

任务 3

选择市场调查的方式



任务目标

知识目标

1. 了解市场普查的特点与步骤；
2. 了解抽样调查的特点和步骤；
3. 掌握抽样调查的基本方法与误差确定。

能力目标

1. 能够正确运用抽样调查方法进行市场调查；
2. 能够合理利用市场调查方式有效地收集市场信息。



案例导入

咖啡杯的市场调查

美国一家公司准备改进咖啡杯的设计，为此进行了市场调查。首先，他们进行咖啡杯选型调查，他们设计了多种形状的咖啡杯，并让 500 位家庭主妇进行观摩评选。研究主妇们干手拿杯子时，哪种形状更方便；湿手拿杯子时，哪种形状不易滑落。调查结果显示，四方长腰型杯子比较好使用。然后对产品名称、图案等进行造型调查。

调研公司重点调查了咖啡杯的颜色对人们饮用咖啡所产生的影响。根据不同颜色的杯子会使人们产生不同口感的特点，选择出最适合咖啡杯的颜色。他们的方法是：首先邀请了 30 多位体验者，让他们每人各喝 4 杯相同浓度的咖啡，但是咖啡杯的颜色则分别为咖啡色、青色、黄色和红色。试饮的结果是：喝过用咖啡色杯子装的咖啡后，30 多位体验者中有 2/3 的人认为口感“太浓了”，而喝过用青色杯子装的咖啡的体验者则异口同声地说味道“太淡了”，接着喝过用黄色杯子装的咖啡的人都说味道“不浓，正好”，而喝过用红色杯子装的咖啡的 30 位体验者中竟有 90% 的人说“太浓了”。根据这一调查结果，公司建议咖啡店里的杯子一律改用红色。咖啡店借助咖啡杯颜色对人们口感的影响，既可以节约咖啡原料，又能提高顾客满意度。公司的这种咖啡杯投入市场后，与市场其他公司的同类产品展开了激烈竞争，结果以销售量稳居市场第一的优势取得了胜利。

(资料来源: 百度, 咖啡杯的设计 <http://www.zybang.com/question/1235f524e14b8a047cf09750d5521eb3.html>)

思考:

1. 本案例中运用的是什么调查方法? 这种方法有什么优缺点?
2. 企业选择市场调查方式时主要考虑的因素是什么?



理论指导

市场调查方式按照调查对象的范围来划分,可以分为全面市场调查和抽样市场调查两种,下面主要介绍这两种市场调查方式的基本内容。

一 全面市场调查方式

全面市场调查也称市场普查,它是指调查者为了收集一定时空范围内调查对象的较为全面、准确、系统的调查资料,对调查对象(总体)的全部个体单位进行逐一的、无遗漏的全面调查。按照实际应用的情况,市场普查可分为宏观、中观、微观三大层次,即分别是全国范围内的,一定地区或一定行业(部门)范围内的,以及企业组织的全面市场调查。

(一) 全面市场调查的特点

1. 专门性

调查者为了特定的目的而专门组织的调查。例如,全国人口普查,国家为了掌握准确的人口规模、结构而开展的调查。

2. 全面性

对调查对象的全部单位都无一例外地进行调查。人口普查是针对所有本国公民而进行的,因此,调查工作量非常大,过程非常复杂。

3. 一次性

对调查对象特定时间的数量表现进行登记,而不是经常性的调查。例如,人口普查数据是反映某个特定时点的人口总数,而不是普查结束或普查某天的人口规模。我们经常看到的人口统计报告的数据其实是一个概数,并不是精确到某天的确切人数。人口普查也不是经常性的,全国人口普查一般10年一次,城市为3年一次,乡镇为1年一次。



4. 准确性

由于对调查总体的全部单位都进行调查,因而不存在抽样调查误差,只要严格控制调查过程,数据的准确性还是比较高的。

5. 标准化程度高

由于普查是统一规定调查项目、时间和方法,以及统一组织、统一标准和统一数据处理,因而获取的数据具有较高的标准化程度。

6. 调查费用比较高

普查由于涉及面广、调查工作量大,因而需要花费较多的调查费用,包括人力、物力、财力和时间的花费。

(二) 全面市场调查的方式

全面市场调查一般有以下两种收集资料的方式。

1. 普查员直接登记式

普查员直接登记式,就是设立专门的组织机构,制订市场普查方案和实施细则,设计调查表或问卷,聘请和培训调查员,然后由调查员深入现场对调查单位的有关情况进行观察、询问和登记。这种方式收集的数据资料必须客观和准确,调查质量在很大程度上取决于调查员的素质,要注意控制调查员工作误差。

2. 被调查者自填式

被调查者自填式,就是设立专门的组织机构,制订市场普查方案和实施细则,设计调查表或问卷,但不派调查员进行直接登记,而是将调查表或问卷下放到企业、事业等基层单位,各基层单位再指定专人根据本单位的原始记录和现成资料进行填报。这种方式要求提高全面市场调查方案、调查表或问卷的设计质量,防止调研者的设计误差,也需减少被调查者的填报误差,其措施是调查表应附有填表说明、指标或项目解释,并做好调查表回收的审核工作。

