

統計譯文專輯

• 抽样法在人口調查中的应用 •

11

統計出版社

出版者說明

本輯的六篇文章，是从“美国統計协会会刊”1952—1955年內若干期上选譯的。这几篇文章以具体的事例和經驗，从不同的方面論述了抽样法在人口調查中的应用，其中所介绍的方法，可供我国統計工作者了解和参考。

目 录

南斯拉夫1953年全国人口普查

- 中的抽样法.....〔南〕S.S.查柯維奇(1)
張世文譯
- 讀写能力資料的抽样檢查.....〔南〕S.S.查柯維奇(23)
王星角譯

* * *

- 中非的抽样調查.....〔英〕J.R.H. 蕭尔(35)
陈嘯武譯

- 論美国1950年人口普查的复查.....〔美〕E.S. 麦克司等(53)
奚言譯

論如何使样本估算数字与人口

- 普查計算数字相符合.....〔美〕F.F. 斯提反等(79)
姚志堅譯

南斯拉夫1953年全国人口 普查中的抽样法

S. S. 查柯維奇

綱 要

在南斯拉夫准备1953年的全国人口普查时，曾經發現兩类問題；这两类問題的解决，需要应用抽样法。第一类問題是所有人口普查所共有的；第二类問題則包括与落后国家特別有關的問題。

在第一类問題中，有一个問題是关于整理資料所需要的时间。1921年、1931年和1948年各次南斯拉夫人口普查資料的制表工作是用手工作的。1921年人口普查的資料整理工作延續了十年之久。1931年人口普查的資料整理工作，至1941年第二次世界大战开始时尚未完竣；而当时所有的資料都丧失无遺。1948年3月的人口普查資料，尽管範圍較狹，并有大量人員参加整理工作，也是到1949年年底才整理完竣的。

这样迟緩的整理工作，使資料的实际价值大部分丧失了。对1953年全国人口普查來說，主要的是，要对最重要的事实尽快地算出数字来，因为社会經濟措施都要以人口普查資料为依据。所以首要的問題是在更好和更完整的資料整理出来以前先編出一套初步估計的数字以資应用。

从下面可以看出，在此次全国人口普查中，是这样应用抽样方法的，即将先行估算数字的工作分为两个部分：

(1)就少数项目十分迅速地编出估计数字；(2)就多数项目以较长的时间编出估计数字。

另一个一般性的问题，是人口计算的全面性问题。对这一点进行研究，在若干方面都有重要意义。第一，要求对人口普查员所取得的1953年全国人口总数的准确程度作一个衡量。其次，和下次全国人口普查有关的某些统计问题，也要求把这种核对工作列入纲要中。核查人口调查的全面性，可以看出在哪些情况下调查工作会遇到困难，以及在一定的调查制度下，什么样的问题是在将来的人口普查工作中应该加以更充分的考虑的。从这里面还可以了解现有普查员的工作能力如何，在人口普查中所采用的各种定义的优缺点如何，采用的调查制度的优缺点如何等等。

又一个问题，是计量人口普查表上各项问题的答案错误的大小。为了正确地拟定人口普查表中的各项问题，为了很好地组织人口普查，也为了便利统计资料的利用者，这是一个非常重要的问题。关于答案中的错误这一问题的一般重要性是很清楚的，但是对这个问题的研究，对于那些文化发展水平较低，文盲人口（他们对于人口普查所使用的概念很不熟悉）比重很大的国家，却具有特殊的意义。在这种情况下，计量错误的大小，能够使我们看出，普查可以作到什么，以及所获得的资料的可靠性如何。如果把这种研究这样设计，使之可以表示出在一个问题上的错误答案所占的百分比，并且表明错误的来源，这种研究的用处便会更大。这种材料对于设法减少错误或改进答案都是有价值的。

这一类問題中的最后一个問題是校訂工作的質量。虽然在校訂过程中查出了人口普查表上的許多錯誤，但是并不是把所有的錯誤都消除了。而大家知道，还留在人口普查表上的許多錯誤从原則上講是可以得到改正的。所以，确定这次普查中的各种錯誤在校訂中被改正到何种程度，更一般地說，校訂工作使工作質量接近于理想到何种程度，是很重要的。如果对这些問題不能获得答案，那末，我們就不能对人口普查結果的可靠程度有充分的了解，也不能为下一次普查設計出完备的校訂程序。

