

第四章 抽样

4.1 抽样问题

抽样指从一个较大的个体集合(总体)中选择部分个体的过程。抽样调查则是从总体中抽选能代表总体的一部分样本, 然后通过样本的情况来对总体情况进行估计和推断。在社会科学研究(以及自然科学)中, 研究的对象在理论上往往包含相当多的个体, 例如对科技人员信息意识的调查, 对大学生信息检索行为的研究, 或对青年人择偶心理的研究等。各方面的限制通常使得我们不能对所有符合定义的个体都进行研究, 而只能从中选取部分个体来实际进行研究。这时, 选取样本的方法影响到所选择样本能否尽可能准确地代表整个总体, 能否满足研究本身的需要, 能否允许有效地估计样本与总体间的误差, 以及选取本身的效率。

4.1.1 抽样对象

抽样中涉及不同层次的抽样对象, 我们在此详细说明:

(1) 研究个体

研究个体指符合研究目的和研究内容规定的具体研究对象, 是研究中实际观察测量和分析的基本单位, 例如在信息用户的信息检索行为研究中, 研究个体就是作为信息用户的个人。在许多研究中研究个体可以是一定的社会群体, 例如在企业信息政策对企业内员工劳动生产率影响的研究中, 研究个体是制定和实施信息政策的企业。还有许多研究中, 研究个体是非生物体, 例如情报学对文献标引度或引文状况的研究中研究个体就是文献或文献引文。研究个体由研究假说明明确规定, 但在某些情况下研究个体可能难以准确界定。例如, 在对信息检索的查全率和查准率进行研究时, 研究个体将是“每一次”信息检索, 但在实际上用户每一次检索可能包含了多个检索过程。或者在对社会项目效用进行研究时, 社会活动的范围就可能较模糊, 从而使研究个体范围难以确定。

(2) 研究总体

符合研究目的和研究内容规定的研究个体之总和被称为研究总体, 又称研究范围。研究总体也由研究假说明明确规定, 体现了研究假说将推广概括的范围, 例如在文科大学生阅读倾向研究中的文科大学生, 在企业科技人员情报需求研究中的企业科技人员, 在企业信息政策研究中的企业, 在社会信息传播活动效用研究中的社会信息传播活动等都界定了研究总体。当然, 从上述范例已可以看出, 研究题目甚至研究假说对研究总体的规定亦有一定模糊度, 需要进一步澄清, 例如需要进一步定义文科大学生的类型和范围, 或在信息用户检索行为研究中也应明确说明什么是信息用户, 或在顾客购物心理研究中的顾客包括哪些人。

(3) 抽样总体

由于种种原因, 研究者实际上不能从所有符合研究假说定义的研究个体中抽样选择实际研究个体, 而只能从某一部分符合定义的研究个体中抽样选择实际研究个体, 那么这一部分被实际抽样选择的研究个体之和就是抽样总体。例如, 在文科大学生阅读倾向研究中, 我们实际上不能对全国所有文科大学生进行抽样选择, 而只能在某城市的某几所主要大学的文科大学生中选择, 这时的抽样总体就是这几所大学的文科大学生。或者, 我们甚至不能对这几所大学的所有文科大学生(应包括研究生, 双学位生, 本科生, 专科生及进修生)进行抽样选择, 而只能对其中的本科生进行抽样选择, 这时的抽样总体就更加缩小了。可以看出, 抽样总体可能与名义上的研究总体有很大差别, 而抽样总体才真正决定和代表了研究假说验证推广的实际范围。但是, 研究假说往往没有说明实际抽样总体, 而留待研究报告对此进行解释。严谨的研究者应在研究假说中尽可能地明确实际抽样总体, 在研究报告中明确说明实际抽样

总体, 并在对研究结果进行推广概括时充分考虑到抽样总体局限性的影响。

(4) 抽样框

抽样框指将抽样总体内所有抽样个体按照一定形式逐一地和唯一地排列起来而形成的抽样个体表, 例如职工名册, 学生登记册等。抽样框可以直接利用研究个体的某种现成自然排列, 例如前述职工名册, 居民电话簿, 用户登记表等。或者, 抽样框可通过专门统计选择而形成, 例如从企业员工中选择曾经从事科学研究或技术改造的职工作为抽样总体, 将其姓名排列而成抽样框。有时, 研究者不能明确规定抽样框。例如在咨询员与信息用户交互作用研究中, 信息用户实际上指所有可能前来进行咨询的用户, 但事先不能确定这些用户的范围, 当然就不能建立他们的名册, 而只能从在研究延续时间内实际进行咨询的用户中选择。无论利用什么形式形成抽样框, 抽样框界定了实际的抽样总体, 是抽样的基础。

抽样框可能存在两方面的问题。一是抽样框可能与名义抽样总体有较大差别。例如, 名义抽样总体是某城市的全体居民, 但使用的抽样框是居民电话簿, 因此很可能许多居民因为经济原因或地区原因没有安装电话, 或者安有电话但因保密等原因没有在电话簿中记载中选择; 这时抽样框只代表了抽样总体的一部分。第二个问题是抽样框中个体可能不唯一, 即同一个体可能在抽样框中出现多次。例如在研究图书馆文献主题分析深度时, 名义抽样总体是某图书馆的全部图书, 抽样框是图书馆主题目录, 这时显然有许多图书由于包含多个主题而在主题目录中出现多次。