(三) 全面市场调查的原则

1. 调查项目和内容统一

全面市场调查必须统一规定调查的内容和项目,不能任意改变或增减,以确保调查内容的一致性,便于数据的汇总和处理,提高数据质量。



2. 调查的标准时点统一

全面市场调查主要用于收集一定时点上的有关调查对象的数据和资料，因此，为了保证调查数据在时间上保持一致，必须统一规定调查的标准时点。如果调查登记时间在这个规定的标准时点之后，也必须按规定的标准时间核实、登记、填报其数据。

3. 调查的各种标准统一

全面市场调查除了统一规定调查内容、项目和调查标准时点之外，还必须统一制定各种标准。如产业或行业划分标准、产品分类目录、调查表式、指标解释、计算方法、数据编码、数据处理程序等，都必须统一规定和实施，才能从各个方面确保调查数据的质量。

4. 调查步骤和方法统一

为确保全面市场调查资料具有一定的时效性和调查质量，调查范围内各个调查点必须统一行动、统一进度、统一方法，力求步骤和方法上协同一致。

(四) 全面市场调查的实施步骤

全面市场调查的实施步骤如图 3-1 所示。

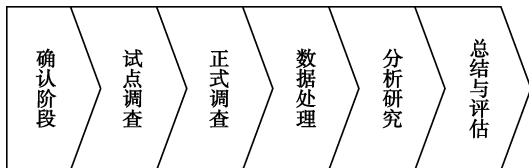


图 3-1 全面市场调查的实施步骤

1. 确认阶段

确认阶段主要是依据决策的信息需求提出全面市场调查的动议，并论证其必要性和可行性。动议获决策者批准后，即可组建调查机构，制订全面市场调查方案，开展宣传鼓励工作。其中全面市场调查方案的制订是确认阶段的重要任务，它是根据全面调查的组织原则，对调查工作所作的全局考虑，全面市场调查方案的内容包括总则，调查对象、范围和调查方法，调查的内容和标准，调查的组织实施，调查经费安排，数据处理与质量控制，数据发布与开发利用，表彰与处罚、附则等。

2. 试点调查

试点调查是根据制订的全面市场调查方案，选择若干调查点作为试点单位，模拟全面调查工作的全过程，用以往组织实施全面市场调查的经验，检验全面市场调查方案的科学性、可行性，发现问题和缺陷，为改进和完善调查方案提供依据。试点调查的具体工作包括制订试点方案、选择试点单位、组织试点调查、评估试点结果、修改和完善正式调查方案等。



3. 正式调查

正式调查阶段的工作包括落实调查方案,各调查点组建调查队伍,认真学习调查方案和有关文件,对调查员进行业务培训,并组织正式的调查登记工作或调查表的填报工作,将调查表按照规定的时间内上报到指定的汇总部门。

4. 数据处理

数据处理阶段包括数据处理程序编制、软件使用培训、调查表人工审核、数据质量抽样复检、数据录入、计算机审核、数据校正、数据汇总、数据备份、数据上报、建立数据库、数据存储和数据管理等工作。

5. 分析研究

分析研究阶段的工作包括数据质量评估确认,经批准后组织对外公布;利用调查数据进行多方面的开发和分析研究,为决策者提供多种形式的信息服务等。

6. 总结与评估

总结与评估即对调查工作进行全面总结与评估,总结经验与教训,并对相关的工作单位或个人进行奖励或惩处。

(五) 全面市场调查的应用

由于全面市场调查在人、财、物等方面的消耗较大,调查时间也长,因此,企业一般只组织小范围的全面市场调查,或者对那些不能或不必要进行经常性调查,但对资料的准确性、全面性要求较高的项目实施全面市场调查,也可利用全国或有关行业或部门组织的普查收集企业市场调查所需要的全面资料。具体应用如下。

(1) 企业人财物普查。企业内部有关人力、物力、财力资源和产供销情况的全面调查,如员工基本情况调查、设备调查和产品产、销、存全面统计等。

(2) 企业全员测评。企业员工满意度、忠诚度全员测评,内部人事制度、分配制度改革的全员测评。

(3) 供应商市场普查。供应商的市场普查,即对企业原材料、零配件等供应商的供货数量、质量、品种、价格、服务、信誉度、运输、配送等情况进行全面调查。

(4) 经销商市场普查。经销商、代理商的市场普查,即对经销、代理本企业产品的全部中间商的经营能力、经营区域范围、服务质量、价格执行、付款信誉、促销能力、财务支付能力、协作状况等进行访问、观察和登记。

企业除了自主组织开展以上项目的全面市场调查之外,还将利用国家统计局或其他部门的普查资料,通过购买、检索或委托开发企业获取市场研究所需的全面资料。

全面市场调查有许多优点,但也存在一定的局限性。全面市场调查通常需要专门的机构来主持,涉及面广、工作量大、时间较长,需要大量的人力和物力、组织工作较为繁重等,因此,它是实际调查中运用较少的一种。而抽样市场调查采用的是以部分推断总体的