在和落后国家有特殊关系的第二类問題中，我們曾对讀写能力資料的准确程度和价值进行了研究。在本刊1954年9月号的一篇文章中*，我們曾說明，这样做的第一个原因是由于以往几次普查所反映的文盲人口的百分数呈現不正常的趋势。这些文盲人口的百分数是：

人口普查	文盲人口的百分数
1921年	50.5
1931年	44.6
1948年	25.4
1953年	24.9 ^①

从上表可以看出，1931年至1948年文盲人口的百分比急剧下降（平均每年降低1%以上），而1948年至1953年則下降極少，这显然是有問題的，因为在前一时期包括战争年代，当时多数学校均已停閉；而后一时期，国家大力发展教育，并且不遺余力地扫除文盲。

除普查資料的可靠性有此种不能确定的情况需要了解

*見本輯，查柯羅奇：“讀写能力資料的抽样檢查”。——編者注

① 这个数字是样本的估計数。

外，对于对讀写能力所作的答案的意义，也还需要有更进一步的了解。在回答自己具有讀写能力的这一类人中，究竟其所代表的熟練程度如何？其中有多少人完全具有讀写能力，有多少人具有中等水平？有多少人位于完全文盲的边緣？

总起来說，在这一次全国人口普查中，我們所应用的抽样方法包括下列各項：

- (1) 初步估計，
- (2) 第二步估計，
- (3) 核对調查的全面性，
- (4) 对答案錯誤的范围进行研究，
- (5) 核对讀写能力資料的准确性，
- (6) 考核校訂工作的質量。

首先，我們將提供执行这一綱要所应用的样本的若干資料；其次，將簡要地說明在各个問題上进行了什么工作，获得了什么結果。

一 般 样 本

对本綱要的所有方面，均以普查区作为初級样本單位，对綱要中(2)和(6)兩部分，則另抽專門样本。其余四部分則应用同一初級样本。关于这个一般样本的情况，我們要在這裡說明；至于(2)和(6)兩部分的專門样本，則在以后說明。

在普查开始的前十天，每一行政單位（如区、县、共和国）中的普查区的数目，中央機構已經知道了。以行政区划为基础，將所有普查区均分为兩層，即乡村層和城市層。从城市層选出 100 个普查区；从乡村層选出 149 个普

查区。作为初級單位的普查区，是以相等的机率抽选的，只是乡村層中样本的成数較城市層中样本的成数为小。

在确定初級單位样本的大小时，有兩点考虑是具有决定意义的：第一种考虑是现有督导員人数。研究綱要是这样安排的，即只指派一个督导員到一个普查区去，在那里他需要做相当复杂的工作。因此，对督导員的訓練必須严格。由于准备情况較差，真正能够胜任的督导員是有限的；因此初級样本單位数必須少于300。第二是与綱要的第(3)項有关的。当时希望，計量普查的全面性时，应当得出一个人口总数的估計数字，其誤差系数不超过0.30%。根据进一步的考察，我們希望人口普查中和复核工作中同一数目的居民所屬的調查区之間，相关系数可高达0.995。因此，可以用一个有250个普查区的样本来滿足上述要求。

应用初級單位的样本，就可以得到綱要中第(1)与第(3)兩項的答案，用不着再抽取子样本。要得到綱要中第(4)与第(5)兩項的答案，則需要从所选出的普查区中再抽选一些个人作为子样本。綱要中的这两部分的第二阶段的样本是不一样的，因为包括在这兩种研究中的人数不一致。有关这些样本的其他材料，將在以下各段加以說明。

第 1 表 关于总体与样本的基本資料

	城 市 層	乡 村 層
1.居民数(千人)	4,852	12,095
2.普查区总数	29,805	89,149
3.样本中的普查区数	100	149
4.样本中普查区所占的百分比	0.30	0.15
5.样本中人数所占的百分比	0.35	0.16

上表資料系根据初步計算。

第1表包括按城市層和乡村層分組的总体中及样本中的居民数和普查区数。关于这个样本的比較詳尽的討論，可參閱文末所列參考資料〔1〕。

初 步 估 計

为执行本綱要所聘用的督導員，于普查日后8至10日去普查区。此时普查登記工作已經完畢，全国每一个人都有一張填好的人口普查表（理論上是如此）。为了获得初步估計数字，原来計劃讓督導員帶着他們区内所有人的人口普查表。將这些普查表加以整理，便可以求得我們所需要的估計数字。用这种作法可以得出結果，而整理过程沒有任何重复，因为样本区的人口普查穿孔卡片以后还可以放在适当的地方，和其他卡片在一起。

