

第三章 社会调查中的问卷 设计

本章内容

第一节 社会调查概述

第二节 选择调查课题

第三节 调查设计

第四节 抽样

第五节 测量与操作化

第六节 问卷与问卷设计

第七节 问卷的结构

第八节 问卷设计的步骤

第九节 问卷设计的原则与标准

第十节 问题和答案的设计

第十一节 问卷设计中的其他问题

第十二节 问卷设计中常见错误分析

第十三节 资料处理

第一节 社会调查概述

一、社会调查的定义和特征

1. 定义

在社会调查概念的具体定义上，不同学者之间的认识不尽相同。比如：

- 有的在比较抽象的层次上把社会调查看作人们认识社会的一种实践活动；而有的则只在十分具体的层次上把它看作社会研究中的一种资料收集方法；
- 有的从非常宽泛意义上把到社会中了解情况的各种不同形式的活动都冠以社会调查的名称；而有的则仅仅只在十分狭义的范围把那种以自填问卷和结构式访问的方法、从一个随机样本那里收集资料的工作称作社会调查；
- 有的认为社会调查的全部工作任务只是收集资料；而有的则认为社会调查既包括收集资料的工作，也包括分析资料的工作……

总而言之，从目前的情况看，社会调查的概念无论是在外延上还是在内涵上都很不一致、很不统一。

本门课所介绍的社会调查，指的是一种采用自填式问卷或结构式访问的方法，通过直接的询问，从一个取自总体的样本那里收集系统的、量化的资料【第三章】，并通过对这些资料的统计分析【第四章】来认识社会现象及其规律的社会研究方式。

第一节 社会调查概述

一、社会调查的定义和特征

2. 基本要素

- ① **抽样**。所谓社会调查，实际上指的就是抽样调查，即调查对象始终只是总体中的一个部分，或者说，是从总体中的一个样本（总体中的一部分个体）那里收集资料，通过对“样本”这个部分进行调查来达到了解总体的目标。
- ② **问卷**。社会调查主要采用自填式问卷与结构式访问两种方法收集资料，而问卷则是这两种资料收集方式中都须采用的、至关重要的工具。正是问卷的使用，决定了社会调查主要依靠对被调查者进行特定方式的询问来收集资料的本质特征，也决定了社会调查所获得的资料是一种易于量化的资料的特征。
- ③ **统计分析**。无论是抽样的使用，还是问卷的方法，在一定意义上都是在为最终的资料分析做准备。而以统计学原理为基础的统计分析，则是社会调查在分析资料这个环节上的主要特征，也是社会调查作为一种定量的社会研究方式的重要标志。

3. 方法论特征

- ① 社会调查在本质上是一种定量的研究方式。这是从方法论角度看待社会调查时应该注意的一个最重要的特征。
- ② 社会调查是一种横剖性的社会研究方式。这是社会调查在时间维度上的基本特征。**【如果条件允许，也可以对（同一）样本进行多轮调查，此时就变成longitudinal而非cross-sectional数据】**

第一节 社会调查概述

二、社会调查的类型和题材

1. 分类

- ① 根据调查对象的范围，可以分为普遍调查与抽样调查
- ② 根据收集资料的方法，可以分为自填问卷调查和访问调查
- ③ 根据调查的目的或作用，可以分为描述性调查和解释性调查

2. 题材

- ① 某一人群的社会背景：有关人们各种社会特征的资料，回答“他们是谁”的问题。这种资料既包括某些人口统计方面的内容，比如性别、年龄、职业、婚姻状况、文化程度等，也包括人们生活环境方面的内容，比如家庭构成、居住形式、社区特点等。这类题材客观性很强，在社会调查中收集这方面的资料往往比较容易，较少出现问题。几乎所有的社会调查都或多或少地包括这一题材中的内容。
- ② 某一人群的社会行为和活动：有关人们“做了些什么”以及“怎样做”等方面的资料。比如人们每天几点钟上班、每周看几次电影等。这类题材也是客观的、事实性的，通常构成大部分社会调查的主体内容。
- ③ **某一人群的意见和态度**：有关人们“想些什么”、“如何想的”或“有什么看法”、“持什么态度”等方面的资料。比如人们怎么看待离婚现象、人们选择对象的标准是什么等。这类题材属于观念性、主观性的，是构成各种民意测验、舆论调查、社会心理调查的主要内容。

第一节 社会调查概述

三、普遍调查与抽样调查

1. 普遍调查

简称普查，指的是对构成总体的所有个体无一例外地逐个进行调查。一般来说，普遍调查的规模往往非常大，属于宏观的调查。普查资料是从总体中的所有个体那里收集的，因此是了解国情、省情、市情和大规模总体概况的最基本的资料，是各级政府部门制定各种政策的重要依据，也是各种科学研究尤其是社会科学的重要参考资料。**普遍调查主要采取两种方式进行。**一种称为统计报表，是由普查部门（通常是国家行政部门，例如国家统计局等）制定普查表，由下级有关部门根据所掌握的资料进行填报。另一种是建立专门的普查机构，组织专门的调查员，制定专门的调查表，对总体中的每一个成员进行直接的调查登记。

2. 抽样调查

是从所研究的总体中按照一定的方式选取一部分个体进行调查，并将在这部分个体中所得到的调查结果推广到总体中去。简而言之，就是调查部分以反映整体，这就是抽样调查的基本思想和逻辑。抽样调查是在20世纪初期，随着抽样理论、统计分析方法、问卷技术，以及计算机技术的发展、完善和普及而逐步发展起来的。抽样方法与问卷法、统计分析一起，成为现代社会调查方法的主要标志。

与普查相比的优点：节省时间、人力和财力；可以比较详细地收集信息，获得内容丰富的资料；应用范围更广泛；准确性高。

第一节 社会调查概述

四、社会调查的一般程序

- 1. 选题：主要任务包括从现实社会大量现象、问题和焦点中恰当地选择一个有价值的、有创新的和可行的调查问题；将比较含糊、笼统、宽泛的调查问题具体化和精确化，明确调查问题的范围，理清调查工作的思路。
- 2. 准备：进行调查设计，包括思路、测量、方式方法和具体技术。
- 3. 实施：执行调查设计，收集资料；如果调查设计与实际情况之间存在偏差，需要根据实际情况进行修正或弥补。
- 4. 分析：将收集到的原始资料进行系统的审核、整理、统计、分析。
- 5. 总结：撰写调查报告（或学位论文、期刊论文）。

选题	准备	调查	分析	总结
选择问题			自填问卷	
	调查设计	邮寄问卷	资料整理	
	抽取样本	集中填答	单变量分析	报告结果
	变量测量	网络调查	双变量分析	
	问卷设计	结构访问	多变量分析	
		电话访问		

第一节 社会调查概述

五、资料收集方法

1. **自填问卷法**指的是调查者将问卷发送给（或者邮寄给）被调查者，由被调查者自己阅读和填答，然后再由调查者收回的资料收集方法。这种方法可以说是现代社会调查中最常用的一种资料收集方法。
2. **结构访问法**是指调查者依据问卷，采取口头询问的方式，向被调查者了解社会情况，收集有关社会现象资料的方法。

这两个大的类别，根据具体操作方法的不同，又可以进一步划分出不同的子类型：

- **自填问卷法**又可分为个别发送法、邮寄填答法、集中填答法和网络调查法；
- **结构访问法**又可分为当面访问法与电话访问法。

第一节 社会调查概述

五、资料收集方法

① 个别发送法

个别发送法是自填问卷法这一大类中最常用的一种。它的具体做法是：研究者将问卷印制好以后，派调查员依据所抽取的样本，将问卷逐个发送到被调查者手中，同时讲明调查的意义和要求，请他们合作填答，并约定收回问卷的时间、地点和方式。

优点：个别发送法既不像邮寄法那样与被调查者完全不见面，又不像当面访问那样与每一个被调查者都交谈相当长的一段时间，而是介于二者之间，较好地处理了调查的质量与数量之间的关系。它比较节省时间、经费和人力；调查员可以向被调查者进行解释和说明；可以保证比较高的回收率；调查具有一定的匿名性；可以减少调查员所带来的某些偏差；被调查者有比较充分的时间对问卷进行阅读和思考，还可以在方便的时候进行填答等。

缺点：调查范围受到一定的限制，不如邮寄填答法那么广泛；问卷的填答质量依然不能完全得到保证等。

总的来说，个别发送法的优点相对多一些，而缺点相对少一些，是一般的社会调查中最应该选用和推广的资料收集方法。

第一节 社会调查概述

五、资料收集方法

② 邮寄填答法

邮寄填答法是社会调查中一种比较特殊的资料收集方法。它的一般做法是：研究者把印制好的问卷连同已写好回邮地址和收信人姓名且贴好足够邮资的信封寄给被调查者，被调查者填答后再将问卷寄回调查机构或调查者。

优点：便宜、方便、快捷。

缺点：它需要有调查对象的地址和姓名，而这在实际调查中常常难以得到；**问卷的回收率难以保证，这是它最致命的弱点。**

第一节 社会调查概述

五、资料收集方法

③ 集中填答法

在条件允许的情况下，我们也可以采取集中填答法来收集调查资料。集中填答法的具体做法是：先通过某种形式将被调查者集中起来，由研究者统一讲解调查的主要目的、要求、问卷的填答方法等事项，然后将问卷发给被调查者，请他们当场填答并当场回收。

集中填答法除了具备一些与个别发送法和邮寄填答法相似的优点外，在某些方面的优点似乎比上述两种方法还要明显和突出。比如它比个别发送法更为节省调查时间、人力和费用，比邮寄填答法更能保证问卷填答的质量和回收率。

集中填答法最主要的局限在于：许多社会调查的调查样本根本不可能集中；同时，将众多被调查者集中在一起，有时会形成某种不利于个人表达特定看法的“团体压力”或“相互作用”，这也是在运用集中填答法时需要注意的一个问题。

第一节 社会调查概述

五、资料收集方法

④ 网络调查法

网络调查法指的是研究者利用互联网向特定对象发送调查问卷，同时也通过互联网将被调查者填答好的问卷收回的调查方法。常见的网络调查方式有三种。

第一种方式是将调查问卷直接链接在网站的网页上。任何一个上网者只要点击该网站的网页就会弹出调查问卷供上网者进行填答。当上网者填答完毕后，问卷数据就自动存入事先设计好的数据文件中。当调查结束时，所有填答者的回答记录就形成了该调查的数据库。这种方式的网络调查虽然十分便利，**但是由于它实际上是一种无特定调查样本和对象的调查方式，同时，上网者是否填答问卷也完全处于一种放任的或完全自愿的状态，因而其调查的对象性质、调查的回收率、调查的质量等均得不到很好的保证，其结果往往具有较大偏差。**所以，这种调查方式较少为学术研究者所利用，较多为非学术研究的大众媒介所采用。

第二种方式也是将问卷链接在特定的网页上，但与上面一种方式所不同的是，它不是一种面向所有上网者自由浏览、无特定对象和总体的调查，**而是一种针对研究者所选定的特定对象进行的调查。**一般情况下，这种做法是先确定调查总体（调查总体中的成员必须有电子邮箱），然后抽取好调查对象的样本，并收集到他们的电子邮箱地址。之后分别给样本中的调查对象发电子邮件，说明调查目的、调查要求，告知调查方法，并附上调查问卷所在的链接地址。被调查者点击链接后就会进入调查问卷并直接在网上填答。填答结束后，问卷数据自动存入事先设计好的数据文件中。全部调查结束后，所有填答好的问卷资料就自动生成调查的数据库。

第一节 社会调查概述

五、资料收集方法

④ 网络调查法

第三种方式是研究者在确定好调查总体（也是有电子邮箱的对象）、抽取完调查样本、收集好被调查者的电子邮箱地址后，直接将调查问卷用电子邮件发送给被调查对象。被调查者打开问卷的电子版在计算机上进行填答，填答完毕后又通过电子邮件将问卷发回给调查者。调查者将所有填答好的问卷下载后进行汇总，形成数据库文件。

后两种方式主要在学术研究中运用。由于这两种方式都是建立在严格的随机抽样基础上形成的样本和调查对象，因此，在问卷回收率得到保证的前提下，其调查效果也与前面介绍的资料收集方法完全一样。而由于事先通过电子邮件（也可以通过电话）与被调查者取得联系，并征得了被调查者同意，因而调查问卷的回收率也有一定保证。

网络调查的最大优点是方便快捷、节省费用。

网络调查的不足主要体现在调查对象的范围有一定的局限上，即它只能调查那些有上网条件、同时也会上网的对象。

第一节 社会调查概述

五、资料收集方法

⑤ 当面访问法

当面访问法的基本做法是：调查员携带着访问问卷分赴各个调查地点，按照调查方案的要求，对所选择的被调查者进行访问，并按照问卷的格式和要求记录被调查者的各种回答。在访问中，调查员严格依据调查问卷提出问题，并严格按照问卷中问题的顺序来提问，调查员不能随意改变问题的顺序和提法，也不能随意对问题作出解释。答案的记录也完全按问卷的要求和规定进行。

优点：

第一，能够对调查过程加以控制，从而提高调查结果的可靠程度。第二，回答率高。第三，它可以对调查资料的效度与信度进行评估。

缺点：

第一，调查费用高。第二，所花时间长。第三，受上述两方面缺点的影响，其调查的范围和规模往往受到很大局限。第四，在当面访问中，由于调查员的不同、提问方式的不同、交谈进展情况的不同等原因，常常会产生出各种访问的偏见，形成一定的访问误差。第五，对某些较敏感问题进行当面访问，很多被调查者的思想压力就可能很大，顾虑就可能比较多。所有这些显然会直接影响到他们回答问题的态度和所提供的答案的真实性及可靠性。