方式，减少了市场调查的工作量，简化了市场调查工作，可以实现以较少的投入而能得到较高质量的数据。抽样市场调查目前已经成为实际调查中运用越来越多的主要调查方法。

二 抽样市场调查方式

（一）抽样市场调查的基本概念

抽样市场调查简称抽样调查，是按照一定的程序，从所研究对象的全体中抽取一部分单位进行调查，并在一定条件下，对研究对象的数量特征进行估计和推断的一种调查方式。与抽样市场调查有关的专业术语包括总体、样本、抽样单位和抽样框等。这些术语的具体含义解释如下。

1. 总体

总体是所要调查对象的全体，即在特定的调查目的或任务条件下的认识客体。在市场中，总体通常都有时间和空间的限制。构成总体的元素称为单位或个体，若这些单位是不能进一步分割的，则称为基本单位，若这些单位还能分解为更小一层次的单位，则称为群体单位。市场调查中，若是研究消费倾向，消费者个人显然是基本单位，若是通过家庭来观察消费者的消费倾向，则家庭是群体单位，它是由消费者个体组成的。

总体中的单位若是不可计数的，则为无限总体，而总体中的单位数是可以计数的，则为有限总体。通常用 N 表示总体单位数，即总体规模的大小。

2. 样本

样本是由总体中抽取的部分个体，是实际的调查对象。例如，某学校有 1 万多名学生，从中抽取 200 名来进行校园生活状况的调查，这 200 名学生就构成了样本。

3. 抽样单位

抽样单位是指样本抽取过程中的单位形式。例如，1 万名学生中抽取 200 名，有不同的抽样方法。如果 1 万名中直接抽取 200 名，就是以个人作为抽样单位。如果 1 万人分布在 300 个班级里，就可以先在 300 个班级里随机抽取 6 个班级，再在 6 个班级中平均每个班级抽取约 33 人，组成 200 名学生作为样本，这种抽样方法的抽样单位就是班级，而不是个人了。

抽样单位是指将总体划分为若干个互不重叠的部分。抽样单位是人为划分的，受到抽样方法的影响。



4. 抽样框

抽样框是指将抽样单位按某种顺序排列编制的名单,是抽样设计人员用来进行抽取样本的工具。其具体表现形式主要包括总体全部单位的名单、电话号码簿、户口档案、企业名录地图等。抽样框在抽样调查中处于基础地位,是抽样调查必不可少的资料。在抽样框中,每个抽样单元都有自己对应的位置或序号,这是抽样调查前通常要做的基础性工作,就是对抽样框中的每个单位进行编号。上面第一种抽样方法中,抽样框是1万名学生的名单,第二种方法中抽样框是300个班级的名单。而对于一些入户调查,抽取的是居民户,抽样框的资料要求比较复杂。

(二) 抽样市场调查的特点

抽样市场调查与全面市场调查相比,具有以下几方面的优点。

1. 时间短、收效快

抽样调查涉及面较小,取得调查结果比较快,能在较短的时间内获得与市场普查大致相同的调查效果,还可以运用抽样调查技术来检验普查及有关资料的准确性,并给予必要的修正。

2. 质量高、可信度高

抽样调查是建立在数理统计基础之上的科学方法,只要由专业调查人士主持抽样调查,严格按照抽样调查的要求进行抽样,就可以确保获取的信息资料具有较好的可靠性、准确性,对那些无法或没有必要进行普查的项目具有很好的适用性。

3. 省费用、易推广

由于抽样调查把调查的对象范围降低到较小的限度,又能保证调查的有效性,从而可以大大减少工作量,降低费用开支,提高经济效益。同时,由于抽样调查只需较少的人力、物力,企业容易承担与组织。

抽样调查也存在着不足,由于所调查的样本只是调查对象中的一部分,抽样调查的结果是从抽样样本中获取的信息资料中推断出来的,因此,抽样调查存在抽样误差。

三

抽样的基本方法

抽样市场调查的方式包括随机抽样调查和非随机抽样调查两种形式。随机抽样方法包括简单随机抽样、等距随机抽样、分层随机抽样、整群随机抽样。非随机抽样方法包括方

便非随机抽样、判断非随机抽样和配额非随机抽样。抽样调查方式如图 3-2 所示。

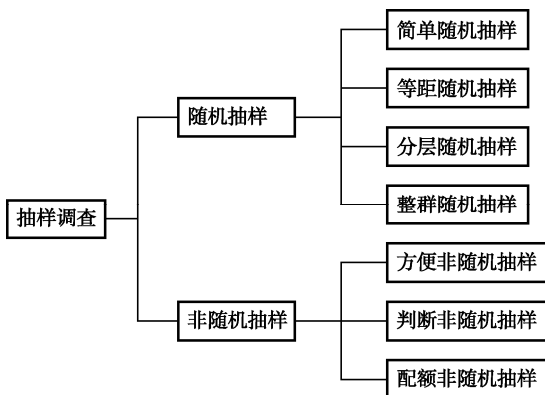


图 3-2 抽样调查方式

（一）随机抽样

随机抽样调查也称概率抽样调查，是从总体中按随机原则抽取一定数目的单位作为样本进行调查，并在一定条件下，对研究对象的数量特征进行估计和推断。随机抽样调查使总体的每一个单位都有同等的机会或可能性被抽中，使被抽中的单位能较好地代表总体。随机抽样调查可在抽样调查之前估计和控制抽样误差，主要包括以下 4 种抽样方法。