由于急需資料的緣故，將計劃作了如下的改变：一个样本区中所有已經調查的个人所作的各項答案，均由派往該区的督導員彙总在一張表上；由此得到的人口总数就用来求出估計数字。普查日为3月31日，这种初步估計数字在4月底即已得出。因此，調查工作一經結束，普查的初步結果便可立刻知道。在这些估計資料中包括下列几类統計表：（1）人口統計資料，（2）族籍，（3）讀写能力，（4）教育程度，（5）宗教，（6）經濟活动人口。

从抽样誤差来看，这种估計方法不及簡單随机抽选个人样本的方法有效，因为某些指标，如产业部門、職業集团、以及一般說来所有的經濟指标，在普查区内均有一种比較高的組內相关。为了比較，我們提出下列的資料：以普查区为抽样單位，用比例估計法所得到的有讀写能力者在总人口中所占比重的相对标准誤，在乡村層为2.5%，

在城市層为 1.9 %。如果随机选出同等数目的个人，則有讀写能力者在总人口中所占比重的簡單的无偏性的估計数的方差，在乡村層为 0.32 %，在城市層为 0.27 %。这两种估計方法的效率是相差頗大的，但是，因为获得前一种估計数字的費用不大，这种方法大概在將來还会采用，唯一的改变是初級單位的数目將要增加。采用較大的初級單位样本来得出这些估計数字是容易的，因为必要的資料的編制很簡單，并不需要一个有訓練的督導員参加。

在类似的場合，我們常常对就普查区进行子抽样感到兴趣。不过在这次工作中，我們未抽取子样本，因为我們發現这样作并不能减少費用。

第二步估計

在需要交叉分类的場合，249 个普查区的样本是不足以得出令人满意的估計数字的。要就六个共和国分別得出估計数，这个数目的样本也是不夠的。因此需要另一个較大的样本。所以需要这第二个样本，也因为在編制第一种估計数字时所用的个人资料以前并没有校訂过，同时还因为普查資料的正规整理工作要在 1955 年年底完成。

为了满足对于較詳細的数字的需要而設計的第二个样本，也是以普查区为初級抽样單位；这些普查区在各个共和国是按不同的抽样分数抽选的，抽样分数的变动是由 0.14 到 0.67。第二阶段的抽选，則选出了 10 % 的住戶。第一阶段的抽选之所以采用不同的抽样分数，是因为要使抽样誤差在各个共和国都位于差不多一样的水平。

以住戶为第二阶段的抽样單位，与簡單随机抽选的个人样本相比，在某些指标方面，效率还是較低的。在一个

共和国曾經作过一个实验，就不同的抽样單位和不同的估計方法的效率进行了比較。实验的結果，就某些选定的項目扼要地列在第2表中，其中以个人簡單随机抽样的方差为100。分層則是按戶的大小进行的。

以这种資料为基础，并且对各种方法所需的費用加以考虑，于是决定以戶为比較經濟的次級單位。

这个样本所采用的工作程序，所获得的結果，以及对所遇到的問題所作的討論，可在文末所列參考資料〔2〕、〔4〕、〔6〕、〔8〕中找到。后面兩種出版物中还包括有一些統計表。这些統計表分別按項目、按共和国和按省列出标准誤的大小。

第2表 不同抽样單位和不同估計方法的效率的比較*

項 目	抽 样 單 位				
	个 人		戶		
	簡單随机抽样	分 層 抽 样	簡單随机抽样	分 層 抽 样	比 例 抽 样
男 子	100	100	70	147	143
文 盲	100	100	71	71	71
按出生年份分组的人数					
1940—1948年	100	108	59	87	103
1900—1909年	100	104	90	100	96
1879年	100	110	91	91	87
就業人口总计	100	102	52	52	71
工 人	100	100	66	73	67
雇 員	100	100	80	81	69
农 民	100	100	30	42	40
其 他	100	100	92	92	85

*上表数字与參考資料〔4〕中的数字在形式上略有不同。

調查的全面性

为了評價普查的全面性，我們根据下面的計劃，利用了249个初級單位的样本，沒有再抽取子样本。

在普查登記工作完竣后，便派督導員往中选的普查区，并供給他們以普查区的地圖，在地圖上画明了普查区的四界。他們的任务是訪問普查区的全部住宅單位，重新計算所有的人口。只把每人的姓名記下来，并把他們分为下列三类：（1）常住本地的，（2）常住本地但暫時不在本地的，（3）暫住本地的。頭兩类居民構成常住人口。进行此次重新計算时，督導員并不了解原来人口普查所得的結果。督導員的任务是把普查日的情況复制出来。