(5) 样本群

样本群有时又简称样本, 指通过抽样所形成的研究个体集合。样本群是研究中实际观察测量和分析的对象。

根据上述讨论, 我们可以用图 4.1 来描述抽样对象及其相互关系。

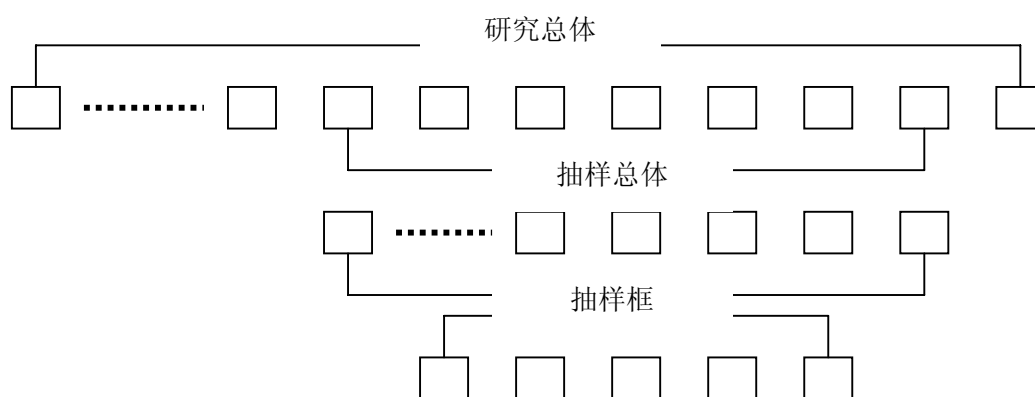


图 4.1

4.1.2 抽样过程与意义

(1) 抽样过程

一般来说, 抽样过程涉及以下几个步骤:

- a. 根据研究目的和内容确定研究个体。
- b. 根据研究目的和内容确定研究总体。
- c. 根据研究目的和内容, 并根据研究条件, 确定抽样总体。
- d. 选择和确定抽样框。
- e. 确定抽样方法。
- f. 实施抽样。

(2) 抽样的意义

一般来说, 人们因为以下种种原因放弃全面调查而采取抽样调查:

a. 节省开支, 简化操作, 提高研究效率。在许多情况下, 研究总体(甚至抽样总体)范围很大, 包括了许多个体, 对他们进行全面调查将花费太多时间和经费, 使得这样的全面调查远远超过调查者的实际能力。例如, 对所有科技人员信息利用情况的调查, 对所有大学生择偶心理的调查等。有时, 即使能对所有个体进行调查, 但调查时间过长, 一方面可能耽误决策, 另一方面不同个体在调查时期发生不同变化, 从而影响研究效度与信度。

b. 总体中可能有些个体难以接触, 无法对之进行调查。例如临时出差或住院的个体, 居住偏僻或地址差误的个体, 没有回答能力的个体(文盲或残疾人), 或无法接触的领导人。

c. 某些研究的具体测量手段可能具有破坏性, 例如产品的耐酸耐压试验; 有些研究的测量过程可能对研究对象的正常活动有妨碍(或占用研究对象的时间精力), 例如对科技人员在学术会议期间个人信息交流活动的追踪观察; 还有些研究可能对研究对象造成不利影响, 例如对青少年进行家庭暴力或吸毒行为的调查等。

d. 科学设计的抽样, 不仅可以使样本具有代表性和能采用严谨的数学方法估计样本结果与总体实际情况的误差范围, 而且可以将精力和经费集中于少量个体(样本)的研究中, 容易获得准确可靠的数据。而全面调查往往在时间和经费压力下容易草率匆忙, 反而难以获得准确可靠的数据。

4.2 概率抽样

如果在某一抽样方法中抽样总体中每一个体被抽中的概率是已知的, 则这个抽样方法属于概率抽样。概率抽样的具体形式有简单随机抽样、系统抽样、分层随机抽样、区域抽样等。概率抽样遵循随机原则, 在抽样过程中样本选取不受任何主观因素及其系统性因素的影响, 从而保证每个个体都有一定的抽中可能。随机原则包括对客观性和随机性的要求。所谓客观性指在整个抽样过程中每个个体是否被抽中不受任何主观因素影响, 所谓随机性是保证每个个体都有被抽中的可能。

随机原则将抽样方法建立在概率论严密的理论基础上, 使我们能够利用概率论原理和方法, 通过样本结果对总体结果进行比较准确的推断, 而且能够比较精确地对样本结果与总体结果的可能误差进行估计。

4.2.1 简单随机抽样(Simple Random Sampling)

简单随机抽样是概率抽样方法中最著名和最基本的一种方法, 其特征是: 抽样总体中每一个个体被抽中的概率都完全相同而且相互独立, 所谓“独立”是指一个个体被抽为样本的概率与其它个体是否被抽为样本完全无关。