1. 简单随机抽样

简单随机抽样是对全部总体不经过任何分组、排队，完全按照随机原则从中抽取样本加以调查的一种抽样技术形式。简单随机抽样的特点是在每次抽选中都能保证总体中每个单位有相等的选中机会，完全排除了抽样中主观因素的干扰。简单随机抽样的具体方法主要有以下 3 种。

（1）直接抽选法。直接抽选法就是从调查总体中直接随机抽取样本进行调查。这种方法适合对集中于某个空间的总体进行抽样，如对存放于仓库中的同一种饼干产品直接随机抽出若干为样本进行霉变状况检查。

（2）抽签法。抽签法就是用抽签方式抽取样本单位。这种方法必须首先给调查总体中的每个单位统一编号，然后将每个编号做成一张签片，再将所有签片混合均匀，最后从中随机抽选，被抽中的号码所代表的单位就作为样本中的一个，直到抽足所需样本单位数目为止。例如，在某学校的 200 个班级里，抽取 10 个班级，面向学生就教学改革效果的问题进行调查。可以制作 200 张卡片，分别标上 1~200 的数字，从中任取 10 张卡片，卡片号码所代表的班级就是所抽取的样本。

（3）随机数表法。随机数表法就是对 0~9 这 10 个数字进行重复抽样，记录每一次的结果，进行成千上万次后，所形成的数表。数表中数字的排列是随机的，因而也称为乱数表。用随机数表法抽取样本单位时，先给总体各单位统一编号，然后按事先规定的顺序读随机乱数表，所读随机乱数表的位数由总体的单位数来确定。若总体为千位数规模，则一



次在随机乱数表上读 4 位数，若总体为百位数，则一次读 3 位数，等等。总体中一个单位的编号与所读出的随机数一致，则该单位入样本，即被抽中。若所读出的随机数大于总体规模数，则跳过该随机数，往下继续读，重复抽样时，遇到已经选用过的数字仍然选用；不重复抽样时，凡已经选用过的数字都不要，直到抽够规定的样本单位数为止。随机乱数表的读法可以是任一起点，采取任一顺序进行。但阅读顺序一旦确定，便不能改变，否则便会破坏随机特征。随机乱数表如表 3-1 所示。

表 3-1 随机乱数表

	(1)	(2)	(3)	(4)	(5)	(6)
1	32044	69037	29655	92114	81034	40582
2	23821	96070	82592	81642	80971	70411
3	82383	94987	66441	28677	95961	78346
4	68310	21792	71635	86089	38157	95620
5	94856	76940	22165	10414	10413	37231
6	95000	61958	83430	98250	70000	50436
7	20764	64638	11359	32556	89822	20713
8	71401	17964	50940	95753	34905	93566
9	88464	75707	16750	61371	10523	69205
10	59442	59247	74955	82835	98378	83513
11	11818	40951	99279	32222	75433	27397
12	65785	60937	96483	10230	58220	90756
13	15933	69834	57402	35168	81438	44850
14	31722	97334	77178	70361	15819	35037
15	95118	88373	26934	42991	10142	90852



案例启示

生产车间需要进行某产品寿命周期的检验，这种检验是破坏性的，因此，只能采取抽样检验方法。现将从 100 个产品中，随机抽取 11 个作为样本，拟用简单随机抽样的乱数表法抽选样本。抽选程序如下。

(1) 将总体单位从 1~100 进行编号；

(2) 然后可确定使用 5 位数字中的哪两位数，可以任意取两位数。前两位、中间两位或后两位都可以，如确定使用后两位数；

(3) 随机确定从某列某行开始，然后按确定顺序抽选。如从第 2 列第 4 行开始，从上至下抽选；

(4) 抽出的数字为 92、40、58、38、34、64、07、47、51、37、34。其中出现了两个 34，把后面一个去掉，再往下补充一个为 73，这样，我们所要抽取的 11 个样本单位的号码就出来了，对应上述 11 个数字标号的产品，就是抽选到的样本单位。如果从上至下，

样本单位数没有抽满规定量，则可以逆时针方向从下至上继续抽取，直到完成规定抽样数量为止。

2. 等距随机抽样

等距随机抽样也称机械随机抽样或系统随机抽样，它是先按某一标志对总体各单位进行排队，然后依一定顺序和间隔来抽取样本单位的一种抽样技术形式。由于这种抽样是在各单位按一定顺序排队的基础上，再按一定间隔取样，这样可以保证所取得的样本单位比较均匀地分布在总体的各个部分，有较高的代表性。

