THESIS

POPULATION SIZE ESTIMATION USING THE MODIFIED HORVITZ-THOMPSON ESTIMATOR WITH ESTIMATED SIGHTING PROBABILITY

Submitted by

Char-Ngan Wong

Department of Statistics

In partial fulfillment of the requirements

for the Degree of Ph.D. in Science

Colorado State University

Fort Collins, Colorado

Summer, 1996

COLORADO STATE UNIVERSITY

May 17, 1996

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY CHAR-NGAN WONG ENTITLED " POPULATION SIZE ESTIMATION USING THE MODIFIED HORVITZ-THOMPSON ESTIMATOR WITH ESTIMATED SIGHTING PROBABILITY " BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF PH.D. IN SCIENCE .

Committee on Graduate Work

Adviser

Department Head

ABSTRACT OF THESIS

POPULATION SIZE ESTIMATION USING THE MODIFIED HORVITZ-THOMPSON ESTIMATOR WITH ESTIMATED SIGHTING PROBABILITY

Wildlife aerial population surveys usually use a two-stage sampling technique. The first stage involves dividing the whole survey area into smaller land units, which we called the primary units, and then taking a sample from those. In the second stage, an aerial survey of the selected units is made in an attempt to observe (count) every animal. Some animals, usually occurring in groups, are not observed for a variety of reasons. Estimates from these surveys are plagued with two major sources of errors, namely, errors due to sampling variation in both stages. The first error may be controlled by choosing a suitable sampling plan for the first stage. The second error is also termed "visibility bias", which acknowledges that only a portion of the groups in a sampled land unit will be enumerated.

The objective of our study is to provide improved variance estimators over those provided by Steinhorst and Samuel (1989) and to evaluate performances of various corresponding interval procedures for estimating population size. For this purpose, we have found an asymptotically unbiased estimator for the approximate variance of the population size estimator when sighting probabilities of groups are unknown and fitted with a logistic model. We have broken down the approximate variance term into three components, namely, error due to sampling of primary units, error due to sighting of groups in second stage sampling and error due

all three components separately in order to get a better insight to error control. Simplified versions of variance estimators are provided when all primary units are surveyed and for stratified random sampling of primary units. Third central moment of population size estimator was also obtained.

Simulation studies were conducted to evaluate performances of our asymptotically unbiased variance estimators and of confidence interval procedures such as the large sample procedure, with and without transformation, for constructing 90% and 95% confidence intervals for the population size. Confidence intervals for the population size were also constructed by assuming that the distribution of $\log(\hat{\tau} - T)$ is normally distributed, where $\hat{\tau}$ is the population size estimate and T is the number of animals seen in a sample obtained from a population survey. From our simulation results, we observed that the population size is estimated with negligible bias (according to Cochran's (1977) working rule) with a sample of at least 100 groups of elk obtained from a population survey when sighting probabilities are known. When sighting probabilities are unknown, one needs to conduct a sightability survey to obtain a sample, independent of the sample obtained from a population survey, for fitting a logistic model to estimate sighting probabilities of sighted groups in the sample obtained from the population survey. In this case, the population size is also estimated with negligible bias when the sample size of both samples is at least 100 groups of elk. We also observed that when sighting probabilities are known, we needed a sample of at least 348 groups of elk from a population survey to obtain reasonable coverage rates of the true population size. When sighting probabilities are unknown and estimated via logistic regression, the size of both samples is at least 428 groups of elk for obtaining reasonable coverage rates of the true population size. Among all these confidence intervals, we found that those approximate confidence intervals constructed based on the assumption that $\log(\hat{\tau} - T)$ is normally distributed and using the delta method have better coverage rates and shorter estimated expected interval widths.

Confidence intervals for the population size using bootstrapping were also evaluated. We were unable to find an existing bootstrapping procedure which could be directly applied to our problem. We have, therefore, proposed a couple of bootstrapping procedures for obtaining a sample to fit a logistic model and a couple of bootstrapping procedures for obtaining a sample to construct a population size estimate. With 1000 pairs of independent samples from a sightability survey and a population survey, each sample of size 107 groups of elk and using 500 bootstrap iterations, we obtained reasonable coverage rates of the true population size.

Our other problem is model selection of a logistic model for the unknown sighting probabilities. We evaluated the performance of the population size estimator and our variance estimator when we fit a simpler model. For this purpose, we have derived theoretical expressions for the bias of the population size estimator and the mean-squared-error. We found, from our simulation results of fitting a couple of models simpler than the full model, that the population size was still well estimated for the fitted model based only on group size but was severely overestimated for the fitted model based only on percent of vegetation cover. For both fitted models, our variance estimator overestimated the observed variance of 1000 simulated population size estimates. We also found that the approximate expression of the expected value of the population size estimator we derived for a fitted model simpler than the full model has negligible bias (by Cochran's (1977) working rule) relative to the average of those 1000 simulated population size estimates. The approximate expression of the variance of the population size estimator we derived for this case somewhat underestimated the observed variance of those 1000 simulated population size estimates. Both approximate expressions apparently give us an idea of the expected size of the population size estimate and its variance when the fitted model is not the full model.

Char-Ngan Wong
Department of Statistics
Colorado State University
Fort Collins, Colorado 80523
Summer, 1996

ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Dave Bowden, for his patience, encouragement and valuable guidance during the two and a half years of my endeavour to complete the Ph.D. thesis. I would also like to thank other members in my Ph.D. committee for their valuable suggestions and comments.

I also wish to thank my family for being extremely understanding and supportive the whole time.

CONTENTS

1 Introduction **1**
1.1 Two-Stage sampling 1
1.2 Population surveys of wildlife population 1
1.3 Early population size estimator 2
1.4 An outline of our study 5

2 Modified Horvitz-Thompson estimator of population size **10**
2.1 Development of the population size estimator 10
2.1.1 When sighting probabilities are known 12
2.1.2 When sighting probabilities are unknown 12
2.2 Development of an unbiased estimator for the variance of the population size estimator 13
2.2.1 When sighting probabilities are known 13
2.2.2 When sighting probabilities are unknown 14
2.3 Estimation of the population size of a finite population 19
2.3.1 Example of how real data are collected and how population size is estimated using modified Horvitz-Thompson estimator . . 19
2.3.2 Some problems of aerial surveys 22
2.3.3 About the sightability model 24
2.4 Ratio estimator 26

3 Modified Horvitz-Thompson estimator of the population size for three special cases **29**
3.1 Introduction 29
3.2 Case 1: Complete census of primary units 30
3.2.1 When sighting probabilities are known 30
3.2.2 When sighting probabilities are unknown 31
3.3 Case 2: Stratified random sampling of primary units 32
3.3.1 When sighting probabilities are known 33
3.3.2 When sighting probabilities are unknown 36
3.3.3 Optimum allocations 40
3.4 Case 3: Population size estimator based on the composite data of several replicated population surveys 40
3.4.1 Simple random sampling of primary units 41
3.4.2 **Stratified random sampling of primary Units** 46
3.4.3 Optimum allocations 52

4 Third central moment of the modified Horvitz-Thompson estimator for population size 53
4.1 Third central moment in terms of conditional moments 53
4.2 Third central moment of the modified Horvitz-Thompson estimator of population size 54
4.2.1 When sighting probabilities are known 54
4.2.2 When sighting probabilities are unknown 57

5 Simulation study of the population size estimator and the unbiased estimator of the variance of the population size estimator 61
5.1 Introduction 61
5.1.1 Measure performances of estimators by bias 62
5.1.2 Construct confidence interval estimate for population size 64
5.2 Simulated populations 67
5.3 Simulation 74
5.4 Results of the simulation 81
5.4.1 Simulation results of the simulated population of 116 groups and the subpopulations with sightability model (5.2.1) as the true model 81
5.4.2 Simulation results of the simulated population of 116 groups and the subpopulations with sightability model (5.2.2) as the true model 92
5.4.3 Simulation results of population *Penn* - the composite data of 107 groups from Pennsylvania, with sightability models (5.2.1) and (5.2.2) as true models 99
5.5 Dominant factor in the variance estimator 106
5.6 Estimation of elk population size in Pennsylvania 107
5.7 Discussion 109

6 Bootstrap confidence interval for the population size 113
6.1 Introduction 113
6.2 Case 1: A complete census of primary units 113
6.2.1 Bootstrapping in phase I 114
6.2.2 Bootstrapping in phase II 115
6.3 Case 2: Simple random sampling without replacement of primary units 117
6.4 Simulation study 118
6.4.1 Simulated population I 118
6.4.2 Phase I 119
6.4.3 Phase II 121
6.4.4 Combining phase I and phase II 123
6.5 Discussion 127

7 Logistic Model Selection for Unknown Sighting Probabilities 132
7.1 Introduction 132
7.1.1 Expectation and covariance matrix of estimated parameters when the fitted model is simpler than the full model 133
7.1.2 Expectation and variance of the modified Horvitz-Thompson estimator when the fitted model is simpler than the full model 137
7.2 A simulation study 139
7.3 Simulation result 139
7.4 Discussion 149

8 Summary, conclusion and topics for future research 164
8.1 Summary and conclusion 164
8.2 Suggested topics for future research 169

9 REFERENCES 173

10 Appendix: Computer codes 182

1. INTRODUCTION

1.1 Two-Stage sampling

Suppose a population can be divided into a number, say N, of smaller, non-overlapping units. A sample of n units is chosen. Suppose further that the i^{th}, $i = 1, 2, ..., n$, selected unit can be subdivided into, say M_i, subunits and a sample of, say m_i, subunits is selected. This technique is called *subsampling* (Cochran, 1977), since the selected unit is not measured completely but is itself sampled. Mahalanobis called it *two-stage sampling,* because the sample is taken in two steps. The selected units are often called primary units and the selected subunits from the chosen primary units are called the second-stage or secondary units.

1.2 Population surveys of wildlife population

In a population survey of a wildlife population, the entire survey area is first divided into smaller, non-overlapping land units and a sample of these primary units is selected using say, simple random sampling. Within each primary unit or selected land unit, groups or clusters of animals are observed, usually via aerial surveys for ungulate populations. Of these observed groups, certain characteristics or measures such as the group size of the group, degree of vegetation or snow cover under which the group is observed, are recorded.

Estimates from these surveys of animals are typically plagued by two major sources of sampling errors. The first major error is due to first stage sampling of the land units. The choice of sampling units and sampling designs are the primary means for controlling this variation. Caughley (1977a) describes basic types of alternative primary sampling units including systematic sampling, strip transects, large quadrats, and small quadrats. In section 3.3, we will consider the special case of using stratified random sampling. The second major source of error is due to second stage sampling of groups within a sampled land unit. In second stage sampling, since some groups are missed during the aerial surveys, the collected data are sometimes called the incomplete or missing data (Cochran, 1983). This type of missing data is called a *unit nonresponse*. The most obvious consequence of nonresponse is that a smaller sample of data is available for making the estimates than was originally planned. There is often reason to believe that nonrespondents differ systematically from respondents (Deming, 1950). Thus, in making estimates from the available respondent data, the sampler may face biases that are essentially unknown in size and direction. This error is hence, also termed "visibility bias" where only a portion of the groups within the selected land unit will be enumerated. In this context, visibility is viewed as analogous to the probability of response in a nonresponse problem where missed groups are viewed as nonresponse.

Readers should note that in this context, a sampling frame such as simple random sampling without replacement for second stage sampling is not available. Clusters of animals, or of any subjects in general, have their respective sighting probabilities or visibilities which could be defined by a sightability model in terms of their group size or characteristics such as degree of vegetation cover.

1.3 Early population size estimator

This paragraph is quoted from section 2.1 of Steinhorst and Samuel, 1989:

> "The application of survey sampling methods to the task of estimating wildlife population parameters was practically nonexistent prior to 1964. At that time, Siniff and Skoog (1964) provided the first signifi-

cant application of sampling techniques by using a stratified random design to estimate the size of a caribou population. Later, Jolly (1969) provided the foundation for applying survey design methods to the special problems of aerial survey. Jolly described the population estimators for stratified sampling based on equal or proportional selection probabilities. Caughley (1977a) described the use of systematic sampling designs and contrasted sampling with or without replacement. This application of survey procedures allowed the calculation of a major component of sampling error–namely, that due to the variation between sample units. However, in the interim, increasing evidence indicated that undercounting of animal groups had a substantial impact on estimates of population size (Caughley, 1974, 1977)."