(1) 简单随机抽样过程

简单随机抽样的基本过程是:

a. 通过一定方式建立一个可靠(无遗漏个体, 无重复个体)的抽样框; 可以利用现成的名册或其它自然排列来建立抽样框, 也可以对抽样总体中实际个体进行统计列举来建立抽样框。

b. 确定将要抽取的样本数。样本数量可以根据有关经验确定, 也可以根据误差范围和置信度的要求通过专门统计公式来计算。我们将在有关章节对此进行专门介绍。

c. 将抽样框中每个个体依次编上数字代号, 形成一个与抽样框完全对应的数字代号表(随机抽样框)。

d. 利用随机数表选取或利用计算机随机数产生程序产生与样本数量相同的随机数。

e. 将获得的随机数作为数字代号, 从随机抽样框中选出对应的个体。所有选出的个体的集合就形成样本。

例如, 我们要研究大学生阅读倾向, 选择某综合大学作为实际抽样总体, 并以学生学籍

管理部门正式名册作为实际抽样框。该名册登记学生 10, 223 名, 将他们按名册排列自然顺序依次编号为 1, 2, 3, ..., 10221, 10222, 10223。样本数确定为 500, 利用计算机随机抽样程序在 1-10223 间产生 500 个随机数, 得到的随机数为 2709, 33, 495, 10103, 6712, ..., 701, 4, 1483。(如果随机数出现重复, 则第二次出现的同样数字不用, 相当于同一个体已被移出抽样框, 不能重复抽选。这时需要继续产生新的随机数)。当获得 500 个不同的随机数后, 根据这些随机数, 我们将抽样框中编在对应位置上的学生抽出作为样本。

(2) 简单随机抽样的若干问题

a. 小样本偏差。在随机抽样中, 如果样本数量较小, 很容易使从样本中获得的结果分布与抽样总体的实际分布出现很大差异, 使抽样调查结果无法代表抽样总体的实际情况。例如, 在一个 500 居民的抽样总体中抽取 20 人样本调查居民的读报倾向。由于随机抽样的“偶然”因素, 有可能“鬼使神差”地恰好抽到 17 名退休职工, 他们空闲时间较多, 每日必读大量报纸(尤其是生活与娱乐新闻)。从这个样本得出的结论就会是: 绝大多数居民每日读报 3-4 小时, 读报内容主要是生活与娱乐性新闻。但实际情况却是: 退休职工只占居民总数的 12%, 而占居民总数 64%的“上班族”和占居民总数 24%的“学生族”每日读报时间只有 0.5-1.0 小时左右, 且更关心政治经济和社会方面的内容。显然, 上述样本的情况远不能真实地代表抽样总体的实际情况。其实, 我们每人可能都有遭遇随机抽样“偶然”性的经验。例如抛掷一枚硬币, 已知国徽在上的概率是 $1/2$, 但如果只抛掷 10 次, 您可能有 7 次都是国徽在上。您如果就此断定国徽在上的概率是 $7/10$, 您就犯了大错误。

随机抽样的偶然性因素可以通过增加样本量来平抑。因为随机抽样中每个个体被抽中的概率是相同的, 因此抽出的个体越多, 抽样总体中数量分布多的类别中被抽为样本的个体就越多, 样本中个体分布就越接近抽样总体的实际分布。这样, 当样本量足够多时, 样本中个体分布就可能足够接近抽样总体中个体分布, 样本就可能足够准确地代表抽样总体。因此随机抽样往往要求较大的样本量。不过, 只有当样本个体数达到抽样总体个体数时, 样本的分布才与抽样总体的分布一致。这就是说, 即使是很大的样本量, 从中所得结论仍然可能与抽样总体实际情况有差异, 即样本与抽样总体间总是可能存在误差。抽样技术的一个重要内容就是估计抽样造成的误差, 我们将在以后讨论该问题。

b. 概率不独立问题。简单随机抽样要求每一个个体被抽中的概率相同且独立, 但实际操作中往往不能满足这个条件。例如在上述抽样中, 在抽样前每个学生都有 $1/10223$ 的概率被抽中。但在抽样过程中, 被抽中的个体(或相应的数字)不能重复使用, 相当于已从抽样框中移出, 抽样总体就逐步地缩小, 余下的学生被抽中的概率就逐步地增大, 例如在抽出了 250 人时余下学生被抽中的概率就成为 $1/9967$, 在抽出了 499 人时余下学生被抽中的概率就成为 $1/9724$ 。也就是说, 在不重复抽样(sampling without replacement)时, 每个个体被抽中的概率实际上是变化的, 并不是完全相同和独立的。如果抽样总体较小, 无论样本量本身大小, 概率变化的影响都可能很大, 以致无法进行必要的抽样误差验证。但如果抽样总体很大, 而样本量相对很小, 每个个体被抽中的概率变化会很小, 可以被视为具有相同概率。在社会科学研究中一般都采用不重复抽样, 因此随机抽样的一个重要条件就是抽样总体要足够大且显著大于样本。

c. 建立抽样框的困难。简单随机抽样依赖一个可靠的抽样框, 然而要建立一个可靠的抽样框可能很困难。一方面, 许多现成的名册存在严重的遗漏或重复, 例如机构或个人电话簿可能登载不全, 职工名册可能没及时记入新职工却保留了许多退休职工。许多名册只登记了在一段时间参加某个活动的人员, 却完全忽略在此时间里恰好没参加该活动的人员, 因此实际上只是理论抽样总体中部分个体(很可能是一部分特殊个体)的名册。例如图书馆咨询记录只记载该季度内接受过咨询服务的用户(有些可能接受多次), 科研成果数据库只记载了通过特定途径申报或公布科研成果的科研人员(有些也可能申报或公布多次), 前者就不适用于

作“用户对情报咨询服务的需求调查”的抽样框, 而后者也不适于用作“科技人员情报意识研究”的抽样框。当然, 绝对完善可靠的抽样框可能并不存在, 但研究者必须考察和保证抽样框满足研究目的和方法的基本要求, 并明确指出抽样框存在的问题及其可能引起的误差。

