杂, 问题要问得科学、具体, 而且要考虑怎样答得全面正确, 需要一定的设计技巧。

二、选择调查对象, 可按随机抽样的方法。

三、培训调查员, 特别是在问卷内容较复杂的情况下, 调查员的严格培训是重要的一环。

四、发放问卷。不同的调查方法有不同的发放形式, 邮寄、专人送发或调查员登门访问。

五、回收问卷, 问卷调查中的一个重要环节。访问问卷的回收率最高, 送发问卷只要及时派人回收, 也问题不大, 但要注意收一次可能不够, 一些人可能忘填了或推脱时间紧没填, 催促他们, 预定时间再去收一次。邮寄问卷的回收率一般比较低, 因此, 应适时向被调查者发一次提示通知或催促信件, 有时需要两、三次的催促。

六、审查问卷。收回的问卷总会有一些不正

确、不完整的回答, 需要及时审查, 根据前后逻辑能改正的改正, 不能改正的作为无效问题处理, 以免直接统计降低调查的可靠性。

问卷一般包括说明、调查问题和结语 3 部分。说明即在问卷开头要阐述一下调查的目的、意义, 对被调查者的希望和要求, 问卷调查的匿名性和保密原则, 回复问卷的时间和方式, 主办调查的单位、组织和个人身分, 问题的填写说明等。语气要谦虚、诚恳、热情, 文字力求简洁。

问卷的结语可以是对被调查者的合作表示真诚的感谢, 也可以征询一下被调查者还有什么意见和补充。有的问卷可以不要结语。

调查问题是问卷的主体, 应重视其编写格式。

在问卷中, 对问题的回答有两种不同的形式, 开放型和封闭型, 开放型的回答即对问题的回答不提供任何具体的答案, 而由被调查者自由填写, 如:

您怎样看待近些年年轻人离婚者增多这一社会现象?

答\_\_\_\_\_。

开放型回答的最大优点是灵活性强, 它适合于回答各种类型的问题, 特别适合于回答那些可以有很多答案, 或答案较复杂, 或尚未弄清各种可能答案的问题。但开放型回答资料的标准程度低, 难以进行整理和分析, 而且往往会出现许多一般化的、不准确的、甚至答非所问的无价值的信息。

封闭型的回答是指列出问题的所有可能答案或几种主要答案, 由被调查者从中选择。封闭型问题的编写格式多种多样, 下面是几种主要形式:

1. 两项式, 只有两个答案可供选择:

您喜欢参加厂团支部通常所组织的业余娱乐活动吗? ( 1 ) 喜欢 \_\_\_\_\_

( 2 ) 不喜欢 \_\_\_\_\_

这种回答方式, 适用于互相排斥的两样一式的这类问题。

2. 多项式, 有若干可供选择的答案, 被访者可选填其一或更多项:

您认为美满的婚姻应主要具备哪些条件 ( 请在您选择的项目后打 V, 可选 3 项 )

( 1 ) 夫妻之间相互信任。

( 2 ) 双方兴趣相投。

( 3 ) 一方勇于为另一方做出任何牺牲。

( 4 ) 经济富裕。

( 5 ) 与对方家庭成员及亲属的关系融洽。



( 6 ) 双方都能在小事情上容忍、谦让。

3. 顺序式, 即列出许多答案, 由被访者填写答案的先后次序。

您在业余学习中遇到哪些困难 ( 请按困难程度排列, 最大的困难写在左边第一格, 依次往右填写。

- ①时间紧
- ②有了问题没人请教

- ③参考资料少
  - ④家里人不支持
  - ⑤购买图书财力紧张
  - ⑥自信心不强
- 

4. 矩阵式: 把两个或两个以上的问题集中在一起, 用一个矩阵来表示:  
您认为做为一名青年应该

	很同意	同意	无所谓	不同意	很不同意
( 1 ) 敢想敢干	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
( 2 ) 乐于冒险	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
( 3 ) 做事应向长辈请教	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
( 4 ) 宁愿做错事而不可不做事	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. 等级式, 即列出不同等级的答案由被调查者选择:

您对您所在单位的团支部是否满意:

- ①很满意 \_\_\_\_\_
- ②比较满意 \_\_\_\_\_
- ③无所谓 \_\_\_\_\_
- ④不满意 \_\_\_\_\_
- ⑤很不满意 \_\_\_\_\_
- ⑥不知道 \_\_\_\_\_

6. 序列式: 用一种连续的回答类型作为替换形式, 并只是在两端作记号:

由学生对课程设 置做最后决定	完全由校方对课 程设置作决定
0	10

你认为学生应该在多大程度上决定课程 的设置? 在你认为合适的标度上划圈。

封闭型的回答方式有许多优点, 有利于对统计结果进行统计分析和对比研究, 同时也有利于询问一些敏感性问题, 但问题设计起来比较困难, 回答起来比较机械, 没有弹性, 被访者可能对自己并不懂、甚至根本不了解的问题任意填写, 从而降低回答的真实性和可靠性。

半封闭半开放的回答方式可在一定程度上弥补封闭型回答方式的缺点, 例如:

你认为目前的社会风气好不好? ( 请在适当的格内打× )

- ①好
- ②难说
- ③不好

为什么 \_\_\_\_\_

将问题汇总组成问卷的时候, 研究者必须决定要问多少问题, 怎样安排其次序。编排问题次序的一般原则如下:

( 1 ) 一些基本的问题、容易回答的问题放在前面, 如年龄、性别、职业、受教育程度等事实性的问题。

( 2 ) 要按问题的时间顺序来排列, 由过去到现在, 再到将来, 也可以先问当前的问题, 然后再追溯过去的情况。在时间顺序上, 问题的排列要有连续性, 不能来回跳跃, 打乱被调查者的思路。

( 3 ) 要按问题的性质或类别来排列, 尽量把同类性质的问题安排在一起, 避免不同性质或类别的问题相互混杂。

( 4 ) 敏感性和开放性的问题在后面。

( 5 ) 处理好列联问题, 列联问题是要回答者对前一过滤问题的回答情况作答的问题。写列联问题可采用不同的格式, 常见的格式是在每一列联问题后面说明回答了某一问题之后应回答哪个问题。如:

10. 你是否经常看电影?

是…… ( 回答A )

否…… ( 跳到第11题 )

A大约一月看几次? \_\_\_\_\_

另一种是带框格的列联问题:

你是否已婚

是 (     )

否 (     )

如果是, 您第一次结婚时的年龄? \_\_\_\_\_

也有的用箭头连结:

25. 您是否有正式职业?

①有\_\_\_\_\_.

②无\_\_\_\_\_.

→26 您从事现在的职业有几年?

① 2 年以内

② 2—5 年

③ 5 年以上

→27. 您已待业几年?