Early population estimators did not adjust for visibility bias. This results in underestimating population size. Caughley (1974) used corrected counts to allow correction of the population estimator for missed groups. Caughley, however, did not consider the case of unknown sighting probabilities. Oh and Scheuren (1983), working on nonresponse in sample surveys, developed procedures for a modified Horvitz-Thompson estimator which relates to our problem.

Steinhorst and Samuel (1989) extended Oh and Scheuren's work to the general case of first stage sampling of primary units and a second stage sighting of clusters or groups within sampled primary units. Samuel and Steinhorst have presented modified Horvitz-Thompson population size estimators for the case of known sighting probabilities and the case of unknown sighting probabilities. In both cases, they showed that their population size estimators are unbiased and they derived expressions for the variance of the population size estimator in both cases.

For the case of unknown sighting probabilities, the sightability model is assumed to be a logistic model (see (2.1.3)). In the context of 'sighting' of groups in the second stage sampling, a group is either sighted (responded) or not sighted (nonresponse) with some sighting probability. Under the assumption that the sighting probability is determined by certain characteristics, such as the group size which is assumed to be recorded without error, during the sampling (or sighting) process, it is natural to consider using logistic regression to fit the sightability model, where the responses are a set of Bernoulli observations. When sighting probabilities are unknown and are estimated using logistic regression, the unbiasness of the modified Horvitz-Thompson population size estimator and the variance of this estimator can be justified based on the asymptotic normality of the maximum likelihood estimators of the parameters of the logistic model (Steinhorst and Samuel, 1989).

For the case of known sighting probabilities, Samuel and Steinhorst (1992) derived an unbiased estimator for the variance of the population size estimator. Steinhorst and Samuel (1992) have also presented a modified Horvitz-Thompson ratio estimator for the population age or sex ratios. For the case of unknown sighting probabilities, using the asymptotic normality of maximum likelihood estimators of the parameters of the sightability model, Steinhorst and Samuel (1989) presented an estimator, not unbiased, for the variance of the population size estimator.

Thompson and Seber (1994) extended Steinhorst and Samuel's work by writing the modified Horvitz-Thompson population size estimator as a function of estimated sighting probabilities instead of reciprocals of estimated sighting probabilities. Their estimated sighting probability was not explicitly stated. Using the delta method, they derived an approximate third error component, i.e. error due to estimation of sighting probability, of the variance of the population size estimator, Like Steinhorst and Samuel, they did not give an unbiased estimator for each error component or for their approximate variance.

1.4 An outline of our study

Our objective is to provide an appropriate estimator for population size as well as the variance of this population size estimator, where a two-stage sampling survey is employed. The examples we will be looking at and our simulation studies in the later chapters will be based on data obtained from elk (cervus elaphus) populations. The procedure we developed, however, is not restricted to wildlife populations. It could be used for other types of populations where it is appropriate (see assumptions listed in section 2.1).

Our work is an extension of Steinhorst and Samuel's modified Horvitz-Thompson estimator for population size. In chapter two, we present Steinhorst and Samuel's population size estimator for cases of known sighting probabilities and unknown sighting probabilities estimated via a logistic model (see notation in (2.1.1) and the definition in (2.1.3)). We also present an unbiased estimator, derived by Steinhorst and Samuel (1992), for the variance of the population size estimator for the case of known sighting probabilities.

As an extension to their work, we will develop an asymptotically unbiased estimator for the variance of the population size estimator for the estimated sighting probabilities case (see section 2.2). With the assumption that we have a perfect count of detected groups, our derivation is a variation of Thompson and Seber's (1994) work and was obtained before their work appeared in print. Without the delta method, the variance of the population size estimator will be written explicitly as a sum of three error components, namely, the error due to sampling of primary units, the error due to sampling of groups within the selected primary units and the error due to estimating sighting probabilities. New results are given in the form of an asymptotically unbiased estimator for each of the three error components. We will explain the difference between Thompson and Seber (1994) and our methods in more detail in section 2.2. In section 2.3, we will begin with a brief description of an example of real field work, explaining how data are collected from a finite wildlife population and how estimators presented in sections 2.1 and 2.2 are applied to the collected data to estimate population size. We will then briefly discuss some problems of aerial surveys and based on the past experience of some elk sightability surveys (Samuel et. al.,1987), we will look at a couple of influential characteristics of groups in fitting a sightability model. In the last section of chapter 2, section 2.4, we will extend Steinhorst and Samuel's (1992) estimator for the variance of the ratio of two modified Horvitz-Thompson population size estimators of known sighting probabilities to the case of estimated sighting probabilities.

In chapter three, under the same assumptions, we look at the modified Horvitz-Thompson population size estimator for three special cases. In section 3.2 of the special case of a complete census of primary units, using simplified notation, we present simplified formulae of the population size estimator and of an asymptotically unbiased estimator of the variance of this population size estimator as a sum of unbiased estimators of the variance components. In section 3.3 of the special case of stratified random sampling of primary units, with an extended set of notation, we give explicit expressions for the population size estimator and an asymptotically unbiased estimator for the variance of this population size estimator, also as a sum of unbiased estimators of the variance components. Optimum allocation of stratum sample sizes for stratified random sampling of primary units is given in subsection 3.3.3. In section 3.4 of the third special case, we present the modified Horvitz-Thompson population size estimator based on the composite data of several replicated population surveys for both simple random sampling and stratified random sampling of primary units. Variance of the population size estimator is again stated as a sum of variance components and an unbiased estimator of each variance component is derived. Optimum allocation of stratum sample sizes in stratified random sampling of primary units for the third special case is given in section 3.4.3. Steinhorst and Samuel (1989) did not discuss the third special case and thus, this is also an extension of their work. In many studies of wildlife populations, biologists have conducted several replicated surveys over the years. If it is reasonable to assume that the sightability model did not change over a period of time, the data collected from several sightability surveys conducted over that period of time can be combined into one composite data set, giving more information. As for combining several replicated population surveys, groups of animals might have restructured over that period of time; however, it should have no sigificant impact on the estimation of the population size. It is therefore useful to present expressions of the population size estimator and the variance of the population size estimator in detail.

In chapter four, we will use a conditional expectation technique to derive explicit expressions for the third central moment of the modified Horvitz-Thompson estimators for both cases of known and unknown sighting probabilities. Results, conclusions and discussions of an evaluation study of the population size estimator and its variance estimator for small populations are described in chapter five. Considering the case of a complete census of primary units and a sightability of three independent variables (see (5.2.1), (5.2.2)), we generated 4 simulated populations. Various cases of this simulated population are studied to evaluate the performances of the population size estimator and the variance estimator we derived. We use procedure LOGISTIC in SAS to fit models. Evaluation of the performance of the population size estimator is based on the bias of the average of a set of simulated population size estimates with respect to the true population size. Evaluation of performance of the variance estimator will be based on the bias of the average of a set of variance estimates with respect to, say the sampling variance of the corresponding set of population size estimates. Confidence interval estimates at 90 and 95 percent confidence levels for population size are constructed assuming the large sample theory (normality) for the sampling distribution of the population size estimates and for the sampling distribution of the transformed population size estimates. Since we observed in earlier trials that the sampling distribution of the population size estimates is positively skewed, we also considered confidence intervals based on four transformations, namely, the natural log, the reciprocal, square root of the reciprocal and power of three halves of the reciprocal. Approximate expectations and the variances of these transformed population size estimators were derived using the delta method.

We also constructed confidence intervals by assuming that the sampling distribution of $(\hat{\tau}-T)$, where $\hat{\tau}$ is a population size estimate and T is the total number of animals or subjects in a phase 2 sample (see section 2.3), is log-normally distributed. In this case, exact expressions for the expectation and the variance of $log(\hat{\tau}-T)$ can be worked out and a Taylor series approximation is not needed to obtain a confidence interval estimate for population size τ.

In chapter six, confidence interval estimates for the population size using bootstrapping are evaluated. We are unaware of any bootstrap method which has been developed for finite population sampling with estimated sighting probabilities. We have therefore proposed modifications of the existing bootstrapping procedures to handle this problem. To study these procedures, we first considered the case of a complete census of primary units and generated a simulated population. This simulated population has a sightability model involving just one independent variable (see (6.4.1)). We also extend our bootstrapping procedures to the case of stratified random sampling of primary units.

Finally, in chapter seven, we will address the problem of (logistic) model selection for the unknown sighting probabilities. Our purpose in this chapter, like all other model selection problems, is to choose an appropriate parsimonious subset of the full model. For simulation, we will use the simulated population we generated in chapter five. We will fit the full model and a couple of subsets of the full model, which we called the short models, to the simulated samples. To check and compare the adequacy of these selected models, we use some conventional measures, namely the Akaike Information Criteria or AIC and the test of the hypothesis that a parameter is set to zero. In addition to those conventional measures, we also evaluate the performances of the population size estimator and its variance estimator the same way we did in chapter five. For this purpose, we have derived expressions to obtain, iteratively, the expected value of estimated parameters of a fitted model simpler than the full model, and an expression for the estimated covariance matrix of the estimated parameters of a fitted short model. We also study two-way scatter frequencies plots of influential predictors.

In the final chapter, chapter 8, we end this report with a summary and conclusion of our work, and some suggested topics for future research.

2. MODIFIED HORVITZ-THOMPSON ESTIMATOR OF POPULATION SIZE

In the first section, we present the modified Horvitz-Thompson population size estimator and estimator of the variance of the population size estimator developed by Steinhorst and Samuel (1989). In the second section, we will provide an improved estimator for the variance of the population size estimator with unknown sighting probabilities estimated using logistic regression. We will then look at how these estimators can be applied to real data set in the third section. Finally, we will extend the results of Steinhorst and Samuel (1992) on estimating population age and sex ratio by providing a ratio estimator and an unbiased estimator of the approximate variance of the ratio estimator when unknown sighting probabilities are estimated via logistic regression.

2.1 Development of the population size estimator

Consider the problem of estimating the size of a wildlife population. An unbiased, modified Horvitz-Thompson estimator for the population size was developed by Steinhorst and Samuel (1989) based on the following five assumptions:

(1) the population is geographically and demographically closed,

(2) groups of animals are conditionally independently observed, given the selection of primary sampling units,

(3) observed groups are completely enumerated without error and observed only once,

(4) the survey design for primary units can be specified,

(5) the probability of observing a group is known or can be estimated with a data set, assuming some model.

The notation given in Thompson (1992) is used in our following presentation. Let

N = the number of primary units in the population,

n = the number of primary units selected,

M_i = the number of groups in the i^{th} primary unit, $i = 1, 2, ..., N$,

m_i = the number of groups observed or sighted in the i^{th} selected primary unit, where $i = 1, 2, ..., n$,

y_{ij} = the number of animals in the j^{th} group in the ith primary unit, where $j = 1, 2, ..., M_i$,

τ_i = the number of animals in the ith primary unit, $\tau_i = \sum_{j=1}^{M_i} y_{ij}$,

τ = the population size, $\tau = \sum_{i=1}^{N} \sum_{i=1}^{M_i} y_{ij}$,

I_i = 1 if the i^{th} primary unit is selected and 0 if the i^{th} primary unit is not selected, thus I_i is an indicator random variable for sampling of the i^{th} primary unit,

$\pi_i = P(I_i = 1)$ is the probability that the i^{th} primary unit is selected,

$\pi_{ii'} = P(I_i = 1, I_{i'} = 1)$ is the probability that the i^{th} and the i'th primary units are both selected,

Z_{ij} = 1 if, in the i^{th} primary unit, the j^{th} group is sighted and 0 if the j^{th} group is not observed, thus Z_{ij} is an indicator random variable as to the sampling of the j^{th} group in selected primary unit i, and

$g_{ij} = P(Z_{ij} = 1 | I_i = 1)$, is the probability that the j^{th} group in the i^{th} selected primary unit is observed or sighted; thus, g_{ij} is the response probability or visibility or sightability or sighting probability of that group,

$\Theta_{ij} = 1/g_{ij}$.

