

## 市场调研中关于敏感性问题调查的统计方法

文/刘春雨

### 一、引言

市场调查现在已被广泛应用于社会经济生活各方面。它省时省力，能比较迅速获得较为准确的结果。这一方面是方法本身具有科学性；另一方面也是前提假定是被调查者的回答都是真实可靠的。在市场调研中，经常会遇到诸如吸毒走私、赌博考试作弊、收入财产、性行为取向等敏感性私密问题，调查过程中若采用直接问答的方式，被调查者为了保护自己的隐私或出于其他目的，往往会拒绝回答或故意做出错误的回答。这样就破坏了数据的真实性。运用随机化回答技术设计问题进行调查可以解决问题。本文就此一问题调查作以简单分析。

### 二、关于沃纳模型

1965年沃纳提出了沃纳模型 (Warner model) 后才发展起来的。基本思想是所论及的总体为一个简单的二元总体，即总体中每个单位或属于A，或不属于A，除此之外，别无它属。根据敏感性特征设计两个相互对立的问题，让被调查者按预定的概率从中选一个回答，调查者无权过问被调查者究竟回答的是哪一个问题，从而起到了为被调查者保密的效果。

#### 1、抽样方法及比例的估计：

制作两类卡片：

卡片1：“具有特征A”

卡片2：“不具有特征A”

卡片1在其中所占的比重以p表示；卡片2所占的比重相应为1-p。

在简单随机有放回抽样下从总体中抽得n个人作样本，然后对这n个样本单位进行调查。让其从混合均匀的两类卡片中任取一张卡片，根据卡片上的问题作答。则此人回答“是”的概率是：

$$\lambda = p(\text{具有特征A的人, 并回答是的}) + p(\text{不具有特征的A的人, 并且回答是}) \\ = p(\text{具有特征A的人}) \cdot p(\text{回答是} | \text{具有特征A的人}) + p(\text{不具有特征A的人}) \\ \cdot p(\text{回答是} | \text{不具有特征A的人})$$

$$= \pi A \cdot p + (1 - \pi A) \cdot p = (1 - p) + (2p - 1) \pi A \quad (1)$$

$$\lambda_1 = p(\text{具有特征A的人, 并回答否的}) + p(\text{不具有特征的A的人, 并且回答否的}) \\ = p(\text{具有特征A的人}) \cdot p(\text{回答否的} | \text{具有特征A的人}) + p(\text{不具有特征A的人}) \\ \cdot p(\text{回答否的} | \text{不具有特征A的人})$$

$$= \pi A (1 - p) + (1 - \pi A) p \quad (\text{回答否的概率})$$

于是由(1)可解出： $(p \neq 1/2)$ 。因此得到 $\lambda$ 的估计量，就可用它来估计 $\lambda$ ，进而可估计 $\pi A$ ，即具有特征A的比例。

为了得到 $\lambda$ 的估计量，从N中用放回方法取一个容量为n的简单随机样本，令n个人分别用放回方法从两类卡片中各取一张，并根据抽中卡片上的问题作出真实回答，则可得到 $\lambda$ 估计量： $\bar{\lambda}$ ，其中为样本中回答“是”的人数。易知满足：

① ②

且是的无偏估计量，证明从略。

以作为的估计量，可得无偏估计量： $(p \neq 1/2)$ ，满足：

a、

b、

c、

在Warner方法当中，p不能取1/2；其抽样误差：  
将代入方差公式，得：

上面等式右端第一项是直接回答时估计量的方差；第二项是随机化回答引起的误差增量。当p=1或p=0时，第二项为0，此时相当于市场调研中的直接回答，但不能有效保护被调查者的隐私，难于得到真实资料；当p→1/2时，对被调查者的保护程度是加强，但估计量的方差却将增大。这是p≠1/2的真正原因。对此问题的解决可以采用西蒙模型(1967)、二次随机化模型(赵俊康1994)、改进的随机化模型。

#### 2、关于卡片比例p与样本量的确定

由上式可以看出，p→1/2，则Var(A)的值越大。当p越靠近0或1时，Var(A)就越小。但当p越接近1或0时，对被调查者的保护程度就会降低，从而降低被调查者的合作程度，使随机化回答

的作用降低，增加市场调研过程中收集真实数据的难度。p的经验取值一般介于0.7—0.8之间效果较好，当然应根据实际市场调研问题的敏感程度适当选取。

由于，所以预先给定精度要求方差不超过 $\alpha$ ，那么市场调研所需最低样本量n的取值大于即可。一般只需取：即可达到要求。其中[x]表示不大于x的最大整数。

3、对于无放回方式抽样下的沃纳模型。

假设总体容量为N，采用无放回简单随机抽样方法抽取容量为n的样本，则估计为：

$$A = (\lambda - (1-p)) / (2p-1), \quad \text{其方差为:}$$

三、改进的随机化回答模型

模型的设计如下：制作一套卡片，由三类卡片组成，第一类卡片上写有“如果你具有特性A，请回答数字1；如果你不具有特性，请回答数字0”。第二类卡片上写上“请直接回答数字1”。第三类卡片上写有“请直接加答数字0”。将这三类卡片各若干混合均匀放入袋中，其比例分别为 $p_1$ ， $p_2$ 和 $p_3$ ，且 $p_1 + p_2 + p_3 = 1$ 。从总体中有放回地抽取容量为n的简单随机样本，由被抽中的人采用有放回的方法随机地从中抽去一张作答，回答结果只有“0”和“1”根据回答“1”的数目m，可估计出的值。

设总体中具有特性A的比例为，样本中回答“1”的概率，则，的无偏估计量为：，其方差为：

又由于，即为的无偏估计量。此方法克服了沃纳模型中不论抽中几号卡片都必须回答敏感性问题的缺点，消除了被调查者的顾虑；且设计简便，利于实际操作。

敏感性调查已成为市场调查中的一个重要的内容，因此对其进行深入的研究很有现实意义(作者单位：中国计量学院管理学院)

#### 相关链接

我国出口产品遭受反倾销的原因分析及对策研究  
城市化进程中农民市民化的现实困境及其突破  
基于制度监管视角的社保基金安全保障分析  
市场调研中关于敏感性问题的统计方法  
论技术性贸易壁垒对我国出口贸易的影响  
发展中国家实现跨越式发展的几点思考  
高薪养廉的经济学方法论批判  
民营企业劳资矛盾成因分析  
论城市规划对城市土地价格的影响

本网站为集团经济研究杂志社唯一网站，所刊登的集团经济研究各种新闻、信息和各种专题专栏资料，均为集团经济研究版权所有。

地址：北京市朝阳区关东店甲1号106室 邮编：100020 电话/传真：(010) 65015547/ 65015546

制作单位：集团经济研究网络中心