

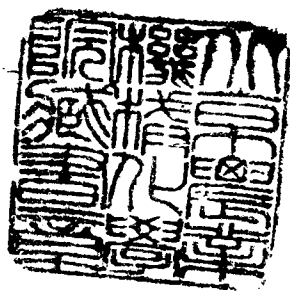
龚 鉴 尧 著

抽样法浅说

中国财政经济出版社

抽 样 法 浅 说

龚 鉴 尧 著



中国财政经济出版社

2436/25

抽样法浅说

龚鉴尧 著

*

中国财政经济出版社出版

新华书店北京发行所发行 各地新华书店经售

通县西定安印刷厂印刷

*

787×1092毫米 32开本 5.875印张 120,000字

1981年7月第1版 1984年6月北京第2次印刷

印数：20,001—30,000

统一书号：4166·254 定价：0.48元

序 言

这本《抽样法浅说》，是在党的十一届三中全会和全国科学大会的鼓舞下，根据一九七八年十二月全国统计教学、科研规划座谈会要求编写的。

从现在起，要把全党全国工作的着重点转移到社会主义现代化建设方面来，在本世纪内把我国建设成为一个具有现代农业、现代工业、现代国防、现代科学技术的社会主义强国。这是一个伟大的历史性转变，各项工作都应当为这个中心服务。抽样法作为一种科学方法，也应当在实现四个现代化中发挥作用。

这本小册子试图以比较通俗的形式，来阐明抽样理论和实际应用中的一一些问题，便于从事统计调查工作的同志，包括基层县、社和企业从事统计工作的同志，在阅读后，对抽样法能有一个初步的了解，并在实际工作中加以运用。因此，本书着重讲清基本的原理和实际的做法；对一些数理统计上的定理和数学证明，尽量少涉及到。在这方面有兴趣探讨和钻研的同志，可以从有关数理统计方面的书籍和教材中得到进一步的答案。为了便于读者进行比较研究，我在第三、四、五章的举例中，故意用了相同的农产量方面的例子。这样做是否会使读者感到更加系统一些呢？

遗憾的是，“文化大革命”期间，手中一些有关这方面的参

考书籍和材料，包括笔者亲自调查取得的一些数据材料，丢失殆尽，动起笔来，要想找点参考材料，每每苦不可得；加上工作条件的限制，也没有能挤出时间去查阅有关材料，错误一定不少，尚希读者指正。

龚 鉴 尧

一九七九年八月于广州

目 录

序 言	(1)
第一章 调查统计与误差	(1)
一、从认识论说起	(1)
二、两种调查统计方法	(2)
三、两种误差	(3)
四、怎样控制误差	(5)
第二章 抽样原理	(7)
一、抽样法的基本概念	(7)
二、抽样法的原理和特点	(8)
三、抽样调查的应用范围	(13)
四、抽样调查的组织形式	(15)
五、多阶段抽样	(21)
六、重复抽样和不重复抽样	(23)
第三章 代表性误差及其计算方法	(25)
一、平均数、成数和实际误差	(25)
二、差异程度	(27)
三、代表性误差	(32)
四、平均数指标代表性误差的计算	(35)
五、成数指标代表性误差的计算	(53)
第四章 抽样调查结果的推算	(64)
第五章 随机抽样的容许误差	(74)

第六章 如何确定抽样数目	(79)
第七章 交叉样本法	(84)
第八章 抽样法的实际应用	(87)
一、农产量抽样实测调查	(87)
二、农民和职工家庭收支抽样调查	(104)
三、工业产品质量抽样检查	(111)
四、数字质量抽样检查	(115)
附录一、抽样误差计算表	(119)
附录二、随机号码表	(139)
附录三、概率表	(142)
附：概率论的基本原理和应用	(148)

第一章 调查统计与误差

一、从认识论说起

马克思主义的认识论告诉我们，人们的认识能力是无穷无尽的，客观事物都是可以认识的。但是，在一定的时间、地点、条件下，人们的认识能力，又都是有限的、不完全的。对自然现象的认识是这样，对社会现象的认识也是这样。