另一方面, 许多时候不存在任何现成的抽样框。例如, 我们欲研究在大型商场, 精品商店和一般商店购物的顾客心理与行为, 事先却并不知道有多少和有哪些顾客会出现; 欲在街头对行人进行民意测验, 也无法事先确定有多少和有哪些人会恰好在调查时出现。对这类动态总体一般无法事先构造抽样框, 更谈不上随机抽样了。

d. 实施调查的困难。在许多大规模调查研究中, 即使有了抽样框并据此进行随机抽样产生出随机样本, 可能因为样本分布太散, 对该样本进行研究将花费太多时间和经费, 在实际上无法按此样本进行研究, 例如对城市居民消费水平的调查, 对全国大学生信息能力的研究等。

4.2.2 系统随机抽样(Systematic Random Sampling)

系统随机抽样是指从抽样框中有规律地(系统地)抽取样本个体的方法, 具体来讲上: 在一个个体随机排列且按顺序编号的抽样框中, 从初值 I 起按间距 K 等距地抽取个体为样本, 抽出的 n 个样本将构成系统随机抽样的样本。为保障系统抽样具有起码的随机性, 一般要求:

- a. 抽样个体在抽样框中是随机排列的;
- b. 初值 I 是在 $1-K$ 范围内随机选取的;
- c. K 值恒定, 其具体值由公式 $K=N/n$ 决定, 其中 N 为抽样总体数量, n 为样本数量。

例如, 我们欲从 10500 名学生中选取 500 同学作为调查对象, 又假定已有符合系统随机抽样的抽样框。这时采用系统随机抽样方法, 首先确定 K 值为 21, 然后在 $1-21$ 间随机选定 13 为初值 I , 则选取的样本个体分别为 $I+n*21$, 其中 n 为 1, 2, 3, ..., 498, 499。这样, 我们获得样本个体 13, 34, 55, 76, ..., 10437, 10459, 10480。

系统随机抽样并不是严格的随机抽样, 只是随机抽样在实践上的近似。因为在系统抽样中, 每个个体被抽中的概率并不完全独立, 除第一个样本的概率等同或近似为 N/n 外, 其它样本的抽取概率就由初值 I 和 K 值共同决定。尽管如此, 由于这种方法简单实用, 工作量少, 而且能保障样本个体在抽样框中“平均分布”, 故实际应用仍很多。

在系统随机抽样中, 要十分注意抽样框中个体排列的随机性。如果抽样框中个体排列呈现任何有逻辑意义的有序排列, 则由系统随机抽样产生的样本有可能也具有这种有序性, 从而破坏样本的代表性。例如, 假定有 15000 学生的名册作为抽样框, 但名册是按班级排列, 每班平均人数为 30 人, 在每班内女同学总是排在前面。这时样本量为 500, 则 K 值为 30。如果初值 I 被选定为 2 或 26, 则抽选出的样本中女同学(或男同学)就会占绝大多数。所以, 一般要求要排除抽样框本身的逻辑有序性, 才能使用系统随机抽样。不过, 如果抽样框的逻辑有序性很弱, 且这种规律性与 K 值完全不重合, 则在初值 I 随机选需的情况下仍可使用系统随机抽样。

4.2.3 分层随机抽样(Stratified Random Sampling)

分层抽样又称分类抽样或类型抽样, 它首先将整个抽样总体中的所有个体按某个特征划分成若干个“层”或“类”, 然后用简单随机方法分别从各个层或类中抽选样本, 并将从所有层中选出的样本之和构成总样本。

分层随机抽样的效率可由下例说明: 假定我们欲调查大学生的业余活动安排, 抽样总体确定为某综合性大学的 5000 本科生, 其中理工科生为 2800 名, 文科生为 2200 名, 样本量为 250 名。如果我们采取简单随机抽样, 理想的样本应包含 140 名理科生和 110 名文科生。但由于随机抽样的“偶然性”因素, 也完全有可能抽出 213 名理工科生和 37 名文科生。又假定

我们已知,理工科生的业余活动安排与文科生差异很大,那么我们完全有理由推断,由该样本所获得的“平均”结果将更接近理工科生的情况,而不能真实反映所有学生的情况。因此,我们可将抽样总体划分为理工科生子总体($N_1=2800$)和文科生子总体($N_2=2200$),再按比例分别从各子总体中随机抽选相应的子样本($n_1=190$, $n_2=110$),这样由 n_1+n_2 构成的总样本就更能代表抽样总体,由此样本获得的“平均”结果会更接近总体平均值。

根据上述范例,我们总结分层随机抽样的要求为:

- a. 总体中存在着若干个按照某种易于鉴别的特征来划分的个体群。
- b. 有证据表明,这些个体群在所关心的测量现象上内部差异较小,但彼此差异较大;
- c. 这些个体群相互不重迭。

按照如此要求将抽样总体划分成若干个“层”,就排除了那些可能引起样本结果与总体结果差别过大的样本组合。有时,还可利用多个特征来分层,例如将大学生按性别和专业分层。

分层随机抽样还被用来处理总体中含有数量小的特殊类别个体时的情况。当总体中存在这样的个体类别时,采用简单随机抽样或系统抽样所产生的样本就很可能将这类个体完全排除在样本之外,使他们的情况在根据样本作出的研究结论中不能得到体现。例如,在某综合性大学中,文科学学生为 1000,其中女生 523;工科学学生 700,其中女生 237;理科学学生 600,其中女生 29;采用简单随机抽样选取 200 人的样本,有可能没有一个理科女生。采用分层随机抽样,至少可保证那些数量少的个体类别有代表出现在总体样本中。