① 1 年以内

② 1—2 年

③ 2 年以上

问卷设计中易犯的错误:

1. 双重意义的问题, 不要把两个或两个以上的问题合成一个, 如你对电影和音乐感兴趣吗? 令那些只对其中一项感兴趣的人很难做肯定或否定的回答。

2. 问题提得模棱两可, 如你所在的协会与××大学有来往吗? 这样的问题将使那些过去与××大学有来往而现在没来往的协会无法回答, 以至同一协会的成员有的说有、有的说无。

3. 问题抽象, 比如你有没有很强的幸福感, 回答可能没有什么可靠性。

4. 问题带诱导性, 比如“许多科学家认为吸烟有害健康, 你同意吗?” 这样的问题通常会增加肯定回答的比例。

问卷调查能突破时空的限制, 在广阔的范围内, 对众多的调查对象同时进行调查, 它为现代社会提供了一种高效率的了解社会情况的途径。如不采用问卷方法, 要想全面、迅速、准确地掌握社会信息, 即将是一件极为困难的事情。其次问卷调查标准统一、形式规范, 易量化、易比较、便于进行定量分析和研究。如用计算机作为统计分析的工具, 问卷法便可成为一种大容量、高效率的定量调查方法。但是, 问卷调查缺乏弹性, 一旦印发, 就不能在调查中根据实际情况予以更改。因为问卷中的问题有限, 所以调查往往缺乏深度, 另外难以指导对问题的填答, 难以全面了解影响回答的各种因素, 难以搜集到不能言表的材料。

## 调查报告的撰写

### 一、调查报告的种类

根据调查报告的内容划分, 可分为综合性调查报告和专题性调查报告。一般来说综合性报告涉及

的问题比较广泛, 反映的情况较复杂, 篇幅也较长, 专题性调查报告的内容较窄, 它们之间没有绝对分明的界限。按调查报告的主要目的分, 可分为研究型报告和应用型报告。研究型报告是以学术研究为目标而撰写的, 学术性强、讲究方法、报告的直接目的是阐明学术理论观点。应用型报告即为了提出某种工作或政策建议而撰写的报告。

### 二、调查报告的基本格式

调查报告由标题、前言、主体、结尾 4 部分构成。

1. 标题即题目。一是用调查对象和主要问题做标题, 二是用一定的判断或评价做标题。也可用提问做标题, 比如×××××的原因在哪里?

2. 前言一般着重于介绍调查的目的、调查工作的具体情况及调查所使用的方法, 其中包括调查范围, 调查对象的情况等。

3. 主体是调查报告的正文。主体部分的写作应考虑 3 个方面: (1) 采取什么样的写法要看想表现什么样的主题。(2) 材料的安排要恰到好处, 恰如其分。因为调查资料很多, 用哪些不用哪些, 用在哪儿, 应根据报告的主题而定。(3) 结构安排要恰当, 要有逻辑性, 使说理层层深入。

4. 结尾: (1) 概括主题、深化主题。

(2) 做出结论。

### 三、调查报告写作的一般程序

#### 1. 确立主题

主题是调查报告的宗旨和灵魂, 是作者说明事物、阐明道理所表现出来的基本思想和观点。精心确立主题, 是写好调查报告的关键。

确立主题应从调查研究的最初目的, 调查中获得的实际材料, 现实中需要回答的问题这 3 个因素出发考虑。主题宜小、宜集中。

#### 2. 选择材料

材料取舍最根本的原则即选取与主题有关材料, 去掉与主题无关的材料, 精选的材料要能深刻地说明问题的本质。应注意选择典型材料, 综合材料、对比材料和统计材料。

#### 3. 拟定写作提纲

写作提纲分为条目题纲和观点题纲两类。条目题纲即从层次上列出调查报告的章、节、目, 观点提纲是在此基础上各章节目所要叙述的主要观点。

#### 4. 撰写报告

## 统计方法在青年研究中的应用

统计方法是通过对事物量的研究来认识其质的



一种方法,是从数量方面综合认识事物规律性的一种工具。统计方法的实施过程是从设计指标及指标体系开始、经过调查、整理、分析,最后达到对客观事物规律性的认识。青年研究是以“青年人”为研究对象的,在研究方法上则表现出综合性和专业性相结合的特征,认识青年现象、研究青年问题,必须从青年的社会群体、青年的婚姻家庭、青年的生理心理等各个方面进行全方位的研究,这就要求除运用综合的观察分析法,还必须运用一定的专业研究法——社会调查和统计,对掌握的第一手材料,必须经过统计的描述和分析,才能对各种青年问题有个科学的数量表述,并预测未来的状况。所以统计方法是使青年研究科学化、数量化的工具,是使青年研究由感性认识走向理性分析的必由之路。

### 描述性统计在青年研究中的应用

描述性统计是一种多变量分析法,在青年研究中,它特别适用于处理两个或两个以上变量的研究。描述性统计一般分为集中趋势和离散趋势两个方面。集中趋势主要从算术平均数、中位数、众数、几何平均数等角度对有关青年问题进行数量的描述,是关于研究对象的某一测量指标的中心值的计量数字;离散趋势是关于研究对象的某一测量指标的离散程度的描述,它包括全距、四分位差、平均差、标准差、变异系数等。在青年研究中,集中趋势和离散趋势往往是同时并用、互相补充,才能更全面、准确地反映青年群体的特征。例如,研究20~35岁青年的工资收入状况,运用集中趋势的方法,我们可知其工资的平均收入水平、众数所在之点等,同时运用离散趋势的统计可知其最高与最低收入的差距、平距差距等,这样二者结合,便把该年龄段青年收入的情况栩栩如生地描绘出来。

### 青年研究中资料的搜集方法

研究青年群体的某一特征,首先要搜集有关此特征的诸多方面的原始资料,这就要求研究者首先熟悉各种资料搜集技术,根据统计学原理,在青年研究中,我们大体可用如下方法搜集有关资料:一、观察法,是由研究者对观察对象所进行的目睹的描述的方法。二、历史文献法,从人口普查资料、警察记录、日记、宣传品等中扑捉有关信息。三、调查法,由研究者对被研究对象进行有目的的问卷调查、电话调查、访问等来搜集有关资料。四、实验法,在实验室或自然环境中搜集有关资料。以上各

种方法均可为青年研究所采用,有时也可二种或多种方法并用,以全面、完整、准确地搜集到有关信息资料。

### 青年研究中资料的整理方法

在青年研究中,运用统计方法可搜集到大量个别单位的、分散的原始材料,但这些材料不可能使我们从量的方面认识事物的规律性,因此必须按照研究问题的需要,或是按事先拟定的汇总提纲进行整理。把搜集到的大量原始材料经过科学的综合、加工,使它系统化,变为能够反映总体特征的综合数字资料。统计资料整理的组织形式基本上有两种:一是逐级整理,另一种是集中整理。前者是从基层取得的调查材料自下而上一级一级地在本系统或本地区范围内汇总起来,如各地团的基层组织的团员登记表经层层上报,最后在团中央汇总;集中整理是把全部调查资料集中在一个机关或集中在最高统计机关进行一次汇总。对于十分重要的或要求时效性极高的调查,往往采用这种汇总方式。有时将以上两种方法同时使用称为综合汇总。在统计资料的汇总技术上可采用手工汇总(包括划记法、过录法、打叠法、卡片法等)和电子计算机汇总两种方法。对某些历史资料的整理,可通过制定各级积累历史资料的统一提纲的形式。青年研究需要横向的比较研究和纵向的历史研究,因此掌握资料整理技术十分必要。

### 青年研究中的统计分组

统计分组是按照某种标志将总体区分为若干部分或若干组的一种统计方法。其目的是区别现象之间存在着质的差别,把总体内不同性质的单位分开,使性质相同的单位归在一个组内。这样才有助于从数量方面揭示现象内部的联系,深入说明总体的特征和规律性。统计分组是基本的统计方法之一,在整个统计研究中具有重要意义,在统计资料整理中有举足轻重的作用。统计分组又分为按品质标志和数量标志二种分组法,品质标志主要指的是被研究对象的质的属性或质的特征,它在总体单位之间表现出的差异是性质上的差异;数量标志指的是按表现总体数量特征的标志所进行的分组,如学生按身高或体重进行的分组。青年研究往往会遇到许多问题,采用统计分组法便会使这些问题简化,达到更清晰的认识。

