This book is an introduction to the model-based approach to survey sampling. It consists of three parts, with Part I focusing on estimation of population totals. Chapters 1 and 2 introduce survey sampling, and the model-based approach, respectively. Chapter 3 considers the simplest possible model, the homogenous population model, which is then extended to stratified populations in Chapter 4. Chapter 5 discusses simple linear regression models for populations, and Chapter 6 considers clustered populations. The general linear population model is then used to integrate these results in Chapter 7. Part II of this book considers the properties of estimators based on incorrectly specified models. Chapter 8 develops robust sample designs that lead to unbiased predictors under model misspecification, and shows how flexible modelling methods like non-parametric regression can be used in survey sampling. Chapter 9 extends this development to misspecification robust prediction variance estimators and Chapter 10 completes Part II of the book with an exploration of outlier robust sample survey estimation. Chapters 11 to 17 constitute Part III of the book and show how model-based methods can be used in a variety of problem areas of modern survey sampling. They cover (in order) prediction of non-linear population quantities, sub-sampling approaches to prediction variance estimation, design and estimation for multipurpose surveys, prediction for domains, small area estimation, efficient prediction of population distribution functions and the use of transformations in survey inference. The book is designed to be accessible to undergraduate and graduate level students with a good grounding in statistics and applied survey statisticians seeking an introduction to model-based survey design and estimation.
Prediction for Small Areas
Raymond L. Chambers and Robert G. Clark

in An Introduction to Model-Based Survey Sampling with Applications
Published in print: 2012 Published Online: May 2012
Publisher: Oxford University Press
DOI: 10.1093/acprof:oso/9780198566625.003.0015
Item type: chapter

Prediction for small areas introduces an important application of sample survey inference, where domain sample sizes are too small to allow domain-specific inference. Typically, these domains are spatially defined, and so are referred to as small areas. Direct and indirect estimation for small areas is discussed, with the latter based on characterising the distribution of the survey variable via a linear mixed model. The empirical best linear unbiased predictor is developed, as are estimates of its mean squared error. An alternative approach, which conditions on differences between the areas, is used to motivate a domain-type linear estimator, the model-based direct estimator, as well as an estimator of its conditional mean squared error. The extension of the indirect approach to where the survey variable can be modelled via a generalised linear mixed model is sketched. The chapter concludes with a discussion of recent developments in small area inference.

Some topics in social statistics
David Firth

in Celebrating Statistics: Papers in honour of Sir David Cox on his 80th birthday
Published in print: 2005 Published Online: September 2007
Publisher: Oxford University Press
DOI: 10.1093/acprof:oso/9780198566540.003.0008
Item type: chapter

This chapter summarizes recent themes and research topics in social statistics, viewed as statistical methods of particular value in substantive research fields such as criminology, demography, economics, education, geography, politics, psychology, public health, social policy, and sociology. Special emphasis is given to multi-level models, small area estimation, models for obtaining measuring instruments, and weighting problems arising in survey data. Particular areas in which further work seems likely to be fruitful are identified by discussing special features connected with incomplete data; policy evaluations; causal inquiries; event history data; aggregate data; macro-level phenomena arising from actions of individuals who influence one another; performance monitoring of public services; and open-source projects for statistical computing.