当然,这又可能产生另一问题:如果按比例在各个“层”中随机抽样,在个体数量少的层中可能就只有极少数个体(例如一,二个)被抽入总体样本,而这极少数个体因种种原因可能与同层的其它个体存在较大差异,使得他们不能真正代表所在层的其它个体和整个情况。例如在上述 23 名理科学学生中按比例抽出一名学生,调查其业余生活习惯;可是该生家长为本校教师,该生每日回家食宿,看电视成为她的主要业余安排,其它学生因缺乏条件一周也难得看一次电视。为避免个别异常个体的影响,可以在这些数量少的层中增大抽样比例,从而增加从这些层中选取的个体数量。不过,在统计总体平均量时必须考虑各层抽样比例不一致的情况。

4.2.4 区域抽样(Cluster Sampling)

区域抽样又称多阶抽样或阶段抽样,可由下例说明。

假定我们欲了解某城市小学生家庭教育情况,并决定采用对小学生访谈为调查手段(因为这样能深入了解情况和避免家长与教师的影响)。如果采取简单随机抽样或上述其它方法,都会在实际操作上遇到很大困难。一是难以获得一个完整的全市小学生名册;二是即使获得这样的名册,抽样结果可能是每一小学中可能只有一,二名学生,整个样本却散布在全市,安排访谈十分困难,相应的时间和经费开支也会很大。为避免上述困难,可设计以下方法:首先将该城市划分为多个社区,以社区为一级单元随机抽取一定数量社区为一级单元,它们构成一级子总体;然后以学校作为二级单元从一级子总体中的社区分别随机抽取若干个学校为二级单元,构成二级子总体;此后,以班级为三级单元从二级子总体中的学校分别随机抽取一定数量的班级为三级单元,构成三级子总体;最后在这个三级子总体中,以所有学生作为最终个体,从各个班级分别随机抽取一定数量的学生作为样本,他们构成这个研究的实际样本。

上述方法即是区域抽样,其步骤可总结如下:

- a. 首先将总体划分为若干个一级单元,每个一级单元中包含若干个二级单元,每个二级单元中包含若干个三级单元,如此直到 N 级单元,每个 N 级单元中包含若干个最终个体。
- b. 同级单元互不重迭,且共同覆盖所有个体。
- c. 在 R 个一级单元中随机抽取 r 个一级单元,再从抽出的一级单元中分别随机抽取 $p(i)$

个二级单元($i=1, 2, \dots, I$), 再从抽出的二级单元中分别随机抽出 $m(j)$ 个三级单元($j=1, 2, \dots, J$), 如此直到从抽出的 N 级单元中分别随机抽出 $n(k)$ 个最终个体($k=1, 2, \dots, K$)。

d. 在区域抽样中, 各级单元中的子单元数可以是相同的或不同的。

e. 在每一级抽样时, 可以采取简单随机抽样, 或者采取分层随机抽样。在不同级次的抽样中可以使用相同的或不同的抽样方法; 即使采用相同的抽样方法, 其抽样比例或分层方法可以是相同的或不同的。

在样本平均值统计中需要将子单元数不同或抽样比例不同的影响考虑在内。

区域抽样并不能严格保证每个个体具有相同的和独立的抽中概率, 因为个体抽中的概率取决于它所在单元被抽中的概率, 某个单元被抽中会大大增加它包含的个体的抽中概率, 而某个单元被排除会取消它包含的个体的抽中机会。因此, 区域抽样与简单随机抽样相比, 更有可能产生与总体实际分布相异的样本(例如最后抽出的学校都处于大专院校科研机构和政府机构集中的社区, 学生生源偏好), 从而增加了产生抽样误差的机会。因此, 一般来说, 区域抽样时的样本量应适当大于简单随机抽样时的样本量, 并应保证每一级抽样时的随机性。

区域抽样简单易行, 在全社会性大规模调查中应用广泛。尤其在中间抽样单元是行政单位时, 一般都采取区域抽样。

4.3 非概率抽样(Non-Probability Sampling)

非概率抽样是在抽样总体中各个个体被抽中的概率无法知道的情况下进行抽样的方法, 包括方便抽样、配额抽样、目的抽样、自愿者抽样等。尽管它们在样本代表性和样本结果误差估计等方面不如概率抽样严谨, 但也在社会科学研究中有较广泛的应用。

4.3.1 方便抽样

方便抽样, 又称偶然抽样(Accidental sampling), 指研究者从抽样总体中选取那些可以方便获得的个体作为样本的抽样方法。例如, 某心理学教师选取自己所教的班级学生作为大学生心理健康调查的样本, 某研究人员选择在某百货商场遇到的头一百名顾客作为顾客购物心理调查的对象, 或在街上行人中每隔十人选取一人作为民意测验对象等, 都属于方便抽样。

方便抽样以简单方便的方式选取一定数量的个体作为样本, 其特征是由于时间地点或其它方面的限制不能对抽样总体进行全面抽样, 而只能对条件允许范围内的个体按某种简单方式进行抽样。所谓“条件”往往是抽样者的时间和经费等。但是, 如此抽得的这些个体完全可能与其它个体不同, 从而使依赖方便抽样得到的结论与总体实际情况差别较大, 而且对样本与总体的误差无法估计。