第一节 社会调查概述

五、资料收集方法

⑥ 电话访问法

电话访问法即通过电话对被调查者进行访问的方法。它与当面访问法在许多方面都比较相似，但二者之间有一点差别十分明显：当面访问中，调查员不仅可以听取被调查者的回答，还可以观察到被调查者的表情、动作，以判断被调查者所提供资料的正确性和真实性；而在电话调查中，调查员必须完全依靠自己的听力来判断这一切。

电话访问法的主要优点是简便易行、省时、省钱。特别是对于内容比较简单的调查，电话访问法的效果更好。

电话访问法的主要问题是调查者的选取及代表性方面的困难。因此，在运用电话访问法时，研究者一定要对总体及样本的情况有清楚的认识，尽可能做到抽样的科学性与代表性。

第二节 选择调查课题

一、调查课题的类型

各种调查课题都可以按其关注点或侧重点，分为理论性课题和应用性课题两大类。

理论性课题（theoretical problem）指的是那些侧重于发展有关社会世界基本知识，特别是侧重于建立或检验各种理论假设的课题。这类课题力图理解和解释社会世界的某一方面（或其中某一部分）是如何运转和相互联系、某一类社会事物或社会现象又是如何发生、发展和变化的。这类课题往往表现出十分明显的理论倾向，其关注点主要在于探索现象之间的因果关系，在于增加对具体社会现象所具有的内在规律的认识。这类课题的出发点和落脚点都是理论知识，或者说，其主要目标是要增加人们对社会现象的内在规律的理解，增加人们对社会事物的认识。

应用性课题（applied problem）是指那些侧重于了解、描述和探讨某种社会现实问题或者针对某类具体社会现象的课题。这类课题的关注点通常比较集中地体现在迅速地了解现实状况，分析现象或社会问题形成的原因，并力图在此基础上有针对性地提供政策建议，以帮助解决社会问题，帮助制定社会政策，以及帮助评估社会后果等。概括地说，这一类课题侧重于通过调查来解决问题。与理论性课题相比，应用性课题的出发点和落脚点都是现实问题，从目前国内的情况看，在社会学以及相关社会科学的研究者中，从事这类课题研究的比例大于从事理论性课题研究的比例。同时，各级政府机构和上述各种实际工作部门所做的各种调查，还有各种类型的市场调查，也基本上都属于这一类调查课题。

第二节 选择调查课题

二、选题的重要性

1. 决定调查的方向（计划做什么、能得到什么成果）
2. 体现调查的水平（理论依据、现象的重要性）
3. 制约调查的过程（应该如何做）
4. 影响调查的质量（是否能做好）

第二节 选择调查课题

三、选题的标准

1. 重要性

重要性是指调查课题所具有的意义或价值。通俗地说，就是指一项调查课题所具有的用途或用处。我们所从事的任何一项调查课题，首先必须具有某种意义或价值，或者说，首先必须是“值得去做”的。当然，对于不同的调查课题来说，这种意义或价值会有大有小。同时，这种意义既可以是理论方面的，也可以是实践方面的，或者是理论与实践两方面兼而有之的。

理论方面的意义或价值主要体现在调查课题对一门学科的发展、对某种理论的形成或检验、对社会规律的认识、对社会现象的解释等所能做出的贡献上；

实践方面的意义或价值主要体现在调查课题对现实社会生活所提出的各种具体问题能否进行科学的回答和能否提供合理的解决办法上。

第二节 选择调查课题

三、选题的标准

2. 创造性

创造性也可以称作创新性或独特性，它指的是调查课题应该具有某种新的东西，具有某种与众不同的地方，具有自己独特的特点。作为一种科学的认识活动，每一项具体调查必须能够在某些方面增加人们对现实世界的认识，能够为人们了解和理解现实社会生活中的各种现象、各种问题、各种规律提供新的东西，而不能总是在同一领域、同一范围、同一层次上重复别人的研究，重提已有的结论。

注意：别人从没做过的课题不必然意味着最具创造性：可能是课题没意义所以大家都不做（没有任何价值）；也可能是有意义但别人没意识到（有价值）。

对于大多数研究者来说，一项课题具有创造性，更经常地是指该课题在研究的思路或者研究的角度、依据的理论、调查的对象、采用的方法、调查的内容等一个方面或某几方面，与前人的调查有所不同，有自己独到的、新颖的地方。

第二节 选择调查课题

三、选题的标准

3. 可行性

可行性指的是研究者是否具备进行或完成某一调查课题的主、客观条件。换句话说，就是指研究者在现有的主、客观条件下去从事这项调查课题行不行得通。在许多情况下，越是具有重要价值和创新性的调查课题，它所受到的主、客观限制往往也越多，这也就是说，它的可行性往往也越差。要进行或完成这样的课题常常十分困难，有时甚至是完全不可能进行。

主观限制是指研究者自身条件方面的限制。它包括调查者在生活经历、知识结构、研究经验、组织能力、操作技术等方面的限制，甚至还包括调查者的性别、年龄、语言、体力等纯粹个体因素方面的限制。

客观限制是指进行一项调查课题时受到的外在环境或条件的限制。例如调查时间不够，调查经费不足，有关文献资料不能取得，所涉及的对象、单位和部门不能给予必要的支持和合作，调查课题违反国家有关政策法规或者违反社会伦理道德，或者与被调查者的生活习俗、宗教信仰相违背等等，都是导致一项调查课题无法进行的客观障碍。

第二节 选择调查课题

三、选题的标准

4. 合适性

合适性指的是所选择的调查课题最适合研究者的个人特点。这种个人特点主要包括研究者对该调查课题的兴趣、研究者对与调查课题相关的社会生活领域的熟悉程度、研究者与所调查的对象之间的相似性程度，以及研究者所具有的各种资源、条件与该课题的要求相符合的程度等等。

以上我们介绍了选择调查课题时人们通常采用的四条标准，需要进一步指出的是，这四条标准之间存在着某种层次上的联系：

- ① 重要性是最基本的标准；
- ② 独特性则是在它的基础上提出的新的标准；
- ③ 可行性在某种意义上可以说是课题选择中的决定性标准；
- ④ 合适性则是在前三条标准的基础上提出的更进一步的标准。

这四条标准层层深入，从几个不同的侧面，将一个理想的调查课题从最初众多不成熟的想法、思路和课题雏形中，逐渐分离出来。

第二节 选择调查课题

四、选题的途径和方法

1. 从现实社会生活中寻找
2. 从现有文献中寻找
3. 从个人经历中寻找

第二节 选择调查课题

五、调查课题的明确化

课题的明确化（focus the problem）指的是通过对调查课题进行某种界定，给予明确的陈述，将最初比较含糊的想法变成清楚明确的调查主题，将最初比较笼统、比较宽泛的研究范围或领域变成特定领域中的特定现象或特定问题。

第二节 选择调查课题

六、文献查阅与选题

二者的关系：

感兴趣的现象或问题领域→较宽泛地检索相关文献→初步确定调查课题→进一步检索更为具体的文献→进一步明确调查课题

文献查阅的两个作用：

1. 帮助研究者熟悉和了解本领域中已有的研究成果，便于确定自己的研究在整个领域中的定位，确定理论贡献
2. 为研究者提供一些可供参考的研究思路和方法

查阅文献的过程：

现在可以依靠AI帮忙检索文献

第二节 选择调查课题

六、文献查阅与选题

文献综述：

文献综述有时也称为文献回顾，就是对文献查阅和分析的结果进行归纳、综合、整理，并以恰当的方式表达出来。在调查报告中，往往会有这样一个文献综述的部分。此外，对某一研究领域或某一研究主题的文献综述还可以成为一篇单独的论文。

撰写文献综述，关键是要抓住研究的主题来组织。文献综述常常分为“述”和“评”两大部分（因此有的文献综述论文的标题往往直接写成“某某问题的研究述评”）：既要介绍该研究领域中的状况，又要在此基础上进行评价和议论。

第三节 调查设计

一、明确调查目的

1. 描述性调查

了解和描述社会现象的状况，是人们深入认识这一现象的基础。因此，社会调查最经常地被用于对某个总体或某种现象进行描述的目的。这种描述性调查通常是要发现总体在某些特征上的分布状况。**描述性调查**所关注的焦点通常在于回答总体的各种结构和分布是怎样的，即它回答的主要是人们的特征、行为或态度“是什么”（what），或者研究现象的特点、分布以及趋势“是什么”的问题，而不在于回答为什么会存在这样的结构或分布。

描述性调查在方法上要求做到系统性和全面性。描述性调查不仅需要采取随机抽样方法来选择调查对象，并且调查样本的规模往往要求比较大，这是因为其调查所得到的众多结果都需要推论到所调查的总体，即要用来自样本的资料去描绘总体的面貌。

对社会现象的描述应当注意两个方面：一是描述的准确性，二是描述的概括性。准确性要求指的是对社会现象的分布状况、基本特征等都要做出定量的和精确的描述与说明。概括性的要求则是指调查结果所描述的不应当是个别的或片面的，而应当是能反映出总体及各个组成部分一般状况的普遍现象或者说，根据样本调查的结果，应能够反映出总体的水平和趋势。总之，描述性调查可以说是一种对现象的全面的“清查”和系统的反映；形象地说，描述性调查所得到的是有关这一现象的“整体照片”，或者说是一幅缩小了的“总体模型”。

第三节 调查设计

一、明确调查目的

2. 解释性调查

人们对事物和现象的认识不会只停留在全面了解其状况的层次上。在认识到现象“是什么”以及其状况“怎么样”的基础上，人们还需要明白事物和现象“为什么”（why）会是这样。因此，社会调查同样常常被用于回答社会生活中许多的“为什么”、说明社会调查现象发生的原因、解释社会现象之间的关系。这样一类社会调查，我们称其为解释性调查（explanatory surveys）。

由于**解释性调查的目标是回答“为什么”，是解释原因，是说明关系**，因而它的理论色彩往往更强。它通常是从理论假设出发，经过实地调查，收集经验材料，并通过对资料的分析来检验假设，最后达到对社会现象进行理论解释的目的。也正因为如此，解释性调查在调查方案的设计上，除了与描述性调查一样，具有系统性和周密性以外，它还比描述性调查显得更为严谨，针对性也更强。它更接近于社会研究方法中的实验设计。解释性调查在内容上不要求具有广泛性，不要求面面俱到，但是，它特别注重调查内容的适用性和针对性，它往往要求调查内容必须紧紧围绕所要验证的理论假设。

在分析方法上，解释性调查往往要求进行双变量和多变量的统计分析。

第三节 调查设计

二、确定分析单位

一项社会调查中所研究的对象（**注意，并非调查对象，也不是研究内容或研究主题**），我们称其为分析单位（units of analysis）。在社会调查中，主要有四类分析单位：个人、群体、组织、社区。无论是哪一类分析单位都具有以下两个特点：

1. 调查所收集的资料直接描述分析单位中的每一个个体；
2. 将这些对个体的描述聚合起来，可以描述由这些个体所组成的群体（调查的样本），以及由这一群体所代表的更大的群体（总体），或者用这种描述的聚合去解释某种社会现象。

在理解分析单位时，需要将其与调查对象和研究内容或主题进行区分。分析单位是一项社会调查中所研究的对象；调查对象则是研究者收集资料时直接询问的对象；研究的内容或主题则是分析单位所具有的属性或特征。需要说明的是，分析单位与调查对象有时可能会由同一种对象承担。

第三节 调查设计

二、确定分析单位

1. 个人

以个人作为分析单位的描述性研究一般旨在描述由那些个人所组成的总体，而那些以个人为分析单位的解释性研究则是为了发现存在于该总体中的社会动力。作为分析单位的个人可以用其所隶属的社会群体来指示其特征。

2. 群体

由若干个人所组成的各种社会群体本身也可以成为社会调查中的分析单位。以各种社会群体为分析单位的调查与那些以个人为分析单位的调查，在描述的对象上有所不同。当以社会群体作为分析单位时，它们的特征有时与群体中个人的特征有关，比如家庭的收入就与个人的收入有关。有时群体的特征可以从其成员的特征中抽取。但在更多的情况下，这种群体的特征则不同于个人的特征。**应当记住的是，当以社会群体做分析单位时，研究和分析就不能下滑到群体层次之下（即不能下滑到个人的层次），研究的群体就是资料集合中的最小单位。**

第三节 调查设计

二、确定分析单位

3. 组织

各种正式的社会组织，比如工厂、公司、机关、学校、商店、医院等，同样可以成为社会调查中的分析单位。由于组织与群体一样，都是由若干个个人组成的，因而作为分析单位的组织所具有的某些特征，往往也在一定程度上与组成它的个人有关。有时，对同一现象的研究，会依据调查的侧重点的不同而使用不同的分析单位，这样就大大地增加了分析单位的复杂性。