当督導員重新清点完竣以后，地方人口普查委员会（这是由一群当地居民組成的，他們負責監督实地的普查工作）便將核对的結果和人口普查的結果进行比較。采用了按戶核对时所制成的名單，来便利这种比較。如果發現有任何差異，就由地方人口普查委员会再进行一次核对。这次新的核对所获得的結果是最后的結果。由此可以看出，並沒有認為督導員所获得的資料是一定正确的。这种第二次核对工作是很必要的，因为以后發現在督導員的工作中有些錯誤。有时督導員少計算了一个人，更多的情况是將人口按上述三类分类时作得不正确。

將兩套結果进行对比，还能消除可能由于下述事实而产生的錯誤，即核对普查的完整性的工作，督導員須在普查日后至少一星期才能进行。這兩次調查的間隔時間，本来可以由普查員和督導員同时工作来縮短。但是，这样并不能得出我們所期望的結果，因为这会使普查員知道要在他

那一个区进行核对，从而將使核对工作的整个目的丧失。所以兩次調查間隔几天是必要的。但是，如果由于人口的变动，督導員不一定总能把普查日的人口情况清查得很确实，那么，上述的对比和下列的核对方法就可以减少他們的錯誤。

以这种調查的結果为基础，用下列方法便能說明普查員的工作質量。在所选出的普查区中，將所有的人都归入下列各組之一中：

- (1) 未經普查員分类的人 (用符号“-”表示)，
- (2) 被普查員和督導員归入同一类的人 (用符号“=”表示)，
- (3) 兩次調查中均被登記，但被归入不同类别的人 (用符号“≠”表示)，
- (4) 普查員曾予以分类，但督導員未予以分类的人 (用符号“+”表示)，

以上各組情况發生的百分比見第3表。这些数字表明普查員工作質量的不同方面。

第3表 四組情况發生的百分比

組 別	百 分 比	
	乡 村 層	城 市 層
-	0.25	0.83
+	0.26	0.45
≠	0.58	1.58
=	98.93	97.34

至于常住人口数，第4表数字是一个粗略的估計。

第 4 表 关于常住人口調查的全面性的估計

	層 別		总 計	
	乡 村	城 市	人口数	百分比
普查数字低于实际数字	27,500	21,500	49,000	0.29
普查数字高于实际数字	40,000	11,500	51,500	0.30
总 誤 差	67,500	33,000	100,500	0.50
净 誤 差	+12,500	-10,000	+2,500	0.10

从上表数字可以看出，城市層虽發生淨負差，乡村層和人口总数則發生淨正差。这种結果是有些不尋常的。汉森、赫尔維茲与普里茲克（参閱参考資料〔7〕）曾經表明，对美国1950年全国人口普查的一次类似的核对，全部都是淨負差。这种情况在統計調查中似乎是一种合乎邏輯的期望。但是我們所得的結果，可以拿下述事实來說明，即許多屬於乡村層的人，他們的工作地点是在他們的家庭所在地以外。这种情况的發生，大致是由于住宅缺乏。这些人有的一个星期回家一次；也有些人只隔若干時間回家一次。普查关于这类人口的指示便弄得杂乱了，結果是重复計算，在家中計算一次，在工作地点又計算一次。同样的困难也發生在学生离家在学校住宿的情况中。

关于普查的全面程度問題，應該在这里加以补充說明。从結果来判断，普查工作从全面性來說是出色的。这可能是由于原来的調查工作很好，也可能是由于核对工作做得不好，不管原来的調查的質量如何。要就这里所發生的問題作一个完滿的答复，几乎是超出实际的可能的，因为这就需要再作一次核对。对于我們的普查結果的一般性質良好的唯一的补充証据，是对21个普查区再作一次核对的結果。从这次核对中，我們看出除了一个人以外，其余