作为总体各单位顺序排列的标志，可以是有关标志或是无关标志。例如，液晶电视需求偏好调查，按照收入多少由低到高进行排列。所谓无关标志是指与调查标志无关的或不起主要影响作用的标志。例如，产品质量抽查按时间顺序取样，农产量抽样调查按田间的地理顺序取样等。抽样间隔（ R ）的大小等于总体数量（ N ）除以样本数量（ n ），计算公式为：

$$R=N/n$$



案例启示

某社区里有商户 100 家，采用等距离抽样方法抽选 10 户进行调查。

第一步，将总体调查对象（100 家商户）进行编号，即从 1 号至 100 号。

第二步，确定抽样间隔。已知调查总体 N 为 100，样本数 n 为 10 家，故抽样间隔为 $100 \div 10 = 10$ （家）。

第三步，确定起抽号数。将 1~10 号编成 10 张卡片，然后从中随机抽取 1 张作为起抽号数。如果抽出的是 3 号，则 3 号为起点。

第四步，确定被抽取单位。从起抽号开始，按照抽样间隔，确定抽取的样本单位数编号。本例从 3 号起每隔 10 个号抽选一个，直至抽足 10 个为止。计算方法为：

$$\begin{aligned} & 3 \\ & 3+10 \times 1=13 \\ & 3+10 \times 2=23 \\ & \dots\dots \\ & 3+10 \times 9=93 \end{aligned}$$

则所抽的单位是编号为 3、13、23、33、43、53、63、73、83、93 的 10 家商户。

与简单随机抽样相比，等距随机抽样更加方便易行，抽样程序简便，对工作人员业务素质要求不高，这样可节省时间与费用。另一方面，在已知总体某些有关信息的条件下，采用等距随机抽样能保证样本单位在总体中均匀地分布，从而可以提高样本对总体的代表性，减少抽样误差。

等距随机抽样也有其缺点。当总体单位排序恰好与抽样间隔周期一致时，可能存在着选取到一个严重偏差的样本的风险。因此，要尽量避免抽样间隔与现象本身周期性的节奏相重合引起的系统性偏差。例如，工业企业产品质量抽查时，抽查时间间隔不宜和上下班



时间一致，以防发生系统性偏差，影响样本的代表。

3. 分层随机抽样

分层随机抽样是先将总体所有单位按某一重要标志进行分类（层），然后在各类（层）中采用简单随机抽样或等距随机抽样方式抽取样本单位的一种抽样技术形式。其可分为等比例分层抽样法与非比例抽样法。

（1）等比例分层抽样法，即按各层（或各类型）中的单位数量占总体单位数量的比例分配各层的样本数量。该方法简便易行，分配合理，计算方便，适用于各类型之间差异不大的分类抽样调查，如果各类型之间差异过大，则不宜采用此方法，而应采用非比例抽样法。



案例启示

某地共有居民 10000 户，按经济收入高低进行分类，其中高等收入的居民为 2000 户，占总体的 20%，中等收入的居民为 6000 户，占总体的 60%，低等收入的居民为 2000 户，占总体的 20%。要从中抽选 150 户进行购买力调查，则各类型应抽取的样本单位数为：

高等经济收入的样本单位数： $150 \times 20\% = 30$ （户）

中等经济收入的样本单位数： $150 \times 60\% = 90$ （户）

低等经济收入的样本单位数： $150 \times 20\% = 30$ （户）

样本单位数 150 个，就是按高、中、低经济收入比例进行抽样选出来的。

（2）非比例抽样法，又称分层最佳抽样法，是根据各层次的样本标准差大小，调整各层次的样本数目的抽样方法。该方法既考虑了各层在总体中所占的比重大小，又考虑了各层标准差的差异程度，有利于降低各层之间的差异，以提高样本的可信度。其计算公式为：

$$n_i = n \times \frac{N_i S_i}{\sum N_i S_i}$$

式中， n_i 代表各类型应抽选的样本单位数； n 代表样本单位总数； N_i 代表各类型的调查单位数； S_i 代表各类型调查单位平均数（成数）的样本标准差。

例如：设上例中高等收入居民的标准差为 5%，中等收入居民的标准差为 2%，低等收入居民的标准差为 3%，则各层应分摊的调查抽样数目为：

高等收入居民抽取样本数 = $150 \times (2000 \times 5\%) / (2000 \times 5\% + 6000 \times 2\% + 2000 \times 3\%)$
 ≈ 54 （户）

中等收入居民抽取样本数 = $150 \times (6000 \times 2\%) / (2000 \times 5\% + 6000 \times 2\% + 2000 \times 3\%)$
 ≈ 64 （户）

低等收入居民抽取样本数 = $150 \times (2000 \times 3\%) / (2000 \times 5\% + 6000 \times 2\% + 2000 \times 3\%)$
 ≈ 32 （户）

分层最佳抽样法因为需要根据各层的差异程度来配置样本，这对于调查单一标志是比较理想的，但对于调查多种标志则难以兼顾；同时，在计算上比较麻烦。因此，如果调查

的总体各层之间的差异程度不过分悬殊，一般采取等比例分层抽样法。

分层抽样比简单随机抽样更为精确，能够通过对较少的抽样单位的调查，得到比较准确的推断结果，特别是当总体较大、内部结构复杂时，分层抽样常能取得令人满意的效果。同时，分层抽样在对总体推断时，还能获得对每层的推断。因此，在我国的社会购买力调查、居民家庭收支调查、商品销售调查、产品产量调查中，经常采用分层随机抽样方法。

当然，分层随机抽样也有其局限性。主要是它要求调查设计人员必须对总体单位的情况有较多的了解，否则难以设计出科学合理的分层样本。而要做到这一点往往是比较困难的，或者必须花费更多的时间和精力。