4. 社区

社区作为一定地域中人们的生活共同体，也可以作为调查中的分析单位。无论是乡村、城市，还是街道、集镇，我们都可以用社区的人口规模、社区异质性程度、社区习俗特点、社区的空间范围等特征对它们进行描述，也可以通过分析社区不同特征之间的关系来解释和说明某些社会现象。如同以个人为分析单位的社会调查中的个人那样，从每一个具体的社区中所收集的资料，既用来描述和反映这一社区自身的具体特征，又作为若干个具体社区的集合中的一个个案，参与到描述整个社区的集合的特征以及解释某些特定的社区现象中去。

第三节 调查设计

二、确定分析单位

层次谬误

层次谬误（ecological fallacy）又称为区群谬误、生态谬误或体系错误，它是一种与分析单位有关的错误。它指的是在社会调查中，研究者用一种比较大的集群分析单位做研究，而用另一种比较小的或非集群分析单位做结论的现象。或者说，研究者在一个比较大的集群分析单位上收集资料，而在一个比较小的或非集群的分析单位上来下结论的现象。

例1. 在以城市为分析单位调查犯罪现象时，研究者发现，流动人口多的城市的犯罪率大大高于流动人口少的城市，呈现出“城市的流动人口越多，城市的犯罪率越高”的趋势。如果研究者根据这一现象得出结论说“流动人口比非流动人口的犯罪率高”，那么他就犯了层次谬误。

为什么？

第三节 调查设计

二、确定分析单位

层次谬误

层次谬误（ecological fallacy）又称为区群谬误、生态谬误或体系错误，它是一种与分析单位有关的错误。它指的是在社会调查中，研究者用一种比较大的集群分析单位做研究，而用另一种比较小的或非集群分析单位做结论的现象。或者说，研究者在一个比较大的集群分析单位上收集资料，而在一个比较小的或非集群的分析单位上来下结论的现象。

例1. 在以城市为分析单位调查犯罪现象时，研究者发现，流动人口多的城市的犯罪率大大高于流动人口少的城市，呈现出“城市的流动人口越多，城市的犯罪率越高”的趋势。如果研究者根据这一现象得出结论说“流动人口比非流动人口的犯罪率高”，那么他就犯了层次谬误。

因为他的调查资料是以城市（分析单位是社区）为单位收集来的，所得出的也只能是有关城市的结论，而不能是有关流动人口和非流动人口（分析单位是群体）的结论。

第三节 调查设计

二、确定分析单位

层次谬误

层次谬误（ecological fallacy）又称为区群谬误、生态谬误或体系错误，它是一种与分析单位有关的错误。它指的是在社会调查中，研究者用一种比较大的集群分析单位做研究，而用另一种比较小的或非集群分析单位做结论的现象。或者说，研究者在一个比较大的集群分析单位上收集资料，而在一个比较小的或非集群的分析单位上来下结论的现象。

例2. 研究者在两个规模相当的小城镇做调查，发现甲城镇高收入居民家庭的比例大大高于乙城镇的比例，同时还发现甲城镇中居民家庭拥有摩托车的比例也大大高于乙城镇中的比例。如果在这些资料的基础上，研究者得出结论说，“收入高的居民家庭拥有摩托车的比例更高”，或者得出结论说，“居民家庭收入与拥有摩托车之间存在正相关”，那么他就同样犯了层次谬误。

为什么？

第三节 调查设计

二、确定分析单位

层次谬误

层次谬误（ecological fallacy）又称为区群谬误、生态谬误或体系错误，它是一种与分析单位有关的错误。它指的是在社会调查中，研究者用一种比较大的集群分析单位做研究，而用另一种比较小的或非集群分析单位做结论的现象。或者说，研究者在一个比较大的集群分析单位上收集资料，而在一个比较小的或非集群的分析单位上来下结论的现象。

例2. 研究者在两个规模相当的小城镇做调查，发现甲城镇高收入居民家庭的比例大大高于乙城镇的比例，同时还发现甲城镇中居民家庭拥有摩托车的比例也大大高于乙城镇中的比例。如果在这些资料的基础上，研究者得出结论说，“收入高的居民家庭拥有摩托车的比例更高”，或者得出结论说，“居民家庭收入与拥有摩托车之间存在正相关”，那么他就同样犯了层次谬误。

因为我们并不知道这两个城镇中的哪些居民家庭收入较高，也不知道哪些居民家庭拥有摩托车，我们所知道的只是对于城镇总体来说的收入分布和摩托车拥有量，也许是甲城镇中的那些中等收入和低收入的家庭而非高收入的居民家庭更多地拥有摩托车（因为高收入的居民家庭都买轿车了）。

第三节 调查设计

三、设计抽样方案

1. 调查对象、研究对象与抽样对象

研究对象就是之前介绍的分析单位，它是我们在社会调查中所要描述和分析的对象。一般情况下，研究对象往往是与我们的调查所要探索 and 回答的问题有关。

调查对象（respondent）即回答者，指的是在调查中（通过自填问卷或结构访问的方式）所询问的对象。他们是我们所收集的调查资料的提供者。尽管不同的调查的研究对象（即分析单位）各不相同，但所有社会调查的调查对象都是且只能是社会中各种类型的个人。调查对象往往只与我们询问的人有关。

抽样对象也称为抽样单位（sampling unit），指的是一次直接抽样时所采用的对象或单位。与研究对象类似，抽样单位既可能是社会中的个人，比如大学生，也可能是其他的对象，比如家庭、居委会、街道、学校、工厂等。抽样对象仅仅只与样本抽取过程以及构成总体的元素有关。但大多数情况下，最终的抽样对象往往与调查对象一致。

弄清楚上述三者的联系和区别，便于我们在调查设计中恰当选择研究对象、调查对象、抽样对象。

第三节 调查设计

三、设计抽样方案

2. 抽样设计

抽样设计的首要任务是明确调查的对象和调查的总体，并根据总体的结构以及研究者所具有的资源、条件等因素选择和决定合适的抽样方法。研究者的任务就是要根据调查研究的目标要求，以及研究者所面对的社会现实条件，设计恰当的抽样方法和程序，获得有最大代表性的样本。

抽样设计的另一个任务是决定样本规模，主要影响因素包括：

- ① 用样本统计值推论总体参数值时的精确程度，即抽样误差的大小。在其他条件不变的情况下，它和样本规模成正比，即精确程度越高（抽样误差越小），所要求的样本规模越大。
- ② 推论的把握性程度（即置信度的大小）。在其他条件不变的情况下，所要求的这种把握性程度越高，所需要的样本规模也越大。
- ③ 研究者所拥有的人力、物力和时间。这是决定样本规模的更为客观的条件。

第三节 调查设计

四、设计收集资料的工具与方法

1. 基本变量确定及其操作化

虽然不同的社会调查在具体内容上可能会千差万别，但有一点却是共同的：**它们都依赖于变量的语言**。无论是以描述为目标的调查，还是以解释为目标的调查，**它们在具体内容上都可以归结为两大方面：一是描述某些变量的特征，二是探讨某些变量相互之间的关系**。因此，在调查中确定基本的研究变量、选择测量这些变量的指标、设计用于这种测量的特殊工具，也就成了调查研究的核心环节之一。

在调查设计阶段，研究者一方面要依据调查研究的目标确立最基本的研究变量，包括研究者最为关注的因变量，也包括作为解释变量的各种自变量；另一方面，研究者还要对这些基本变量进行操作化，即研究者既要构建起一套需要进行测量的变量体系，以满足调查研究目的的需要，又要将这些变量操作化为一系列更为具体的测量指标，以便于在调查中收集相应的资料。前一方面的工作与调查的总目标密切相关，与调查的基本内容相关；而后一方面的工作则是为调查设计中的下一项具体任务——问卷设计——提供指引。

第三节 调查设计

四、设计收集资料的工具与方法

2. 设计调查问卷

问卷是调查研究中用来收集资料主要工具。无论是采用自填问卷的方式，还是采用结构访问的方式，都离不开问卷的使用。因而问卷的设计始终是调查设计中一项既相对独立、同时又与调查目标，特别是与概念操作化紧密相连的工作。问卷设计的实质是以书面的形式将研究者希望了解的信息、希望收集的资料、希望询问的问题编制成一份询问表，以便于系统地、客观地并且是定量地收集所有被调查者的情况，为探讨和回答研究问题收集所需的经验材料和依据。

调查问卷的设计可以看成是前面研究变量操作化的后续工作。它在内容上与变量操作化具有一致性。研究者应该依据前面所确定的变量框架和操作化指标，并遵循问卷设计的基本原则和方法技巧，将这些变量和测量指标转化为适合被调查对象回答的具体问题。在此基础上，最终合成一份与调查方式相配合的调查问卷。

第三节 调查设计

四、设计收集资料的工具与方法

3. 设计资料的收集方法

社会调查中的资料收集方法有多种不同的形式，每一种具体的资料收集方法都有其特定的优点和不足，它们分别适用于各种不同的条件和场合。**研究者的任务就是要根据自己所从事的调查课题的具体情况，从中进行选择，以达到最好的调查效果。**从大方面看，研究者要考虑是采用自填的方式还是采用结构访问的方式；从更具体的方面看，则又要进一步在个别发送、集中填答、邮寄调查、网络调查、当面访问、电话访问等多种不同方法中进行选择。为此，研究者在调查设计阶段，必须依据多种因素综合进行考虑。这些因素包括调查总体及其调查对象的性质、调查样本规模的大小、调查内容的敏感性程度、调查的目标和重点、调查课题完成的时间要求、调查者拥有的人力和物力是否充足等。**研究者要在调查正式实施之前，全面地分析和考虑上述各项因素，做出恰当的选择，并做好相应的计划和安排。**

与资料收集方法的选择有关的另一个问题是对调查的过程以及调查的组织与实施安排进行考虑，包括调查人员的组织、培训、管理，经费的使用，时间的安排等等。特别的，研究者在设计阶段要对所选择的资料收集方法在实施过程中所可能遇到的各种困难、所面临的各种障碍有一定的分析，并提前准备好对应的措施和处理的方法。

第三节 调查设计

五、制订实施方案

1. 说明调查课题的目的和意义

在研究计划书中应说明为什么要进行这项调查研究，从事这项调查研究在理论上或在实践上有什么样的价值。当然，要对这些方面进行清楚说明的前提条件是，调查研究者必须首先对自己的调查课题有一个清楚明确的认识。这种认识既包括对调查课题本身含义的理解，即该调查究竟要探讨和回答什么问题，也包括对调查课题在人们认识社会、改造社会中所具有的作用的理解。

这种在调查方案设计中调查课题的目的和意义的说明，表面上看起来与实际调查中的操作过程并不相关，但事实上，它既是对调查者选择这一课题的动机、意图、方向、价值等是否明确的一种检验；同时也是进一步地帮助调查者强化和突出这一课题的总目标，加强这一目标对整个调查研究过程的影响。如果调查者本人对调查课题的目标和意义都说不清楚，那么，这一课题是否值得去搞，以及是否能够真正做好，显然是值得怀疑的。

第三节 调查设计

五、制订实施方案

2. 说明调查的内容

调查内容是对调查目的的具体分解和细化。在调查设计中，详细说明调查的内容，是落实调查目标的十分重要的一环。如前所述，调查课题的确定只是指出了我们所研究现象的大致范围或基本方向，至于在这个题目下究竟应该调查研究哪些具体现象，则是在调查设计中所要解决的问题，所应完成的工作任务。

3. 说明调查范围、调查对象和分析单位

调查范围的界定，有助于明确调查结果所推论的总体；调查对象的确定，有助于选择合适的调查方法和测量工具；指明调查课题的分析单位，则可以帮助研究者有针对性地收集研究所需的资料，同时也可以使研究者避免犯层次谬误。

4. 说明调查的理论假设

尽管不是每一类调查都必须有理论假设，但对于那些必须有理论假设的解释性调查来说，则应该在调查方案中对理论假设进行一番陈述和说明。一般来说，描述性调查的主要目的是全面描述某种社会现象的状况和特点，为进一步分析和探讨不同现象之间的联系打下基础，因此，它一般不需要建立明确的理论假设。只有在解释性调查中，才必须事先建立起明确的理论假设。所以，在解释性调查的具体方案中，不能缺少对理论假设的陈述和说明。

第三节 调查设计

五、制订实施方案

5. 说明调查方案

这一部分主要用来说明调查过程中所面临的具体技术问题，包括调查对象的抽取（即总体与抽样方法）、变量测量（从概念操作化到问卷设计）、资料收集方法以及资料分析方法等几个大的方面。在方案中需要对这方面逐一进行简明扼要的说明。比如，在调查对象抽取方面，需要说明调查的总体是什么，研究者将采用什么样的抽样方法和程序，样本规模如何确定及其依据等。同样的，对于调查研究中的基本变量要进行界定和操作化的说明，对所使用的资料收集方法、实施步骤以及资料分析技术也都要做出既符合研究目标同时也符合研究者所拥有的条件的说明。

6. 说明调查人员的组成、组织结构及培训安排

要完成一项较大规模的调查课题，往往需要很多研究者的共同努力，同时还会涉及挑选、培训调查员的问题。因此，在调查方案设计中，必须对调查课题的组成人员及其在调查中所承担的任务进行全盘考虑，明确分工，制定相应的组织管理办法。对调查员的挑选、培训工作也要事先进行规划，制订出切实可行的培训方案，以保证调查工作的顺利进行。