(2.1.1)

2.1.1 When sighting probabilities are known

Using the modified Horvitz-Thompson approach, and assuming g_{ij}'s are known, an unbiased population size estimator is

$$\hat{\tau}_\pi = \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} y_{ij} \Theta_{ij} = \sum_{i=1}^{n} \frac{1}{\pi_i} \sum_{j=1}^{m_i} y_{ij} \Theta_{ij}. \tag{2.1.2}$$

2.1.2 When sighting probabilities are unknown

This section summarizes results given in Steinhorst and Samuel (1989) when the g_{ij}'s are unknown. A logistic model for the g_{ij} is used,

$$\frac{g_{ij}}{1 - g_{ij}} = e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}}, \tag{2.1.3}$$

where $\boldsymbol{x}'_{ij}\boldsymbol{\beta} = \beta_0 + \beta_1 x_{1,ij} + ... + \beta_{p-1} x_{p-1,ij}$ is a suitable regression function, with those $x_{k,ij}$ for $k = 1, 2, ..., p-1$, being the covariates such as group size, vegetation cover, snow cover, etc. From (2.1.1),

$$\Theta_{ij} = \frac{1}{g_{ij}} = 1 + e^{-\boldsymbol{x}'_{ij}\boldsymbol{\beta}}.$$

Define

$$\widetilde{\Theta}_{ij} = 1 + e^{-\boldsymbol{x}'_{ij}\widehat{\boldsymbol{\beta}} - \boldsymbol{x}'_{ij}\Sigma \boldsymbol{x}_{ij}/2},$$

where $\widehat{\boldsymbol{\beta}}$ is an estimator of $\boldsymbol{\beta}$. Assume that the distribution of $\widehat{\boldsymbol{\beta}}$ is the p-dimensional normal distribution $N_p(\boldsymbol{\beta}, \Sigma)$ with Σ being the p x p covariance matrix for $\widehat{\boldsymbol{\beta}}$. Then $E(\widetilde{\Theta}_{ij}) = \Theta_{ij}$. Thus, $\widetilde{\Theta}$ is asymptotically unbiased if $\widehat{\boldsymbol{\beta}}$ is asymptotically (multivariate) normally distributed.

For the subsequent results, it is assumed that $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \Sigma)$. Then the variance of $\widetilde{\Theta}_{ij}$ is

$$\text{var}(\widetilde{\Theta}_{ij}) = e^{-2\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij}\Sigma \boldsymbol{x}_{ij}} - 1),$$

and the covariance between two $\widetilde{\Theta}$'s is

$$\text{cov}(\widetilde{\Theta}_1, \widetilde{\Theta}_2) = e^{-(\boldsymbol{x}_1+\boldsymbol{x}_2)'\boldsymbol{\beta}}(e^{\boldsymbol{x}_1'\Sigma\boldsymbol{x}_2} - 1),$$

where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are vectors of covariates of two different groups whose sighting probabilities are Θ_1 and Θ_2. We also assume that all covariates are recorded without error. If in addition, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\Sigma}$ are independent and $\widehat{\Sigma} \approx \Sigma$ for large samples, the following are (asymptotically) unbiased estimators of the preceding quantities,

$$\widehat{\Theta}_{ij} = 1 + e^{-\boldsymbol{x}_{ij}'\widehat{\boldsymbol{\beta}} - \boldsymbol{x}_{ij}'\widehat{\Sigma}\boldsymbol{x}_{ij}/2};$$

$$\widehat{\text{var}}(\widehat{\Theta}_{ij}) = e^{-2\boldsymbol{x}_{ij}'\widehat{\boldsymbol{\beta}} - 2\boldsymbol{x}_{ik}'\widehat{\Sigma}\boldsymbol{x}_{ik}}(e^{\boldsymbol{x}_{ik}'\widehat{\Sigma}\boldsymbol{x}_{ik}} - 1);$$

$$\widehat{\text{cov}}(\widehat{\Theta}_1, \widehat{\Theta}_2) = e^{-(\boldsymbol{x}_1+\boldsymbol{x}_2)'\widehat{\boldsymbol{\beta}} - (\boldsymbol{x}_1+\boldsymbol{x}_2)'\widehat{\Sigma}(\boldsymbol{x}_1+\boldsymbol{x}_2)/2}(e^{\boldsymbol{x}_1'\widehat{\Sigma}\boldsymbol{x}_2} - 1).$$

With the assumption that $\widehat{\boldsymbol{\beta}}$, I_i and $Z_{ij}, i = 1, 2, ..., N, j = 1, 2, ..., M_i$, are independent, an (asymptotically) unbiased population size estimator is given by

$$\hat{\tau}_{LR} = \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} y_{ij} \widehat{\Theta}_{ij} = \sum_{i=1}^{n} \frac{1}{\pi_i} \sum_{j=1}^{m_i} y_{ij} \widehat{\Theta}_{ij}, \qquad (2.1.4)$$

where the subscript LR stands for logistic regression.

2.2 Development of an unbiased estimator for the variance of the population size estimator

We will now proceed to use conditional expectation techniques to write the variance of the population size estimator explicitly as a sum of error components.

2.2.1 When sighting probabilities are known

The variance of $\hat{\tau}_\pi$ of (2.1.2) when sighting probabilities are known can be written as a sum of two error components,

$$\text{var}(\hat{\tau}_\pi) = \text{var}_I(E_{Z|I}(\hat{\tau}_\pi)) + E_I(\text{var}_{Z|I}(\hat{\tau}_\pi)), \qquad (2.2.1)$$

where

$$v_1 = \text{var}_I(E_{Z|I}(\hat{\tau}_\pi)) = \text{var}(\sum_{i=1}^{N} \frac{I_i}{\pi_i}\tau_i)$$
$$= \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i}\tau_i^2 + \sum_{i\neq i'}^{N} \frac{\pi_{ii'} - \pi_i\pi_{i'}}{\pi_i\pi_{i'}}\tau_i\tau_{i'},$$

is the first error component, the error due to sampling of primary units, and

$$v_2 = E_I(\text{var}_{Z|I}(\hat{\tau}_\pi)) = E(\sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j=1}^{M_i} Z_{ij}\frac{y_{ij}^2}{g_{ij}^2}) = \sum_{i=1}^{N} \frac{1}{\pi_i} \sum_{j=1}^{M_i} \frac{1-g_{ij}}{g_{ij}} y_{ij}^2,$$

is the second error component, the error due to the sampling of groups within a primary unit such that $\text{var}(\hat{\tau}_\pi) = v_1 + v_2$. An unbiased estimator of the first error component v_1 is

$$\hat{v}_1 = \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i\hat{\tau}_i^2 + \sum_{i\neq i'}^{N} \frac{\pi_{ii'} - \pi_i\pi_{i'}}{\pi_{ii'}\pi_i\pi_{i'}} I_iI_{i'}\hat{\tau}_i\hat{\tau}_{i'} - \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \sum_{j=1}^{M_i} \frac{1-g_{ij}}{g_{ij}^2} Z_{ij}y_{ij}^2,$$

and an unbiased estimator for the second error component v_2 is

$$\hat{v}_2 = \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j=1}^{M_i} \frac{1-g_{ij}}{g_{ij}^2} Z_{ij}y_{ij}^2.$$

Then, an unbiased estimator for the variance (Samuel et. al., 1992) of the population size estimator is

$$\widehat{\text{var}}(\hat{\tau}_\pi) = \hat{v}_1 + \hat{v}_2$$
$$= \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i\hat{\tau}_i^2 + \sum_{i\neq i'}^{N} \frac{\pi_{ii'} - \pi_i\pi_{i'}}{\pi_{ii'}\pi_i\pi_{i'}} I_iI_{i'}\hat{\tau}_i\hat{\tau}_{i'} + \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} \frac{1-g_{ij}}{g_{ij}^2} Z_{ij}y_{ij}^2, \qquad (2.2.2)$$

where $\hat{\tau}_i = \sum_{j=1}^{M_i} \frac{Z_{ij}y_{ij}}{g_{ij}}$.

2.2.2 When sighting probabilities are unknown

When the sighting probabilities are unknown and are assumed to satisfy a logistic model, the variance of the population size estimator of (2.1.4) can be expressed as

$$\text{var}(\hat{\tau}_{LR}) = \text{var}(\hat{\tau}_\pi) + E_{I,Z}\text{var}(\hat{\tau}_{LR}|\mathbf{I}, \mathbf{Z}),$$

where $\hat{\tau}_\pi$ denotes the population estimator with sighting probabilities g_{ij} known. Steinhorst and Samuel (1989) did not give an unbiased estimator for each of the the variance components. An estimator for the variance, not unbiased but consisting of components related to the three variance components, was given by Steinhorst and Samuel (1989) as

$$\widehat{\text{var}}(\hat{\tau}_{LR}) = \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \hat{\tau}_i^2 + \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_{ii'} \pi_i \pi_{i'}} I_i I_{i'} \hat{\tau}_i \hat{\tau}_{i'}$$

$$+ \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j=1}^{M_i} (1 - \frac{1}{\widehat{\Theta}_{ij}})(y_{ij} \widehat{\Theta}_{ij})^2 + \sum_{i=1}^{N} \sum_{i'=1}^{N} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} y_{ij} y_{i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}).$$

As as extension to their work, we will express the variance explicitly as a sum of three error components and find an (asymptotically) unbiased estimator for each variance component. This is the major contribution of our thesis. Using conditional expectation techniques, we have

$$\begin{aligned}
\text{var}(\hat{\tau}_{LR}) &= \text{var}_I(E_{Z,\widehat{\Theta}|I}(\hat{\tau}_{LR})) + E_I(\text{var}_{Z,\widehat{\Theta}|I}(\hat{\tau}_{LR})) \\
&= \text{var}_I(E_{Z|I}(E_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))) + E_I[\text{var}_{Z|I}(E_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR})) + E_{Z|I}(\text{var}_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))] \\
&= \text{var}_I(\sum_{i=1}^{N} \frac{I_i \tau_i}{\pi_i}) + \sum_{i=1}^{N} \frac{1}{\pi_i} \text{var}_{Z|I}(\sum_{j=1}^{M_i} Z_{ij} y_{ij} \Theta_{ij}) + E_I(E_{Z|I}(\text{var}_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))) \\
&= v_1 + v_2 + v_3, \qquad (2.2.3)
\end{aligned}$$

where those three variance components, starting from the left, correspond to errors due to primary units sampling, sampling of groups within primary units, and estimation of g_{ij}. We will develop unbiased estimators for those three components.