例如前例中的教师任课班级很有可能在学生构成和专业知识等方面与其它班级不同, 而某个时间出现在某百货商场的一百名顾客也可能出现在其它商场或其它时间的顾客在购物心理与行为方面不一致(例如全是退休职工, 至少排除了平时不在该商场购物或不在此时购物的顾客)。

还应注意, 有些抽样方法虽叫作其它名称(甚至可能叫随机抽样), 但实质上属于方便抽样, 只是其“方便抽样”性质较隐蔽或体现在具体操作中。例如, 我们欲进行信息用户的信息需求和信息检索行为调查, 请某省所有省地市县图书馆从本馆用户中随机抽选若干名用户进行调查。但在实际选取调查对象时, 各馆可能只是选择那些工作人员熟悉的且预先知道能够合作的(甚至是估计能给出预期结果的)用户, 从而使许多个体是由“方便”因素抽中的。在通过行政单位(例如大专院校及其各系)进行的许多调查中, 由于实施调查者对最终结果没有责任, 更多地关心如何取得配合以尽快完成任务, “方便”抽样情况就很普遍。

当然, 我们并不一律排斥方便抽样。在有些“初步了解情况”的探索性预研究中, 研究

者的目的是对研究对象有个感性认识, 以便为详细研究提供线索和框架。这时为了节省时间和经费, 往往采取方便抽样或其它非概率抽样。在有些以行政单位为子抽样总体, 且参加的行政单位数量多种类复杂时, 要编制统一抽样框很不容易, 而严格控制各单位具体抽样方法也很难, 可能只有依赖方便抽样, 以便能在较短时间内完成研究。

4.3.2 配额抽样

配额抽样(Quota Sampling)力求克服一般方便抽样时忽略个体差异分布的缺点, 希望能在抽样操作简单方便的同时尽量获得具有代表性的样本。配额抽样的具体方法是: 将抽样总体按照某种特征划分为多个个体群(抽样子总体), 然后从这些个体群中分别抽取一定数量的个体作为样本, 从而使每类个体在总样本中都有代表。这种方法与分层随机抽样类似, 个体群类似于“层”, 都是按照某种特征划分的个体组合。相对于要研究的现象, 这些个体群的内部差异可能较小, 但它们彼此间差异可能较大。因此, 每个个体群都应在抽样总体中有体现, 才能使样本分布接近抽样总体的实际分布。

与分层随机抽样不同的是, 配额抽样在每个个体群抽样时仍采取方便抽样。例如, 在前述对顾客购物心理行为的调查中, 可以每隔两小时对遇到的头十名顾客进行调查, 或每隔两小时对遇到的头五名女顾客和头五名男顾客进行调查, 或对上午和下午来商场的顾客中 30 岁以下, 30-40 岁, 40-50 岁, 50-60 岁, 60 岁以上各年龄段的头五名男顾客和头五名女顾客进行调查。在涉及许多行政单位的大规模调查中, 往往也会要求各具体抽样单位按类按比例抽取研究对象。

严格来说, 配额抽样中的配额额度应与各个个体群在抽样总体中所占比例一致。例如, 某高校男女学生比例为 6: 4, 1-4 年级学生比例为 3: 3: 2: 2, 样本量为 200, 则配额应是 1-4 年级男生分别为 36, 36, 24, 24, 1-4 年级女生分别为 24, 24, 16, 16。当然, 在许多时候我们并不知道总体中个体群的比例。例如在调查长途旅行者的旅途生活习惯时, 我们也知道其中男女数量不一样, 而且年龄比例也不一样, 但我们往往并不知道男女和年龄的真正比例, 可能只能根据经验来决定比例, 或通过小范围普查来近似表达比例。

4.3.3 目的抽样

目的抽样(Purposive sampling)是研究者按照研究目的的需要, 根据主观判断来选择“合适”的个体为样本的一种抽样方法。所谓“合适”, 一方面指总体中那些能在欲研究的现象方面典型地代表这个现象的个体。例如, 我们欲研究学习成绩与生活习惯的影响, 我们就专门从学习成绩好、成绩一般、成绩差的学生中分别选择若干个作为样本。又如, 在研究乡镇企业与社区发展的关系时, 我们也有意识地选取乡镇企业发达、一般、不发达的若干典型地区进行分析调查。这时我们认为, 这些具有明确特征的个体能更直接地和清楚地反映我们所要研究的现象, 能更直接和更突出地考验研究假说。而且我们不必对整个抽样总体进行抽样, 只需选取和考察少量的个体。

有时, 研究目的要求被调查者对于被调查现象具有一定知识和经历, 从而能够作出准确的判断和回答。因此, 这时的“合适”个体将是那些具有相应知识和经历的人。例如, 在了解人们对未来信息技术发展趋势的预测时, 我们需要被调查者对信息技术及其发展现状有所了解, 才能保证他们所作预测有起码的有效性。因此, 许多预测类调查往往采用专家调查法, 而不是从所有公民中随机抽样进行调查。有些调查并不是面对专家, 但实际上仍要求被调查者具备一定专门知识, 例如关于中国企业经营机制改革的民意调查, 关于空调器技术质量的民意调查等。如果这些调查不仅是测验公众对有关问题的了解程度, 而且要求由此对有关问题的具体性质或程度等作出判断, 则适宜用目的抽样, 以保证判断的质量。