的人的情况都是一致的。其他的比較主觀的証据，是普查工作的一般实地督導員关于这一点的报告：他們也認為正式的普查員的工作很好。

为了改进普查方法和这种研究的技术，分析普查結果与核对結果不一致的事例是非常重要的。在这里看到，不一致的情况很少是因为普查員沒有把某人进行分类而發生的。大多数的差異是与將定义应用于个别情况的困难有关。为了說明問題，我們談一下兩個普查区的例子。有一个普查区内設有一座肺病疗养院；另一个普查区有一个新工厂的建筑工地。這兩個普查区都經常發生人口变动；有些人在那里住一个月，有些人住几个月，甚至于几年。在这种情况下，一个普查員几乎不可能依据人口普查所下的定义去进行普查工作。同时，这些定义本身也不能常常規定得适合于這兩個普查区的情况。关于这方面的核对工作的詳細說明，參閱參考資料[10]。

答案中的錯誤

人口普查表答案的錯誤，可以在任何被調查者身上發生，不論他們的年齡大小，或其他个人特征如何，也不論普查表是由被調查者本人填写（如指示所規定的），或由普查員代填（代文盲或小孩等填写），或由第三者代填（这是常有的事）。因此，在研究綱要的这一部分所使用的第二阶段的样本，对 249 个选出的初級單位中所調查的一切个人都要加以考虑。在这种情况下，次級單位的样本是由督導員在实地抽选的。他們用填好了的普查表作为研究的基础，并且采用了下列的抽样比例：乡村層为 1:8，城市層为 1:10。因此从乡村層中选出了 2,470 人，从城市層

中选出了1,684人。城乡兩層采用不同的抽样比例，是为了使督導員工作的分配接近相等。

实地調查工作是依下列程序进行的。督導員并不知道以前在人口普查中的答案，他們要和样本中的每个人进行接触，就普查表上比較重要的問題再向他們取得答案。为了达到这样的目的，他們曾經利用一些“互相核對的問題”，并且請求出示文件以証明所得答案的正确性（如果有文件，同时回答人也願意拿出来的話）。然后以所得結果与原有普查結果核對。在这次核對中，彼此一致的答案就被認為是正确的。在大多数的情况下，这可以認為是合理的假設；但是，它不一定是說答案是絕對正确的。举例來說，如果有一个人，由于某种原因，在“職業”這個問題上曾經給了一个錯誤的答案，可是后来他向督導員所作的答案与以前相同，那么，在这次核對工作中，他的答案也就被認為是正确的，而实际上是錯誤的。这里又出現了这个問題：用各种核對方法究竟可以做到什么程度？为了簡單和經濟，我們的办法是如前所述的那樣，指出所得結果在正确性上的相应的限度。

与此相反，有些答案發現不一致时，便要再进行一次新的核對。这一次要采用更严格的方法。督導員和地方人口普查委员会的工作人員要負責去寻找証件，如果有地方可以找到的話，要和了解某人情况的人們去接談，以便从他們那里获得关于这个人的材料，或者再找这个人，再詢問些問題，以便对所考慮的問題得到更多的說明。这样所得的答案，就看作是正確的。

为了作最后的分析，采用了下面对錯誤所下的定义。普查表上的各个答案，如与核對所得相应問題的答案不同，

就是錯誤的。除上述這一類錯誤外，形式上的錯誤也要算。下面就是這一類形式上的錯誤的例子：如果有人人在“教育程度”這一問題後面畫一短橫綫符號，這就算為錯誤，因為填表說明清楚地要求寫出完全的答案。這樣，我們在分析中所用的“錯誤量”這個概念，就包括這兩類錯誤。

首先把普查表看作一個整體。凡是核對時沒有發現什麼差異的，以及普查表上沒有什麼形式上的錯誤的，都應該算是正確的。經過了這樣的核對，我們就可以找出正確無誤的普查表。這樣我們得到下列第5表中各項百分比，在表中按六個共和國正確的普查表所占百分比的大小的次序加以排列。我們看到，各共和國排列的次序，是和它們的文化發展水平的高低相符合的。所說的文化發展水平，是以識字人口所占的百分比和教育程度來計量的。對於專業統計人員來說，這是重要的材料。

第5表 各共和國正確的普查表所占的百分比

共和國名稱	百分比
斯洛文尼亞	59.1
克羅地亞	54.0
塞爾維亞	45.9
門的內哥羅	40.9
波斯尼亞—黑塞哥維那	36.8
馬其頓	24.8

如果根據在普查表中所發現錯誤數目的多少把普查表加以分類，我們就得到第6表第1欄的百分比。如果按性別對所得結果加以分析，便可發現男子的錯誤答案比女子的錯誤答案為少。