## 频数分布

所谓“频数分布”，就是把我们所研究的数量资料分成几组或几类，并分别记录各组事件的次数的系统化材料。研究青年问题，常常要运用频数分布的统计方法，如要分析某一年龄段青年工人的工资，首先得搜集这些工人的工资资料，然后可用两个步骤把原始资料组织成一种合乎逻辑的形式。首先是安排一个序列。这种序列，是把数值从最小到最大，或者按相反次序进行排列。第二步是把所有要研究的青年工人的工资划记在序列中。这样便可很快地通过低工资、高工资和工资的集中部分，对序列的情况进行分析。对数量较大的分析单位，可采取分组的方式通过每组内的次数分配来完成。为编制次数分配，必先决定每组的界限，然后根据分组的情况把所计数目直接列入相应的组别，就得出频数分配表。利用频数分配表，可计算出各种平均数，来代表整个数值，并且能够计量离差，来描述资料数值的离中趋势。在编制频数分配表时，表中的组距宽度应尽可能使之相当，并尽量避免出现开口组。编制组限，应该使各组的组中值符合原始资料的集中点，并且所选的组距应使编制次数分配的结果仅有一个集中处或单峰。频数分布也可用各种图示如：直方图、多边形图等来表示。

## 统计表

是表示统计资料的一种形式。大量调查单位的资料，经过统计整理，将资料系统化，形成一种统计数列并填在相应的表格内便形成一定的统计表。统计表的结构从内容上看主要是两部分：一部分是总体及其分组（通称作表的主词），另一部分就是说明总体的统计指标（一般称作宾词）。此外，还有相应的计量单位。从表的组成要素来看，主要有标题、标目、纵、横栏组成的表格及表中的数字。另外还应有必要的附注和注明资料来源。统计表的种类按用途可分为调查表、整理表或汇总表、计算表；按表的总体分组情况分类，可分为：简单表、分组表、复合表等。统计表是实际工作部门比较常用的统计方法，它能使统计资料条理化，便于比较各项目之间的关系，并易于检查数字的完整性。掌握统计表，对于青年研究和各类青年工作都具有重大意义。

## 集中趋势

是描述性统计的一个组成部分，它是关于中心值的计算数字。最常用的是算术平均数、中位数、众数及几何平均数。还有两种不常用的计量集中趋势数字，即调和平均数和修正平均数。集中趋势的这些计量数字较之任何其它的统计计量数字适用得更为广泛。集中趋势研究是对定距尺度作描述性统计的新方法。青年研究中，常常要运用集中趋势的研究方法。如比较二个群体青年的身高、体重状况，我们不可能用总和进行比较，而只能用平均身高和平均体重进行比较。

## 算术平均数

所谓算术平均数就是分值之和除以样本的总数所得之值。在社会统计学中常用 $\bar{X}$ 来表示，计算算术平均值的公式是，

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{N} = \frac{\sum_{i=1}^N X_i}{N}$$

其中 $X_1$ 表示第一个个体的分值， $X_2$ 表示第二个， $X_i$ 是泛指个体分值。N表示样本总数。对于未分组资料，算术平均数通常又分为简单算术平均数和加权算术平均数二种。简单算术平均数就是直接将总体各单位的标志值相加，除以总体单位数而求得平均数。公式如上。加权算术平均数又分为单变量数列和组距数列二种。当我们研究的对象包括了较大量的单位，而它们在各个组所出现的次数又不相同，首先需要将个体单位的资料编成变量数列，然后采用加权算术平均数的方法计算算术平均数。其公式是：

$$\bar{X} = \frac{\sum xf}{\sum f}, \text{ 其中 } f \text{ 是频数。}$$

利用组距数列计算算术平均数与用单项变量计算的算术平均数的方法基本相同，所不同的只是要利用组的平均数作为代表标志值进行计算。算术平均数适合于代数方法的演算。应用非常广泛，但它最容易受极端变量值的影响，而且受极大值的影响大于受极小值的影响。并且当组距数列有空端组时，组中点往往难以确定。在青年研究中，算术平均数的方法得到广泛应用，但也应注意其优缺点，灵活运用。



### 调和平均数

调和平均数是各个变量值倒数的算术平均数的倒数。在青年研究中，有时由于资料的原因，不能直接计算算术平均数，而以调和平均数的形式却能迂回算出算术平均数，得出同采用算术平均数法相同的效果。调和平均数法的计算公式是：假设N个变量值： $X_1, X_2, X_3 \dots X_n$ ；MH为调和平均数，则

$$MH = \frac{N}{\sum \frac{1}{x}}$$

如果我们遇到的是一个变量数列的资料，那么计算调和平均数就要以次数加权，

$$MH = \frac{\sum \frac{f}{x}}{\sum f}$$

调和平均数的应用特点是：容易受极端数值的影响，而且受极小数值的影响比受极大数值的影响还要大。当组距数列有开口组时，所遇到的问题与计算算术平均数是一样的，调和平均数应用的范围比较狭小。

### 几何平均数

几何平均数是若干项变量值的连乘积开项数方根的结果。其公式是：设有n项变量值 $X_1, X_2, \dots, X_n$ ，其几何平均数为mg，则有

$$mg = \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \dots X_n} = \sqrt[n]{nX}$$

在多项变量值的情况下，需要开多次方，为计算上的方便，可用对数， $lgmg = \frac{\sum lgm}{n}$ 。再由对数找出真数，即为几何平均数。几何平均数在应用时，受极端数值的影响较算术平均数和调和平均数为小；只要在被平均的数列中的数值有一项为零，就不能计算几何平均数；几何平均数主要适用于具有等比或近似等比关系的数列。我们进行青年问题的研究，会遇到各种各样的问题，故针对不同的问题，该采用不同的解决办法。

### 中位数

所谓“中位数”，就是位于数值序列中间位置的一个数值，其地位是，在此点之上恰有总次数的一半，而在此点之下有总次数的一半。当数组为奇数时，中位数即是这个数组中间的那个数；当数组为偶数时，中位数即是居中两个数的平均数。中位数也是一种代表值，在标志变异度较大的情况下，它可以避免极端数值的影响。对于未分组资料，正

数值已按大小排列好的情况下，中位数的项次等于  $\frac{n+1}{2}$ ，对分组资料则方法不同，由于有次数分布

的问题，所以需要采用累积次数制的方法。在此情况

下，中位数需要采用  $\frac{n}{2}$  或  $\frac{\sum f}{2}$  的公式来确定。中

位数的计算公式为：

$$Me = L \pm \frac{\frac{\sum f}{2} - S_{m-1}}{f_m} \times i$$

其中Me代表中位数

L代表中位数所在组的上（或下）限

$f_m$ 代表中位数所在组的次数

$S_{m-1}$ 代表中位数所在组以下（或以上）的累积次数

$\sum f$ 代表总次数

i代表中位数所在组的组距

由于中位数仅用次数而不用两端出现的开口组的组限，所以它能用于计算开口组的次数分配。此外，如果给的是百分比的分配，而不是实际数字的分配，中位数也能确定。所以，在青年研究中有时遇到不可能计算平均数或有极端数字的情况，中位数将会提供许多方便。