无论“合适”个体的定义是什么, 目的抽样要求事先准确定义“合适”个体。对于符合

条件的个体,可采取简单随机抽样来建立样本,也常常采取方便抽样或配额抽样来建立样本。

4.3.4 自愿者抽样

自愿者抽样(Volunteer sampling)是通过被研究对象自愿参加研究来建立样本的一种方法,在社会科学研究中,尤其在心理学、社会学、图书情报学、甚至医学中等到广泛应用。一般方法是:研究者通过一定的宣传媒介介绍自己的研究,介绍参加研究的条件和参加方式,招募自愿者参加研究。许多时候自愿者将获得一定报酬,例如免费赠品或某种特别便利,甚至可能是现金报酬。例如,某医学院欲研究某减肥新药的效力,在院刊上登出广告,邀请体重在70公斤以上的女性和85公斤以上的男性自愿参加,凡参加者免费获得三个疗程的药品和一个疗程的治疗,并免费获得减肥专家的咨询。或者,某情报机构将研究两个采用不同检索技术的情报检索系统的效率,因此在公告牌上宣传,邀请感兴趣的用户报名参加,凡参加者将获得《情报检索技术》一本,或免费进行国际联机检索一次。当然有时研究者将详细描述该研究对科学和对被研究者的意义,以激起人们的责任感或兴趣,真正自愿地参加研究。

自愿者抽样往往用于以下类型的研究:要求被研究者投入较多时间和精力,或被研究者知道自己在被动地接受研究,或需要研究者采取某些他们平时不一定采取的行为。在后面讨论的实验研究往往属于这一类。这时,按一般方法选择的样本可能不愿意参加,即使勉强参加研究也不能积极配合。自愿者抽样能在一定程度上保证参加者对研究有兴趣,从而能够参加完整个研究和给予积极配合。但是,由于这时谁参加研究不是由研究者决定,而实际上由被研究者自己决定,因此最后形成的样本很可能是一个特殊样本,例如在前述检索技术研究中参加者可能都是经常使用情报机构的用户。不过,有时研究本身的目的就限制了被研究者范围,自愿者抽样将具有目的抽样的功用,将且只将那些符合条件又感兴趣的个体吸引来参加研究。而且,在许多实验研究中,被研究者的行为受到严格控制,并不因被研究者不同而产生很大差异。

另外,需要指出,自愿者抽样有时是很隐蔽的。例如在街头进行某项民意测验,调查者将首先简介调查目的,这时有些行人会停下来接受调查,有些行人却会因种种原因不接受调查,例如“家中有病人”、“要去幼儿园接小孩”、“要参加会议”、或“对此问题实在不了解”等。其中有些人可能确实有事,但肯定相当多的人是因为不想被调查而借故离去。这样,被调查者中可能很多属于对此感兴趣,与此相关,或强烈希望表达自己意见者,而没接受调查者可能很多属于对此不感兴趣,对此丧失信心,不习惯或不愿意表达暴露自己意见者。这时,是否参加调查实际上也是由被调查者自己决定的,所形成的样本可能主要由自愿者组成。问卷调查中也有这个问题,不感兴趣者不返回问卷,而返回问卷者可能都具有相当大的自愿性,这一点在返回率很低时影响很大。

以上所有各类非概率抽样方法,都因其抽样操作简单方便节省而在实际研究中得到不少应用。当然,正如前面已经指出,非概率抽样产生的样本对抽样总体的代表性无法估计,样本结果可能与总体情报差异很大,致使通过样本获得的结果不能推广到总体,也无法通告统计验证来确定抽样误差。因此,在外在效度要求严格的研究中,一般不采用非概率抽样。不过,如果所进行的研究需要及时获取少量的有用样本时,或研究目的并不要求严格的外在效度时,或目前研究主要是另一大规模研究的前期探索性预研时,或者研究目的对被研究者有特殊要求时,我们不能简单排除非概率抽样。而且在社会科学研究中,有时很难在规定时间内完成严格随机抽样;即使有了严格随机样本,也可能因种种原因不能样本进行细致的调查研究,以致最后结果由于其它方面原因达不到应有的效度和信度。因此,权衡整个研究的效率,非概率抽样有时可能是更好的方法。所以,我们应对非概率抽样有充分的了解。

4.4 非生物现象的抽样

在科学研究(包括社会科学研究)中, 往往涉及对非生物现象的抽样, 常见的这类现象包括时间、机构、活动和概念现象。

4.4.1 对时间的抽样

一般来说, 时间并不是直接研究对象。但是由于直接研究对象可能是在一定时间范围内陆续出现而且可能是不相同的或具有不同行为方式(例如到商场购物的顾客, 到情报机构查阅文献的用户, 在若干月内接受过某种服务的人等), 那么研究者不能只选择某一时刻出现的对象来进行研究, 必须要对各个时间出现的对象都进行研究, 这就需要确定哪些时间属于“合法”的研究时间, 从而通过一定方法从这些时间内出现的对象中抽取实际研究对象。

例如, 在一项利用直接观察法研究顾客购物行为的研究中, 我们有理由相信一年中不同时期人们的购物行为是不同的, 而且一天中不同时间人们的购物行为也是不同的。为了准确了解人们的实际购物行为, 我们需要对一年内各个时期和一天内各个时间进行购物的顾客都进行观察研究。然而, 由于人力和经费等方面的原因显然不可能一年 365 天每天 12 小时(假定商家开门营业 12 小时)都进行观察。因此, 我们需要选择某些天和这些天内的某些时候作为时间样本。