We see that the first component is

$$v_1 = \text{var}_I(\sum_{i=1}^{N} \frac{I_i \tau_i}{\pi_i}) = \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i} \tau_i^2 + \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_i \pi_{i'}} \tau_i \tau_{i'}.$$

To develop an unbiased estimator for this component, consider

$$\sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \hat{\tau}_i^2 + \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_{ii'} \pi_i \pi_{i'}} I_i I_{i'} \hat{\tau}_i \hat{\tau}_{i'}, \qquad (2.2.4)$$

where $\hat{\tau}_i = \sum_{j=1}^{M_i} Z_{ij} y_{ij} \widehat{\Theta}_{ij}$. The expectation of expression (2.2.4) is

$$E[(2.2.4)] = \text{var}(\sum_{i=1}^{N} \frac{I_i \tau_i}{\pi_i}) + \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i} \text{var}(\hat{\tau}_i)$$

$$+ \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_i \pi_{i'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} g_{ij} g_{i'j'} y_{ij} y_{i'j'} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'})$$

$$= \text{var}(\sum_{i=1}^{N} \frac{I_i \tau_i}{\pi_i}) + \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i} \text{var}(\hat{\tau}_i) + \text{cov}(\hat{\tau}_i, \hat{\tau}_{i'}).$$

We therefore need to develop unbiased estimators for $\text{var}(\hat{\tau}_i)$ and $\text{cov}(\hat{\tau}_i, \hat{\tau}_{i'})$ so that unbiased estimators of their corresponding terms can be subtracted from expression (2.2.4). We note that

$$\text{var}(\hat{\tau}_i) = \text{var}_Z(E_{\widehat{\Theta}|Z}(\hat{\tau}_i)) + E_Z(\text{var}_{\widehat{\Theta}|Z}(\hat{\tau}_i)),$$

where

$$\text{var}_Z(E_{\widehat{\Theta}|Z}(\hat{\tau}_i)) = \text{var}(\sum_{j=1}^{M_i} Z_{ij} y_{ij} \Theta_{ij}) = \sum_{j=1}^{M_i} y_{ij}^2 (\Theta_{ij} - 1), \tag{2.2.5}$$

and

$$E_Z(\text{var}_{\widehat{\Theta}|Z}(\hat{\tau}_i)) = E_Z[\sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 \text{var}(\widehat{\Theta}_{ij}) + \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \text{cov}(\widehat{\Theta}_{ij} \widehat{\Theta}_{ij'})]. \tag{2.2.6}$$

Thus

$$\widehat{\text{var}}_Z(\sum_{j=1}^{M_i} Z_{ij} y_{ij} \Theta_{ij}) = \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 [\widehat{\Theta}_{ij}^2 - \widehat{\Theta}_{ij} - \widehat{\text{var}}(\widehat{\Theta}_{ij})],$$

provides an unbiased estimator for expression (2.2.5). An unbiased estimator for expression (2.2.6) is given as

$$\widehat{E}_Z(\text{var}_{\widehat{\Theta}|Z}(\hat{\tau}_i)) = \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 \widehat{\text{var}}(\widehat{\Theta}_{ij}) + \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'}).$$

Addition of the two unbiased variance component estimators gives $\widehat{\text{var}}(\hat{\tau}_i)$ where

$$\widehat{\text{var}}(\hat{\tau}_i) = \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 (\widehat{\Theta}_{ij}^2 - \widehat{\Theta}_{ij}) + \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'}),$$

is an unbiased estimator of $\text{var}(\hat{\tau}_i)$. Now,

$$\text{cov}(\hat{\tau}_i, \hat{\tau}_{i'}) = \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} g_{ij} g_{i'j'} y_{ij} y_{i'j'} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}),$$

and

$$\widehat{\text{cov}}(\hat{\tau}_i, \hat{\tau}_{i'}) = \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} y_{ij} y_{i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'})$$

is a corresponding unbiased estimator. Hence,

$$\hat{v}_1 = \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \hat{\tau}_i^2 + \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_{ii'} \pi_i \pi_{i'}} I_i I_{i'} \hat{\tau}_i \hat{\tau}_{i'} - \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 (\widehat{\Theta}_{ij}^2 - \widehat{\Theta}_{ij})$$

$$- \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i^2} I_i \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'})$$

$$- \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_{ii'} \pi_i \pi_{i'}} I_i I_{i'} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} y_{ij} y_{i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}), \quad (2.2.7)$$

is an unbiased estimator for the first component v_1 of expression (2.2.3).

The second component v_2 of expression (2.2.3) is

$$v_2 = E_I(\text{var}_{Z|I}(E_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))) = \sum_{i=1}^{N} \frac{1}{\pi_i} \text{var}(\sum_{j=1}^{M_i} Z_{ij} y_{ij} \Theta_{ij}) = \sum_{i=1}^{N} \frac{1}{\pi_i} \sum_{j=1}^{M_i} y_{ij}^2 (\Theta_{ij} - 1). \quad (2.2.8)$$

An unbiased estimator of this expression is

$$\hat{v}_2 = \widehat{E}_I(\text{var}_{Z|I}(E_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))) = \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 [\widehat{\Theta}_{ij}^2 - \widehat{\Theta}_{ij} - \widehat{\text{var}}(\widehat{\Theta}_{ij})]. \quad (2.2.9)$$

As for the third component, we have

$$v_3 = E_I(E_{Z|I}(\text{var}_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))) = E_{I,Z}(\text{var}_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR}))$$

$$= E_{I,Z}[\sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 \text{var}(\widehat{\Theta}_{ij}) + \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'})$$

$$+ \sum_{i \neq i'}^{N} \frac{I_i I_{i'}}{\pi_i \pi_{i'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} y_{ij} y_{i'j'} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'})] \quad (2.2.10)$$

$$= \sum_{i=1}^{N} \frac{1}{\pi_i} \sum_{j=1}^{M_i} \frac{y_{ij}^2}{\Theta_{ij}} \text{var}(\widehat{\Theta}_{ij}) + \sum_{i=1}^{N} \frac{1}{\pi_i} \sum_{j \neq j'}^{M_i} \frac{y_{ij} y_{ij'}}{\Theta_{ij} \Theta_{ij'}} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'})$$

$$+ \sum_{i \neq i'}^{N} \frac{\pi_{ii'}}{\pi_i \pi_{i'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} \frac{y_{ij} y_{i'j'}}{\Theta_{ij} \Theta_{i'j'}} \text{cov}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}), \tag{2.2.11}$$

An unbiased estimator is found by replacing variance and covariances terms for the $\widehat{\Theta}_{ij}$ by their unbiased estimators and dropping the expectation operator in expression (2.2.10).

Thus

$$\hat{v}_3 = \widehat{E}_{I,Z}(\text{var}_{\widehat{\Theta}|I,Z}(\hat{\tau}_{LR})) = \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j}^{M_i} Z_{ij} y_{ij}^2 \widehat{\text{var}}(\widehat{\Theta}_{ij})$$

$$+ \sum_{i=1}^{N} \frac{I_i}{\pi_i^2} \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'})$$

$$+ \sum_{i \neq i'}^{N} \frac{I_i I_{i'}}{\pi_i \pi_{i'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} y_{ij} y_{i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'})]. \tag{2.2.12}$$

Finally, we combine all three unbiased variance component estimators in expressions (2.2.7), (2.2.9), (2.2.12) to obtain

$$\widehat{\text{var}}(\hat{\tau}_{LR}) = \sum_{i=1}^{N} \frac{1 - \pi_i}{\pi_i^2} I_i \hat{\tau}_i^2 + \sum_{i \neq i'}^{N} \frac{\pi_{ii'} - \pi_i \pi_{i'}}{\pi_{ii'} \pi_i \pi_{i'}} I_i I_{i'} \hat{\tau}_i \hat{\tau}_{i'}$$

$$+ \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 (\widehat{\Theta}_{ij}^2 - \widehat{\Theta}_{ij})$$

$$+ [\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{jj'}^{M_i} Z_{ij} Z_{ij'} y_{ij} y_{ij'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'})$$

$$+ \sum_{i \neq i'}^{N} \frac{I_i I_{i'}}{\pi_{ii'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} y_{ij} y_{i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'})], \tag{2.2.13}$$

an unbiased estimator for $\text{var}(\hat{\tau}_{LR})$.

Thompson and Seber (1994) wrote $\hat{\tau}_{LR}$ as a function of $\hat{g}_{ij}$ instead of its reciprocal, $\widehat{\Theta}_{ij}$. They did not assume any particular model for estimating the unknown

g_{ij} and just let $\hat{g}_{ij}$ be some estimate of g_{ij}, without stating it explicitly. By applying the delta method, they have obtained an approximate third error component for the variance of the population size estimator in terms of $\hat{g}_{ij}$. They presented an estimator, not unbiased, for the approximate variance and like Steinhorst and Samuel (1989), they did not give an unbiased estimator for each of the three error components.

We, on the other hand, saw that in our logistic regression setting, it is natural to write $\hat{\tau}_{LR}$ as a function of $\widehat{\Theta}_{ij}$. In this case, the delta method was not needed for our derivation. Unlike Thompson and Seber (1994), we presented the variance of the population size estimator as a function of $\widehat{\Theta}_{ij}$. We then derived an unbiased estimator for each of the three variance components. By taking the sum of unbiased estimators of the variance components, we gave an unbiased estimator for the variance of the population size estimator.

2.3 Estimation of the population size of a finite population

In this section, we will begin with a brief description of a field example, explaining how data are collected from a finite, wildlife population and how estimators presented in sections 2.1 and 2.2 are applied to the collected data to estimate population size. We will then discuss some problems of aerial surveys. Then, based on the past experience of some elk sightability surveys (Samuel et. al., 1987), we will look at a couple of influential characteristics of groups in fitting logistic sightability models.

2.3.1 Example of how real data are collected and how population size is estimated using modified Horvitz-Thompson estimator

Two techniques have been used by researchers to collect data to estimate elk population size. The first technique breaks both survey and estimation process into two phases. The sightability survey, or phase I, is to collect data for estimating the sightability model. In phase I, some elk in the study area were fitted with radio collars. A fixed-wing flight is conducted to locate radio-collared animals. A helicopter survey of the primary units containing these animals is then conducted. Helicopter observers are informed of the (primary) units containing specified radio-collared animals and their radio frequencies. However, they do not know the location of collared animals within the unit. Typically a single radio-collared animal might be searched for in 2 units or 2 animals might be searched for in 2-3 units. The helicopter observers survey the required units and record group size, vegetation cover, etc. for each group seen that contains a radio-collared animal. Upon completion of the search, radio-collared animals not seen are located using a receiver in the helicopter. Group size, vegetation cover, etc. for each group that contained a missed radio-collared animal are recorded. This entire process, from radio-collaring of the animals to recording data of missed animals, is called the sightability survey or referred to as phase I. The data collected, of seen and missed animals, will be referred to as a phase I sample. A sightability model is fitted based on a phase I data. Researchers can enlarge the size of a phase I sample by conducting a sightability survey several times over a certain period, for example three years, possibly with different collared animals at different times.

After an estimated sightability model is obtained, a population survey or phase II is conducted to collect data for estimating the population size. In phase II, data are collected using a two-stage sample design as we described in the first paragraph of section 1.2. This data set, or the phase II sample, is independent of the phase I sample. Similarly, researchers can enlarge the size of a phase II sample by conducting a population survey several times, also over a certain period. Based on the phase II sample and the estimated sightability model obtained in phase I, a population size estimate is computed using (2.1.4). A single population size estimate obtained using the first technique requires 2 independent samples, namely, phase I and phase II samples.

Note that for the first technique, a sightability survey and a population survey are conducted independently at different times. For example, a population survey may be conducted a few days or months after conducting the sightability survey. So it is likely that the phase II sample is collected under some underlying conditions different from those related to estimating the sightability model. Researchers will have to assume that both the sightability survey and population survey are conducted under the same underlying conditions.

An alternative technique overcomes this problem by conducting both sightability survey and population survey simultaneously. For this technique, some animals are again assumed to be selected from the study area and radio-collared. Then the sample frame of primary units comprising the survey area is constructed. A helicopter survey is conducted only on the sample of selected primary units. Regardless of whether there is a collared animal contained in an observed group, group size, vegetation cover, etc. are recorded for the group. Upon completion of this survey, the helicopter returns to the sample survey primary units to locate the groups containing radio-collared animals not observed, to record their group size, vegetation cover, etc. Therefore, using this technique, there is only one data set. From this data set, data of groups that contained a collared animal are used to fit a sightability model. Based on this estimated sightability model and the entire data set, a population size estimate is obtained using (2.1.4). If the study area is censused, the population size of groups with collars is known. We would only estimate the number of elk in groups with no collars.

One problem with this alternative technique is that there may be an insufficient number of groups containing a collared animal in the sample of primary units. Thus sightability data sets would need to be collected over several surveys before a population size estimate could be constructed.