### 众数

在总体中出现次数最多的那个变量值既为众数。确定众数的方法，在单项变量数列中，出现次数最多的数值就是众数；对组距数列资料确定众数，从原则上讲依然是根据次数分配的多少来决定，在等距数列中，众数在次数最大的一组里面。如果是不等组距，则应先将它换算成标准组距的次数，经过换算后，次数最大的一组才是众数组。计算众数的公式为：

$$m_o = L + \frac{f - f_{-1}}{(f - f_{-1}) + (f - f_{+1})} \times i \text{ 下限公式}$$

$$m_o = u - \frac{f - f_{+1}}{(f - f_{-1}) + (f - f_{+1})} \times i \text{ 上限公式}$$

式中 $m_o$ 代表众数

L代表众数组下限

u代表众数组上限

f代表众数所在组的次数

$f_{-1}$ 代表众数所在组以下一组的次数

$f_{+1}$ 代表众数所在组以上一组的次数

$i$ 代表组距

众数的特点是最不受极端数值的影响；组距数列中出现开口组时，对众数也无影响；众数往往是不容易确定平均数。

### 全距

所谓全距是指数组中最高数值与最低数值之差。全距可以做为计量离差的一种标志。例如：将一群年轻人的年龄加总并计算出其平均年龄为28岁，同样另一群人的年龄加总计算出其平均年龄也是28岁，仅根据这二个平均数我们下结论说这二组年龄相似或事实上是等同的。但对二组详细分析就会发现，第一组青年群最小年龄可能20岁，最大可能32岁，其差距为12，而第二组最小年龄可能8岁，而最大可能58岁其差距为50岁。显然这二组虽平均年龄一样，但其实质年龄构成是不同的，经全距的计算便可分析其内部离差程度。另外，对于分组资料，在计算全距时应注意。第一种方法是求最高组与最低组下限之差；另一种方法是用最高组的组中值减去最低组的组中值。在实际工作中，这二种方法均可采用。全距是计量离差最简单的方法。但因其值仅取决于两个观察值，因而容易受极端数值的影响。

### 四分位差

四分位差(Q、D)，也叫做四分位数间距半值，类似全距，是一种按位置测定离差的计量方法。四分位差只取决于中间占次数50%的那部分差异距离，其实质是指第一个和第三个四分位数之间距离的半值。计算四分位差，需要测定两个计量数字：第一个四分位数(Q1)与第三个四分位数(Q3)。(Q1)是按次数定于使25%的观察值位于该点之下那一点的数值；(Q3)是按次数定于使25%的观察值位于该点之上那一点的数值。所以中间50%的观察值位于Q1与Q3之间。对于分组资料，计算四分位差的公式为：

$$Q_1 = L + \frac{\frac{n}{4} - CF}{f} \cdot i \quad (i)$$

$L$ 为第一个四分位数组的下限

$n$ 为总次数

$CF$ 含有第一个四分位数组之前所有各组次数的累积数

$f$ 为含有第一个四分位数组的次数

$i$ 为第一个四分位数所在组的组距。

$$Q_3 = L + \frac{\frac{3n}{4} - cf}{f} \cdot i$$

四分位数间距 =  $Q_3 - Q_1$

由于各样本的 $Q_3$ 和 $Q_1$ 变换不定，故而它比全距测量要稳定得多。而对不等组距的数据，四分位数可以比较容易地求得，也能对开口组的变量进行测量。

### 平均差

平均差是每一数值与平均值的绝对值差的算术平均数。它首先测定每一观察值距离样本或总体的平均数的偏差有多大，然后将那些与平均数(或者在很少时将那些与中位数)的离差进行总和(总和时不区别其代数的正负值)，并加以平均。对未分组资料，其公式为：

$$A. D. = \frac{\sum |X_i - \bar{X}|}{n} \quad \text{或} \quad \frac{\sum |X_i - Me|}{n}$$

对于分组资料，其公式为：

$$A. D. = \frac{\sum f |X_i - \bar{X}|}{n}$$

其中 $X_i$ 代表每组的组中值

$\bar{X}$ 指算术平均数

$f$ 指每组观察值的数目

$n$ 为观察值的总数

使用平均差来测量离散程度可避免全距和四分位差只考虑二个数值而丢失的信息，它采用了每一项观察值的信息。但它也有二个明显的弊端：一是绝对值的计算较麻烦，二是人们很难从理论上对此进行说明。因此我们从事青年研究，在使用平均差测量离散程度时，应考虑这一特点。

### 方差和标准差

方差是与总体平均数的离差平方的算术平均数。总体方差常用的符号是 $\sigma^2$ ，读作Sigma，其公式为：

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

其中的 $X$ 代表总体的所有数值， $\mu$ 为总体的平均数， $N$ 为总体观察值的总数。同样，来自总体的一个样本的方差为：

$$S^2 = \frac{\sum (X - \bar{X})^2}{n}$$



标准差是平均离差平方的算术平均值的开方。即：方差的平方根。它分为总体标准差 $\sigma$ 和样本标准差 $S$ ，其公式为：

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}}; S = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$$

对于分组资料，可令 $X_i = X_i - \bar{X}$ ， $X_i$ 表示平均数的商差，因此便可用如下公式计算标准差：

$$S = \sqrt{\frac{\sum_{i=1}^n X_i^2}{N}}$$

利用方差和标准差可以描述顺序或比率资料的离散程度。方差有很重要的统计理论价值，它与标准差的关系是可以相互导出的。

### 标准分数 (Z 分值)

当已知一组资料的平均数和标准差时，就可以表示出与平均数和标准差相关的单分数。这种分数变换叫做观测值的标准分数Z数值。通常，观测值i的标准分是：

$$Z_i = \frac{X_i - \bar{X}}{S}$$

使用标准分数的一个重要原因是：若我们把所有分数都换成了标准分数，所得的新分数中就有一个0平均数和一个标准差1，这样就不必在一组资料中找集中趋势和离散趋势了，用Z分数的方法也可以把资料标准化，使研究范围之列的变数的集中趋势和离散趋势都得到一个0平均数和一个标准差1。即可使与它们的平均数和标准差联系着的，并且已经标准化了的各个变数可以在“相同基点”上加以比较。

### 变异系数

变异系数是用标准差除以平均值，用V表示，

$$V = \frac{S}{\bar{X}} \cdot (100)$$

其用途在下例中便可见到。例如青年工人组平均月工资为105元，标准差为10，而老年工人组平均月工资为800元，标准差为50元。因为老年工人标准差50元远远大于青年工人的标准差10元，很容易使人认为老年工人的收入有更大的差异。但事实上，相反的结论才是正确的。

$$V_{青} = \frac{10}{105} (100) = 0.095 \times 100 = 9.5\%$$

$$V_{老} = \frac{50}{800} (100) = 0.063 \times 100 = 6.3\%$$

因青年工人的变异系数为9.5%，而老年工人仅为6.3%，说明青年工人工资的分配具有更大的离差。同样，当比较不同单位的分配时，比如工资和年龄，也可利用变异系数，青年研究中常常会遇到这类问题，变异系数会给我们提供良好的方法。