4. 整群随机抽样

整群随机抽样是指将市场调查的总体按一定的标准分成若干群，然后从中随机抽取部分群单位进行调查的方式。整群随机抽样法同分层随机抽样法的内容要求不同，分层随机抽样要求所分层之间有差异性，分层内部的基本单位具有相同性，而分群随机抽样恰恰相反，要求各群体之间具有相同性，每一群体内部的基本单位具有差异性。如图 3-3 和图 3-4 所示。

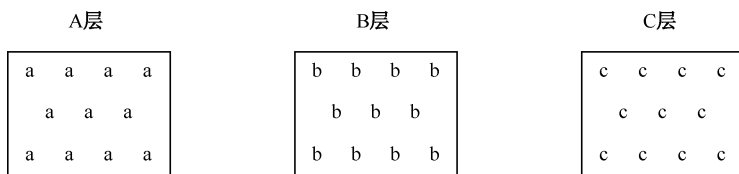


图 3-3 分层抽样后的各层

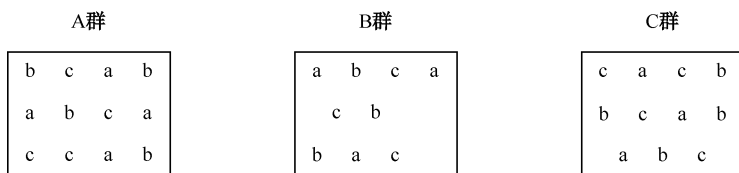


图 3-4 分群抽样后的各群

整群随机抽样是对选中群的全面调查，调查单位集中，抽样简便，可节省经费开支。整群抽样的可靠程度取决于群与群之间的差异大小，当各群间差异越小时，整群抽样的调查结果就越准确。因此，整群抽样适用于抽样调查中没有总体单位的原始记录可利用时，或群内各单位间的差异较大，而各群之间差异较小时。例如，我们欲从一座住宅套数大约为 N 为 12000 的城市中采用简单随机抽样抽取一个 n 为 400 的样本。个别住宅的抽样比较困难，缺乏所有住宅情况的完整资料，而准备这样的资料花费较大。作为替代方案，我们可选择一个街区的样本。在城市的地图上把整个区域划分为形状规则或不规则街区（一般以道路、街道或河流等自然的界线为边界），然后抽取 $1/30$ 的街区作为样本，把位于样本街内的住宅都作为样本点。这样任何一个住宅被抽取的概率就是其对应街区被抽中的概率，即 $1/30$ ，这与理论上的取样概率 $400/12000$ 数值上相等。表 3-2 列举了一些适合整群抽样的实例。



表 3-2 适合整群抽样的部分实例

总 体	变 量	基 本 单 元	群
某个城市	住户特征	住宅	街区
某个城市	购买衣物	人	住宅
机场	旅游信息	离开旅客人数	航班
农村	社会态度	成人	村
通过桥梁的年交通量	发车地和到达地	机动车	40 分钟间隔
城市土地所有者档案	税务信息	土地所有者	分类台账的页数
健康保险档案	医疗数据	卡片	连续 10 张卡片为一组

（二）非随机抽样

非随机抽样调查是从调查者方便出发或根据调查者主观的选择来抽取样本、推算总体的抽样方法。非随机抽样不能估计和控制抽样误差，无法确定抽样推断的概率保证程度。与随机抽样调查相比，非随机抽样较为简单易行，适宜于探索性的调查研究。

在市场调研中，采用非随机抽样通常是出于下述几个原因：受客观条件的限制，无法进行严格的随机抽样；调查时效性要求高，要迅速取得调查结果；在调查对象不确定或无法确定时，可采用非随机抽样；总体各单位的标志值差异不大，且调查者具有丰富的调查经验时，可采用非随机抽样。非随机抽样主要包括以下 3 种方法。

1. 方便非随机抽样

方便非随机抽样又称任意抽样，是根据调查者的方便与否，以尽可能使调查对象对总体具有代表性为原则的一种抽样调查方法。

运用方便抽样技术进行抽样，一般由调研人员从工作便利出发，在调研对象范围内随意抽选一定数量的样本进行调查。方便抽样适用于非正式的探测性调查，或调查前的准备工作。一般在调查总体中每一个个体都是同质时，采用此类方法。但是实践中并非所有总体中每一个个体都是相同的，所以抽样结果偏差较大，可行程度较低，样本不具代表性，因此，在正式市场调查时，很少采用任意抽样法。

2. 判断非随机抽样

判断非随机抽样又称立意抽样，它是指市场调查者根据自己的主观判断、经验和知识，从总体中选择具有典型代表性的单位作为样本的一种抽样方法。

判断抽样选取样本单位一般有两种方法：一种是选择最能代表普遍情况的单位作为样本，常以“平均型”或“多数型”为标准，应尽量避免选择“极端型”；另一种是利用调查总体的全面统计资料，按照一定的标准，主观选取样本。