第三节 调查设计

五、制订实施方案

7. 确定调查的时间进度和经费使用计划

一项社会调查从定下题目到完成报告，往往有时间上的限定或要求。为了在规定的时间内保质保量地完成调查任务，顺利达到预定的调查目标，研究者应该在课题研究开始之前，对整个调查工作的时间分配和进度进行安排。每一阶段所分配的时间要合适，还要留有一点余地。特别要注意给调查研究的设计和准备阶段多安排一些时间，不要匆匆忙忙开始收集资料。此外，对于调查课题的经费使用，也应有一个大致的考虑和合适的分配，以保证调查的各个阶段工作都能顺利进行。

第四节 抽样

一、抽样的意义与作用

1. 抽样的概念

- ① **总体 (population)** 通常与构成它的元素共同定义：总体是构成它的所有元素的集合，而元素 (element) 则是构成总体的最基本单位。在社会调查中，最常见的总体是由社会中的个体组成的，这些个体便是构成总体的元素。一个总体中所包含的元素数目常用大写字母N表示。
- ② **样本 (sample)** 就是从总体中按一定方式抽取出一部分元素的集合。在社会调查中，资料的收集工作或者说调查的实施正是在样本中完成的。样本中的元素数目通常用小写字母n表示。
- ③ **抽样 (sampling)** 指的是从组成总体的所有元素的集合中，按一定的方式选择或抽取一部分元素（总体的一个子集）的过程，或者说，抽样是按一定方式从总体中选择或抽取样本的过程。
- ④ **抽样单位 (sampling unit)** 就是一次直接抽样时所使用的基本单位。抽样单位与构成总体的元素有时是相同的，有时又是不同的。
- ⑤ **抽样框 (sampling frame)** 又称作抽样范围，指的是一次直接抽样时总体中所有元素的名单。

第四节 抽样

一、抽样的意义与作用

1. 抽样的概念

- ⑥ **参数值 (parameter)** 也称为总体值，它是关于总体中某一变量的综合描述，或者说是总体中所有元素的某种特征的综合数量表现。简单地说，是某一变量的总体值。需要说明的是，参数值只有对总体中的每一个元素都进行调查或测量才能得到。
- ⑦ **统计值 (statistic)** 也称为样本值，它是关于样本中某一变量的综合描述，或者说是样本中所有元素的某种特征的综合数量表现。统计值是从样本中计算出来的，它是相应的参数值的估计量。参数值和统计值之间有一个重要的区别：参数值是确定不变的、唯一的，并且通常是未知的；而统计值则是变化的，即对于同一个总体来说，不同样本所得的统计值是有差别的；同时，对于任一特定的样本来说，统计值是已知的，或者说是可以通过计算得到的。从样本的统计值来推论总体的参数值，正是社会调查的一项重要内容。
- ⑧ **置信水平 (confidence level)** 又称为置信度，指的是总体参数值落在样本统计值某一区间（即置信区间）内的概率，或者说是总体参数值落在样本统计值某一区间中的可能性程度。它反映的是抽样的可靠性程度，也是我们从统计值推断参数值时的把握性程度。
- ⑨ **置信区间 (confidence interval)** 是指在一定的置信水平下，样本统计值与总体参数值之间的误差范围。置信区间反映的是抽样的精确性程度。置信区间越大，即误差范围越大，抽样的精确性程度就越低；反之，置信区间越小，即误差范围越小，抽样的精确性程度就越高。

第四节 抽样

一、抽样的意义与作用

2. 抽样的作用

从抽样的定义中不难看出，抽样主要涉及总体与部分之间的关系问题。抽样作为人们从部分认识整体这一过程的关键环节，其基本作用是向人们提供一种实现“由部分认识总体”这一目标的途径和手段。

在社会调查中，抽样主要解决的是调查对象的选取问题，即如何从总体中选出一部分对象作为总体的代表的问题。

民调机构	里根	蒙代尔
《时代》	64	36
《今日美国》	63	37
哥伦比亚广播公司/《纽约时代周刊》	61	39
盖洛普民意测验	59	41
实际投票结果	59	41
盖洛普民意测验/《新闻周刊》	59	41
美国广播公司/《华盛顿邮报》	57	43
哈里斯民意测验	56	44
罗珀民意测验	55	45

第四节 抽样

一、抽样的意义与作用

3. 抽样的类型

各种抽样都可以归为概率抽样与非概率抽样两大类，这是两种有着本质区别的抽样类型：

- ① 概率抽样是依据概率论的基本原理，按照等概率原则进行的抽样，因而它能够避免抽样过程中的人为误差，保证样本的代表性；
- ② 非概率抽样主要是依据研究者的主观意愿、判断或是否方便等因素来抽取对象，它不考虑抽样中的等概率原则，因而往往产生较大的误差，难以保证样本的代表性。

概率抽样	简单随机抽样
	系统抽样
	分层抽样
	整群抽样
	多段抽样
	PPS抽样
非概率抽样	偶遇抽样
	判断抽样
	定额抽样
	滚雪球抽样

第四节 抽样

二、非概率抽样

1. 偶遇抽样 (accidental or convenience sampling)

又称作方便抽样或自然抽样，是指研究者根据现实情况，以自己方便的形式抽取偶然遇到的人作为调查对象，或者仅仅选择那些离得最近的、最容易找到的人作为调查对象。

2. 判断抽样 (judgmental or purposive sampling)

又称为立意抽样（或目的抽样），它是调查者根据研究目标和自己主观分析来选择和确定调查对象的方法。这种抽样首先要确定抽样标准。由于标准的确定带有较大的主观性，所以此法的运用结果如何往往与研究者的理论修养、实际经验和对调查对象的熟悉程度有很大关系。判断抽样的主要优点在于可以充分发挥研究人员的主观能动作用，特别是当研究者对研究的总体情况比较熟悉，研究者的分析判断能力较强、研究方法与技术十分熟练、研究经验比较丰富时，采用这种方法往往十分方便。但是由于它仍然属于一种非概率抽样，所以，其所得样本的代表性往往难以判断。

3. 定额抽样 (quota sampling)

又称为配额抽样，它是一种比偶遇抽样复杂一些的非概率抽样方法。进行定额抽样时，研究者要尽可能地依据那些有可能影响研究变量的各种因素来对总体分类，并找出具有各种不同特征的成员在总体中所占的比例，然后依据这种划分以及各类成员的比例，采用偶遇抽样或判断抽样的方法去选择调查对象，使样本中的成员在上述各种因素、各种特征方面的比例尽量接近总体情形。如果把各种因素或各种特征看作不同的变量的话，那么，定额抽样实际上就是依据这些变量的组合。

第四节 抽样

二、非概率抽样

3. 定额抽样

我们以性别、年级和专业3个因素来解释这种变量的组合及其定额抽样的实施办法。假设某高校有2000名学生，其中男生占60%，女生占40%；文科学生和理科学生各占50%；一二三四年级学生分别占40%、30%、20%和10%。现要用定额抽样方法依上述3个变量抽取一个规模为100人的样本。依据总体的构成和样本规模，我们可得到下列定额表。

	男生60人								女生40人							
	文科30人				理科30人				文科20人				理科20人			
年级	一	二	三	四	一	二	三	四	一	二	三	四	一	二	三	四
人数	12	9	6	3	12	9	6	3	8	6	4	2	8	6	4	2

第四节 抽样

二、非概率抽样

3. 定额抽样

定额抽样方法的两个主要不足在于：一方面，研究者总是只能对少数几个方便控制的特征进行分类和定额，而无法对许多同样影响着、体现着总体特征的其他因素进行分类和定额；另一方面，由于定额抽样中调查员可以在保证各种类型定额的情况下自由地选取自己方便获得的调查对象，这种自由度很大的选择方式同样极易造成样本与总体之间的偏差。

表面上看起来，定额抽样与**分层抽样（一种概率抽样）**十分相似，实际上，二者却有本质差别。二者虽然都根据某些特征对总体进行分层或分类，但抽样方法不同。定额抽样是通过主观的分析或依据方便性来确定和选择组成这种模拟物的成员。也就是说，定额抽样注重的是样本与总体在结构比例上的表面一致性。而分层抽样的方法则是完全依据概率原则，排除主观因素，客观地、等概率地在各层中进行抽样，这与定额抽样中那种“按事先规定的条件，采用偶遇方式和主观判断方式来寻找样本成员”的做法是完全不同的。

4. 滚雪球抽样（snowball sampling）

有人认为它是偶遇抽样的一种。在特定总体的成员难以找到时，滚雪球抽样是最适合采取的一种抽样方法。这种抽样方法是先搜集目标群体少数成员的数据，然后再向这些成员询问有关信息，找出他们认识的其他成员。所谓滚雪球，就是根据既有研究对象的建议，找出其他研究对象的积累过程。由于用这种方法产生的样本在代表性上可疑，因此，通常用于探索性研究。

第四节 抽样

三-一、概率抽样：原理与程序

1. 概率抽样的基本原理

由于现实社会中的绝大多数群体通常都存在着程度不同的异质性，即它们所包含的个体之间总是存在着这样或那样的差别。在社会各种总体都普遍存在异质性的现实面前，严格的概率抽样程序与方法就必不可少。**而概率样本所要反映的正是总体所具有的那种内在的异质性。**抽样的最终目的在于通过对样本统计值的描述来勾画出总体的面貌，概率抽样的方法可以帮助我们实现这一目标，并且可以对这种勾画的准确程度做出估计。在概率抽样的过程中，我们总是要求保证总体中的每一个个体都有同等的机会入选样本，而且，任何一个个体的入选与否，与其他个体毫不相关、互不影响。或者说，每一个个体的抽取都是相互独立的，是一种随机事件。**概率抽样之所以能够保证样本对总体的代表性，其原理就在于它能够很好地按总体内在结构中所蕴含的各种随机事件的概率来构成样本，使样本成为总体的缩影。**

在讨论概率抽样问题前，先对**放回抽样与不放回抽样**问题略做说明。严格地说，由于研究者在实际抽样中所做的基本上都是不放回抽样，因而并没有完全满足抽样的独立性要求。这种独立性要求指的是任何一个元素的抽取都不会影响到其他元素被抽取的概率。但不放回抽样中，每抽出一个元素，总体就减少了一个元素，其他元素被抽到的概率就会增加。**然而，只要总体相对于样本来说要大得多，我们就可以忽略这种不放回抽样所产生的微小改变。**因为事实上对于一个相当大的总体来说，缺少一个元素基本上不改变总体中其他众多元素被抽中的概率；即使将抽中的元素放回总体，它也基本上不会有第二次被抽中的机会。

第四节 抽样

三-一、概率抽样：原理与程序

2. 抽样的一般程序

- ① **界定总体。**界定总体就是在具体抽样前，首先对从中抽取样本的总体范围与界限作明确的界定。这一方面是由抽样调查的目的所决定的，因为抽样调查虽然只对总体中的一部分个体实施调查，但其目的却是为了描述和认识总体的状况与特征，所以必须事先明确总体的范围；另一方面，界定总体也是达到良好的抽样效果的前提条件。如果不清楚明确地界定总体的范围与界限，那么即使采用严格的抽样方法，也可能抽出对总体严重缺乏代表性的样本来。
- ② **制定抽样框。**这一步骤的任务就是依据已经明确界定的总体范围，收集总体中全部抽样单位的名单，并通过对名单进行统一编号来建立起供抽样使用的抽样框。需要注意的是，当抽样是分几个阶段、在几个不同的抽样层次上进行时，则要分别建立起几个不同的抽样框。
- ③ **决定抽样方案。**具体的抽样方法有好几种，不同的抽样方法都有自身的特点和适用范围。因此，对于具有不同研究目的、不同调查范围、不同调查对象和不同客观条件的社会调查来说，所适用的抽样方法也不一样。这就需要在具体实施抽样之前，依据研究的目的和要求，依据各种抽样方法的特点和其他有关因素来决定具体采用哪种抽样方法。除了抽样方法的确定以外，还要根据调查的要求确定样本的规模和主要目标量的精确程度。

第四节 抽样

三-一、概率抽样：原理与程序

2. 抽样的一般程序

- ④ **实际抽取样本。**实际抽取样本的工作就是在上述几个步骤的基础上，严格按照所选定的抽样方法，从抽样框中抽取一个个的抽样单位，构成调查样本。依据抽样方法的不同，以及依据抽样框是否可以事先得到等因素，实际的抽样工作既可能在实地调查前就完成，也可能需要到实地后才能完成。即既可能先抽好样本，再下去直接对预先抽好的调查对象进行调查，也可能一边抽取样本一边开始调查。
- ⑤ **评估样本质量。**完整的抽样过程还应包括样本抽出后对样本进行评估的工作。所谓样本评估，就是对样本的质量、代表性、偏差等进行初步的检验和衡量，其目的是防止由于样本的偏差过大而导致调查的失误。评估样本的基本方法是对研究者实际抽取样本的具体方法和程序进行分析与检查，看其是否保证了（做到了）总体中的每一个个体都有已知且相等的概率被选入样本。同时，也可以辅之以比较的方法，将可得到的、反映总体中某些重要特征及其分布的资料与样本中的同类指标的资料进行对比。若二者之间的差别很小，则我们对样本的代表性就有较大的信心；反之，若二者之间的差别十分明显，那么样本的质量和代表性就一定不会很高。

