



Australian
National
University

ES SIS/LIBRARY
MENZIES LIBRARY BUILDING NO:2
AUSTRALIAN NATIONAL UNIVERSITY
ERRA ACT 0200 AUSTRALIA

TELEPHONE: +61 2 6125 4631
FACSIMILE: +61 2 6125 4063
EMAIL: library.theses@anu.edu.au

USE OF THESES

This copy is supplied for purposes
of private study and research only.
Passages from the thesis may not be
copied or closely paraphrased without the
written consent of the author.

ESTIMATION IN FINITE POPULATION SAMPLING: ROBUSTNESS AND OPTIMALITY

Siu Ming Tam

A thesis submitted for the degree of
Doctor of Philosophy
of the Australian National University
July 1988

DECLARATION

The results in Chapters 1 to 5 of this thesis have been published as follows:

Chapter 1	Tam (1988a)
Chapter 2	Brewer, Hanif & Tam (1988)
Chapter 3	Tam (1986a), Tam (1987a) and Tam (1988b)
Chapter 4	Tam (1986b)
Chapter 5	Tam (1987b)

Excepted where otherwise stated in the text, the work presented in this thesis is my own.



Siu Ming Tam

ACKNOWLEDGEMENTS

I am grateful to my supervisors, Professor E. J. Hannan, for offering me an ANU scholarship for the first year of my study, and Professor C. C. Heyde, for his encouragement during the preparation of this thesis.

I would also like to acknowledge the many useful discussions I had with Drs. K. R. W. Brewer and R. L. Chambers of the Australian Bureau of Resource and Agricultural Economics on my research work during my stay there.

Above all, my work would not have been completed without the support and love of Annie, Tiffany and Timothy during the past years of research. I like to dedicate the thesis to them.

ABSTRACT

Considerable progress has been made in the past two decades on research into survey sampling methodologies. A significant trend in this development is that the classical survey sampling techniques are formulated using an inference framework, which is made possible by assuming that the given finite population is a realization of a superpopulation model with unknown parameters. One attraction of this method of formulation of survey sampling problems is that not only the classical tools of inference can usefully be brought to bear on the problem, but more importantly, the optimal properties of certain familiar estimators in classical survey sampling are more and more better understood.

A central issue in the use of superpopulation models for designing sampling strategies is robustness of the procedures. Under the design-based approach, it is claimed in the literature that robustness is provided by requiring that the estimators are design-unbiased, a property which is independent of the assumed superpopulation model. Chapter 1 examines the claim in detail and provide sufficient conditions for the sampling strategy to attain

the Godambe-Joshi lower bound to the expected variance (Godambe & Joshi, 1965).

Chapter 2 and Chapter 3 examine the issues of reconciling design-based estimators with model-based predictors. The approach adopted in Chapter 2 is to retain the asymptotic design unbiasedness of the predictors, but modify the form of the estimators for the superpopulation model parameters in constructing the predictors of population totals. On the other hand, the approach adopted in Chapter 3 is to retain the model-based properties of the predictors but modify the selection probabilities for selecting the sample.

The discussion for Chapters 1 to 3 is restricted to the design and analysis of ad hoc surveys, that is, surveys which are conducted once on a given finite population. Chapter 4 and Chapter 5 examine issues related to the analysis of repeated surveys of the finite population over time.

Chapter 6 looks at specification of superpopulation models for survey sampling from the point of view of performing regression analysis on the survey data from ad hoc surveys. It is argued that under certain circumstances, the uniform intra-class correlation assumed in the covariance matrix of the superpopulation model is unnecessary.

CONTENTS

DECLARATION	(ii)
ACKNOWLEDGEMENTS	(iii)
ABSTRACT	(iv)
CHAPTER 1. Robust design-based estimators under a general linear model	
1.1 Summary of main results	1
1.2 Asymptotic design unbiasedness	2
1.3 Notation and some well known results	5
1.4 Some key results in robust estimation	10
1.5 Misspecification of the design matrix	18
1.6 Robustness of certain model-based predictors	19
1.7 Misspecification of the covariance matrix	20
1.8 An example revisited	21
CHAPTER 2. Model-based prediction and design-based estimation: how nearly can they be reconciled?	
2.1 Summary of main results	25
2.2 The three estimator classes	26
2.3 Conforming to the generalized regression estimator	29

2.4	Imposing design-based properties on the BLU predictors	37
-----	--	----

CHAPTER 3. Robust model-based predictors under a general linear model

3.1	Summary of main results	39
3.2	Some key results	41
3.3	Robust estimation	45
3.4	Asymptotically design-unbiased predictors	47
3.5	Optimality of Royall's predictor under the Gaussian assumption	53
3.6	A lower bound for the model-based mean squared error	55

CHAPTER 4. Analysis of repeated surveys I: use of stochastic regression models

4.1	Summary of main results	58
4.2	Brief review of literature on repeated surveys	58
4.3	Motivating the model	63
4.4	Model specification	64
4.5	Estimation of regression coefficients	67
4.6	Prediction of random variables	71
4.7	Analysis of repeated surveys - extension of classical results	72
4.8	Analysis of repeated surveys - minimum mean squares linear unbiased predictors	72

CHAPTER 5. Analysis of repeated surveys II: use of state-space models

5.1	Summary of main results	74
5.2	Introduction	75
5.3	Motivating the state-space model	80
5.4	Prediction results with known initial state	86
5.5	Prediction results with unknown initial state	90
5.6	Maximum likelihood estimation	94
5.7	Square-root filtering	97
5.8	An example	97
5.9	Conclusion	103
Appendix		
5.A	Selected list of notation	106
5.B	Proofs of main theorems	108
5.C	Square-root filtering algorithm	113

CHAPTER 6. Survey data analysis: the uniform intra-class correlation assumption is unnecessary

6.1	Summary of main results	115
6.2	Introduction	115
6.3	Maximum likelihood estimation	117
6.4	A special case	121
6.5	Relaxing the assumption of multivariate normality	128
6.6	Conclusion	128

REFERENCE	130
-----------	-----