对时间抽样的基本方式有几种:

a. 随机抽样, 即将所有时间划分为等量的时间片, 对这些时间片进行随机抽样。

b. 按照研究对象时间分布规律的“分层”随机抽样。许多人类活动都存在一定的时间规律, 例如商场顾客在一周中集中在周六周日, 周一或周五也有较多顾客, 但周三周四最少; 高校图书馆借阅量在一学期中开学和期末阶段最多, 而在其它时间较少。因此, 对时间的分层随机抽样考虑到这些分布规律, 将所有时间片分为高峰片, 低谷片和其它片, 并按照各类时间片人流分布比例确定各类时间片的抽样比例, 然后在各类时间片中随机抽样。

c. 目的抽样和配额抽样, 也是考虑到研究对象随时间分布的规律, 有意识地从各类时间片中按一定数量抽取时间片, 或专门选择某些有特点的时间片作为调查时间。

由于时间只是一个间接抽样对象, 因此在确定了“样本”时间后还需确定如何对样本时间内的实际对象进行抽样。这时主要有两类: 或者将该时间内所有对象作为实际样本, 适用于样本时间不长且该时间内对象不多的情况; 或者在该时间内按某种方法抽样, 多采取方便抽样(因为事先往往不知对象数量)。另外, 在时间抽样中还有时间点与时间段抽样之分, 我们将在讨论观察研究方法时详细论及。

4.4.2 对机构和活动的抽样

在许多社会科学研究中, 研究对象是机构, 例如关于企业信息政策的研究, 关于信息机构管理方式的研究等。一般来说研究者都能事先确定机构数量, 获得机构名录也不太难, 在某种意义上可以采取简单随机抽样来选择研究对象。但是, 许多时候这些机构存在很多(可能是“太多”)差别, 简单随机抽样很可能形成与总体分布相异很大的样本, 使样本结果的外在效度受到局限。因此, 在机构抽样中更常见的是分层随机抽样, 在抽样总体较小时常采取配额抽样或目的抽样。还要注意, 在针对机构的调查中, 有些实际上是针对机构中某一部门或活动的调查。这时可能有些机构没有此类部门或活动, 因此可能需要将它们排除在抽样总体外, 或针对它们设计特定调查方法。

许多研究的对象是某种活动, 例如科技人员在学术会议期间的信息交流活动结构研究, 企业广告促销活动成效研究, 社区群防群治活动调查等。在这类研究中最棘手的问题之一是如何确定“活动”范围和确定有多少“活动”, 然后才谈得上抽样和研究。

4.4.3 对概念现象的抽样

我们在讨论构成效度时就已指出, 一个概念往往包含了复杂的现象, 或称为有多个现象化定义。例如, “好丈夫”就可能表现在诚恳、能干、体贴、勤俭、健康、英俊、勇敢、事业发达、收入丰厚等; 往往每一个现象化定义都只能部分地反映这个概念, 只有众多现象化定义的集合才能完整地体现这个概念。问题在于, 这个概念的现象化定义可能很多, 在实际研究中不可能一一用来进行观察测量, 因此必须从众多现象化定义中抽取一部分作为代表来构成实际观察测量手段。例如“好丈夫”的现象化定义可以列举近百个, 但在实际研究中至多用 20 个来进行测量。这种从众多现象化定义中抽取部分现象化定义也是一种抽样, 被称为概念域抽样(Domain sampling)。在概念域抽样中, 也希望抽出的现象化定义样本能尽量代表现象化定义总体。

实际上我们都遇到过此类抽样。每一门课程都有成百(或更多)条具体内容, 它们对应于成百(甚至更多)条题目。但在实际考试中不可能将所有题目都作为正式考题, 只能选择其中一部分来构成实际考题, 即从构成该课内容的所有问题中选择部分问题, 并以它们的集合来近似地代表整个课程内容, 以对它们的回答水平来近似地代表学生掌握整个课程内容的水平。如果这个抽样比较合理, 考试成绩就能较好地反映学生掌握整个课程内容的水平; 如果这个抽样偏差较大, 则考试成绩就只能反映学生对该课部分内容的掌握程度, 使考试失真。

在社会科学研究中这个问题也存在, 而且往往更严重。许多研究涉及到许多概念, 而这些概念代表了人或人所构成的群体在某一方面的复杂行为, 这些行为都可能具有众多现象化定义, 例如现代化意识, 信息意识, 社会责任感, 社会安全感, 社会地位感, 企业营销活动, 群体信息管理, 社会信息资源等, 这些概念或行为都不能用少数几个现象化定义简单地代替。而且, 很多时候我们并不确切知道概念或行为的边界, 即不知道全部现象化定义, 从而难以建立一个“抽样总体”。减轻这个问题影响的一个办法是尽可能详细地列举所有可能的现象化定义, 即尽可能地构造完整的概念域, 具体操作上可采用 Delphi 方法, 有多名专家即独立又交互作用地列举所有现象化定义; 另一办法是构造较大的现象化定义样本, 因此在其它情况相同时, 一个较大的样本(项目较多的问卷, 试题数量多的考卷等)在数量上可以更接近总体, 从而使样本结果更接近总体情况。当然, 样本本身的选取和构造也涉及许多问题, 我们将在下一章进行讨论。对构成的样本还可采用上一章讨论的构成效度评判方法来评价。