第四节 抽样

三-一、概率抽样：原理与程序

3. 抽样设计的原则

- ① **目的性原则**是指在进行抽样方案设计时，要以课题研究的总体方案和研究的目标为依据，以研究的问题为出发点，从最有利于研究资料的获取以及最符合研究的目的等因素来考虑抽样方案和抽样方法的设计。
- ② **可测性原则**是指抽样设计能够从样本自身计算出有效的估计值或者抽样变动的近似值。在研究中通常用标准差来表示。这是统计推断必需的基础，是样本结果与未知的总体值之间客观、科学的桥梁。通常，只有概率样本在客观上才是可测的，即概率样本可以计算出有效的估计值或抽样变动的近似值。
- ③ **可行性原则**是指研究者设计的抽样方案必须在实践上切实可行。它意味着研究者设计的方案能够预料实际抽样过程中可能出现的各种问题，并设计了处理这些问题的方法。由于在理论上设计抽样方案 and 在实际中执行这一方案是两码事，因而可行性是抽样设计的一条重要标准。
- ④ **经济性原则**主要指的是抽样方案的设计要与研究的可得资源相适应。这种资源主要包括研究的经费、时间、人力等。

相对而言，这四条标准中，**目的性原则和可行性原则是首要的**。抽样设计要服务于研究的目标，这是设计的出发点和基本目的。而可行性则是设计方案得以实现的前提和保证。研究者应该在优先考虑这两条标准的基础上，去进一步增加方案的可测性，同时减少方案所需的资源。

第四节 抽样

三-二、概率抽样：方法

1. 简单随机抽样 (simple random sampling)

又称纯随机抽样，是概率抽样的最基本形式。它是按等概率原则直接从含有 N 个元素的总体中抽取 n 个元素组成样本 ($N > n$)。常用的办法是抽签：把总体的每一个元素都编号，将这些号码写在一张张纸条上，然后放入一个容器（例如纸盒、口袋）中，搅拌均匀后，从中任意抽取，直到抽够预定的样本数目。这样，由抽中的号码所代表的元素组成的就是一个随机样本。

对于总体单位很多的情形，我们则采用随机数表来抽样。随机数表中的数码和排列都是随机形成的，没有任何一点规律性，故也称为乱数表。利用随机数表进行抽样的具体步骤如下，而按照步骤选择出来的元素集合就是所需的样本：

- ① 先取得一份调查总体所有元素的名单（即抽样框）；
- ② 将总体中所有元素一一按顺序编号；
- ③ 根据总体规模是几位数来确定从随机数表中选几位数码；
- ④ 闭着眼睛用铅笔点中随机数表中的一个数字，以此处作为选择数码的起点；
- ⑤ 以总体的规模为标准，对随机数表中的数码逐一进行衡量并决定取舍；
- ⑥ 根据样本规模的要求选择出足够的数码个数；
- ⑦ 依据从随机数表中选出的数码，到抽样框中去找它对应的元素。

现在不用这么麻烦，不用所谓的随机数表，用计算机生成伪随机数，直接确定名单即可。

第四节 抽样

三-二、概率抽样：方法

2. 系统抽样 (systematic sampling)

又称等距抽样或间隔抽样。它是把总体的元素进行编号排序后，再计算出某种间隔，然后按这一固定的间隔抽取元素来组成样本的方法。它和简单抽样一样，需要有完整的抽样框，样本的抽取也是直接从总体中抽取元素，而无其他中间环节。系统抽样的具体步骤如下：

- ① 给总体中的每一个元素按顺序编上号码，即制定出抽样框，这与简单随机抽样的做法一样。
- ② 计算出抽样间距，方法是用总体的规模除以样本的规模。假设总体规模为 N ，样本规模为 n ，那么抽样间距 K 就由下列公式求得： $K = N / n$ ；
- ③ 在最前面的 K 个元素中，采用简单随机抽样的方法抽取一个元素，记下这个元素的编号（假设所抽取的这个元素的编号为 A ），它称为随机的起点。
- ④ 在抽样框中，自 A 开始，每隔 K 个元素抽取一个元素，即所抽取元素的编号分别为 A 、 $A+K$ 、 $A+2K$ 、...、 $A+(n-1)K$ 。
- ⑤ 将这 n 个元素合起来，就构成了该总体的一个样本。

第四节 抽样

三-二、概率抽样：方法

2. 系统抽样 (systematic sampling)

值得注意的是，系统抽样的一个十分重要的前提条件是总体中元素的排列，相对于研究的变量来说应是随机的，即不存在某种与研究变量相关的规则分布。否则，系统抽样的结果将可能产生极大的偏差。因此，在使用系统抽样方法时一定要注意抽样框的编制方法。特别要注意下列两种情况：

- ① 在总体名单中，元素的排列具有某种次序上的先后、等级上的高低
- ② 在总体名单中，元素的排列有与抽样间隔相对应的周期性分布

第四节 抽样

三-二、概率抽样：方法

3. 分层抽样 (stratified sampling)

又称类型抽样，它是先将总体中的所有元素按某种特征或标志（例如性别、年龄、职业或地域等）划分成若干类型或层次，然后再在各个类型或层次中采用简单随机抽样或系统抽样的办法抽取一个子样本，最后，将这些子样本合起来构成总体的样本。

优点：

- ① 在不增加样本规模的前提下降低抽样误差，提高抽样的精度。
- ② 非常便于了解总体内不同层次的情况，便于对总体中不同的层次或类别进行单独研究，或者进行比较。

第四节 抽样

三-二、概率抽样：方法

3. 分层抽样 (stratified sampling)

在实际运用分层抽样方法时，研究者需要考虑下列两个方面的问题。

一是分层的标准问题。同一个总体可以按照不同的标准进行分层，或者说，根据不同的标准可以将一个总体分成不同的类别或层次。那么，在实际抽样中究竟应该按什么标准来分层呢？通常采用的原则如下。

- ① 以调查所要分析和研究的主要变量或相关的变量作为分层的标准。
- ② 以保证各层内部同质性强和各层之间异质性强、突出总体内在结构的变量作为分层变量。
- ③ 以那些已有明显层次区分的变量作为分层变量。

二是分层的比例问题。分层抽样中有按比例分层和不按比例分层两种方法。按比例分层抽样是指按总体中各种类型或层次的比例来抽取子样本的方法。采取按比例分层抽样的方法，可以确保得到一个在某种特征上与总体结构完全一样的样本。但是在有些情况下，又不宜采用这种方法。例如，有时总体中有的类型或层次的单位数目太少，若以按比例分层的方法抽样，则有的层次在样本中个案太少，不便于了解各个层次的情况，也不利于进行不同层次之间的比较，这时往往要采取不按比例抽样的方法。但需要注意的是，我们采用不按比例分层抽样的方法，主要是便于对不同层次的子总体进行专门研究或进行相互比较，但若要用样本资料推断总体时，则需要先对各层的数据资料进行加权处理，即通过调整样本中各层的比例，使数据资料恢复到总体中各层实际的比例结构。

第四节 抽样

三-二、概率抽样：方法

4. 整群抽样 (cluster sampling)

是从总体中随机抽取一些小的群体，然后由所抽出的这些小群体内的所有元素构成调查样本的方法。**整群抽样与前几种抽样的最大差别在于，它的抽样单位不是单个的元素，而是成群的元素。**这种小的群体可以是居民家庭、学校中的班级，也可以是工厂中的车间班组，还可以是城市中的社区等。当构成总体的这些小群体的规模相同或者相差不大时，整群抽样的方法特别适用。整群抽样中对小群体的抽取可采用简单随机抽样、系统抽样或分层抽样的方法。

采取整群抽样的方法，不仅可以简化抽样的过程，更重要的是可以降低调查中收集资料的费用，同时还能相对地扩大抽样的应用范围。但是，应该看到，**整群抽样具有的简便易行、节省费用的优点，是以其样本的分布面不广、样本对总体的代表性相对较差等缺点为代价的。**由于整群抽样所抽样本中的个体相对集中，而涉及面相对缩小，故在许多情况下会导致样本的代表性不足，使得调查结果的偏差较大。

整群抽样方法的运用，尤其要与分层抽样方法相区别。当某个总体是由若干个有着自然界限和区分的子群（或类别、层次）组成，同时，不同子群之间差别很大而每个子群内部的差异不大时，则适合分层抽样方法；反之，当不同子群之间差别不大而每个子群内部的异质性较大时，则特别适合采用整群抽样方法。

第四节 抽样

三-二、概率抽样：方法

5. 多段抽样（multistage sampling）

又称多级抽样或分段抽样，它是按抽样元素的隶属关系或层次关系，把抽样过程分为几个阶段进行。在社会调查中，当总体的规模特别大，或者总体分布的范围特别广时，研究者一般采取多段抽样的方法来抽取样本。

多段抽样的具体做法是：先从总体中随机抽取若干大群（组），然后再从这几个大群（组）内抽取几个小群（组），这样一层层抽下来，直至抽到最基本的抽样元素为止。

在运用多段抽样方法时，有一点要注意，就是要在类别和个体之间保持平衡，或者说，保持合适的比例。如何确定每一级抽样的单位数目呢？主要考虑三方面因素：一是各个抽样阶段中的子总体同质性程度，二是各层子总体的个体数，三是研究者拥有的人力和经费。一般来说，抽取的类别相对较多、每一类中抽取的个体相对较少的做法样本代表性较好。当然，此种抽样所得样本在调查时的代价则相对较大。

多段抽样适用于范围大、总体对象多的社会调查。由于它不需要总体的全部名单，各阶段的抽样单位数一般较少，因而抽样比较容易进行。但由于每级抽样时都会产生误差，故这种抽样方法的误差较大，这是它的主要不足。**在同等条件下减少多段抽样误差的方法是：相对增加开头阶段的样本数并适当减少最后阶段的样本数。**

第四节 抽样

三-二、概率抽样：方法

6. PPS抽样

多段抽样其实暗含了一个假定，即每一个阶段抽样时，其元素的规模是相同的。但这一假设很可能不成立，导致每个个体被抽中的概率不相等。

为了解决这一问题，在社会调查中，学者们常常采用一种叫作“概率与元素的规模大小成比例的抽样”（sampling with probability proportional to size, 简称PPS抽样）的不等概率抽样方法。其原理可以通俗地理解成以阶段性的（或暂时的）不等概率换取最终的、总体的等概率。

其做法是：在第一阶段，每个群按照其规模（其所含元素的数量）被给予大小不等的抽取概率（群越大，其被抽中的概率也越大）。但到了抽样的第二阶段，从每个抽中的群中都抽取同样多的元素（实际上也是不等概率的）。

正是通过这样两个阶段上的不等概率抽样，使得总体中的每一个元素最终都具有同样的被抽中的概率。其实质是：在第一个阶段中，大的群被抽中的概率大，而小的群被抽中的概率小；到了第二阶段，被抽中的大的群中的元素被抽中的概率小，而被抽中的小的群中的元素被抽中的概率大。正是这“一大一小”平衡了由于群的规模的不同带来的概率差异。

第四节 抽样

三-二、概率抽样：方法

6. PPS抽样

PPS抽样原理的公式表示：

每一个元素被抽中的概率=所抽取的群数 $\times$ (群的规模/总体的规模) $\times$ (平均每个群中所要抽取的元素/群的规模)

PPS抽样的方法可以扩展到多阶段的情形，只要在中间的每一个阶段都同样依据概率与规模成比例的原则，除了最后一个阶段以外。

第四节 抽样

三-二、概率抽样：方法

6. PPS抽样

具体操作方法（用PPS方法抽取第一阶段样本的举例）

序号	规模	所占比例(‰)	累计(‰)	选择号码范围	所选号码	进入样本的元素
企业1	3000	15	15	000-014	012	元素1
企业2	2000	10	25	015-024		
企业3	16000	80	105	025-104	048、095	元素2、3
企业4	200	1	106	105		
企业5	1200	6	112	106-111		
企业6	6000	30	142	112-141	133	元素4
企业7	800	4	146	142-145		
企业8	600	3	149	146-148	148	元素5
企业9	1400	7	156	149-155		
企业10	4200	21	177	156-176	171	元素6
.....						
企业98	400	2	988	986-987		
企业99	1800	9	997	988-996	995	元素20
企业100	600	3	1000	997-999		

第四节 抽样

四、总体与抽样框

抽样框是总体要素的列表或准列表。通过正确抽样获得的样本信息，只适合描述构成抽样框的要素组成的总体，不能再扩展。抽样框必须和所要研究的总体是一致的。之所以要强调这一点，是因为存在一种普遍现象，研究者往往从某个抽样框抽选样本，而后对一个近似、但并不等同于抽样框的总体做结论。

例如，可以看看下面这个报告，讨论的是美国医生最常开的处方药物：

有关处方药物零售的相关信息不易取得。不过纽约州联合大学奥本尼药学院的德纳索教授已经用调查邻近药房的方式追踪了处方药物的零售情况达25年之久。他将研究结果公布在工业贸易杂志上。德纳索教授最近的调查资料（包含了整个80年代）来自纽约以及新泽西州48个社区中的66个药房。如果没有特别的情形，这个调查结果就可以代表整个美国其他药房的情况。

第四节 抽样

四、总体与抽样框

总结：

1. 根据样本所得的研究成果只能代表组成抽样框各个要素的集合。
2. 通常，抽样框并未真正包含所有要研究的要素。因此，省略几乎是不可避免的做法。因此，研究人员的首要任务是先评估被省略的内容，继而在可能的情况下进行更正。
3. 即使总体与抽样框是一致的，为了说明组成抽样框的总体，所有要素必须具有同等的代表性：基本上，所有要素应该只在名单中出现一次。在名单中出现多次的要素，会有较大的机会被抽到，因此，这样的调查结果将会有偏差。