2.3.2 Some problems of aerial surveys

Failure to observe all animals or "visibility bias" is generally a major cause of underestimating the population size (Caughley, 1974, 1977). Samuel et. al., (1987) has discussed some problems of aerial surveys that might influence the visibility. Samuel et. al. (1987) also suggested some ways to correct the visibility bias. The following paragraphs are summarized from Samuel et. al. (1987), describing some problems of aerial surveys and some suggestions to correct the visibility bias:

> "Aerial surveys are an important method for estimating populations. However, aerial surveys underestimate animal abundance (Caughley 1977). A major goal for the improvement of aerial survey estimates is to determine the number of animals missed during surveys. Failure to observe all animals, as stated in section 1.2, is called visibility bias and is generally a major cause of inaccuracies in aerial surveys (Caughley, 1974, 1977). The magnitude of visibility bias depends on numerous factors, including habitat characteristics such as animal behavior and dispersion, weather (for example, migration, animals scatter more widely during light snow, etc.), vegetation cover, and the observers and equipments. Visibility bias also may confound the estimation of age and sex ratios when males, females, or young have different visibility factors.
>
> Recognition of and correction for visibility bias have been subjectively and quantitatively used to adjust wildlife population surveys. Graham and Bell (1969) contended that visibility of animals was influenced

by environmental conditions and by factors attributable to the observers. However, assessment has generally focussed on those factors that can be controlled during the survey: observer factors (Caughley 1974, LeResche and Rausch 1974) and aircraft factors (Caughley et. al. 1976). Even when these factors are rigorously standardized considerable flight-to-flight variability is possible. In part, this variability is due to environmental factors that cannot be controlled by the survey technique, including cover type (LeResche and Rausch 1974, Floyd et. al. 1979, Biggins and Jackson 1984, Gasaway et. al. 1985), animal group size (Cook and Martin 1974, Samuel and Pollock 1981, Gasaway et. al. 1985), animal behavior (Gasaway et. al. 1985), snow conditions (Lovaas et. al. 1966, LeResch and Rausch 1974, Biggins and Jackson 1984), and weather (Anderson 1958). In some recent aerial surveys of the elk population for example, there is considerable evidence that counts from the ground or from the air do not enumerate all the elk.

Increasing the intensity of search effort is commonly believed to increase the proportion of animals seen and may reduce the effects of some environmental factors such as weather, lighting and snow conditions. Thus, there appears to be a distinct advantage of using intensive helicopter quadrat surveys over transect surveys with fixed-wing aircraft (Gasaway et. al. 1985). Helicopters further enhance the observers' ability to minimize terrain and cover problems (Kufeld et al. 1980). In addition, the noise of the helicopter generally causes animals to move (perhaps increasing visibility) and may be used to flush animals from patches of cover (Kufeld el al. 1980). Helicopter also may reduce observer fatigue problems by minimizing disorientation and airsickness (Kufeld et. al. 1980) and by allowing more frequent breaks (for more frequent refuelling). There is little doubt that sightability will vary due to the experience levels of observers and pilots or to the type of aircraft used.

In some surveys, white radio collars were used to enhance the identification of marked animals from the helicopter once the animal is located and to avoid the possibility of visually locating an animal due to a bright-colored collar. In some situation, it was found (Gasaway et. al. 1985) that bedded radio-collared moose were more likely to missed during intensive fixed-wing surveys.

There is therefore a need to design an appropriate sampling strategy, such as the stratified random sampling, for conducting surveys. Applications of a sightability model in conjunction with a sampling approach should also consider the minimum sampling area or the number of groups required to produce reliable estimates. Of primary importance in any survey design will be a clear formulation of objectives, a careful choice of the required level of precision, and a knowledge of population distribution. For example, optimum counting conditions for elk are likely to occur during winter (or early spring) when animals occur in larger groups and in areas of open habitat."

2.3.3 About the sightability model

During the sighting process in second-stage sampling, a group is either sighted (responded) or not sighted (nonresponse) with some sighting probability. The responses are thus a set of Bernoulli observations. In addition, as observed in previous studies that visibility bias depends on numerous habitat characteristics of the groups as stated above, it is therefore natural to consider using logistic regression to fit sightability model.

Past experience (Samuel et. al., 1987) with elk populations has found that group size, vegetation cover and snow cover were primary factors influencing observability. The study also found that many factors influencing sightability are interrelated using a multivariate approach. Group size is generally believed to be an important factor influencing sightability. The inability of some studies to measure group size may have caused related factors to appear as important influences on sightability. In some cases, the natural log transformation of the group size might give a better fit. The importance of vegetation cover has been stressed by nearly every study that has attempted to examine visibility. Previous authors have suggested dividing a study area into discrete habitats or cover types and estimating visibility in each type (Floyd et. al. 1979). Such an approach might require substantially more data to estimate a separate sightability function for each habitat. In addition, problems may arise from assuming uniform vegetative cover within a habitat. Both of these problems are avoided by estimating percent vegetative cover directly. Further, taking this approach offers the possibility that sightability functions developed in one area may be applicable over broad areas. Further research in other habitats might broaden the applicability of the models or identify habitats that require a more complex approach. As for snow cover, there were controversial findings for different animal populations (Samuel et. al., 1987). Snow cover may interact with other factors such as group size and behavior. Further, the effect of snow cover may be reduced during the intensive quadrat surveys conducted. Another reason that the snow cover factor was not significant in some studies may be partially due to the limited range over which these factors were evaluated.

2.4 Ratio estimator

Samuel et. al. (1992) developed a modified Horvitz-Thompson estimator for a population age or sex ratio. Given the same assumptions as section 2.1, define

τ_a = population size of a type animals,

τ_b = population size of b type animals,

$\hat{\tau}_{a\pi}$ = population size estimator of a type animals when sighting probabilities are known,

$\hat{\tau}_{b\pi}$ = population size estimator of b type animals when sighting probabilities are known,

$\hat{\tau}_{aLR}$ = population size estimator of a type animals when sighting probabilities are unknown and are estimated via a logistic model,

$\hat{\tau}_{bLR}$ = population size estimator of b type animals when sighting probabilities are unknown and are estimated via a logistic model,

a_{ij} = the number of a type animals in the jth group of ith primary unit,

b_{ij} = the number of b type animals in the jth group of ith primary unit,

and with the notation defined in the last section, an estimator for the ratio $R = \tau_a/\tau_b$, using (2.1.2) when sighting probabilities g_{ij} are specified, is

$$\widehat{R}_\pi = \frac{\hat{\tau}_{a\pi}}{\hat{\tau}_{b\pi}} = \frac{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} \Theta_{ij}}{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} b_{ij} \Theta_{ij}}. \tag{2.4.1}$$

Using (2.1.4) when the $\Theta_{ij} = 1/g_{ij}$'s are unknown and are estimated via a logistic model, the ratio estimator is

$$\widehat{R}_{LR} = \frac{\hat{\tau}_{aLR}}{\hat{\tau}_{bLR}} = \frac{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} \widehat{\Theta}_{ij}}{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} b_{ij} \widehat{\Theta}_{ij}}. \tag{2.4.2}$$

By applying the delta method,

$$\frac{\hat{\tau}_a}{\hat{\tau}_b} \doteq \frac{\tau_a}{\tau_b} + \frac{1}{\tau_b}(\hat{\tau}_a - \tau_a) - \frac{1}{\tau_b{}^2}(\hat{\tau}_b - \tau_b). \tag{2.4.3}$$

In (2.4.3), $\hat{\tau}_a$, $\hat{\tau}_b$ are $\hat{\tau}_{a\pi}$, $\hat{\tau}_{b\pi}$ (see (2.1.2)) if sighting probabilities are known and they are $\hat{\tau}_{aLR}$, $\hat{\tau}_{bLR}$ (see (2.1.4)) if sighting probabilities are unknown. Thus, we obtain an approximate variance of a ratio estimator as

$$\mathrm{var}(\widehat{R}) \doteq R^2[\frac{\mathrm{var}(\hat{\tau}_a)}{{\tau_a}^2} + \frac{\mathrm{var}(\hat{\tau}_b)}{{\tau_b}^2} - \frac{2\mathrm{cov}(\hat{\tau}_a, \hat{\tau}_b)}{\tau_a \tau_b}]. \tag{2.4.4}$$

To obtain an (approximate) estimator of $\mathrm{var}(\widehat{R})$ in (2.4.4), we replace the ratio, the population sizes, variances and covariances of the population size estimator by their respective unbiased estimators. We have already derived (approximate and/or asymptotic) unbiased estimators for the ratio, the population sizes (see (2.1.2) and (2.1.4)) and variances of the population size estimator (see (2.2.2) and (2.2.13)). Therefore, to extend the results of Samuel et. al. (1992), we only need to find an unbiased estimator for the covariance term. To do so, we write

$$\mathrm{cov}(\hat{\tau}_a, \hat{\tau}_b) = E[\hat{\tau}_a \hat{\tau}_b] - E[\hat{\tau}_a]E[\hat{\tau}_b] = E[\hat{\tau}_a \hat{\tau}_b] - \tau_a \tau_b, \tag{2.4.5}$$

Then, the expression $\hat{\tau}_a \hat{\tau}_b - \hat{\tau}_{ab}$, where $\hat{\tau}_{ab}$ is an unbiased estimator of $\tau_a \tau_b$, is an unbiased estimator of the covariance term in (2.4.5). To find an unbiased estimator of $\tau_a \tau_b$, we note that

$$\tau_a \tau_b = (\sum_{i=1}^{N}\sum_{j=1}^{M_i} a_{ij})(\sum_{i=1}^{N}\sum_{j=1}^{M_i} b_{ij}) = \sum_{i=1}^{N}\sum_{j=1}^{M_i} a_{ij}b_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}} a_{ij}b_{ij'} + \sum_{i\neq i'}^{N}\sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}} a_{ij}b_{i'j'}.$$

Hence, an unbiased estimator for the covariance term in (2.4.5) is

$$\widehat{\mathrm{cov}}(\hat{\tau}_a, \hat{\tau}_b) = \hat{\tau}_{a\pi}\hat{\tau}_{b\pi} + \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} \frac{Z_{ij}a_{ij}b_{ij}}{g_{ij}} + \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j\neq j'}^{M_i} \frac{Z_{ij}Z_{ij'}a_{ij}b_{ij'}}{g_{ij}g_{ij'}}$$

$$+ \sum_{i\neq i'}^{N} \frac{I_i I_{i'}}{\pi_{ii'}} \sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}} \frac{Z_{ij}Z_{i'j'}a_{ij}b_{i'j'}}{g_{ij}g_{i'j'}}, \tag{2.4.5}$$

for the case of known g_{ij}'s.

For the case of unknown Θ_{ij}'s estimated via a logistic model, an unbiased estimator for the covariance term in (2.4.5) is

$$\widehat{\text{cov}}(\hat{\tau}_a, \hat{\tau}_b) = \hat{\tau}_{aLR}\hat{\tau}_{bLR} + \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} b_{ij} \widehat{\Theta}_{ij}$$

$$+ \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} a_{ij} b_{ij'} [-\widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'}) + \widehat{\Theta}_{ij}\widehat{\Theta}_{ij'}]$$

$$+ \sum_{i \neq i'}^{N} \frac{I_i I_{i'}}{\pi_{ii'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} a_{ij} b_{i'j'} [-\widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}) + \widehat{\Theta}_{ij}\widehat{\Theta}_{i'j'}]. \qquad (2.4.6)$$

3. MODIFIED HORVITZ-THOMPSON ESTIMATOR OF THE POPULATION SIZE FOR THREE SPECIAL CASES

3.1 Introduction

In the last chapter, we presented formulae for estimating the population size when the sighting probabilities are known, $\hat{\tau}_\pi$ of (2.1.2), and when the unknown sighting probabilities are assumed to satisfy a logistic model defined in (2.1.3), $\hat{\tau}_{LR}$ of (2.1.4), for a general two-stage sampling design. We also derived explicit expressions (2.2.1) and (2.2.3) of the variances of these population size estimators in terms of error components. We developed (asymptotically) unbiased estimator for those error components separately and combined them to obtain (asymptotically) unbiased estimator of the variances (see (2.2.2) and (2.2.13)).

