

## 蒙特卡罗模拟对多分类敏感问题随机应答技术模型下整群抽样信度与效度的评价

施佳琛<sup>1</sup>, 陈向宇<sup>1</sup>, 高歌<sup>1\*</sup>, 周云华<sup>1</sup>, 傅颖<sup>1</sup>, 王磊<sup>2</sup>

(<sup>1</sup>苏州大学公共卫生学院流行病与卫生统计学系, 江苏 苏州 215123; <sup>2</sup>上海市徐汇区疾控中心, 上海 200030)

**[摘要]** **目的:**对多项选择敏感问题随机应答技术结合整群抽样的调查方法和公式进行效度和信度评价。**方法:**通过蒙特卡罗模拟,用西昌市女性性工作者敏感问题调查结果作为模拟总体的参数,以模拟直接真实应答值作为相对准则,分别对 100 对样本作两组构成分布比较的卡方检验,进行信度和效度评价。**结果:**100 个卡方检验的  $P$  值中有 99 个大于 0.05,按  $\alpha = 0.05$  水准,100 次假设检验的结果中两组构成分布差异无统计学意义的为 99 次。**结论:**本项目研究的调查方法和计算公式有较高的信度及效度,可用于今后大规模的敏感问题抽样调查。

**[关键词]** 多分类敏感性问题;随机应答技术;蒙特卡罗模拟;效度;信度

**[中图分类号]** O212

**[文献标志码]** A

**[文章编号]** 1007-4368(2013)07-1007-05

**doi:**10.7655/NYDXBNS20130733

## Assessing reliability and validity of cluster sampling on multichotomous sensitive question with randomized response technique by monte carlo simulation

Shi Jiachen<sup>1</sup>, Chen Xiangyu<sup>1</sup>, Gao Ge<sup>1\*</sup>, Zhou Yunhua<sup>1</sup>, Fu Ying<sup>1</sup>, Wang Lei<sup>2</sup>

(<sup>1</sup>Public Health School, Soochow University, Suzhou 215123; <sup>2</sup>Xuhui Center for Disease Control and Prevention, Shanghai 200030, China)

**[Abstract]** **Objective:** To assess the reliability and validity of randomized response technique (RRT) model of cluster sampling survey on multichotomous sensitive question investigation method and statistical formulae. **Methods:** A total of 404 female sex workers were selected randomly by cluster sampling in Xichang during May 2011 to July 2011. RRT was applied to obtain their sensitive characters. With the sensitive characters taken as parameter of simulated overall, simulated direct questioning results taken as criterion, results got from multichotomous RRT model were compared with ones from simulated direct questioning by using chi-square test with Monte Carlo simulation to access validity and reliability. **Results:** There were 99 results have no statistically significant different from criterion ( $\alpha=0.05$ ). **Conclusion:** The method and corresponding formulae have high validity and reliability, and can be used for sample survey of large-scale sensitive issues.

**[Key words]** multichotomous sensitive question; randomized response technique; Monte Carlo simulation; validity; reliability

[Acta Univ Med Nanjing, 2013, 33(7): 1007-1011]

敏感性问题是指被调查者出于安全、个人形象等原因拒绝真实回答的隐私性问题,例如卖淫、同性恋、艾滋病等<sup>[1]</sup>。敏感性问题可分为数量特征敏感性问题和属性特征敏感性问题;属性特征敏感性问题又可进一步分为两分类敏感问题和多分类敏感问题<sup>[2]</sup>。在采用传统的直接问答方式调查敏感问题时,被调查者常

常会故意误答,产生信息偏倚<sup>[3]</sup>。为了确保调查结果的真实性,发展出了许多专门用于敏感问题调查的方法,其中随机应答技术(randomized response technique, RRT)被认为是最为可靠的方法<sup>[4]</sup>。

本文采用多分类敏感问题 RRT 模型结合整群抽样的调查方法对西昌市女性性工作者这类性病、艾滋病高危人群展开调查,并以实际调查结果作为模拟总体的参数,采用蒙特卡罗模拟方法,实现从同一模拟总体 100 对重复样本的模拟调查,以期对本