第四节 抽样

五、样本规模与抽样误差

1. 样本规模 (sample size)

又称为样本容量，是指样本中所含个案的多少。确定样本规模也是每一项具体的社会调查所必须解决的问题之一。**统计学中通常以30为界，把样本分为大样本（30个个案及以上）和小样本（30个个案以下）。**之所以这样区分，是因为当样本规模大于30时，无论总体的分布如何，其平均数的抽样分布将接近正态分布，从而可以运用许多统计学的公式，也可以用样本的资料对总体进行推断。**但需要注意的是，30个个案的样本对于社会调查来说却常常是不够的。统计学中的大样本与社会调查中的大样本并不是一回事。**

根据一些社会调查专家的看法，社会调查中的样本规模至少不能少于100个个案。这是因为在社会调查中，研究者不仅仅需要以样本整体为单位来计算平均数、标准差、相关系数等统计量，同时更经常地需要将样本中的个案按不同的指标划分为不同的类别，进而分析不同类别之间的差别、分析不同变量之间的关系。因此，要保证所划分出的每个子类别中都有一定数量的个案，就必须扩大整个样本的规模。

在简单随机抽样中推断总体平均数的样本规模的计算公式为 $n = \frac{t^2 \times \sigma^2}{e^2}$ ，其中t为置信水平对应的临界值， σ 为总体的标准差，e为容许的抽样误差；推断总体百分比的样本规模的计算公式为 $n = \frac{t^2 \times p(1-p)}{e^2}$ ，其中p为总体的百分比。

第四节 抽样

五、样本规模与抽样误差

1. 样本规模 (sample size)

根据 $n = \frac{t^2}{4e^2}$ 可得下表:

容许的抽样误差比例e (%)	样本规模	容许的抽样误差比例e (%)	样本规模
1.0	10000	5.5	330
1.5	4500	6.0	277
2.0	2500	6.5	237
2.5	1600	7.0	204
3.0	1100	7.5	178
3.5	816	8.0	156
4.0	625	8.5	138
4.5	494	9.0	123
5.0	400	9.5	110

第四节 抽样

五、样本规模与抽样误差

2. 影响样本规模确定的因素

① 总体的规模

按一般的想法，总体规模越大时，则样本规模也要越大，这样才能保证一定的精确度。但是，这种想法只在一定的程度上是正确的。当总体规模大到一定程度时，样本规模的增加与它并不保持同等的增长速度。

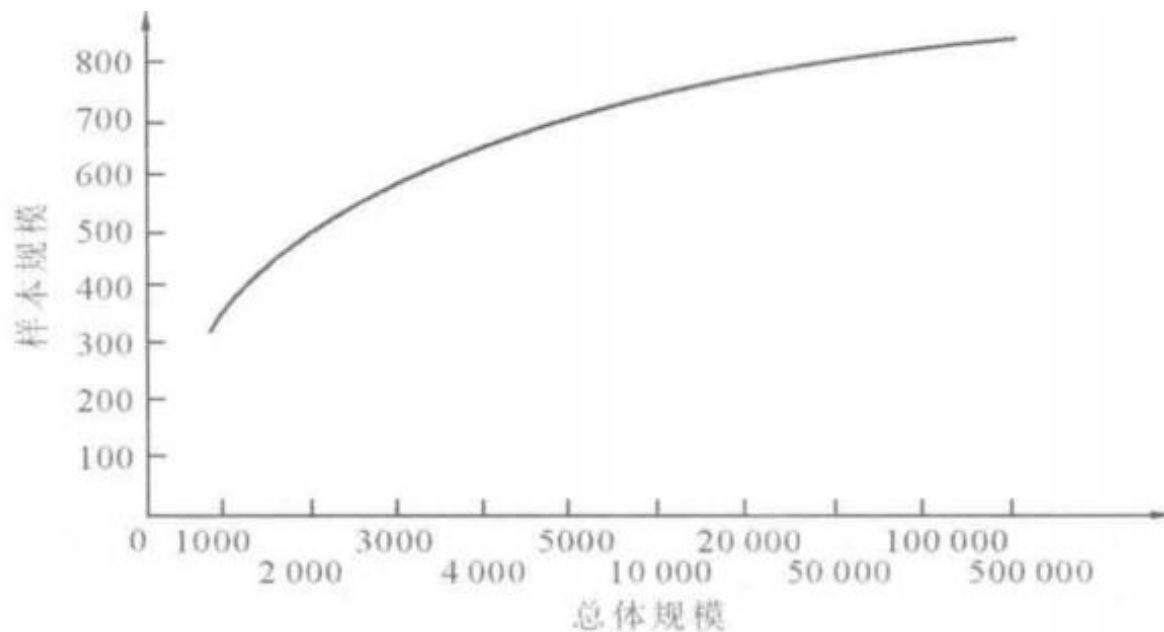


图 4-8 不同的总体规模所需要的样本量

(相对于 95% 的置信水平、 $\pm 3\%$ 的置信区间和总体参数值以 50% 对 50% 比例均分的假定而言)

第四节 抽样

五、样本规模与抽样误差

2. 影响样本规模确定的因素

② 推断的把握性与精确性

在社会调查中，用置信水平与置信区间这两个概念来说明样本规模与抽样的可靠性及精确性之间的关系。一般来说，在其他条件一定的情况下，置信水平越高，即推论的可靠性越大，则所要求的样本规模就越大。

③ 总体的异质性程度

一般来说，要达到同样的精确性，在同质程度高的总体中抽样时，所需要的样本规模就小一些；而在异质程度高的总体中抽样时，所需要的样本规模就大一些。与总体异质性程度有关的另一个因素是，当总体中的大部分成员对某个问题的回答或选择与小部分成员的回答或选择不同时，比如70%的成员选择甲，30%的成员选择乙，则所需要的样本规模要小一些；而当选择两种不同回答的成员比例相差无几时，比如说选择甲、乙的比例都为50%左右时，则所需要的样本规模为最大。

第四节 抽样

五、样本规模与抽样误差

2. 影响样本规模确定的因素

③ 总体的异质性程度

容许的抽样 误差比例/(%)	所期望的给予特定回答的总体比例/(%)					
	5或95	10或90	20或80	30或70	40或60	50/50
1	1900	3600	6400	8400	9600	10000
2	479	900	1600	2100	2400	2500
3	211	400	711	933	1066	1100
4	119	225	400	525	600	625
6	—	100	178	233	267	277
7	—	73	131	171	192	204
8	—	—	100	131	150	156
9	—	—	79	104	117	123
10	—	—	—	84	96	100

第四节 抽样

五、样本规模与抽样误差

2. 影响样本规模确定的因素

④ 研究者拥有的经费、人力和时间

从样本的代表性、抽样的精确性考虑，样本规模当然是越大越好。但抽样所得到的样本是要用来进行调查的，样本规模越大，同时也意味着调查需要投入的人力、物力和时间越多，意味着调查可能受到的限制和障碍也越多。因此，从调查的可行性、简便性考虑，样本规模又是越小越好。究竟选择多大规模的样本，调查者往往需要做出选择。而这种选择的一个重要砝码，就是研究者拥有的经费、人力和时间。总之，样本规模的确定需要综合考虑各方面因素，没有一成不变的规定。一个经验标准是：

小型调查，样本规模在100~300之间；

中型调查，样本规模在300~1000之间；

大型调查，样本规模在1000~3000之间。

第四节 抽样

五、样本规模与抽样误差

3. 样本规模与抽样误差

抽样误差就是样本的统计值与总体的参数值之间的误差。它是由于抽样本身的随机性所引起的误差。无论采取什么样的抽样方式，这种误差都是不可避免的。但是，抽样误差的大小是可以在样本设计中事先限定的。除了抽样误差以外，抽样调查中还存在另一种误差，即测量误差，也称为非抽样误差，它是指在调查、测量、询问、记录、填答、汇总等工作中所出现的误差。**抽样误差主要取决于总体的分布方差和抽样规模。**

有关抽样规模与抽样误差之间的关系问题，还应该注意两点。一是对于比较小的样本来说，样本规模上的很小的一点增加，便会带来精确性方面很明显的增加；对于比较大的样本来说，同样增加这么多个案，却收效甚微。二是扩大样本规模虽然可以减少抽样误差，但同时它又会增加非抽样误差。所以，并非样本越大调查结果就肯定会越精确。

第五节 测量与操作化

一、测量的概念

所谓测量，就是根据一定的法则，将某种物体或现象所具有的属性或特征用数字或符号表示出来的过程。测量的主要作用，在于确定一个特定分析单位的特定属性的类别或水平。它不仅可以对事物的属性作定量的说明，也能对事物的属性作定性的说明。在社会调查中，研究者进行的大部分测量往往是定性的测量。

测量的四个要素：

- ① **测量客体即测量的对象**。它是客观世界中存在的事物或现象，是我们要用数字或符号来进行表达、解释和说明的对象。测量客体对应的是“测量谁”的问题。
- ② **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，我们测量的对象虽然是某一客体，但测量的内容却并不是客体本身，而是这一客体的特征或属性。测量内容对应的是“测量什么”的问题。
- ③ **测量法则即用数字和符号表达事物各种属性或特征的操作规则**。也可以说，它是某种具体的操作程序和区分不同特征或属性的标准。测量法则对应的是“怎么测”的问题。
- ④ **数字和符号即用来表示测量结果的工具**。数字和符号对应的是“如何表示”的问题。

第五节 测量与操作化

一、测量的概念

首先，人一方面作为测量的客体或对象，另一方面又作为测量过程的主体，这给社会现象的测量带来了无法回避的主客观矛盾。无论是作为测量主体的人，还是作为测量客体的人，都具有主观意识、思想感情、思维能力和价值观念，都会对测量的过程和方式作出种种反应；人与人之间还存在着各种各样、错综复杂的社会关系。这些都使得社会现象的测量在很大程度上受到人们的认识水平和价值取向的影响，带有明显的主观色彩。

其次，社会测量的内容常常是社会中人们的行为，以及由人们的行为所构成的各种社会现象。与此同时，人们对各种社会现象所进行测量的活动本身，也是一种社会行为，也是一种社会现象。二者相互联系，也相互影响。

最后，在社会科学中，由于测量的对象十分复杂，因而测量的量化程度比较低，可重复性也比较差。对许多社会现象，例如人的智力、社会群体的凝聚力、社会职业的声望、人们的教育成就、社会参与程度等，广大的社会科学家们还没能（或者根本就不可能）建立起某种公认的、适合多种不同情况的测量单位和测量标准，以及与之相应的测量工具和测量方法。

第五节 测量与操作化

二、测量层次

1. 定类测量

也称为类别测量或定名测量，它是测量层次中最低的一种。定类测量在本质上是一种分类体系，即将研究对象的不同属性或特征加以区分，标以不同的名称或符号，以确定其类别。定类测量的数学特征主要是等于和不等于（或者属于和不属于）。而每一个被研究者则分别属于或者不属于其中某一类别。由于定类测量实质上是一种分类体系，因而必须注意所分的类别既要具有穷尽性，又要具有互斥性。即所分的类别既要相互排斥，互不交叉重叠，又对各种可能的情况包罗无遗。这样，我们测量的每一个对象都会在我们的分类体系中占据一个类别，且只会占据一个类别。

由于对任何一门科学来说，分类都是基础。其他几种层次的测量，也都把分类作为其最低限度的操作。或者说，其他的几种测量层次中，都无一例外地包含着定类测量的分类功能。而在社会现象的测量中，大量的变量又都是定类变量，分类更是最基本的目标和最经常性的操作。因此，运用好定类测量，发挥其应有的作用，是社会调查人员的一项重要任务。

第五节 测量与操作化

二、测量层次

2. 定序测量

也称为等级测量或顺序测量。定序测量的取值可以按照某种逻辑顺序将调查对象排列出高低或大小，确定其等级及次序。或者说，定序测量可以按某种特征或标准将对象区分为强度、程度或等级不同的序列。

定序测量不仅能够像定类测量一样，将不同的事物区分为不同的类别，而且还能反映事物或现象在高低、大小、先后、强弱等序列上的差异。它的数学特征是大于或小于（ $>$ 或 $<$ ），比定类测量的数学特征高一个层次，所得到的信息比定类测量所得的多。

关于定序测量有一点需要略作说明。通常，研究者为了统计分析的需要，总是将这种有高低、大小、强弱不同的序列转化成大小不等的数字。比如，将“文盲与半文盲”、“小学毕业”、“初中毕业”、“高中或中专毕业”、“大专毕业及以上”等类别，分别用数字“1”、“2”、“3”、“4”、“5”来代表。但是，要特别注意的是，这种数字并非真正意义上的“数字”，即这些数字并不具有数学中数字的实际内涵和功能。它们不能用来进行数学运算，比如，它们不能用来进行加、减、乘、除运算，而只是一种单纯表示“大”“小”的符号。

第五节 测量与操作化

二、测量层次

3. 定距测量

也称为等距测量或区间测量。它不仅能够将社会现象或事物区分为不同的类别、不同的等级，而且可以确定它们相互之间不同等级的间隔距离和数量差别。在定距测量中，我们不仅可以说明哪一类别的等级较高，而且能说明这一等级比那一等级高出多少单位。也就是说，定距测量的结果相互之间可以进行加减运算。需要注意的是，定距测量的值虽然可以为0，但这个0却不具备数学中我们所熟悉的0的含义。