### 正态分布

正态分布通常称为高斯曲线，或叫做误差正态曲线。其公式为：

$$Y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-u)^2}{2\sigma^2}}$$

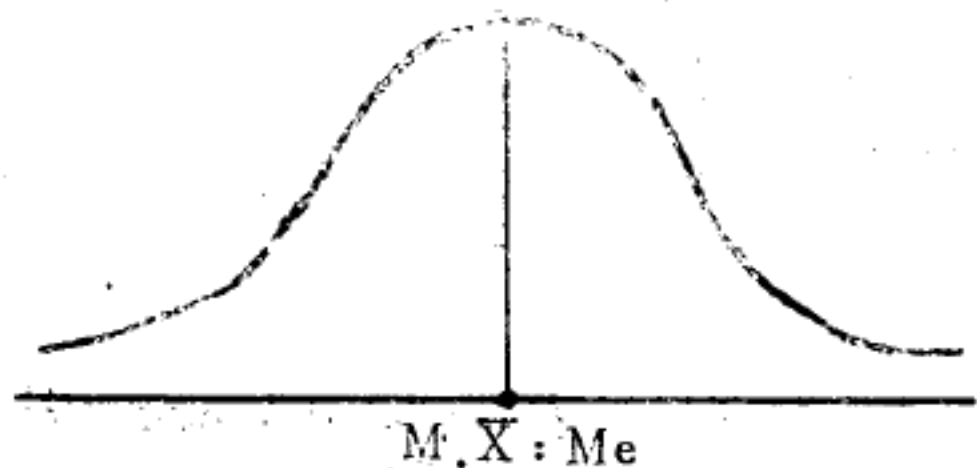
这里 $u$  = 平均数， $\sigma$  = 标准差， $\pi = 3.14159$ 。

$$e = 2.71828$$

只要有平均数 $u$ 与标准差，就有一个正态分布，因此，每一对平均数与标准差，就有一个正态分布。正态曲线是对称平滑曲线的一个特殊类型，又称为钟形曲线。因为其既对称又是单峰，它的平均值、中位数和众数重合一处，如下图。对于平均值为 $\bar{X}$ ，标准差为 $S$ 的X量的正态分布，通过用

$$Z = \frac{X - \bar{X}}{S}$$

的公式转换，便可得到平均值为0，标准差为1的正态分布，即标准正态分布。由于X量是无限的，利用公式求正态曲线非常困难，因此通常经过标准化，将变量变为标准值Z，然后根据Z表来求有关量。正态曲线下的面积是一个单位(1000)，曲线之下从平均数向左、右各占50%，曲线下的面积为曲线各高度所代表的次数的总和。从理论上讲，正态曲线两端是渐近的，但又永远不与X轴相交。总之，不管总体的形状如何(正态或偏态)，在青年研究中，选自总体的许多样本平均数的分布将为正态分布。正态分布将广泛应用于青年研究中。



## 统计推断在青年研究中的应用

统计推断是根据样本的统计量对总体的参数进行估计或作出决策,它可以帮助我们从一个别的调查结果中去认识规模巨大的总体。统计推断发挥着以下两个重要功能:估计总体的参数;检验总体中各变数之间的关系。由于仅是利用样本,所以推断中的风险和误差是不可避免的。这里需要的问题就在于明确地知道和把握这种风险,以便确定我们对于参数估计和变数关系估计的结论可以抱多大的信心。

统计推断依赖于关于常态曲线的知识。常态曲线是一条数理曲线,它由二个参数决定:一、平均值;二、标准离差。曲线下的面积就构成这些参数值发生的概率。当我们反复从总体中抽取规模固定的样本时,不论总体是否服从正态分布,样本指标构成的抽样分布总会近似地服从常态分布,其平均值就等于总体的参数值。在进行参数估计时,参数值本身总是未知的。假设检验需要构造理论假设和虚无假设,理论假设是虚无假设的备择假设。构造虚无假设的目的是为了否定它。当虚无假设被否定时,理论假设就更为可信了。重要的问题是使虚无假设尽可能地与理论假设相关,以便出现一个可能包含参数值的有限的区间。同样,假设也存在风险。若虚无假设为真而我们否定了它并肯定理论假设,就犯了弃真的错误;若虚无假设为假而我们决定接受它,就犯了纳伪的错误。所以重要的是如何将两类错误减少到最小程度。

### 点估计

点估计是从样本数据中选出一个被认为可能与参数值最为接近的样本指标值,由于参数和样本指标都有许多种,因此研究人员需要具备一些指导原则。一个估计若满足下列3个标准,就被认为是一个好的估计,即:无偏性;一致性;有效性。

如果抽样分布的平均值恰好等于被估计的参数值,那么这一估计便可被认为是无偏估计;样本规模越大,参数的估计值和参数就越趋于一致。这样的估计就是“一致估计”,某些有偏估计值,如标准离差和相关系数,都是一致估计,因为当样本规模增大时,他们就越来越与参数值接近;当不同样本的抽样指标值的变化范围相对较小时,参数估计就是有效的。一个估计的有效性是指抽样分布在真实的参数值周围集中的程度。如果估计是无偏

的,这一集中程度可由估计的标准误的平均值测得;标准误越小,估计的有效性就越强。在青年研究中,采用点估计的方法进行参数估计时,必须考虑以上3种原则。

### 区间估计

区间估计是一种最有用的推断方法,它能精确地告诉我们一个参数值被“框”在统计指标值的某一区间的可能性,并能告诉我们在这一区间之中的危险程度。区间一般是在点估计(如样本平均值)沿两个方向相对于选定的置信区间的标准误的乘积。区间估计是一个由样本指标的一系列邻近值构成的一个区间,而我们知道这个区间中可能存在参数值,同时也知道这种可能性的风险程度或不确定程度。前面已经指出,平均值标准差可以作为样本平均值抽样分布的标准离差, $\mu \pm 1\sigma \bar{X}$ 的意思是Z百分比的把握样本平均值落在这区间内。反过来,如果我们已经构造了一个关于样本平均值 $\bar{X}$ 两边对称但不一定关于总体平均值 $\mu$ 也两边对称的区间,如 $\bar{X} \pm 1.96\sigma \bar{X}$ ,我们有理由预测总体平均值有95%可能性会落入这一区间。然而,平均值的标准差 $\sigma \bar{X}$ 通常是不知道的,必须对其进行估计,平均值的标准差的点估计公式是:

$$S \bar{X} = \frac{S}{\sqrt{n}}, \text{ 其中 } S = \text{样本的标准离差, } n = \text{样本}$$

规模。于是我们可以利用平均值 $S \bar{X}$ 的标准误差估计值来构造一个关于样本平均值两边对称的区间,称为可信区间。其公式是:

$$\bar{X} + Z(1 \pm C)/2 \cdot \frac{S}{\sqrt{n}}. \text{ 此外,也可做相关系}$$

数的区间估计,首先利用“FisherI值换算”法把r的标准误差转化为Z的标准误差,可信区间的公

$$\text{式为: } Zr + Z(1 + C)/2 \cdot \frac{1}{\sqrt{n-3}} \text{ 再利用Z换}$$

利用区间换算出r的标准误差。

### 中心极限定理

中心极限定理是统计学中最重要的定理之一,它说明:

一、如果从一个总体中抽选出规模为n的全部可能的随机样本,并且计算出样本平均数,样本分布的平均数等于总体的平均数。即, $\bar{X} = \mu$ 。

二、从非常偏态的总体中抽选出来的样本平均数的分布,随着样本大小的增大,它会越来越接近



态分布。

三、从正态分布的总体中抽选出来的样本平均数的分布是正态分布。

四、平均数的抽样分布的标准误差是

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

在抽样调查中，往往利用样本标准差 $S$ 来估计总体标准差 $\sigma$ ，平均数的标准误差为 $S_{\bar{X}} = \frac{S}{\sqrt{n}}$ ， $S$ 代