**[基金项目]** 国家自然科学基金项目(81273188)

\*通信作者(Corresponding author), E-mail: gaoge@suda.edu.cn

文所研究的多分类敏感问题 RRT 模型下整群抽样调查方法及其统计公式的效度和信度做出评价,为今后大规模的敏感性问题的抽样调查提供科学的方法和公式。

## 1 多分类敏感问题 RRT 模型下整群抽样调查的统计方法

### 1.1 整群抽样方法

设包含  $N$  个群的总体中第  $i$  个群内有  $M_i$  个观察对象,先从  $N$  个群中随机抽取  $n$  个群;再使用 RRT 模型对抽中群的全部观察对象进行调查。

### 1.2 多分类敏感问题的 RRT 模型

假设某多分类敏感性问题有  $t$  种互斥的类别,将写有数字  $0, 1, 2, \dots, t$  的  $t+1$  种相同重量、质感、大小的小球放入袋中作为随机化装置,小球的数量比例为  $P_0:P_1:P_2:\dots:P_t (P_0+P_1+P_2+\dots+P_t=1)$ 。每个被调查者随机从袋中有放回地取出一球,如果取出的球上数字是  $0$  则填写自己所属敏感特征的序号;如果取出的球上数字不是  $0$  则直接填写该数字。回答过程没有其他人在场,以免对被调查者产生影响。

### 1.3 计算公式

$\hat{\pi}_{ij}$  为第  $i$  个群第  $j$  类敏感问题总体比例的估计值,  $m_{ij}$  记样本中第  $i$  个群回答数字  $j$  的频数,  $a_{ij}$  为第  $i$  个群调查对象属于类别  $j$  的估计频数。总体中属于类别  $j$  的总体比例的估计量为  $\hat{\pi}_j$ 。根据相关文献<sup>[5-6]</sup>:

$$\hat{\pi}_{ij} = \frac{m_{ij}/M_i - P_j}{P_0}, i=1, 2, \dots, n, j=1, 2, \dots, t \quad (1)$$

$$a_{ij} = M_i \hat{\pi}_{ij} = M_i \frac{m_{ij}/M_i - P_j}{P_0}, i=1, 2, \dots, n, j=1, 2, \dots, t \quad (2)$$

$$\hat{\pi}_j = \frac{\sum_{i=1}^n a_{ij}}{\sum_{i=1}^n M_i}, j=1, \dots, t \quad (3)$$

$\hat{\pi}_j$  的估计方差为:

$$v(\hat{\pi}_j) = \frac{1-f}{n\bar{M}^2} \frac{\sum_{i=1}^n a_{ij}^2 - 2\hat{\pi}_j \sum_{i=1}^n a_{ij} M_i + \hat{\pi}_j^2 \sum_{i=1}^n M_i^2}{n-1}, j=1, \dots, t \quad (4)$$

其中  $\bar{M} = \frac{\sum_{i=1}^n M_i}{n}$  是样本中每群包含的平均观

察对象数,  $f = \frac{\sum_{i=1}^n M_i}{N}$  是抽样比。

### 1.4 调查实例

本次调查采用整群抽样方法,调查对象为西昌市的女性工作者。于 2011 年 5~7 月,本项目组成员对西昌市皮肤病性病防治所确认存在有偿性服务的商业活动区的女性工作者进行了调查。样本含量的确定参考 Wang 等<sup>[7]</sup>给出的计算结果,调查总体按街道与社区人为划分为 57 个群,并从中随机抽取 19 个群(抽样比为 1/3),对 19 个群中的全部对象均采用多分类敏感问题随机应答技术进行调查,共调查 404 人,平均每个群约有 21 名调查对象。本次对女性工作者调查的多分类敏感问题包括:最近 1 年到正规医疗机构进行性病检查情况的类别;最近 1 年到正规医疗机构进行艾滋病检测情况的类别;最近 1 次性服务时安全套破损情况的类别;最近 1 个月全程使用安全套情况的类别。以性病检查情况类别的调查为例,由上述公式(1)~(3)计算得到: $j=1$ (从未到正规医疗机构检查过)、 $j=2$ (检查过但结果不确定)、 $j=3$ (检查过但不知道结果)、 $j=4$ (检查结果无性病)、 $j=5$ (检查结果有性病)的比例分别为 0.448 894、0.043 131、0.110 406、0.335 066、0.062 502。