4. 定比测量

也称为等比测量或比例测量。定比测量除了具有上述三种测量的全部性质之外，还具有一个有实际意义的零点（绝对零点）。所以，它测量得到的数据既能进行加减运算，又能进行乘除运算。是否具有实际意义的零点（绝对零点），是定比测量与定距测量的唯一区别。

第五节 测量与操作化

二、测量层次

数学性质	定类测量	定序测量	定距测量	定比测量
类别区分：=、≠	有	有	有	有
次序区分：>、<		有	有	有
距离区分：+、-			有	有
比例区分：×、÷				有

明确不同的测量层次具有的不同数学性质，这一点十分重要。因为在社会调查资料的整理和统计分析中，需要根据不同测量层次具有的数学特性采用不同的统计方法。另外，在对社会现象进行测量时，有一个重要的规则：**在条件许可的情况下，要尽可能对它们进行高层次的测量。**即凡是能够用定比测量或定距测量的，就一定不要只用定序测量甚至只用定类测量。因为高层次测量包含的信息更多，且高层次测量的结果很容易转化为低层次测量的结果，反之则不行。

第五节 测量与操作化

三、测量的信度与效度

所谓信度 (reliability)，即测量结果的一致性、稳定性或可靠性，反映测量工具能否稳定地测量所测度的事物或变量。它指的是采取同样的方法对同一事物进行重复测量时，所得结果一致性的程度。一致性程度越高，即信度越高；反之则信度越低。大部分信度指标都以相关系数 r 来表示，主要有以下三种。

再测信度 (test-retest reliability)。对同一群对象采用同一种测量，在不同的时间点先后测量两次，根据两次测量的结果计算出相关系数，这种相关系数就叫作再测信度。使用这种方法时，两次测量所采用的方法、所使用的工具是完全一样的。再测信度的缺点是容易受到时间因素的影响。

复本信度 (parallel-forms reliability)。复本信度采取的是另一种思路：如果一套测量可以有两个以上的复本，则可以根据同一群研究对象同时接受这两个复本测量所得的分数来计算其相关系数。复本信度虽然可以避免再测信度的缺点。但是，它要求使用的复本在形式、内容等方面都完全一致。

折半信度 (split-half reliability)。将研究对象在一次测量中所得的结果，按测量项目的单双号分为两组，计算这两组分数之间的相关系数，这种相关系数就叫作折半信度。通常，研究者为了采用折半信度来检验测量的一致性，需要在测量表中增加一倍的测量项目。这些项目与另一半项目在内容上是重复的，只是表面形式不同而已。如果被调查者在前后两部分项目上的得分之间高度相关，则可以认为这次测量是可信的。这种方法与复本信度的情况类似，它要求前后两个部分的项目的确是在测量同一个事物或概念。一旦二者测量的并不是同一个事物或概念，那么，研究者就无法用它来评价测量的信度了。

第五节 测量与操作化

三、测量的信度与效度

所谓效度 (validity)，也称为测量的有效度或准确度，指的是测量工具或测量手段能够准确测出所要测量的概念的程度，或者说能够准确、真实地度量事物属性的程度。测量的效度具有三种不同的类型，即表面效度、准则效度和构造效度，它们分别从不同的方面反映测量的准确程度。

表面效度 (face validity)，也称为内容效度或逻辑效度，是指测量内容或测量指标与测量目标之间的适合性和逻辑相符性。也可以说是测量所选择的项目在内容上是否“看起来”符合测量目的和要求。评价一种测量是否具有表面效度，首先必须知道所测量的概念是如何定义的，其次需要知道这种测量所收集的信息是否和该概念密切相关，然后评价者才能尽其判断能力之所及，做出这一测量是否具有表面效度的结论。

准则效度 (criterion validity)，也称为实用效度、预测效度或共变效度，指的是用一种不同以往的测量方式或指标对同一事物或变量进行测量时，将原有的一种测量方式或指标作为准则，用新的方式或指标所得到的测量结果与原有准则的测量结果做比较，如果新的测量方式或指标与原有的作为准则的测量方式或指标具有相同的效果，那么，我们就说这种新的测量方式或指标具有准则效度。

建构效度 (construct validity)，涉及一个理论的关系结构中其他概念（或变量）的测量。如果理论在两个概念之间建立了逻辑关系，而且实际测得的构念之间符合理论预测，则建构效度高。

第五节 测量与操作化

三、测量的信度与效度

测量的信度与效度之间存在着一定的关系。这种关系主要体现为三种情形：

- ① 有效度的测量一定是有信度的测量，简单地说就是有效必可信；
- ② 有信度的测量既可能是有效度的，也可能是无效度的；
- ③ 无信度的测量一定也是无效度的。

第五节 测量与操作化

四、概念的操作化

1. 概念、变量和指标

① 概念

概念是对现象的抽象，它是一类事物的属性在人们主观上的反映。由于现实世界中事物和现象的类型、结构的不同，复杂程度的不同，所以概念的抽象程度也有高有低。抽象层次高的概念往往包含多个抽象层次低的概念，并且往往是难以直接观察和描述的。这是因为概念的抽象层次越高，其涵盖的面就越大，特征也就越含糊；相反，一个概念的抽象层次越低，其涵盖面也就越小，特征也越明确。

概念是理论意义上的：概念是理论模型中的元素。

第五节 测量与操作化

四、概念的操作化

1. 概念、变量和指标

② 变量

许多概念往往包括若干个范畴、值或亚概念。正是因为概念具有这种多值的特性，因此，在社会调查中，人们借用了一个数学术语，把所研究的概念称作“变量”（或“变数”）。所谓变量，就是指具有一个以上取值的概念，或者说，就是包括一个以上范畴的概念。而那些只有一个固定不变的值的概念，则叫做常量。

变量具有两个重要的性质：首先，构成变量的各个值必须是穷尽的，即每个被调查者的情况都应能归于某个取值中。其次，构成变量的取值必须是互斥的，即每个被调查者的情况仅属于一个取值，而不能同时属于两个或多个取值。

根据变量取值的性质不同，又可以分为定类变量、定序变量、定距变量和定比变量，对它们的测量分别构成前述的定类测量、定序测量、定距测量和定比测量。

【对变量的上述说明不敢苟同】变量是实证意义上的：概念的测度方式和测量结果是变量。上述将概念称作变量的做法实际上假设了变量能完全测度概念。

第五节 测量与操作化

四、概念的操作化

1. 概念、变量和指标

③ 指标

表示一个概念或变量含义的一组可观察到的事物，称作这一概念或变量的一组指标。概念是抽象的，而指标则是具体的；概念是人们的主观印象，而指标则是客观存在的事物。因此，概念只能想象，而指标则可以观察和辨认。

由于指标是变量在经验层次上的一种体现，因而它同样具有变量的特征。也即是说，一个指标也会有若干个不同的取值。

指标也是实证意义上的：一个概念可以用多个指标（合成为一个变量）进行测度。

第五节 测量与操作化

四、概念的操作化

2. 操作化的含义与作用

所谓操作化，就是将抽象的概念转化为可观察的具体指标的过程，或者说，是对那些抽象层次较高的概念进行具体测量时所采用的程序、步骤、方法、手段的详细说明。

操作化在社会调查中有着极为重要的作用。第一，存在于研究者头脑中的各种概念、意识，研究者用以构建其理论模型的各种基本概念，都只有经过了合适的操作化之后，才会在普通人可以看得见、摸得着的现实社会中显现出来。换句话说，操作化就是要把我们无法得到的有关社会结构、制度或过程，以及有关人们行为、思想和特征的内在事实，用代表它们的外在事实来替换，以便于通过后者来研究前者。第二，操作化也是具有定量取向的社会调查的关键一环。尤其是在解释性研究中，若要对任何有关社会现实的理论假设进行检验，则操作化往往是不可回避的前提：只有通过操作化的过程，将思辨色彩很浓的理论概念转变成、“翻译”成经验世界中那些人人可见的具体事实，假设检验才成为可能。

第五节 测量与操作化

四、概念的操作化

3. 操作化与命题

在社会调查中，研究者有时需要通过收集问卷资料来检验某种命题或假设。而问卷设计对这种假设验证的目标同样有着重要的作用。

第五节 测量与操作化

五、操作化的方法

1. 澄清与界定概念

在具体操作上，我们首先要弄清概念定义的范围。在采用或给出某个具体的定义之前，可以先看看其他的研究者对这一概念所下的定义是怎样的。而对于那些并未对概念下正式的定义的研究来说，我们就需要从其对概念的运用中来确定他对这一概念的界定。当我们通过收集和查询，了解到有关这一概念的各种不同的定义，从而对这一定义的大致范围有所理解以后，便可对这些定义进行分类。

其次，我们要决定一个定义。列出了有关这一概念的各种类型的定义，或者总结出各种定义中最具共同性的元素后，就该决定采取哪一种定义方式了。我们面临各种不同的选择：既可以直接采用一个现成的定义（即从现有的各种定义中确定一个我们认为比较科学、比较确切的），也可以在现有定义的基础上自己创造出一个新的定义。采用现成的定义时，既可以用一个十分经典的定义，也可以用一个比较新的定义。需要特别注意的是，这种选择应该以研究者进行具体社会调查的需要为标准，哪种定义方式最适合研究的目的，就应该重点考虑哪种定义方式。

第五节 测量与操作化

五、操作化的方法

2. 发展（开发）测量指标

首先，列出概念的维度。如前所述，许多比较抽象的概念往往具有若干不同的方面或维度，或者说，一个抽象的概念往往对应现实生活中的一组复杂的现象，而不仅仅对应一个单纯的、可直接观察到的现象。因此，我们在界定概念的定義的同时，指出概念所具有的不同维度，对于概念的操作化，对于概念的测量指标的选择，以及对于综合的理论思考与分析，都是十分有用的。

其次，建立测量指标。通常可以采取两种方式。第一种方式是寻找和利用前人已有的指标。尤其是一些测量人格、态度方面的量表，其指标往往经过多次的运用和修改，常常可以成为我们可用的指标。当然，许多前人的指标不一定完全适合我们的概念，需要做一定的修改和补充。第二种方式是研究者先进行一段时间的探索性研究。研究者可采用实地观察和无结构式访问的方式，进行资料收集的初步工作。尤其是与被研究者中的关键人物进行比较深入的交谈，从这些人那里获得符合实际的答案。

第五节 测量与操作化

六、指标选择的多样性

有些抽象概念往往很难，甚至不可能在具体现象中找到与其一一对应的指标。而且，在许多情况下，一组操作化指标也往往不能够完全代表一个概念。这也是社会调查有时受到批评和指责的原因之一。

同时，对抽象概念进行操作化时，往往在具体方法和测量指标方面存在多种不同的选择。也就是说，对同一个概念进行测量时，可能会产生出不同的测量指标。而一项具体社会调查的结果，又与它采用的操作化方式及其产生的测量指标密切相关。由于一个概念常常需要多个指标来测量，所以指标的选择就成了概念能否被正确地和完整地测量的关键。

第五节 测量与操作化

六、指标选择的多样性

以对“夫妻权力”的测量为例，研究者通常将这一概念界定为“夫妻在家庭中的决策权”。对其进行操作化时，不同研究者所用的指标不尽相同。**1960年，美国社会学者布拉德和沃尔夫**在他们的研究中选取了他们认为既能涉及丈夫又能涉及妻子在家庭中的决策权的8个指标：丈夫的职业选择、买什么样的汽车、是否买人寿保险、到什么地方度假、买什么样的房子、妻子是否应该参加社会工作、家里有人生病时应去看哪位医生、全家每周在食品方面应花多少钱。他们根据调查测量的结果得出结论说，丈夫和妻子在家庭决策中各有侧重，二者的权力相当。这一研究结果被广泛引证并大大促进了后来的许多研究。但是**1971年森特斯等研究者**在对同一问题进行研究时，将“家庭中决策权”的概念操作化为14个指标，除上述8个以外，增加了6个指标：请谁来做客和与谁一起出门、怎样装饰房间和摆设家具、收看什么电视节目和广播节目、家庭的正餐吃什么、买什么样的衣服、配偶应买什么样的衣服。在他们的研究中，前8个问题的调查结果与布拉德等人的研究结果几乎完全一致，但加上6个新问题时，丈夫的权力下降了。这种结果表明，夫妻权力的对比和分布随着所考察的决策方面的不同而不同。中国台湾的社会学者也曾研究过夫妻权力的问题。**陈明穗1986年**通过15项指标进行测量：丈夫的职业选择，妻子的职业选择，家人生病该看哪个医生，家庭生活费的支配，度假、旅游及休闲活动，生育子女数目，购买贵重物品，置产（买房子、土地等），房子布置及购买家具，订阅报刊、选择电视节目等，子女的教育，谁代表参加婚丧喜庆，谁决定送礼与回赠礼物，是否买保险，何种家电用品该淘汰换新。**伊庆春、蔡瑶玲1988年**用15项指标进行测量：丈夫的职业或工作选择，妻子是否外出工作或改变工作，家用支出分配，储蓄、投资、保险等，婚丧喜庆时贺礼的数额，买房子或房地产，要不要搬家，是否与上一代同住，生几个孩子，用哪种避孕方法，小孩上哪所学校，小孩教育，请哪些客人吃饭，娱乐休闲活动的计划，家中的布置及买家具。