In this chapter, we look at three special cases - two special cases of the first stage sampling design and the special case of combining data of several replicated population surveys. Sections 3.2 and 3.3 give simplified versions of those formulae, under the same assumptions listed in section 2.1, for two particular first stage sampling designs, namely, the case of a complete census of primary units and the case of stratified random sampling of primary units. Using the technique of conditional expectations, we express the variance of the population size estimator in terms of error components and develop unbiased estimators for each error components. Optimum allocations for stratified random sampling of primary units is given in subsection 3.3.3.

In section 3.4, we look at the special case of obtaining a modified Horvitz-Thompson population size estimate based on the composite data of several replicated population surveys. This is important because researchers sometimes conduct several replicated population surveys over a period of time. For this special case, we consider simple random sampling and stratified random sampling of primary units. Optimum allocations for stratified random sampling of primary units is stated in subsection 3.4.3.

3.2 Case 1: Complete census of primary units

If all primary units are sampled, we see that the notation N, subscript i, variable I_i and, thus, the probabilities π_i, $\pi_{ii'}$ defined in (2.1.1) are no longer needed. The notation for this special case is:

M = the number of groups in the entire survey area,

y_j = the number of animals in the j^{th} group, $j = 1, 2, ..., M$,

τ = the population size $\sum_{j=1}^{M} y_j$,

$Z_j = 1$ if the j^{th} group is sighted and 0 if the j^{th} group is not observed, thus, Z_j is an indicator random variable as to the sampling of the j^{th} group, and

$g_j = P(Z_j = 1)$, is the response probability or visibility or sightability or sighting probability of the j^{th} group.

3.2.1 When sighting probabilities are known

When the g_j's are known, the unbiased estimator for the population size is given as

$$\hat{\tau}_\pi = \sum_{j=1}^{M} \frac{Z_j y_j}{g_j}. \tag{3.2.1}$$

The variance of $\hat{\tau}_\pi$ is

$$\text{var}(\hat{\tau}_\pi) = \sum_{j=1}^{M} \frac{1 - g_j}{g_j} y_j^2 \, , \tag{3.2.2}$$

which is the error due to sampling of groups (that is, only sighting some of the groups) in the second stage of sampling. An unbiased estimator for the variance is

$$\widehat{\text{var}}(\hat{\tau}_\pi) = \sum_{j=1}^{M} \frac{1-g_j}{g_j^2} Z_j y_j^2 \,. \tag{3.2.3}$$

3.2.2 When sighting probabilities are unknown

When the sighting probabilities are unknown and are assumed to satisfy a logistic model,

$$\frac{g_j}{1-g_j} = e^{\boldsymbol{x}_j'\boldsymbol{\beta}} \tag{3.2.4}$$

as in (2.1.3), where $\boldsymbol{x}_j'\boldsymbol{\beta} = \beta_0 + \beta_1 x_{1,j} + \ldots + \beta_{p-1}x_{p-1,j}$ is a suitable regression function, with $x_{k,j}$ for $k = 1, 2, \ldots, p-1$, being the covariates such as group size, vegetation cover, etc. Define

$$\Theta_j = \frac{1}{g_j} = 1 + e^{-\boldsymbol{x}_j'\boldsymbol{\beta}}$$

and

$$\widetilde{\Theta}_j = 1 + e^{-\boldsymbol{x}_j'\widehat{\boldsymbol{\beta}} - \boldsymbol{x}_j'\Sigma\boldsymbol{x}_j/2}$$

where $\widehat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\beta}$. Thus, $\widetilde{\Theta}_j$ is asymptotically unbiased. With the same assumption that $\widehat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \Sigma)$ as before, expressions for the variance of $\widetilde{\Theta}_j$ and covariance between two different $\widetilde{\Theta}$'s are similar with the subscript i removed (see section 2.1.2).

If, as before, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\Sigma}$ are independent and $\widehat{\Sigma} \approx \Sigma$ for large samples, expressions for an asymptotically unbiased estimator $\widehat{\Theta}_j$ of Θ_j, and asymptotically unbiased estimators of $\text{var}(\widehat{\Theta}_j)$ and $\text{cov}(\widehat{\Theta}_1, \widehat{\Theta}_2)$, the covariance between two different Θ's, are similar, with the subscript i removed (also see section 2.1.2). The unbiased estimator for the population size is then given as

$$\hat{\tau}_{LR} = \sum_{j=1}^{M} Z_j y_j \widehat{\Theta}_j, \tag{3.2.5}$$

where $\widehat{\Theta}_j$ is the unbiased estimator of Θ_j. The variance of $\hat{\tau}_{LR}$ can be written as the sum of two error components, namely, the variance due to sampling of groups, given by

$$e_1 = \text{var}_Z(E_{\widehat{\Theta}|Z}(\hat{\tau}_{LR})) = \text{var}_Z(\sum_{j=1}^{M} Z_j y_j \Theta_j) = \sum_{j=1}^{M} y_j^2(\Theta_j - 1), \tag{3.2.6}$$

which is (asymptotically) unbiasedly estimated by

$$\hat{e}_1 = \widehat{\text{var}}_Z(E_{\widehat{\Theta}|Z}(\hat{\tau}_{LR})) = \sum_{j=1}^{M} Z_j y_j^2[\widehat{\Theta}_j^2 - \widehat{\Theta}_j - \widehat{\text{var}}(\widehat{\Theta}_j)], \tag{3.2.7}$$

and the variance due to estimating Θ_j, given as

$$e_2 = E_Z(\text{var}_{\widehat{\Theta}|Z}(\hat{\tau}_{LR})) = \sum_{j=1}^{M} \frac{y_j^2}{\Theta_j}\text{var}(\widehat{\Theta}_j) + \sum_{j \neq j'}^{M} \frac{y_j y_{j'}}{\Theta_j \Theta_{j'}}\text{cov}(\widehat{\Theta}_j, \widehat{\Theta}_{j'}), \tag{3.2.8}$$

which is (asymptotically) unbiasedly estimated by

$$\hat{e}_2 = \widehat{E}_Z(\text{var}_{\widehat{\Theta}|Z}(\hat{\tau}_{LR})) = \sum_{j=1}^{M} Z_j y_j^2 \widehat{\text{var}}(\widehat{\Theta}_j) + \sum_{j \neq j'}^{M} Z_j Z_{j'} y_j y_{j'} \widehat{\text{cov}}(\widehat{\Theta}_j, \widehat{\Theta}_{j'}). \tag{3.2.9}$$

3.3 Case 2: Stratified random sampling of primary units

The area of interest is divided into primary sampling units (land areas). These primary units are then assigned uniquely to one of L strata. A stratified random sample of primary units is then selected. The following notation is used :

L = the number of strata in the population,

N = the number of primary units in the population,

N_h = the number of primary units in the h^{th} stratum, so that

$$N_1 + N_2 + ... + N_L = N,$$

n_h = the number of primary units sampled in the h^{th} stratum, $h = 1, 2, ..., L$,

M_{hi} = the number of groups in the i^{th} primary unit in the h^{th} stratum, $i = 1, 2, ..., N_h$,

m_{hi} = the number of groups observed or sighted in the i^{th} selected primary unit in the hth stratum, $i = 1, 2, ..., n_h$,

y_{hij} = the number of animals in the j^{th} group in the i^{th} primary unit in the h^{th} stratum, $j = 1, 2, ..., M_{hi}$,

τ_{hi} = the size number of animals in the i^{th} primary unit in the h^{th} stratum, $\tau_{hi} = \sum_{j=1}^{M_{hi}} y_{hij}$,

τ_h = the number of animals in the h^{th} stratum, $\tau_h = \sum_{i=1}^{N_h} \tau_{hi}$,

τ = the population size, $\tau = \sum_{h=1}^{L} \tau_h$,

I_{hi}, Z_{hij} are indicator random variables where

I_{hi} = 1 if, in the h^{th} stratum, the i^{th} primary unit is selected and 0 if the i^{th} primary unit is not selected,

Z_{hij} = 1 if, in the i^{th} primary unit in the h^{th} stratum, the j^{th} group is sighted and 0 if the j^{th} group is not observed,

such that

$\pi_{hi} = P(I_{hi} = 1) = \frac{n_h}{N_h}$, is the probability that the i^{th} primary unit in the h^{th} stratum is selected,

$\pi_{hii'} = P(I_{hi} = 1, I_{hi'} = 1) = \frac{n_h}{N_h}\frac{n_h - 1}{N_h - 1}$, is the probability that both i^{th} and i'th primary units in the h^{th} stratum are selected, and

$g_{hij} = P(Z_{hij} = 1 | I_{hi} = 1)$, is the probability that the j^{th} group in the i^{th} selected primary unit in the h^{th} stratum is observed or sighted. It is the visibility or sightability or response probability or sighting probability of that group.

3.3.1 When sighting probabilities are known

An unbiased estimator for the population size when g_{hij}'s are known is given as,

$$\hat{\tau}_\pi = \sum_{h=1}^{L} N_h \sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h} = \sum_{h=1}^{L} N_h \sum_{i=1}^{n_h} \frac{\sum_{j=1}^{m_{hi}} y_{hij} \Theta_{hij}}{n_h},$$

where $\hat{\tau}_{hi} = \sum_{j=1}^{M_{hi}} Z_{hij} y_{hij} / g_{hij}$ is an unbiased estimator of τ_{hi}. The variance of $\hat{\tau}_\pi$ is

$$\text{var}(\hat{\tau}_\pi) = \sum_{h=1}^{L} N_h^2 \text{var}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}).$$

Now

$$\begin{aligned}
\text{var}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}) &= \text{var}_I(E_{Z|I}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h})) + E_I(\text{var}_{Z|I}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h})), \\
&= \text{var}(\sum_{i=1}^{n_h} \frac{\tau_{hi}}{n_h}) + E_I((\text{var}_{Z|I}(\sum_{i=1}^{N_h} \frac{I_{hi}\hat{\tau}_{hi}}{n_h})) \\
&= (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h^2} \left[\sum_{i=1}^{N_h} E_I(I_{hi}^2) E_I(\text{var}_{Z|I}(\hat{\tau}_{hi})) + \sum_{i \neq i'}^{N_h} E(I_{hi} I_{hi'}) E_I(\text{cov}_{Z|I}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})) \right] \\
&= (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} \\
&= \sum_{h=1}^{L} \frac{1}{n_h} \left(S_h^2 + \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} \right) - \sum_{h=1}^{N_h} \frac{1}{N_h} S_h^2 \, ,
\end{aligned}$$

where

$$\text{var}(\hat{\tau}_{hi}) = \sum_{j=1}^{M_{hi}} \frac{1 - g_{hij}}{g_{hij}} y_{hij}^2 \text{ (see (3.2.2))} \, , \quad S_h^2 = \sum_{i=1}^{N_h} \frac{(\tau_{hi} - \bar{\tau}_h)^2}{N_h - 1}, \quad \bar{\tau}_h = \sum_{i=1}^{N_h} \frac{\tau_{hi}}{N_h} \, ,$$

and

$$\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'}) = \sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{hi'}} \frac{y_{hij} y_{hi'j'}}{\Theta_{hij} \Theta_{hi'j'}} \text{cov}(Z_{hij}, Z_{hi'j'}) = 0,$$

as $Z_{hij}, Z_{hi'j'}$ are independent. Let

$$v_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2$$

and

$$v_2 = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{N_h} \text{var}(\hat{\tau}_{hi}) = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} \frac{1 - g_{hij}}{g_{hij}} y_{hij}^2 \, .$$

Then $\text{var}(\hat{\tau}_\pi) = v_1 + v_2$. The first error component v_1 is due to the stratified random sampling of primary units. The second error component v_2 is associated with sampling of groups within primary units. We obtain an unbiased estimator of $\text{var}(\hat{\tau}_\pi)$ by constructing unbiased estimators of each variance component.