## 2 蒙特卡罗模拟

### 2.1 基本原理及思路

蒙特卡罗模拟法(Monte Carlo simulation, MCS)即概率论与数理统计学中以统计试验作为手段,利用电子计算机模拟实验,实现计算某个指标或总体参数的随机试验法;该方法与一般的数值直接计算方法有着本质区别,属于试验数学的一个分支<sup>[8]</sup>。它最早起源于用频率近似概率的数学思想,利用随机数求得待解问题近似值。MCS 的理论基础就是利用计算机软件产生的独立随机变量来进行模拟,本研究使用 Windows 时间函数生成随机数的方法<sup>[9]</sup>,命令计算机根据生成的随机数和预先设定的参数将所研究的多分类敏感问题的取值随机地赋予模拟总体中的每个对象,从而模拟得到 1 个总体。随后使用整群抽样法随机地在模拟总体中抽得 1 个样本。最后根据实际随机化装置的设定来模拟使用了随机应答技术的具体调查过程。对模拟调查的数据通过公式(1)~(4)计算结果。由于本研究在于获得比例指标,故模拟样本数在 100 次以内为佳<sup>[10]</sup>,因此通过计算机编程自动实现同样条件下 100 个样本的重复抽样调查并计算结果。对每个模拟样本,同时对应产生 1 个整群随机抽样下非 RRT 的直接真实应答模拟样本并以此作为相对准则,对 100 对样本(每对样本中,

一个是多分类 RRT 模型下整群抽样模拟样本,另一个是其对应产生的整群抽样非 RRT 直接真实应答模拟样本的平均水平)作 100 次两个构成比比较的卡方检验,以评价多分类敏感问题 RRT 模型结合整群抽样的调查方法和公式的效度和信度。

2.2 模拟过程

2.2.1 模拟总体的构建

本次模拟总体设定为 60 个群,包含 1 200 人,接近真实总体的大小(1.4 中群的抽样比为 1/3,19 个样本群包含 404 人);以女性性工作者性病检查情况类别调查为例,依据 1.4 中实际调查结果对模拟总体的参数进行设定:近一年到正规医疗机构进行性病检查各情况类别( $j=1\sim 5$ )的人群比例分别设定为 0.448 894、0.043 131、0.110 406、0.335 066、0.062 502。在 .net 编程平台下通过计算机生成的 0~1 范围内的随机数对所有模拟个体的排布次序进行随机再排布,以完成每个模拟个体数值的设定(设定每个模拟个体所属的群号及敏感特征

类别的序号)。

2.2.2 抽样过程的模拟

根据模拟调查的内容对抽样的要求和随机装置进行设定,每个模拟样本中样本群数设定为 20 个,女性性工作者的数量近似为 400,采用整群抽样;随机装置中贴有数字 0 的小球数量为 5 个,贴有数字 1~5 的小球各有 1 个。依照比例对新生成的变量赋值(该值为各敏感特征的应答值),变量个数为样本量,并利用随机数重新排布变量次序(变量的随机化)。

2.2.3 模拟结果

进行 100 个样本的模拟试验后,通过公式(1)~(4)计算统计量;同时对整群随机抽样非 RRT 模型下直接真实应答的 100 个模拟样本,按整群随机抽样相关统计公式计算统计量。以女性性工作者性病检查情况类别调查为例,在检查情况类别  $j=1\sim 5$  下,两种模拟调查各 100 个样本共计 1 000 个模拟调查的部分结果见表 1。

表 1 女性性工作者性病检查情况  $j$  类别下两种调查的模拟结果

Table 1 Results of two kinds of simulated investigations of female sex workers'  $j$ th STD test

