

DOCTORAL THESIS

Reliable techniques for survey with sensitive question

Wu, Qin

Date of Award:
2013

[Link to publication](#)

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

Reliable Techniques For Survey With Sensitive Question

WU Qin

A thesis submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Principal Supervisor: Dr. Tang Manlai

Hong Kong Baptist University

August 2013

Abstract

Refusal and providing untruthful answers from respondents are very common in the survey of collecting highly personal or sensitive issues. The randomized response model is the first interview technique designed to eliminate the non-response and response bias when sensitive questions are being asked. The idea of this technique is to hide the real status of a respondent's answer by a deliberate contamination of the data via a randomizing device. However the application of this technique is restricted by its low efficiency and high cost induced by the introduction of the randomizing device.

Non-randomized response (NRR) models have recently been developed for analyzing sensitive questions. Unlike traditional randomized response models, non-randomized response models do not require randomizing device which limits the survey format and the reproductivity of results. On the other hand, non-randomized response models introduce a non-sensitive question with binary outcome in the questionnaire in order to protect privacy and encourage cooperation from respondents. Unfortunately, the proportion of subjects who possess the non-sensitive characteristic is assumed to be known, and the non-sensitive and sensitive questions are assumed to be independent. However the two assumptions might be invalid in practical application and it can be shown that an incorrect prior knowledge about the non-sensitive question could lead to severely biased estimate of the sensitive proportion. Furthermore it is assumed that the respondents will provide their true answers applying with the survey design, but it is reported in various experimental studies that some respondents still choose to provide untruthful answers to demonstrate their positive image when they belong to the sensitive group. In this thesis several non-randomized response techniques will be proposed to solve the known- p , independence assumption and noncompliance problem.

For the item count technique, it is noteworthy that all respondents only need to report the total number of statements that apply without telling which specific statements apply to them. This technique prohibits interviewers from getting the

information of the sensitive question at the level of individual subject. Despite the various attractive advantages, existing item count techniques possess three severely fatal drawbacks. Firstly, if an item count technique design consists of m non-sensitive questions and one sensitive question, respondents who give an answer of $m + 1$ (i.e., answers to all questions are “yes”) will automatically expose their sensitive characteristic. In this case, these respondents may refuse to answer or provide untruthful answers. This definitely weakens the applicability and credibility of the technique. Secondly, the proportion estimate and confidence interval provided in the existing papers may not be reliable as they sometimes lie outside the interval $[0, 1]$. Thirdly, the Wald-type confidence interval obtained from the technique may not be reliable for small sample designs and more reliable confidence intervals are desired. A new technique called modified item count technique is introduced to solve the above problems in which the m binary non-sensitive questions are replaced by one single question which follows Poisson distribution. This technique not only protects the privacy of respondents who possess the sensitive characteristic but also provides reliable proportion estimate, shorter confidence interval, and useful sample size formulas.

Keywords: randomized response model, non-randomized response model, triangular model, item count technique, noncompliance.

Table of Contents

Declaration	i
Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	viii
Chapter 1 Introduction	1
1.1 Background	1
1.2 Models For Surveys With Sensitive Characteristics	2
1.2.1 Randomized Response Models	2
1.2.2 Item Count Techniques	4
1.2.3 Non-Randomized Response Models	5
1.3 Arrangement of Thesis	7
Chapter 2 Flexible Non-Randomized Response Models For Survey With Sensitive Question	10
2.1 Introduction	11
2.2 Three New Non-Randomized Response Triangular (NRRT) Models .	13
2.2.1 Brief Review On Non-Randomized Response Triangular (NRRT) Model	13

2.2.2	NRRT Model With Unknown p Being Estimated By A Supplement Table (NRRT-ST)	15
2.2.3	NRRT Model Using Alternating Tables (NRRT-AT)	18
2.2.4	Dependence NRRT Model Using Alternating Tables (DNRRT-AT)	21
2.3	Simulation Studies	23
2.4	Real Examples	25
2.4.1	Real Example Based On NRRT-ST	26
2.4.2	Real Example Based On NRRT-AT	27
2.4.3	Real Example Based On DNRRT-AT	27
2.4.4	Conclusion	27
2.5	Discussion	28

Chapter 3 Non-Randomized Response Model For Sensitive Survey With

	Non-compliance	30
3.1	Introduction	31
3.2	Non-Randomized Response Triangular Models With Noncompliance .	34
3.2.1	The Original Non-Randomized Response Triangular Model . .	34
3.2.2	The Dual Non-Randomized Response Triangular Model (DNRRT)	35
3.2.3	The Alternating Non-Randomized Response Triangular Model (ANRRT)	39
3.3	Sample Size Determination	42
3.4	Simulation Studies	43
3.5	Real Example: Pre-marital Sex Experience	45
3.6	Conclusion	47

Chapter 4 Modified Item Count Techniques For Surveys With Sensitive

	Question	48
4.1	Introduction	49

4.2	Modified Item Count Technique With Known Parameter	52
4.2.1	Illustration	52
4.2.2	Estimate	53
4.2.3	Confidence Interval Estimates	54
4.3	Modified Item Count Technique when λ is unknown	55
4.3.1	The Proposed Technique	56
4.3.2	Estimate	57
4.3.3	Confidence Interval Construction	58
4.4	Selection of λ	59
4.5	Sample Size Determination	60
4.6	Simulation Studies	62
4.7	Real example	65
4.7.1	MICT with known λ_0	66
4.7.2	MICT with unknown λ	67
4.8	Discussion	67
	Chapter 5 Conclusion and Future Work	69
	Bibliography	72
	Curriculum Vitae	78