根据这个认识论的道理，我们进行调查统计，只要工作做好，是可以取得符合客观实际的结果的；但是，一般说来，我们调查统计的结果只能是接近实际，而不能达到百分之百的准确，即等于实际。换句话说，就是我们通常所说的调查统计的准确性，具有相对的意义，或多或少总会有一些误差。调查统计数字要做到百分之百的准确，只有在范围极小、对象极为简单的情况下，才能做到。

二、两种调查统计方法

调查统计方法是多种多样的，但如果按调查统计包括的范围来讲，概括起来，不外是两种：一种是全面调查，一种是非全面调查。

全面调查，就是对调查对象所包括的全部单位，都要毫无遗漏地进行调查统计。全面调查最普通的形式是：统一制发报表，全面统计，层层汇总，逐级上报。有的全面调查，则是采取普查的形式。如人口普查，往往是全国成立专门的普查机构，拟定一个调查方案，规定在同一个时间，每一个城镇、农村、工矿企业、机关、学校、街道和单位，都要进行调查。

非全面调查，就是对调查对象的部分单位进行调查。通常讲的非全面调查有以下几种：

（一）重点调查。就是对调查对象的部分重点单位进行调查。这些重点单位在调查对象中为数不多，但调查的标识在数量上占整个调查对象的绝大比重；通过对这部分重点单位的调查，就可以对整个调查对象有一个基本的了解。如重点钢铁企业调查，大型煤矿调查，大型水利工程调查，重点产棉区调查等等，就属于这类调查。

（二）典型调查。就是在对被调查对象进行全面分析的基础上，有计划地抓住几个典型单位，作系统的周密的调查研究。

（三）抽样调查。就是从被调查对象的所有单位中，抽取部分样本单位进行调查，用以推算总体。

采取什么调查方法，取决于调查的目的、要求、对象、力量、经费和时间等条件。

搞普查，是为了取得某些重要项目在一定时点的比较接近实际的全面资料，它的范围大，要求高，涉及面广，需要人力、物力、财力多，一般不轻易采用。如人口普查，世界

各国差不多都是十年、二十年才搞一次，我国建国以来才搞过两次；耕地普查，许多国家几十年、甚至几百年才搞一次。

对一些国民经济的基本指标，如计划执行情况、生产管理情况等，需要经常或定期提供的，一般采用发表调查的形式。

但不管采用什么形式，全面调查所需要的人力、物力、财力都是比较大的。发表调查必须有健全的统计机关和统计人员进行，而且从基层收集整理资料逐级汇总上报，中间环节多，花的时间比较长，容易出差错。

非全面调查则具有节省人力、物力、财力和时间等许多优点。

搞重点调查，适用于调查标识在数量上集中于少数单位的调查对象，即具有明显的重点单位的调查对象。通过少数重点单位的调查，就可以达到对全面情况有一个基本的了解；对一些次要的数目很多，而在总量中所占比重却很小的单位，则可以不花力量去调查。

搞典型调查的目的，是通过深入实际，解剖麻雀，发现问题，观察趋势，指导一般。有时，虽然也通过划类选点调查，用以推算全面，但也只用于观察一般趋势，而不要求精确计算总体的数量。

抽样调查的目的，则在于通过非全面调查推算取得全面资料。

三、两种误差

通常调查统计工作上讲的误差，和前面讲的调查方法一

样，也有两种。一种是调查误差，一种是代表性误差。

调查误差，是在调查统计工作过程中，由于种种工作上的原因而产生的误差，也叫做工作误差。产生调查误差的原因有：

（一）调查方案的设计有缺点，指标含义不清，口径不一致，调查方法不够完善，目的要求不够明确；

（二）遗漏或重复调查单位；

（三）被调查人理解问题、回答问题有错误或出入；

（四）被调查人或单位记忆不清，或是提供的情况和材料不实；

（五）调查人员了解情况不深入、不准确；

（六）调查人员或单位有意弄虚作假，虚报浮夸或瞒报；

（七）测量或测量工具不准确；

（八）登记、计算、汇总、抄写上有差错；

（九）调查统计工作中的其他原因，等等。

这种调查工作过程中产生的误差，是全面调查和非全面调查都会发生的。调查的范围愈广，规模愈大，内容愈复杂，参加调查的人员愈多，发生调查误差的可能性也就愈大；反之，调查的范围和规模愈小，内容愈简单，参加调查的人员愈精干，发生调查误差的可能性也愈小。