编号	$j=1$		$j=2$		$j=3$		$j=4$		$j=5$	
	RRT	直接真实应答								
1	0.425	0.435 0	0.045	0.045 0	0.115	0.122 5	0.345	0.335 0	0.070	0.062 5
2	0.460	0.457 5	0.045	0.037 5	0.095	0.107 5	0.320	0.330 0	0.080	0.067 5
3	0.500	0.480 0	0.040	0.040 0	0.115	0.107 5	0.305	0.317 5	0.040	0.055 0
4	0.490	0.440 0	0.025	0.050 0	0.090	0.112 5	0.350	0.352 5	0.045	0.045 0
5	0.460	0.475 0	0.045	0.040 0	0.150	0.140 0	0.290	0.290 0	0.055	0.055 0
6	0.450	0.462 5	0.030	0.032 5	0.145	0.122 5	0.335	0.325 0	0.040	0.057 5
7	0.430	0.440 0	0.040	0.047 5	0.135	0.122 5	0.310	0.320 0	0.085	0.070 0
8	0.430	0.440 0	0.050	0.047 5	0.115	0.127 5	0.350	0.330 0	0.055	0.055 0
9	0.410	0.432 5	0.025	0.050 0	0.130	0.122 5	0.355	0.322 5	0.080	0.072 5
10	0.480	0.435 0	0.035	0.040 0	0.115	0.107 5	0.300	0.342 5	0.070	0.075 0
...	...	...	...	...	...	...	...	...	...	...
99	0.485	0.512 5	0.020	0.027 5	0.115	0.102 5	0.330	0.312 5	0.050	0.045 0
100	0.495	0.512 5	0.040	0.042 5	0.090	0.090 0	0.305	0.300 0	0.070	0.055 0

2.2.4 信度和效度评价

信度指调查结果的一致性 or 稳定性,即多次调查的结果是否存在一致性。效度指调查结果与客观实际相吻合的程度。效度越高,即表明调查结果越能反映出所要测量对象的真正特征。对一项调查来说,效度是其首要条件,而信度是效度的必要条件。

本文以模拟整群随机抽样非 RRT 模型直接真实应答的各类比例的平均水平作为准则比例(接近

模拟总体的各比例真值),分别对 100 对样本作两组各类别比例构成分布比较的卡方检验,SAS 计算得 100 次卡方检验的  $P$  值见表 2。

从表 2 可以看出,100 次假设检验中仅有 1 次  $P < 0.05$ ,99 次  $P$  均  $> 0.05$ ,说明各类别样本比例与各类别准则样本比例的差别无统计学意义,可认为本文研究的多分类敏感问题随机应答技术结合整群抽样的调查方法和公式具有较高的效度;又因 100 次重复模拟调查的计算结果均接近同一准则,说明

表2 信度与效度评价结果

Table 2 Results of assessment of reliability and validity

编号	P 值	编号	P 值	编号	P 值	编号	P 值
1	0.964 5	26	0.095 3	51	0.183 2	76	0.521 2
2	0.825 1	27	0.458 1	52	0.278 3	77	0.738 4
3	0.453 2	28	0.188 1	53	0.564 7	78	0.760 1
4	0.334 7	29	0.753 1	54	0.674 6	79	0.416 3
5	0.403 1	30	0.833 7	55	0.310 9	80	0.271 3
6	0.312 9	31	0.914 0	56	0.264 4	81	0.986 0
7	0.559 0	32	0.556 1	57	0.232 9	82	0.976 5
8	0.939 5	33	0.352 9	58	0.692 2	83	0.618 3
9	0.364 8	34	0.725 7	59	0.116 9	84	0.284 3
10	0.797 4	35	0.621 7	60	0.968 4	85	0.636 1
11	0.518 0	36	0.859 7	61	0.503 7	86	0.882 2
12	0.180 9	37	0.658 4	62	0.535 0	87	0.215 0
13	0.060 2	38	0.952 3	63	0.879 8	88	0.974 3
14	0.299 5	39	0.000 3	64	0.919 7	89	0.110 2
15	0.643 0	40	0.187 2	65	0.269 8	90	0.815 8
16	0.723 1	41	0.388 9	66	0.192 9	91	0.934 5
17	0.678 5	42	0.842 7	67	0.823 4	92	0.258 1
18	0.179 5	43	0.072 5	68	0.631 1	93	0.275 3
19	0.485 6	44	0.102 7	69	0.850 5	94	0.484 1
20	0.198 9	45	0.987 4	70	0.635 8	95	0.174 2
21	0.975 5	46	0.562 1	71	0.910 9	96	0.388 6
22	0.954 5	47	0.942 7	72	0.526 8	97	0.357 1
23	0.427 1	48	0.190 0	73	0.352 7	98	0.659 1
24	0.757 1	49	0.227 0	74	0.260 3	99	0.110 2
25	0.513 5	50	0.781 9	75	0.230 0	100	0.275 3

多分类敏感问题随机应答技术结合整群抽样的调查方法和公式具有较高的信度。

### 3 讨论

随着社会发展,各类敏感性问題不断涌现。在医疗卫生领域,近年来我国艾滋病的流行呈快速上升趋势,形势严峻。因此,对艾滋病这类健康危害巨大的疾病的防控刻不容缓。其中对艾滋病传播高危人群总体特征的准确把握,是国家制订相关疾病防治政策的前提。