表样本的标准差， $S_{\bar{X}}$ 代表平均数的标准误差，这里把 $S$ 看成是一个大样本，通常包括30个或30个以上的观察单位。

### 样本的容量

样本的容量就是一个样本所含的样本单位的数量，又叫样本大小，为使所研究的样本具有一定的精确性，必须掌握确定样本数量的规则。在研究中，有4个因素主要影响着样本的数量，它们是：总体的规模；估计的可信间距；可信系数 $t$ 和方差 $\sigma^2$ 。首先从总体规模看，当其它3个因素已知时，总体规模增大，样本的必需数量也有所增大，但是，当总体规模达到足够大时，样本的必需数量相对于总体来说只是受到微小的影响。再看可信间距和可信系数（ $C = 1 - \alpha$ ），如果已知总体规模和特征方差，那么所需样本的多少完全随可信系数的变化而变化，较大的可信系数要求较大的样本，样本量还受期望的可信间距的影响。总体特征的方差也影响样本的必需量。异质分子较多的总体需要较多的样本。可见，当我们知道以上4个条件时，便可正确地计算出一个多大量的样本是精确的，其计

算公式为： $n = \frac{t^2 \sigma^2}{\Delta \bar{X}^2}$ 其中 $n$ 为必要的样本数； $t$ 为

概率度对应的数值， $\sigma^2$ 为方差， $\Delta \bar{X}^2$ 为平均数的标准误。在青年研究中，常常要运用抽样调查的方法，那么究竟需抽多大的样本，便可用此公式的计算方法来确定，这样便可避免样本过小造成无代表性和样本过大造成浪费。

### 假设检验

假设检验的目的，是检验由观察得到的、表现为样本数据资料的各种关系在总体中的适用程度。这样推断是包含风险的，我们能做到的只是指出风险程度并说明其意义。判断一个理论假设成立与否总是包含着失误的风险，为了稳妥，把风险减至

最小程度，人们采用间接的方法进行假设检验。先是做出与理论假设相反的假设，然后为此而构造一个抽样指标的抽样分布。检验过程中要将样本的平均值（即观察值）与抽样分布的平均值（即预测值）加以比较。当样本的平均值与抽样分布的平均值相差很远，以至观测值在虚无假设下构造的抽样分布中发生机率不超过5%，那么我们就有理由认为虚无假设是不成立的，应予抛弃，因为在此假设下，所讨论事件发生的机率是如此之小。观测值的情况将导致人们采取另一种抽样分布，而如果理论假设恰是这样的分布，则对虚无假设的抛弃，能使我们更加相信理论假设的可靠性。假设检验的步骤大致有如下几种：

一、做出理论假设。这由 $H_1$ 代表。这种理论假设是由某种理论体系或一些理论观点推演出的。

二、作出虚无假设。用 $H_0$ 表示。构造这个虚无假设的目的是为了能够否定它。

三、在虚无假设的基础上为样本指标构造一个样本分布。

四、在做出虚无假设不能成立结论的过程中，要指明这种结论在多大程度上可能是错的。在区间估计中要确定一个可信系数 $C(C = 1 - \alpha)$ ，但在假设检验中，我们更注意 $\alpha$ ，称 $\alpha$ 为“显著度”，本质上，显著程度就是一种机率，即假如某一虚无假设能成立，某一特定样本指标指被实际观测到的机率，因此，如果规定的显著度水平很低，如0.05或0.01，并在实际上样本观测的结果也确实与此机率相适应，那么这种事件就被称为“小概率事件”。即在虚无假设下构造的样本分布中仍可能发生的事件；但是，由于其实际发生和被观测到的机率如此之小，我们就有理由相信观测值不是来自以虚无假设为主的抽样分布，并以此来判断估计失误的危险。又由于我们有理由认为最好是由理论假设去代替虚无假设，于是我们就倾向于更加相信理论假设的真实性。

五、分析和检验样本数据和决策过程。我们要收集资料、计算样本指标值，将观测值与虚无假设成立下的预期值进行比较，利用给定的显著度来决定是否抛弃虚无假设而更加相信理论假设的可靠性。

### 卡方检验

卡方检验，是在某些理论假设下，评价经验中所获得的频数是否显著地有别于期望值的一般性检验。这一检验应用于多种经验研究，尤其适用于两

个交叉跨类的定类尺度的列联问题。当然，在某些特定条件下，定序和定距的尺度也可利用 $X^2$ 检验，但必须是：第一，样本间是互斥的；第二，样本数量较大。 $X^2$ 检验的显著性水平是0.05。计算 $X^2$ 的公式为：

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

其中 $f_o$ 是观察次数， $f_e$ 是期望次数。其计算步骤

为：

一、算出每一对 $f_o$ 与 $f_e$ 的差（注：这差别的总和在逻辑上必须等于0）。

二、将其差平方。

三、每一差的平方被期望次数 $f_e$ 除。把这些商的结果加总，其总和既 $X^2$ 的计算值。计算出的 $X^2$ 值与 $X^2_{0.05}(k-1)$ 的值进行比较，（其中的 $k-1$ 为自由度），如下图：



注：由 $X^2_{0.05}(k-1)$ 查表得到临界值。

若 $X^2 > X^2_{0.05}(k-1)$ ，则说明结果落在否定域内，因此拒绝虚无假设 $H_0$ ，接受与 $H_0$ 相反的假设。若 $X^2 \leq X^2_{0.05}(k-1)$ ，则接受虚无假设。利用 $X^2$ 检验也有一定的限定性，若一个单元中期望次数特别小，应用卡方检验可能会出现错误结论。因此对于单元次数必须遵循：第一，若只有二个单元，则每个单元的期望次数至少是5及5以上；第二，若是有两个以上的单元，如果有20%的单元次数小于5，则不能应用 $X^2$ 。在青年研究中，我们时常会遇到检验经验资料符合期望值的程度，因此经常要利用 $X^2$ 检验进行分析。

## 方差分析

如果在比较中包含着3个或更多组别，t检验就有困难，因为包含着个别样本方差总数的样本分布预计误差会由于组数之多变大。选择检验程序时一般用F分布，这叫“方差分析”，它需要t检验中应用同样假定：独立的和随机的样本，在每一总体中的依变数为常态分布，方差的变方共型。方差分析的特点是用总体中依变数的方差信息来检验虚无假设。它的基本原理是：如果理论假设正确，那么，如果自变数使依变数产生差别，我们应该看到在依变数个人之中的误差（方差）小于总人数之间的误差（方差）。但是虚无假设要求不应如此，因为自变数和依变数没有不同。根据研究者的观点，

自变数应使依变数发生变异，因此，平均值 $V_b$ 间的方差，表现出那些对应依变数的自变数的效果，这时，运用平均值之间的方差就能测定自变数对依变数的“解释程度”。平均值间方差愈大，自变数解释能力也愈大，直到依变数受到重视为止。使用两个方差比例的F检验公式为：

$$F = \frac{MS_b}{MSW} \text{ 对应 } J-1 \text{ 和 } n-1 \text{ 自由度。}$$

这里的 $J$ =组数， $n$ =样本总体数。

方差分析主要用于定类（或定序尺度）同一个定距尺度之间的关系，也可用于对一个定距尺度同两个以上的定类尺度关系的检验。

## 相关分析

相关分析是研究相关关系的。相关关系是客观现象之间相互关系的一种形式。它有两个特点：

一、相关关系是指现象之间确实存在的数量上的相互依存关系。其中作为根据的变量叫做自变量，发生对应变化的变量叫做因变量。其中这种依存关系的具体关系值不是固定的。相关分析的目的在于帮助我们对关系的密切程度和变化的规律性有一个具体的数量上的认识，以便对改进工作做出某种判断，并且用于各种推算和预测。其主要内容是：第一，确定现象之间有无关系，确定相关关系的表现形式。这是相关分析的出发点。有相互依存关系才能用相关方法进行分析。第二，确定相关关系的