判断抽样方法在样本量小及样本不易分门别类挑选时有其优越性。但由于其精确性依赖于研究者对调查对象的了解程度、判断水平和对结果的解释情况，因此判断抽样方法的



结果的客观性常受到人们的怀疑。

3. 配额非随机抽样

配额非随机抽样又称定额抽样，它与分层随机抽样有相似之处，它是先按照市场调查对象的某种特征，如被调查者的姓名、年龄、收入、职业、文化程度等，将总体分为若干类，然后按一定的比例在各类中分配样本单位数额，并按各类数额在每类中任意或主观抽样。

配额抽样是最流行的一种非随机抽样方式。配额抽样方法简单易行，可以保证总体的各个类别都能包括在样本之中，使样本具有较高的代表性。

按照配额的要求，配额抽样可以分为独立控制和交叉控制两种。独立控制配额抽样是根据调查总体的不同特征性，对具有某个特性的调查样本分别规定单独分配数额。交叉控制配额抽样是对调查对象的各个特性的样本数额交叉分配，再进行抽选。

独立控制配额抽样方法简单易行，费用较低，但调查者为了方便，往往在选择样本时过于偏重某一特征的组别，从而影响到样本的代表性，而交叉控制配额抽样可以弥补这一缺陷。

案例启示

某市进行化妆品消费需求调查，确定样本量为 300 名，选择消费者年龄、性别、月收入三个标准分类。其中，18~34 岁占 16.7%，35~44 岁占 33.3%，45~60 岁占 36.7%，60 岁以上占 13.3%；男女比例为 1:1；收入 4500 元以上占 20%，2500~4500 元占 33.3%，2000~2500 元占 33.3%，2000 元以下占 13.4%。按照上述三个变量用配额抽样法抽取样本，可编制配额表 3-3。

表 3-3 年龄、性别、月收入交叉控制配额抽样分配表

单位：名

		月 收 入								合计
		2000 元以下		2000~2500 元		2500~4500 元		4500 元以上		
性别		男	女	男	女	男	女	男	女	
年 龄	18~34 岁	4	5	7	7	9	3	10	5	50
	35~44 岁	7	6	10	16	23	17	10	11	100
	45~60 岁	5	5	20	28	19	20	4	9	110
	60 岁以上	3	5	8	4	6	3	5	6	40
小计		19	21	45	55	57	43	29	31	
合计		40		100		100		60		300

四

抽样误差及样本容量确定

(一) 抽样误差的确定

抽样误差是指随机抽样调查中必然发生的代表性误差，即平均误差，通常用 μ 表示。因为抽样调查是以样本代表总体，以样本综合指标推断总体综合指标，所以平均误差是不可避免的，但这种误差一般不包括技术性误差，即调查过程中的工作误差。

1. 抽样误差的影响因素

(1) 总体各单位标志值的差异程度。总体单位之间标志差异程度大，则抽样误差就大，反之则小，因此，抽样误差大小与总体标准差大小成正比例关系。

(2) 样本单位数量。样本单位数目较多，则抽样误差较小，反之，则抽样误差较大。因此，样本单位数的多少与抽样误差的大小成反比例关系。

(3) 抽样方法。一般情况下，简单随机抽样比分层、分群抽样误差大，重复抽样比不重复抽样误差大。

2. 抽样误差的估算

重复抽样是指样本抽出后再放回去，有可能第二次抽中。而不重复抽样是指样本抽出后不再放回总体中，每个单位只能抽中一次。

(1) 平均数重复抽样误差

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$$

式中， $\mu_{\bar{x}}$ 为抽样平均误差； σ^2 为总体平均方差； n 为样本单位数。

σ^2 一般需进行换算， $\sigma^2 = \frac{\sum(x - \bar{x})^2}{n}$ 或以样本标准差代替，也可以采取经验估算等。

(2) 平均数不重复抽样误差

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$$

式中， N 为总体单位数； $1 - \frac{n}{N}$ 为修正系数。



(3) 成数重复抽样误差

$$\mu_p = \sqrt{\frac{P(1-P)}{n}}$$

式中, μ_p 为成数(相对数)抽样误差; P 为成数(相对数); n 为样本单位数。

(4) 成数不重复抽样误差

$$\mu_p = \sqrt{\frac{P(1-P)}{n} \left(1 - \frac{n}{N}\right)}$$

(二) 样本容量的确定

样本容量又称为抽样单位数, 抽样单位数的确定也是抽样调查中需正确解决的问题。抽样数目过多, 则成本增加, 形成浪费; 如果抽样数目过少, 又会使调查结果存在较大误差, 达不到要求的精度。因此, 有必要确定必要的样本数目。必要的样本数量是指在所给定的抽样误差范围内所确定的、能够达到对调查结果精确度要求的样本单位数。

1. 影响必要样本数量的因素

(1) 总体中各单位之间标志值的变异程度。变异程度越大, 需抽取的样本数目越多; 反之, 需要抽取的样本数目越少。

(2) 允许误差的大小。允许误差指极限误差或最大可能误差, 是指抽样误差的范围。在其他条件一致的情况下, 允许误差小, 抽样数目就相应多一些; 反之, 允许误差大, 抽样数目就可能少一些。允许误差的选择主要受到调查目的的要求、调查经费和时间等影响。如果调查的准确度要求高、调查力强、调查经费充足, 则允许误差可定小一点; 反之允许误差就只能放大。