To obtain an unbiased estimator of v_1, we need an unbiased estimator of S_h^2. Consider

$$s_h^2 = \sum_{i=1}^{n_h} \frac{(\hat{\tau}_{hi} - \hat{\bar{\tau}}_h)^2}{n_h - 1} = \frac{\sum_{i=1}^{n_h} \hat{\tau}_{hi}^2 - n_h \hat{\bar{\tau}}_h^2}{n_h - 1} \text{ , where } \hat{\bar{\tau}}_h = \sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h} .$$

The expected value of s_h^2 is

$$E(s_h^2) = \frac{1}{n_h - 1} E(\sum_{i=1}^{n_h} \hat{\tau}_{hi}^2) - \frac{n_h}{n_h - 1} E(\hat{\bar{\tau}}_h^2).$$

Now

$$\begin{aligned} E(\sum_{i=1}^{n_h} \hat{\tau}_{hi}^2) &= E_I(E_{Z|I}(\sum_{i=1}^{N_h} I_{hi} \hat{\tau}_{hi}^2)) \\ &= \sum_{i=1}^{N_h} E_I(I_{hi} E(\hat{\tau}_{hi}^2 | I)) = \sum_{i=1}^{N_h} \frac{n_h}{N_h} E(\hat{\tau}_{hi}^2) \\ &= \sum_{i=1}^{N_h} \frac{n_h}{N_h} [\text{var}(\hat{\tau}_{hi}) + (E(\hat{\tau}_{hi}))^2] \\ &= \sum_{i=1}^{N_h} \frac{n_h}{N_h} [\text{var}(\hat{\tau}_{hi}) + \tau_{hi}^2] , \end{aligned} \quad (3.3.1)$$

and

$$\begin{aligned} E(\hat{\bar{\tau}}_h^2) &= E_I(E_{Z|I}((\sum_{i=1}^{N_h} I_{hi} \frac{\hat{\tau}_{hi}}{n_h})^2)) \\ &= \sum_{i=1}^{N_h} \frac{1}{n_h^2} E_I(I_{hi}^2 E(\hat{\tau}_{hi}^2 | I)) + \frac{1}{n_h^2} \sum_{i \neq i'}^{N_h} E_I(I_{hi} I_{hi'} E(\hat{\tau}_{hi} \hat{\tau}_{hi'} | I)) \\ &= \sum_{i=1}^{N_h} \frac{1}{n_h N_h} E(\hat{\tau}_{hi}^2) + \frac{n_h - 1}{n_h} \frac{1}{N_h(N_h - 1)} \sum_{i \neq i'}^{N_h} E(\hat{\tau}_{hi}) E(\hat{\tau}_{hi'}) \\ &= \sum_{i=1}^{N_h} \frac{1}{n_h N_h} E(\hat{\tau}_{hi}^2) + \frac{n_h - 1}{n_h N_h (N_h - 1)} \sum_{i \neq i'}^{N_h} \tau_{hi} \tau_{hi'} . \end{aligned}$$

Therefore,

$$\begin{aligned} E(s_h^2) &= \sum_{i=1}^{N_h} [\frac{n_h}{N_h(n_h - 1)} - \frac{1}{N_h(n_h - 1)}] E(\hat{\tau}_{hi}^2) - \frac{1}{N_h(N_h - 1)} \sum_{i \neq i'}^{N_h} \tau_{hi} \tau_{hi'} \\ &= \frac{1}{N_h} \sum_{i=1}^{N_h} \tau_{hi}^2 - \frac{1}{N_h(N_h - 1)} \sum_{i \neq i'}^{N_h} \tau_{hi} \tau_{hi'} + \frac{1}{N_h} \sum_{i=1}^{N_h} \text{var}(\hat{\tau}_{hi}) \\ &= S_h^2 + \frac{1}{N_h} \sum_{i=1}^{N_h} \text{var}(\hat{\tau}_{hi}) = S_h^2 + E(\frac{1}{n_h} \sum_{i=1}^{n_h} \widehat{\text{var}}(\hat{\tau}_{hi})), \end{aligned}$$

where $\widehat{\text{var}}(\hat{\tau}_{hi}) = \widehat{\text{var}}(\hat{\tau}_{\pi}) = \sum_{j=1}^{M_{hi}} \frac{1-g_{hij}}{g_{hij}^2} Z_{hij} y_{hij}^2$ (see (3.2.3)). Hence, an unbiased estimator for S_h^2 is given by

$$\widehat{S}_h^2 = s_h^2 - \frac{1}{n_h} \sum_{i=1}^{n_h} \widehat{\text{var}}(\hat{\tau}_{hi}) \ .$$

Then an unbiased estimator of v_1, the error component due to stratified random sampling of primary units, is

$$\hat{v}_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) \widehat{S}_h^2 \ .$$

Now the second error component v_2 due to sampling of groups within the selected primary units is

$$v_2 = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{N_h} \text{var}(\hat{\tau}_{hi}) = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} \frac{1-g_{hij}}{g_{hij}} y_{hij}^2 \ .$$

An unbiased estimator can be written as (see (3.3.2))

$$\hat{v}_2 = \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{i=1}^{N_h} I_{hi} \sum_{j=1}^{M_{hi}} \frac{1-g_{hij}}{g_{hij}^2} Z_{hij} y_{hij}^2 \ .$$

Hence,

$$\begin{aligned}
\widehat{\text{var}}(\hat{\tau}_{\pi}) &= \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) \widehat{S}_h^2 + \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{i=1}^{n_h} \sum_{j=1}^{M_{hi}} \frac{1-g_{hij}}{g_{hij}^2} Z_{hij} y_{hij}^2 \\
&= \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) \widehat{S}_h^2 + \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} \frac{1-g_{hij}}{g_{hij}^2} y_{hij}^2
\end{aligned}$$

is an unbiased estimator of $\text{var}(\hat{\tau}_{\pi})$.

3.3.2 When sighting probabilities are unknown

In the case where a logistic model (see (2.1.3)) is assumed for the unknown g_{hij}'s, and with all the assumptions for and between $(\widehat{\boldsymbol{\beta}}, \widehat{\Sigma})$ and I_{hi}, Z_{hij} identical to those stated in section 2.1.2, an unbiased estimator for the population size τ is

$$\hat{\tau}_{LR} = \sum_{h=1}^{L} N_h \sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h} \ ,$$

where $\hat{\tau}_{hi} = \sum_{j=1}^{M_{hi}} Z_{hij} y_{hij} \widehat{\Theta}_{hij}$ is an unbiased estimator of τ_{hi}. The variance of $\hat{\tau}_{LR}$ can be written as

$$\text{var}(\hat{\tau}_{LR}) = \sum_{h=1}^{L} N_h^2 \text{var}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}) + \sum_{h \neq h'}^{L} N_h N_{h'} \text{cov}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}, \sum_{i'=1}^{n_{h'}} \frac{\hat{\tau}_{h'i'}}{n_{h'}})$$

where

$$\begin{aligned} \text{var}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}) &= (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} + \sum_{i \neq i'}^{N_h} \frac{1}{n_h^2} E_I(I_{hi} I_{hi'}) \text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'}) \\ &= (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} + \frac{n_h - 1}{n_h} \sum_{i \neq i'}^{N_h} \frac{\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})}{N_h(N_h - 1)}, \end{aligned}$$

and

$$\begin{aligned} \text{cov}(\sum_{i=1}^{n_h} \frac{\hat{\tau}_{hi}}{n_h}, \sum_{i'=1}^{n_{h'}} \frac{\hat{\tau}_{h'i'}}{n_{h'}}) &= \sum_{i=1}^{N_h} \sum_{i'=1}^{N_{h'}} \frac{1}{n_h n_h'} E_I(I_{hi} I_{h'i'}) \text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{h'i'}) \\ &= \sum_{i=1}^{N_h} \sum_{i'=1}^{N_{h'}} \frac{1}{N_h N_{h'}} \text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{h'i'}). \end{aligned}$$

Therefore,

$$\text{var}(\hat{\tau}_{LR}) = \sum_{h=1}^{L} \frac{N_h^2}{n_h} \left[S_h^2 + \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} - \sum_{i \neq i'}^{N_h} \frac{\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})}{N_h(N_h - 1)} \right]$$

$$- \sum_{h=1}^{N_h} \frac{N_h^2}{N_h} \left[S_h^2 - N_h \sum_{i \neq i'}^{N_h} \frac{\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})}{N_h(N_h - 1)} \right] + \sum_{h \neq h'}^{L} \sum_{i=1}^{N_h} \sum_{i'=1}^{N_{h'}} \text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{h'i'})$$

Now (see (3.2.6) and (3.2.8))

$$\begin{aligned} \text{var}(\hat{\tau}_{hi}) &= \textstyle\sum_{j=1}^{M_{hi}} y_{hij}^2 (\Theta_{hij} - 1) \\ &+ E_Z[\sum_{j=1}^{M_{hi}} Z_{hij} y_{hij}^2 \text{var}(\widehat{\Theta}_{hij}) + \sum_{j \neq j'}^{M_{hi}} Z_{hij} Z_{hij'} y_{hij} y_{hij'} \text{cov}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hij'})] \end{aligned} \tag{3.3.2}$$

which can be unbiasedly estimated by (see (3.2.7) and (3.2.9))

$$\begin{aligned} \widehat{\text{var}}(\hat{\tau}_{hi}) &= \textstyle\sum_{j=1}^{M_{hi}} Z_{hij} y_{hij}^2 [\widehat{\Theta}_{hij}^2 - \widehat{\Theta}_{hij} - \widehat{\text{var}}(\widehat{\Theta}_{hij})], \\ &+ \sum_{j=1}^{M_{hi}} Z_{hij} y_{hij}^2 \widehat{\text{var}}(\widehat{\Theta}_{hij}) + \sum_{j \neq j'}^{M_{hi}} Z_{hij} Z_{j'} y_{hij} y_{hij'} \widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hij'}). \end{aligned}$$

Also

$$\begin{aligned}
\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'}) &= \sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{hi'}} \frac{\text{cov}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hi'j'})}{\Theta_{hij}\Theta_{hi'j'}} y_{hij} y_{hi'y'} \\
&= E[\sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{hi'}} Z_{hij} Z_{hi'j'} y_{hij} y_{hi'j'} \widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hi'j'})],
\end{aligned} \tag{3.3.3}$$

and

$$\begin{aligned}
\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{h'i'}) &= \sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{h'i'}} \frac{\text{cov}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{h'i'j'})}{\Theta_{hij}\Theta_{h'i'j'}} y_{hij} y_{h'i'j'} \\
&= E[\sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{h'i'}} Z_{hij} Z_{h'i'j'} y_{hij} y_{h'i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{h'i'j'})],
\end{aligned}$$

where $\widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hi'j'})$, $\widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{h'i'j'})$ are unbiased estimators of $\text{cov}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{hi'j'})$, $\text{cov}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{h'i'j'})$ respectively. The unbiased estimators of the covariances between population size estimators of

$$\widehat{\text{cov}}(\hat{\tau}_{hi}, \hat{\tau}_{h'i'}) = \sum_{j=1}^{M_{hi}} \sum_{j'=1}^{M_{h'i'}} Z_{hij} Z_{h'i'j'} y_{hij} y_{h'i'j'} \widehat{\text{cov}}(\widehat{\Theta}_{hij}, \widehat{\Theta}_{h'i'j'}) \quad \text{for } h = h' \text{ or } h \neq h'.$$

To find an unbiased estimator of $\text{var}(\hat{\tau}_{LR})$, we expressed it as a sum of three error components, $\text{var}(\hat{\tau}_{LR}) = v_1 + v_2 + v_3$. The first error component v_1, due to stratified random sampling of primary units, is

$$v_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2$$

which is exactly the same as the first error component of $\text{var}(\hat{\tau}_\pi)$ when sighting probabilities are known. If, as before, we consider s_h^2, we obtained the same expression for $E(\sum_{i=1}^{n_h} \hat{\tau}_{hi}^2)$ (see (3.3.1)) but

$$E(\hat{\bar{\tau}}_h^2) = \sum_{i=1}^{N_h} \frac{1}{n_h N_h} E(\hat{\tau}_{hi}^2) + \frac{n_h - 1}{n_h} \frac{1}{N_h(N_h - 1)} \sum_{i \neq i'} [\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'}) + \tau_{hi}\tau_{hi'}].$$