密切程度。相关关系是一种数量关系不严格的相互依存关系，相关分析的一个目的就是在这种不严格的关系中想办法来判断它们之间关系的密切程度。其主要方法是绘制相关图和计算相关系数。第三，测定两个变量之间的一般的关系值。自变量和因变量的数量关系本来是不严格的，自变量发生变化，会有许多个因变量的数值与它相对应。相关分析的一个目的就是要测定它们之间的一般性关系。这种分析要使用数字方法，如果现象之间的关系表现为直线相关，则采用配合直线的方法。使用这种方法能使我们找到现象之间相互依存关系的数量上的规律性。这是进行判断、推算、预测的根据。第四，测定因变量估计值和实际值之间的差异，用来反映因变量估计值的可靠性。相关分析的作用表现为以下几点：第一，通过相关图、相关系数等方法可以帮助我们判断现象之间有没有关系，密切程度如何，哪些因素是主要的、哪些是次要的，一个现象的数量发生变化，另一现象将会对应地发生什么变化等等，而且所有内容都用数量表现出来。第二，用于预测和预报，这是相关分析的主要作用。相关分析是根据过去的实际资料所进行的概括和总结，但一旦找到它们之间数量上变化的规律性，便可用来预测未来。第三，用于补充缺少的资料。在应用相关分析时，应注意以下几个方面：相关方法本身不能判断现象之间有没有关系，也不能解释相关关系产生的原因，同时要注意相关关系发生作用的范围；在使用相关方法时，最好是回归方程和相关系数或估计标准误差同时使用；同时，运用相关方法进行推算或预测时要注意条件的变化。

### 相关系数

相关系数是直线相关条件下说明两个现象之间密切程度的统计分析指标。根据简单相关表计算相关系数的方法为：a：计算自变量数列的标准差，

公式为  $\sigma_x = \sqrt{\frac{1}{n} \sum (X - \bar{X})^2}$ ；b：计算因变

量数列的标准差  $\sigma_y = \sqrt{\frac{1}{n} \sum (y - \bar{Y})^2}$ ；C：两

个数列的协方差  $\sigma^2_{xy} = \frac{1}{n} \sum (x - \bar{X})(y - \bar{Y})$ ；

根据以上材料便可求出相关系数r，公式为：

$$r = \frac{\sigma^2_{xy}}{\sigma_x \sigma_y}$$

$$= \frac{\frac{1}{n} \sum (x - \bar{X})(y - \bar{Y})}{\sqrt{\frac{1}{n} \sum (x - \bar{X})^2} \cdot \sqrt{\frac{1}{n} \sum (y - \bar{Y})^2}}$$

根据计算的r值，判断相关关系的密切程度，大致标准为：0.3以下算无相关，0.3以上算有相关，具体为0.3~0.5是低度相关；0.5~0.8是显著相关；0.8以上是高度相关。

### 相关关系的种类

相关关系有不同的种类，这和研究方法有关系。从相关关系涉及的因素多少，可分为单相关和复相关。两个因素之间的相关关系叫单相关，3个或3个以上因素的相关关系叫复相关。从相关关系的表现形式讲，可以分为直线相关和曲线相关。直线相关按照变化的方向又可分为正相关和负相关，当自变量X的数值增大或减少时，因变量y的数值若随其增大或减小，叫做正相关，若反之，y的数值减少或增大，则为负相关。

### 回归分析

回归分析是根据一个变量的适当的精确性来预计另一变量。它要求被分析的变量为定距尺度。回归分析与相关分析密切相联，是在相关分析的基础上，确认两变量具有相关关系，为进一步分析两变量的关系而采用。为找出两相关变量之间的规律性，一般多用配合一条直线的方法，以此代表现象之间的数量关系，这条直线在数学上叫回归直线，表现这条直线的数学公式叫回归方程。回归直线运用的前提条件是：首先，现象之间确实存在数量上的相互依存关系；其次，现象之间存在直线相关关系；最后，要有一定数量的自变量与因变量的对应资料，而且要确定哪个是自变量、哪个是因变量。当具备这些条件后，我们便可采用回归直线法。回归直线的确认一般是采用“最小二乘法”，其标准是我们所发现的直线具有所有y值偏离直线的离差平方和为最小的性质，即正负距离之和为零，并且各点到直线的标准差最小。回归方程的数学表达式为： $y_c = a + bx$ ，其中a为直线的起点值，即坐标轴纵轴的截距；x代表自变量及其变量值；b代表自变量增加一个单位时，因变量的平均增加值，数学上一般称其为斜率或回归系数； $y_c$ 就是推算出来的直线上的估计值。回归方程求得的关键是找出a、b二个值，方法很多，最基本的公式是：

$$\begin{cases} b = \frac{n \sum xy - (\sum x)(\sum y)}{n \cdot \sum x^2 - (\sum x)^2} \\ a = \bar{y} - b\bar{x} \end{cases}$$

在具体研究中, 根据已掌握的资料, 可灵活运用各种关系式推导的简捷公式, 在此不一一列举。相关和回归是社会科学领域研究二变量或多变量关系的使用最普遍的方法, 要使青年研究数量化、科学化, 这种研究分析问题的方法是不可忽视的。

### 估计标准误差

估计标准误差是用来说明回归方程推算结果的准确程度的统计分析指标, 或者说是反映回归直线代表性大小的统计分析指标。因为回归直线是在直线相关条件下, 反映两个变量之间一般数量关系的平均线, 以此所推算的因变量仅是一个估计值, 和实际值有出入, 便产生了估计标准误差。其计算方法有两种, 其一是根据因变量的实际值和估计值的

离差计算, 公式为:  $S_{yx} = \sqrt{\frac{\sum (y - y_c)^2}{n}}$ ,

其中  $S_{yx}$  为估计标准误差,  $y$  是因变量数列的实际值,  $y_c$  是根据回归方程推算出来的估计值。  $n$  为因变量的次数。另一计算方法是根据  $a$ 、 $b$  两个参数值计算

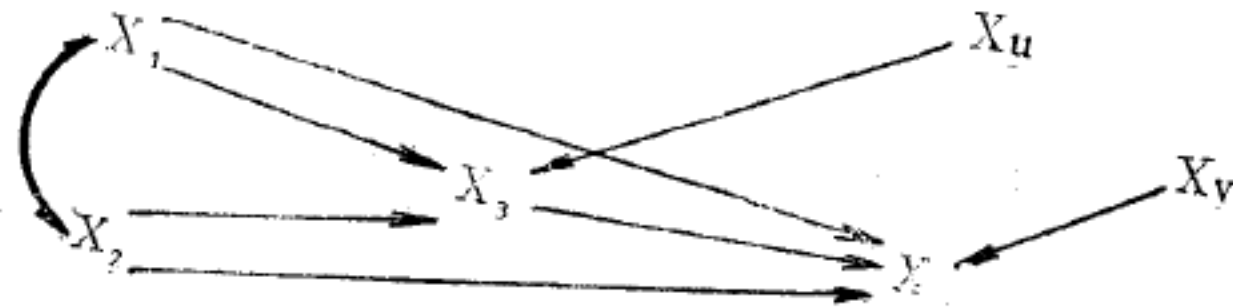
估计标准误差。公式为:

$$S_{yx} = \sqrt{\frac{\sum y^2 - a(\sum y) - b(\sum xy)}{n}}$$

估计标准误差是结合回归直线而使用的, 其作用表现为: 第一, 说明以回归直线为中心的所有相关点的离散程度。第二, 说明回归直线的代表性的大小。第三, 在抽样调查条件下, 它是计算回归抽样误差的一个根据。在实际研究中, 我们也可以用相关系数在推算估计标准误差, 其公式为:  $S_{yx} = \sigma_y \cdot \sqrt{1 - r^2}$ , 其中  $\sigma_y$  是因变量数列的标准差。估计标准误差和相关系数表现为相反的关系, 即  $r$  值越大,  $S_{yx}$  值越小。因为  $r$  值越大, 说明相关程度越密切, 也就是相关点距回归直线比较近, 自然  $S_{yx}$  越小了。相反,  $r$  值越小,  $S_{yx}$  值则越大。当  $r = 0$  时,  $S_{yx} = \sigma_y$ , 即估计标准误差等于  $y$  数列的标准差。可见, 估计标准误差从另一角度说明了相关关系的密切与否。