代表性误差，是指用部分来代表总体、推算全面时所产生的误差，即调查样本和其所代表的总体之间的误差。一般地说，任何部分和全体之间，总是或多或少存在一些差别，部分顶多只能近似全体，而不能等于全体。这种误差，就叫

做代表性误差。代表性误差是全面调查所没有的，重点调查和典型调查，不要求推算全面，也不存在这个问题，只是在抽取部分样本单位来代表总体、推算全面时，才有这种误差。所以，通常把代表性误差也叫做抽样误差。

四、怎样控制误差

误差一定会有，调查统计数字的准确性一般不可能达到百分之百。这符合认识论的客观规律。我们要认识这个规律，承认这个规律，按照客观规律办事，并且利用它来为我们的工作服务。

调查统计工作人员的一条重要任务，就是提高调查统计资料的准确性，换句话说，就是尽可能地控制和减少调查统计资料的误差。

怎样才能做到这一点，即把调查统计资料的误差控制和减少到最低限度，从而提高调查统计资料的准确性呢？

调查误差是调查统计工作过程中产生的，因而，要控制和减少调查误差，必须做好以下几方面的工作：

一是调查方案的设计，力求完善，做到指标含义清楚，口径一致，要求明确，计算方法简便易行；

二是要有健全的调查统计机构或专业调查队伍；

三是培训好调查统计人员，提高调查统计人员的政治思想和业务水平，使每个调查统计人员都有高度的责任感，熟悉调查业务，作风深入，工作细致，善于发现矛盾，分析问题，能坚持原则，如实反映情况；

四是向被调查人员或单位做好宣传解释工作，解除他们对调查存在的思想顾虑，取得他们的积极支持和合作，使他们清楚地了解调查的目的、要求，正确地提供有关的数字和情况；

五是完善测量和计算工具，做好资料的登记、审查、计算、汇总、抄写等方面的工作。

通常，进行调查统计工作的单位，在调查统计工作开始前，要召开调查统计人员会议，举办调查统计人员训练班；调查统计工作开始后，要加强调查统计工作过程中的检查督促，对调查统计资料采取集中审查汇编，等等。这些都是控制和减少调查误差的有效措施。

下面，将要着重探讨做好抽样调查，控制和减少代表性误差的问题。

第二章 抽 样 原 理

一、抽样法的基本概念

抽样法是一种用来了解全面情况的非全面调查观察的方法。它是按照科学的原理和计算，从若干单位组成的事物总体中，抽取部分样本来进行调查观察，用以推算总体，了解总体的基本情况。

通常，我们所要调查观察的全部事物，叫做全及总体，简称总体；抽取出来调查观察的单位，叫做抽样总体，简称样本。总体中的单位，叫总体单位；样本中的单位，叫样本单位。总体单位的平均数，叫总体平均数；样本单位的平均数，叫样本平均数。

例如，从一个堆放上百万斤水稻的仓库中，抽取部分水稻样品来化验和称量，根据这部分样品的含水量和杂质，推算整个仓库水稻的含水量和杂质；或是从一批布匹中，抽出若干匹来检验有无竹节、挑纱、瑕疵、破眼等，用以推算这批布匹的质量。这样的方法，就叫做抽样法，也叫做抽样调查法或抽样观察法。例子中，仓库的全部水稻和整批布匹就是我们说的总体，抽取出来调查观察的水稻样品和布匹就是我们说的样本。

但是，要做到抽取出来调查观察的样本能够代表事物的

总体，根据样本的调查观察结果能够推算总体的数量，在理论和实践上都有许多问题需要研究解决。

从仓库中抽取水稻样品，情况比较简单，抽样比较容易。因为在一般情况下，水稻在收割、干燥、扬净、入仓的过程中，拌和是比较均匀的；从仓库中任意抽取少量样品，对整个仓库的水稻都有很强的代表性。但是，抽样调查观察通常是在事物总体中各单位不相似的情况下进行的，尤其在社会经济调查中，差异程度更大。例如，对职工家庭收支情况进行抽样调查，由于各个职工家庭的收支，受到地区、行业、工种、家庭成员数和就业人口等许多因素的影响，情况复杂，差别较大，要使抽选出来的少部分样本单位对总体有较强的代表性，就不那么容易了。因此，必须有一套科学的办法。否则，调查推算出来的结果，就可能对总体缺乏足够的代表性，甚至出现很大的偏差。