第五节 测量与操作化

六、指标选择的多样性

可以看出，上述这些测量指标都具有表面效度——它们的确是在测量夫妻在家庭中的决策权。美国和中国台湾的两组指标也有许多相似和相同的地方，但也有一些小的差别。如果我们根据中国大陆的情况对“夫妻权力”的概念重新进行操作化，所用的指标可能与上述两者都不完全一样。相对而言，中国大陆学者所用的指标可能会比较接近中国台湾学者所用的指标，但肯定还会有一些与其不同的地方。因为，中国大陆的社会现实与中国台湾的不同，而与美国的差别就更大了。

上述例子告诉我们，就像语言学家在翻译其他语言的作品时，常常会碰到同一部作品由不同的人翻译，译文互不相同、有优劣之分一样，对同一概念的操作化结果也存在着优劣之分。社会调查中众多理论概念的操作化的结果也不是唯一的。**各种不同的操作化结果，相互之间只是在反映概念内涵的准确性和涵盖性上存在着程度上的差别，唯一的、绝对准确、绝对完善的操作化指标是不存在的。**

第五节 测量与操作化

七、操作化的几个例子

1. 人的现代性

著名美国社会学家英克尔斯及其合作者在研究“现代人”时，需要对“人的现代性”这一概念进行测量。为此，他们进行了非常周密细致的操作化工作。最终将“人的现代性”操作化为具有24个维度的个人现代性综合量表，并在此基础上形成了一份总共包括438个具体问题的访问问卷。这24个维度是：

积极参与公共事务、年老者的角色、教育期望与职业期望、可依赖性、对变革的认识与评价、公民权、消费态度、对尊严的评价、效能、家庭大小、意见的增多、与国家的认同、信息、大众传播媒介、亲属义务、社会阶级分层、新经验、妇女权利、宗教、专门技能、对时间的评价、计划、工作信念、了解生产。

每一个维度下面，又分解成若干个更为具体的指标。比如，第一个维度“积极参与公共事务”下面，就又分解成下述6个指标：是否属于某一个组织、所参加的组织的数目、哪一个组织在政治上持有自己的观点、是否用谈话或书信方式向政府官员表明自己的观点、参加投票的次数、是否曾高度关心某件公共事务。

第五节 测量与操作化

七、操作化的几个例子

2. 父母投资

陈皆明1998年发表于《中国社会科学》的文章分析了父母投资与子女赡养的关系，将“父母投资”定义为父母为子女所做的各种帮助，并操作化为早期家庭帮助、较近期的帮助和目前正在给予的帮助以及一系列具体指标。其基本的操作化框架如下表所示。

早期家庭帮助	被访者早期对家庭帮助的依赖（5个指标：是否在家做作业、在升学/就业/换工作/分房上是否得到帮助）
	父母为子女结婚所花的费用
	被访人的教育水平
	婚后的居住情况
较近期的帮助	照看孩子
	做家务
	提供经济上的帮助
正在给予的帮助	照看孩子
	做家务
	提供经济上的帮助

第五节 测量与操作化

七、操作化的几个例子

3. 现代化的生存环境

李银河博士1996年在研究浪漫爱情时，试图检验这样一个假设：人们的生存环境越接近现代化，则越会看重浪漫爱情。为此，她需要对“现代化的生存环境”进行操作化。在研究中，她将这一概念操作化为三个不同的维度和四个主要的指标，如表所示。

概念	维度	指标
现代化的生存环境	时间维度	年龄
	空间维度	城乡社区
	社会经济地位维度	教育程度
		职业地位

当然，同一领域的其他研究者还可以增加其他维度指标，比如增加“生活水平”的维度，以及诸如“人均收入”、“是否有电话”、“是否有电脑”等指标。

第五节 测量与操作化

七、操作化的几个例子

4. 溺爱孩子

风笑天1992年在研究独生子女的教育问题时，需要了解和比较独生子女父母与非独生子女父母在溺爱孩子方面的行为表现和心理状态。他对“溺爱孩子”的概念进行了如下的操作化处理：

首先，根据“溺爱孩子”概念的名义定义，将其操作化为四个不同的维度，以反映这一概念的四个不同层面。这四个维度是：不注意培养孩子的生活自理能力、不注意培养孩子的劳动习惯、对孩子过分迁就、物质上对孩子尽量满足。

然后，在每个维度上选择若干具体指标。比如，第一个维度选择的指标有：是否自己洗头、洗澡、穿衣、收拾书包、整理床铺、自己上下学；第二个维度选择的指标有：在家是否抹桌子、扫地、洗碗、洗手绢、上街买东西等。

第五节 测量与操作化

八、量表

指标和**量表**经常被不正确地使用或混用。指标：一种复合测量，包含多个具体的观察，并代表着一些更一般的维度；量表：一种复合测量，由多个具有逻辑结构或经验结构的访题组成。

1. 相同之处

- ① 指标和量表都是对变量的定序测量，即依变量（例如信仰虔诚度、疏离性、社会经济地位、偏见、智商等）的分析单位进行排序。
- ② 指标和量表都是变量的复合测量，即基于一项以上数据的测量。

2. 不同之处

通过分值分配方式对指标和量表进行区别。指标往往通过单个属性的分值累积来建立对变量的测量。量表则是通过对访题的不同反应模式赋予相应的分值，且不同选项反映的相对强弱程度不同。量表利用了任何存在于相同变量的各种属性之间的强度结构来识别反应模式的差异。**指标根据人们的反应打分，而量表根据人们的反应模式打分：我们确定应答的逻辑模式，并且根据最接近于他们反应的模式来打分。**

第五节 测量与操作化

八、量表

1. 总加量表（summated rating scales）【也称为总和量表或总合评量】

它由一组反映人们对事物的态度或看法的陈述构成，回答者分别对这些陈述发表意见，根据回答者同意或不同意分别计分，然后将回答者在全部陈述上的得分加起来，就得到了该回答者对这一事物或现象的态度的得分。这个分数是其态度的量化结果，它的高低就代表了个人在态度量表上的位置。

贝利生育量表（Baily K D. Methods of Social Research. Third Edition. Free Press.1987: 344.）

序号	态度	同意	不同意
1	结婚的主要原因之一是要生孩子。	1	0
2	只生一个孩子是错误的，因为独生子女是在孤独中成长，且由于无兄弟姐妹而忧郁。	1	0
3	生育孩子是一个妇女所能具有的最深刻的经历之一。	1	0
4	两种性别的孩子至少都有一个比仅有一种性别的孩子好。	1	0
5	没有孩子的妇女决不会感到完全的满足。	1	0
6	男人直到他业已证明自己成为孩子的父亲时，才算是“真正的男人”。	1	0
7	（由于生育控制、绝育或年老等因素）不能导致怀孕的性活动是不道德的。	1	0
8	未结婚的或者结了婚而没有孩子的男人可能是同性恋者。	1	0
9	妇女的首要职责是做母亲，只有在不影响其母亲职责时，才谈得上她的事业。	1	0
10	没有孩子的夫妇实在可怜。	1	0

第五节 测量与操作化

八、量表

1. 总加量表 (summated rating scales) 【也称为总和量表或总合评量】

上述这类总加量表有一个潜在的假设或前提：**每一个态度陈述都具有同等的效果**，即它们在反映人们的态度方面是“等值的”，不同的陈述之间不存在数量的差别（它们的“分值”都一样）。只有在这样的假定下才能说，那些总得分为2的回答者具有同样程度的不主张生育的态度；同样，也只有在这样的假定下才能分辨出同意其中8条陈述的回答者比同意其中2条陈述的回答者有更强烈的生育义务感。

然而在实际应用中却有一个很大的困难，这就是指标（即表中的陈述）的测量效度问题。我们想测量人们对生育孩子的态度，但却往往难以保证用来进行这种测量的所有指标（本例中的10条陈述）都是在完全地测量着同一事物。比如，第五条陈述和第九条陈述所测量的或许是对“女性角色”的态度，第六条和第八条所测量的或许是对“男性角色”的态度，而第二条和第四条则可能是在测量人们对孩子“社会化”的态度，因而，对于同样的2分，这一量表所实际测量的并不一定是同样的态度和在这种态度上完全一样的程度。

第五节 测量与操作化

八、量表

2. 李克特量表（Likert scaling）

表1 请您对下列看法发表意见（请在每一行选择一个方框打“√”）

序号	看法	非常同意	同意	无所谓	不同意	很不同意
1	婚事应该尽量办得简单一些。	☐	☐	☐	☐	☐
2	结婚是人生一件大事，婚事应该办得隆重、热闹，花再多钱也值得。	☐	☐	☐	☐	☐
3	就是有钱，婚事也不应大操大办。	☐	☐	☐	☐	☐
4	为了不让别人笑话，就是借钱也要把婚事办得像个样子。	☐	☐	☐	☐	☐

表2 四种看法的赋值

看法	非常同意	同意	无所谓	不同意	很不同意
1	1	2	3	4	5
2	5	4	3	2	1
3	1	2	3	4	5
4	5	4	3	2	1

第五节 测量与操作化

八、量表

2. 李克特量表 (Likert scaling)

李克特不仅改进了总加量表的形式，更重要的是提出了一种帮助研究者从量表中消除有问题的项目 (item, 即陈述) 的方法。这种方法成为研究者设计总加量表时确定量表项目的主要依据，其基本程序如下。

- ① 围绕要测量的态度或主题，以赞成或反对的方式写出与之相关的看法或陈述若干条（一般为20-30条）。对每一陈述都给予5个答案：非常同意、同意、无所谓、不同意、很不同意，并根据赞成或反对的方向分别赋以1、2、3、4、5分。
- ② 在所要测量的总体中，选择一部分对象（一般不能少于20人）进行试测。
- ③ 统计每位受测者在每条陈述上的得分以及每人在全部陈述上的总分。
- ④ 计算每一条陈述的分辨力，删除分辨力不高的陈述，保留分辨力高的陈述，形成正式的量表。

分辨力的计算方法是：先根据受测对象全体的总分排序；然后取出总分最高的25%的人和总分最低的25%的人，并计算这两部分人在每一条陈述上的平均分；将这两个平均分相减，所得出的就是这一条陈述的分辨力系数。该系数的绝对值越大，说明这一陈述的分辨力越高。

被调查者		题目												个人总分
		1	2	3	4	5	6	7	8	9	10	11	12	
总分最高的25%的人	1	4	5	5	4	3	5	4	4	3	5	2	5	49
	2	5	4	4	5	5	4	3	2	5	4	1	4	46
	3	5	4	3	3	4	5	4	3	4	4	2	5	45
	4	4	4	4	4	5	3	3	3	4	5	1	4	45
	5	5	5	3	2	4	4	3	4	5	2	2	4	43
中间50%的人	6	4	3	2	5	4	5	4	4	2	3	1	5	42
	7	4	4	4	4	2	3	3	4	4	3	2	4	41
	8	3	3	4	4	2	3	5	4	2	3	2	5	40
	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	14	2	3	2	3	3	2	3	4	4	3	2	4	36
	15	2	4	2	3	2	2	3	4	4	3	1	4	34
总分最低的25%的人	16	2	2	4	2	3	3	2	1	4	2	2	5	32
	17	2	2	2	3	4	2	4	1	3	3	2	4	32
	18	1	3	2	4	1	3	3	2	1	2	2	5	29
	19	1	1	2	2	2	3	2	3	4	1	1	4	26
	20	1	1	1	2	1	2	1	2	3	2	2	3	21
总分最高的25%的人的平均分		23/5=4.6	4.4	3.8	3.6	4.2	4.2	3.4	3.2	4.2	4.0	1.6	4.4	
总分最低的25%的人的平均分		7/5=1.4	1.8	2.2	2.6	2.2	2.6	2.4	1.8	3.0	2.0	1.8	4.2	
分辨力系数		3.2	2.6	1.6	1.0	2.0	1.6	1.0	1.4	1.2	2.0	-0.2	0.2	

第五节 测量与操作化

八、量表

3. 语义差异量表 (semantic differential)

也称为语义分化量表，主要用来研究概念对于不同的人所具有的不同含义。在社会学、社会心理学和心理学的研究中，语义差异量表主要用于文化的比较研究、个人及群体间差异的比较研究，以及人们对周围环境或事物的态度、看法的研究等。

语义差异量表的形式由处于两端的两组意义相反的形容词构成，每一对反义形容词中间分为七个等级，每一等级的分数从左至右分别为7、6、5、4、3、2、1，也可以计为+3、+2、+1、0、-1、-2、-3。被测量的概念或事物（比如某一群体、某种问题、某个国家等）放在量表的顶端，调查时要求被调查者根据自己的感觉在每一对反义形容词构成的量尺中的适当位置画记号，比如画×号。研究者通过对这些记号所代表的分数的统计和计算，来研究人们对某一概念或事物的看法或态度，或者进行个人或团体间的比较分析。

语义差异量表所采用的数对形容词要能够考察被调查者对研究对象的感觉和态度的各种要素或各种维度。许多研究者认为，这种形容词中通常包括三个一般的维度：即评价（比如好与坏、善良与残酷、重要与不重要等）、力量（比如强与弱、硬与软、刚与柔等）和行动（比如主动与被动、快与慢等）。

第五节 测量与操作化

八、量表

3. 语义差异量表 (semantic differential)

姐妹							
	+3	+2	+1	0	1	2	3
热情的							冷漠的
主动的							被动的
强 的							弱 的
快 的							慢 的
大 的							小 的
慈善的							残忍的