### 路径分析

路径分析是作为回归模型的一种简单扩大, 使两个或更多的依变数具体化。如研究者要估计青年职业选择中其性别、受教育程度的影响, 及这 3 个因素其对收入的影响, 便可用以下路径图表示:



其中  $X_1$  = 性别;  $X_2$  = 受教育程度;  $X_3$  = 职业;  $X_4$  为收入。  $X_u = X_3$  的剩余变数;  $X_v = X_4$  的剩余变数。显然变量  $X_3$  依赖于  $X_1$ 、 $X_2$ 、 $X_u$ , 而  $X_4$  又依赖于  $X_1$ 、 $X_2$ 、 $X_3$ 、 $X_v$ 。说明  $X_1$ 、 $X_2$  为自变量,  $X_3$  为双重角色, 既是  $X_1$ 、 $X_2$ 、 $X_u$  的因变量, 又是  $X_4$  的自变量, 而  $X_4$  在此为因变量。路径分析中, 各种关系的确定常与因果分析相一致, 在很大程度上依靠理论思维。对于一个完整的路径模型来说, 对其路径系数的计算与对其回归系数的计算是一样的。路径系数实质就是“标准回归系数”。在一个理论模型的构造中, 要排除模型大小的统计影响, 其规则是: 当路径系数小于 0.1 时, 这条路径将被排除, 估计要重新计算。

### 对统计资料的解释

进行一次青年问题的调查, 资料分析完毕后,

研究人员要根据研究的理论目的与实践目的, 对资料加以解释, 在解释分析结果过程中, 必须考虑到进行研究的各种限制条件以及所采用的统计技术的各种假设。对于没有向总体进行推断的描述性资料, 解释仅限于对研究所涉及到的那一组被调查者的资料进行全面的、精确的说明。一般描述性分析结果的解释是很简单的, 而推断性分析结果的解释却十分复杂。需要指出的是, 研究人员必须对预料到和预料不到的差误以及在研究过程中出现的差异对研究结果的影响, 给予客观的评价和真实的报告, 并将此公布与众。同时, 对涉及到各种变量之间关系的问题也必须在解释中阐述明了, 如组合、联附、关联或功能关系; 共变关系与因果关系; 重要相关关系与显著度等。对能否确定变数间关系还需作出经验性检查, 如在考虑到自变数对依变数作用后, 对依变数未得到解释的变化或残差作出评价



等,即使在经验资料已经肯定了变数间关系的情况下,概念间关系的相关性也还需进一步的证实。最后,在解释调查结果的过程中,调查者必须清楚地

介绍测量及调查过程、资料搜集以及资料分析所采用的步骤,以使读者可以判断出变数与概念间的已知相关关系适当或不适当。

## 青年研究机构

### 全国青年思想教育研究中心

1987年5月19日成立,由中共中央宣传部、共青团中央主管。该中心是由社会各界青年教育工作者组成的,进行青年思想教育、宣传和理论研究的民间团体。研究中心继承和发扬党的思想政治工作的优良传统,自觉地服从党的政治领导,和接受马列主义、毛泽东思想的指导,始终把培养和造就有理想、有道德、有文化、有纪律的一代新人作为研究和解决青年思想问题的准则。该中心成立以来,组织了各种专题研讨活动,并组织研究报告员开展演讲报告、对话座谈、调查研究、学术交流等活动。

全国青年思想教育研究中心主任是团中央第一书记宋德福,副主任是刘延东、朱伯儒、曲啸、刘吉、李燕杰、张海迪、赵喜明、柳斌杰、徐良、魏英敏等28位专家、学者、英雄人物、青年工作者被聘为兼职研究报告员。

### 共青团中央研究室

1983年创建,由中国共青团中央委员会主管。研究室是在共青团中央书记处的领导下,从事指导、协调全国青年和青年工作理论研究、政策研究的学术性研究机构。研究室的主要任务是:围绕青年的思想动向,进行经常性的调查研究,为青年的健康成长创造良好的条件;围绕团组织的中心工作,抓好当前重大青年研究课题的研究调查,为团组织的发展服务;围绕团的改革,开展政策性研究,为团中央书记处提供有关的决策依据。研究室多年来先后开展了农村改革与团的工作;企业团组织发展对策;观念更新与思想工作;大学生全方位透视;青年结婚消费;共青团体制改革研究等问题的系列调查。出版了《青年工作调查文集》、《中国青年工作年鉴》、《共青团工作手册》等。研究室编辑发行《中国共青团》、《团的情况》等刊物和信息材料。

共青团中央研究室现在研究人员17名,其中高级职称1人,中级职称10人。现任负责人:吴黔生。

办公地址:北京市前门东大街10号楼

电话:757201—275、276

邮政编码:100051

### 中国社会科学院社会学研究所青少年研究室

1985年1月成立,原为青少年研究所。1985年并入社会学研究所。该室是专门从事国内外青少年教育和青少年问题研究的科学研究机构。自成立以来,与日本青少年研究所合作开展了《中日青年职工劳动意识》的调查研究,并在北京举行了第二次青年研究学术讨论会进行交流。近年来该室与全国各地青少年研究机构联合召开了多许专题研讨会,对青少年教育和研究起到了推动作用。现在,该所承担了国家哲学、社会科学“七·五”重点科研课题《改革开放中的青年价值观和社会问题》。自1981年以来,研究室主办《青年研究》一刊,在理论界、青年界有广泛的社会影响。

中国社会科学院社会学研究所青少年研究室现有研究人员10人,其中高级职称4人,中级职称5人。现任研究室主任:楼静波,副主任:单光肅。

办公地址:北京建国门内大街5号中国社会科学院

电话:5137744—2580

邮政编码:100732

### 辽宁省青少年研究所

1988年5月成立,由共青团辽宁省委员会和辽宁省社会科学院主管。研究所是从事青少年教育和青少年社会问题研究的科研机构。该所以共青团改革和战略研究;青少年成长规律探索;青年学科理论建设为主要研究任务。研究所成立以来,参与了有关《辽宁省未成年人保护条例》的制定工作,创办了青年理论研究刊物《当代青年思潮》,进行了《辽宁省青少年状况调查》和《辽宁省青年企业家专题调查》,组织发表了十几篇各类专业论文,编发了《青年研究信息》等。

辽宁省青少年研究所现有研究人员12人,其中中级职称5人。现任研究所负责人:李有光。

办公地址:辽宁省沈阳市沈和区文二路二段四