(3) 不同的抽样组织方式和抽样方法。通常情况下, 简单随机抽样和分群随机抽样比等距随机抽样和分层随机抽样所需的样本单位数要多, 重复抽样则比不重复抽样的样本单位数要多。

2. 必要样本数量的确定

(1) 平均数指标重复抽样数目

$$n = \frac{t^2 \sigma^2}{\Delta \frac{2}{x}}$$

式中, n 为样本单位数; σ^2 为总体方差; t^2 为概率度平方; $\Delta \frac{2}{x}$ 为平均数允许误差平方。

(2) 平均数指标不重复抽样数目



$$n = \frac{Nt^2\sigma^2}{N\Delta_x^2 + t^2\sigma^2}$$

式中， N 为总体单位数。

(3) 成数指标重复抽样数目

$$n = \frac{t^2 p(1-p)}{\Delta_p^2}$$

(4) 成数指标不重复抽样数目

$$n = \frac{t^2 Np(1-p)}{N\Delta_p^2 + t^2 p(1-p)}$$

式中， t 为概率； p 为成数； Δ_p 为成数允许误差； n 为抽样数目。

t 值越小，允许误差范围越小，把握程度越大，反之越小。因为抽样调查时，每100次抽样中有95.45次的正负误差不超过2，其余4.55次则会超过，其概率度 t 值为2，可信程度为95.45%。它们之间的关系可参照表3-4。

表 3-4 概率表

概 率 度	可信程度 F	允许误差 $\Delta = t\mu$
1.00	0.6827	1.00 μ
1.50	0.8864	1.50 μ
1.96	0.9500	1.96 μ
2.00	0.9545	2.00 μ
3.00	0.9973	3.00 μ

五

抽样设计的程序

抽样设计的任务，就是要依据调查目的，在给定的人力、物力、经费、时间要求等条件下，设计一个精确度，能够由样本正确推断总体的良好的抽样调查方案。如图3-5所示，抽样设计通常包括以下环节。

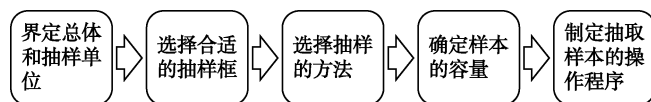


图 3-5 抽样设计的程序



（一）界定总体和抽样单位

（1）调查总体的界定是指确定在实施抽样时哪些对象应包括在内，哪些对象不应包括在内。调查总体的界定主要根据个体、抽样单位、范围和时间来界定，可从地域特征、人口统计学特征、产品或服务使用情况等角度进行描述（表 3-5）。只有对总体进行正确界定和精确了解，才能获得代表性较高的样本。例如，1936 年美国《文学摘要》关于总统选举民意调查的失败，就是因为对总体的错误界定。当时是以电话簿和汽车登记簿为抽样框进行抽样调查。但是，当时的美国正处于经济萧条时期，很多人既没有汽车也没有电话，从而使得样本缺乏代表性。尽管样本量达到了 200 万，得出的结论是阿尔夫·兰登将胜过富兰克林·罗斯福，但与后来的事实大相径庭。

（2）抽样单位的界定是指需要明确调查对象是个体、居民户还是企业等。这必须根据调查目的和内容来确定。如果需要调查 90 后年轻消费者的消费行为和态度，抽样单位就是个体，如果需要了解某地区家庭平均消费水平，抽样单位就是这个地区的长住居民户。

在总体和抽样单位的界定过程中，还需考虑抽样精度的问题，即调查结果需要达到什么样的精确程度。抽样调查的结果不可能也不必达到百分百的精确，因此，只要调查结果的精确度满足需要就行。抽样精度主要是根据用户的要求和决策的需要来确定的。

表 3-5 调查总体界定的一般描述

	地域特征	人口统计学特征	产品或服务使用情况
调查总体的描述	按照总体单位活动的范围或区域进行界定。可以是城镇，也可以是总体单位的户籍所在地等	考虑人口统计学变量方面具有某些特征的总体单位	同质产品的共同特征通常依据产品或服务的使用情况来定义
举例说明	调查长沙居民家庭购房的需求特征，则长沙就是此次活动的地理区域	卷烟市场调查，被调查者为男性，年龄为 18~50 岁。其他年龄段受访者意见相对意义不大	调查某企业产品满意度，被调查者应是产品的使用者，或是根据其使用产品的频率、次数来定

（二）选择合适的抽样框

只有选择了合适的抽样框，才可以采用抽签的方式或按照随机数表来抽选必要的单位数。一般情况下，抽样框的评判标准有以下方面：一是抽样框应具有识别性。抽样框能识别每一个单位，有一份高质量的标识名单，避免出现抽样单位混淆，而破坏了原先设定的总体单位被抽中的概率，从而在估计中引入了偏差。二是抽样框应具有完整性。抽样框应包括每个可能要被调查的单位，所有的总体单位都不能被遗漏，否则会影响抽样调查的精确度。三是抽样框应具有唯一性。每个总体单位不能在抽样框中出现两次及以上，或是使用两次及以上的标识。否则，同样会在估计中出现偏差。例如，在调查某品牌液晶电视机