二、抽样法的原理和特点

抽样法的首要目的，是获得一个规模有限、能够代表总体的样本。这个样本应当按照随机原则来抽取。随机样本应当成为总体的缩影，要尽可能地使总体的特征在样本中得以再现出来。

抽样法所依据的原理，是概率论和大数定律。就是说，按照随机原则从事物总体中抽取的每一个样本单位的数值，对总体的平均数来说，可能大，也可能小，而当抽样数目达到足够多时，则抽样误差是遵从正态分布的。因此，抽取

的样本单位数目愈多，样本平均数接近总体平均数的概率也愈大。

抽样法的特点，可以归纳如下：

第一，从总体中抽取出来调查观察，并用以推算总体的样本单位，是不加任何选择，按随机原则抽取的；总体中每一个单位都有被抽中的可能性。因此，能够保证被抽中的单位在总体中均匀分布，对总体具有充分的代表性，不致出现倾向性误差。

第二，抽样法是以抽取的全部样本单位作为一个“代表团”，用整个“代表团”来代表总体，而不是用个别单位来代表总体。个别单位有高有低，当抽取的样本单位数足够多时，高低单位趋向于互相抵销，因而“代表团”的平均数接近总体平均数。

第三，抽样误差和总体各单位之间的差异程度成正比，和抽样数目成反比例（确切地说，是与抽样数目的平方根成反比例）。抽样调查观察的误差，可以根据总体中各单位的差异程度和抽取样本单位的数目，事先通过计算控制在一定范围之内，从而使调查观察的准确程度比较有把握。

由于抽样法具有这些特点，因此，它被公认为是非全面调查方法中用来推算全面的最完善、最有科学根据的方法。

抽样法必须遵循的首要原则是随机原则。它是抽样法的基础。只有在坚持随机原则的前提下，样本的代表性才有可能得到保证；随机原则一旦遭到破坏，抽样法的其它原则和特点也就不稳固了，从而抽取出来的样本单位的代表性也就得不到保证。

曾经流行过一种说法：根据随机原则，按照抽签方式来抽选调查单位，抽好抽坏，全靠“碰运气”，很难保证调查结果具有充分的代表性。甚至有人认为抽样法就是“抽签算卦”，搞“迷信”，“不依靠人的主观能动性”，“见物不见人”。这完全是对抽样法的误解。“抽签算卦”固然是迷信，但用抽签方式来抽取调查观察的样本单位，却不能认为不是科学的；何况这只是抽样方法的一种形式。在实际工作中，更多地是用等距抽样或类型抽样，以有效地保证样本在总体中的均匀分布。但是，不论抽签方式也好，其它方式也好，道理是一样的，其目的在于排除人们的主观偏见，使各个单位被抽中的机会相等。例如，建国初期我国发行的国家经济建设公债，就是采用抽签还本的办法，规定每年抽还公债总额的5—20%。为什么要用抽签而不用其它办法呢？原因就是在这个办法最公平合理，也最简单易行；每张公债券都有可能中签，机会均等。用这种办法，对于个别人来说，什么时候中签是带有偶然性的，可能有些“碰运气”的感觉；但是，对于整个国家来说，每年还本的数额是规定了的，总有那么多债券中签，这却是一种必然性。如果不用抽签办法来还本，还有什么别的办法能保证这种最大的公平合理和机会均等呢？同样，抽签办法在世界的体育比赛中也是广为应用的，有谁能说这不是合理和科学的办法呢？

当然，对于社会经济现象来说，情况比公债还本、体育比赛要复杂得多；但是，这个公平合理、机会均等的原则，同样是适用的。任何事物，任何调查总体的共性，都包含在它的各个单位的个性之中。如果只从抽出的个别单位来看，它