随机应答技术是由美国统计学家 Warner 首创,后经众多国内外统计学家发展完善而形成的一系列专门针对敏感性问題进行调查并推算总体特征的有效方法。由于这类方法使用特定的随机化应答装置,根据概率论知识计算出敏感问題特征在人群中的分布,故其能避免调查对象在没有任何保护的情况下直接回答问题,从而能够起到保护调查对象的隐私,

提高调查结果真实性的作用。

信度和效度是反映研究方法是否可靠有效的重要指标。本项目组成员对多种 RRT 模型结合复杂抽样方法及其统计公式,进行过重测信度、和谐信度及准则效度的评价<sup>[11-13]</sup>,但评价方法仅限于对少量调查结果进行分析。国外学者还使用 Meta 分析<sup>[14]</sup>和实际调查<sup>[15]</sup>的方法比较 RRT 和传统调查方法在调查敏感问題上的优劣,也仅仅是对少量实际调查结果的一种综述。本文采用蒙特卡罗方法通过计算机模拟调查 100 对 200 个样本,在随机应答技术的效度与信度评价方法上进行了一次新的尝试,具有方法学上的创新意义。

评价结果显示整群抽样结合多项选择敏感问題随机回答模型的调查研究方法和计算公式具有良好的信度与效度,可用于今后较大规模敏感问題的抽样调查,对制定艾滋病防治规划、政策及措施具有较高的实际价值。

[参考文献]

- [1] 朱宏儒,高歌,于明润,等. 数量特征敏感问题分层两阶段整群抽样的统计方法[J]. 南京医科大学学报:自然科学版,2009,29(6):909-912
- [2] 高歌,范玉波. 敏感问题 Simmons 模型的(分层)整群抽样研究[J]. 中国卫生统计,2008,25(6):562-565
- [3] 张倩倩,还锡萍,羊海涛,等. 江苏省部分地区暗娼人群性病感染率及相关危险因素分析[J]. 南京医科大学学报:自然科学版,2012,32(4):473-478
- [4] Gupta S,Shabbir J,Sehra S. Mean and sensitivity estimation in optional randomized response models [J]. J Stat Plan Inference,2010,140:2870-2874
- [5] Cochran WG. Sampling Techniques [M]. 3rd ed. New York:Wiley,1977:87,93-95,432,491
- [6] 苏良军. 高等数理统计 [M]. 北京:北京大学出版社,2007:3-4
- [7] Wang JF,Gao G, Fan YB, et al. The estimation of sample size in multi-stage sampling and its application in medical survey [J]. Appl Math Comput,2006,178:239-249
- [8] 布斯连科. 统计试验法(蒙特卡罗法)及其在电子计算机上的实现[M]. 王毓云,译. 上海:科学技术出版社,1964:256
- [9] 戴颖,计奎. 生成正态分布随机数的一种新方法——基于 Windows 时间函数[J].地矿测绘,2004,20(2):7-8
- [10] Bonate PL.A brief introduction to Monte Carlo Simulation [J]. Clin Pharmacokinet,2001,40(1):15-22
- [11] 高歌,范玉波. 敏感问题改进的随机应答技术模型分层整群抽样研究及应用[J]. 苏州大学学报:医学版,2008,28(5):750-754
- [12] 王冕,高歌. 数量特征敏感问题的整群抽样调查方法研究及应用[J]. 中国卫生统计,2008,25(6):586-589
- [13] 王磊,高歌,于明润. 敏感问题双无关问题模型分层二阶段整群抽样的统计方法及应用[J]. 中国卫生统计,2011,28(1):37-39
- [14] Lensvelt-Mulders GJ,Hox LM,van der Heijden JJ,et al. Meta-analysis of randomized response research[J]. Sociol Methods Res,2005,33(3):319-348
- [15] Krumpal I. Estimating the prevalence of xenophobia and anti-Semitism in Germany: A comparison of randomized response and direct questioning[J]. Soc Sci Res,2012,41(6):1387-1403

[收稿日期] 2013-01-09

本刊邮发代号 28-61

网址: <http://jnmn.njmu.edu.cn>