Therefore,

$$
\begin{aligned}
E(s_h^2) &= S_h^2 + \frac{1}{N_h}\sum_{i=1}^{N_h}\operatorname{var}(\hat{\tau}_{hi}) - \frac{1}{N_h(N_h-1)}\sum_{i\neq i'}^{N_h}\operatorname{cov}(\hat{\tau}_{hi},\hat{\tau}_{hi'}) \\
&= S_h^2 + E[\frac{1}{n_h}\sum_{i=1}^{N_h} I_{hi}\widehat{\operatorname{var}}(\hat{\tau}_{hi}) - \frac{1}{n_h(n_h-1)}\sum_{i\neq i'}^{N_h} I_{hi}I_{hi'}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{hi'})].
\end{aligned}
$$

Hence, an unbiased estimator of S_h^2 is

$$
\widehat{S}_h^2 = s_h^2 - \frac{1}{n_h}\sum_{i=1}^{n_h}\widehat{\operatorname{var}}(\hat{\tau}_{hi}) + \frac{1}{n_h(n_h-1)}\sum_{i\neq i'}^{n_h}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{hi'}).
$$

Thus, an unbiased estimator of the first error component is

$$
\hat{v}_1 = \sum_{h=1}^{L} N_h^2(\frac{1}{n_h} - \frac{1}{N_h})\widehat{S}_h^2.
$$

The second error component v_2, due to sampling of groups within selected primary unit, is

$$
v_2 = \sum_{h=1}^{L}\frac{N_h}{n_h}\sum_{i=1}^{N_h}\sum_{j=1}^{M_{hi}} y_{hij}^2(\Theta_{hij}-1).
$$

An unbiased estimator for v_2 is

$$
\hat{v}_2 = \sum_{h=1}^{L}\frac{N_h^2}{n_h^2}\sum_{i=1}^{N_h} I_{hi}\sum_{j=1}^{M_{hi}} Z_{hij}y_{hij}^2[\widehat{\Theta}_{hij}^2 - \widehat{\Theta}_{hij} - \widehat{\operatorname{var}}(\widehat{\Theta}_{hij})].
$$

The third error component, v_3, due to estimation of Θ_{hij} is

$$
\begin{aligned}
v_3 = \sum_{h=1}^{L}\frac{N_h}{n_h}\sum_{i=1}^{N_h}[\sum_{j=1}^{M_{hi}}\frac{y_{hij}^2}{\Theta_{hij}}\operatorname{var}(\widehat{\Theta}_{hij}) + \sum_{j\neq j'}^{M_{hi}}\frac{y_{hij}y_{hij'}}{\Theta_{hij}\Theta_{hij'}}\operatorname{cov}(\widehat{\Theta}_{hij},\widehat{\Theta}_{hij'})] \\
+ \textstyle\sum_{h=1}^{L}\frac{N_h(n_h-1)}{n_h(N_h-1)}\sum_{i\neq i'}^{N_h}\operatorname{cov}(\hat{\tau}_{hi},\hat{\tau}_{hi'}) + \sum_{h\neq h'}^{L}\sum_{i=1}^{N_h}\sum_{i'=1}^{N_{h'}}\operatorname{cov}(\hat{\tau}_{hi},\hat{\tau}_{h'i'})\ .
\end{aligned}
$$

An unbiased estimator for v_3 is

$$
\begin{aligned}
\hat{v}_3 = \sum_{h=1}^{L}\frac{N_h^2}{n_h^2}\sum_{i=1}^{N_h} I_{hi}[\sum_{j=1}^{M_{hi}} Z_{hij}y_{hij}^2\operatorname{var}(\widehat{\Theta}_{hij}) + \sum_{j\neq j'}^{M_{hi}} Z_{hij}Z_{hij'}y_{hij}y_{hij'}\operatorname{cov}(\widehat{\Theta}_{hij},\widehat{\Theta}_{hij'})] \\
+ \textstyle\sum_{h=1}^{L}\frac{N_h^2}{n_h^2}\sum_{i\neq i'}^{N_h} I_{hi}I_{hi'}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{hi'}) + \sum_{h\neq h'}^{L}\frac{N_hN_{h'}}{n_hn_h'}\sum_{i=1}^{N_h}\sum_{i'=1}^{N_{h'}} I_{hi}I_{h'i'}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{h'i'})
\end{aligned}
$$

$$
\begin{aligned}
= \textstyle\sum_{h=1}^{L}\frac{N_h^2}{n_h^2}\sum_{i=1}^{n_h}[\sum_{j=1}^{m_{hi}} y_{hij}^2\operatorname{var}(\widehat{\Theta}_{hij}) + \sum_{j\neq j'}^{m_{hi}} y_{hij}y_{hij'}\operatorname{cov}(\widehat{\Theta}_{hij},\widehat{\Theta}_{hij'})] \\
+ \textstyle\sum_{h=1}^{L}\frac{N_h^2}{n_h^2}\sum_{i\neq i'}^{n_h}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{hi'}) + \sum_{h\neq h'}^{L}\frac{N_hN_{h'}}{n_hn_h'}\sum_{i=1}^{n_h}\sum_{i'=1}^{n_{h'}}\widehat{\operatorname{cov}}(\hat{\tau}_{hi},\hat{\tau}_{h'i'})
\end{aligned}
$$

3.3.3 Optimum allocations

Optimum allocation of sampling effort assuming that survey costs are a linear function of stratum sample sizes (n_h), satisfies (Cochran, 1977)

$$\frac{n_h}{n} = \frac{N_h S_h^* / \sqrt{c_h}}{\sum_{h=1}^{L} N_h S_h^* / \sqrt{c_h}},$$

where $n_1 + n_2 + ... + n_h = n$ and c_h is the cost of sampling a primary unit, which may vary from stratum to stratum, and S_h^* depends on whether sighting probabilities are known or estimated via a logistic model.

3.3.3.1 When sighting probabilities are known

We have

$$S_h^{*2} = S_h^2 + \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h},$$

where $\text{var}(\hat{\tau}_{hi}) = \sum_{j=1}^{M_{hi}} y_{hij}^2 (\Theta_{hij} - 1)$.

3.3.3.2 When sighting probabilities are unknown

When sighting probabilities are unknown and are assumed to satisfy a logistic model defined in (2.1.3) with all the assumptions described in section 3.3.2, we have

$$S_h^{*2} = S_h^2 + \sum_{i=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hi})}{N_h} - \sum_{i \neq i'}^{N_h} \frac{\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})}{N_h(N_h - 1)},$$

where expressions of $\text{var}(\hat{\tau}_{hi})$ and $\text{cov}(\hat{\tau}_{hi}, \hat{\tau}_{hi'})$ can be found in (3.3.2) and (3.3.3).

3.4 Case 3: Population size estimator based on the composite data of several replicated population surveys

If a phase II sample is the composite data of r replicated population surveys, there are three possible scenarios. The first scenario is the case of a complete census of primary units. In this case, we simply divide the population size estimate by r and the estimates of error components by r^2 to obtain the required estimates.

The second scenario is when sample primary units of r population surveys are selected without replacement and are thus, all different. In this case, we could simply treat the composite data as a data set from one population survey and use the population size estimator presented in chapter 2 or in section 3.3. Population size estimation based on population surveys conducted on primary units sampled without replacement will probably be more precise than the population size estimation based on population surveys conducted on primary units sampled with replacement, which is the third scenario described in the following paragraph.

The third scenario is when same sample primary units are surveyed r times. In this case, we will be combining data of replicated population surveys. One example of this case is the Colorado elk data collected from 2 population surveys conducted on 132 miles^2 of winter range in Game Management Unit (or GMU) 42, South of Rifle and Newcastle, Colorado on 15-20th January and from 29th Frebruary to 3rd March 1996. In the following subsections, using the notation defined in section 2.1 and in section 3.3, we present the population size estimator for simple random sampling and stratified random sampling of primary units under the third scenario. Under the third scenario, we allow the likely possibility that a replicated sample primary unit may consist of different number of groups during different population surveys.

In all three scenarios of this special case, it is possible that the population (or groups of animals) has restructured over a period of time. For example, frequencies of group size could change over time. However, combining data collected over a period of time should not have a big impact on the population size estimation.

3.4.1 Simple random sampling of primary units

The modified Horvitz-Thompson population size estimator in this case is

$$\hat{\tau} = \sum_{i=1}^{r} \frac{\hat{\tau}_i}{r} = \sum_{i=1}^{r} \sum_{j=1}^{N} \frac{I_j}{\pi_j} \frac{\hat{\tau}_{ij}}{r} = \sum_{j=1}^{N} \frac{I_j}{\pi_j} \sum_{i=1}^{r} \frac{\hat{\tau}_{ij}}{r} = \sum_{j=1}^{N} \frac{I_j}{\pi_j} \bar{\hat{\tau}}_j,$$

where $\hat{\tau}_{ij}$ is the size estimator of the j^{th} primary unit based on the i^{th} population survey and $\bar{\hat{\tau}}_j = \sum_{i=1}^{r} \frac{\hat{\tau}_{ij}}{r}$. We will show that $\text{var}(\hat{\tau}) = v_1 + v_2 + v_3$ when sighting probabilities are unknown and estimated via logistic regression. If sighting probabilities are known, we have $\text{var}(\hat{\tau}) = v_1 + v_2$.

First,

$$
\begin{aligned}
\text{var}(\hat{\tau}) &= \text{var}(\sum_{j=1}^{N} \frac{I_j}{\pi_j}\bar{\hat{\tau}}_j) \\
&= \sum_{j=1}^{N} \text{var}(\frac{I_j}{\pi_j}\bar{\hat{\tau}}_j) + \sum_{j \neq j'}^{N} \frac{1}{\pi_j \pi_j'}\text{cov}(I_j\bar{\hat{\tau}}_j, I_{j'}\bar{\hat{\tau}}_{j'}).
\end{aligned}
$$

Now,

$$
\begin{aligned}
\text{var}(\frac{I_j}{\pi_j}\bar{\hat{\tau}}_j) &= \frac{1}{\pi_j^2}\text{var}(I_j\bar{\hat{\tau}}_j) \\
&= \frac{1}{\pi_j^2}\left[\text{var}(E(I_j\bar{\hat{\tau}}_j|I_j)) + E(\text{var}(I_j\bar{\hat{\tau}}_j|I_j))\right] \\
&= \frac{1}{\pi_j^2}\left[\text{var}(I_j\bar{\tau}_j) + E(I_j\text{var}(\bar{\hat{\tau}}_j))\right] \\
&= \frac{1}{\pi_j^2}\left[\pi_j(1-\pi_j)\bar{\tau}_j^2 + \pi_j\text{var}(\bar{\hat{\tau}}_j)\right].
\end{aligned}
$$

But

$$
\begin{aligned}
\text{var}(\bar{\hat{\tau}}_j) &= \frac{1}{r^2}\text{var}(\sum_{i=1}^{r}\hat{\tau}_{ij}) \\
&= \frac{1}{r^2}\left[\sum_{i=1}^{r}\text{var}(\hat{\tau}_{ij}) + \sum_{i \neq i'}^{r}\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j})\right].
\end{aligned}
$$

So,

$$
\text{var}(\frac{I_j}{\pi_j}\bar{\hat{\tau}}_j) = \frac{1-\pi_j}{\pi_j}\bar{\tau}_j^2 + \frac{1}{r^2\pi_j}\left[\sum_{i=1}^{r}\text{var}(\hat{\tau}_{ij}) + \sum_{i \neq i'}^{r}\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j})\right],
$$

where

$$
\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j}) = \text{cov}(\sum_{k=1}^{M_{ij}} Z_{ijk}y_{ijk}\widehat{\Theta}_{ijk}, \sum_{k'=1}^{M_{i'j}} Z_{i'jk'}y_{i'jk'}\widehat{\Theta}_{i'jk'})
$$

$$
\begin{aligned}
&= \sum_{k,k'} y_{ijk}y_{i'jk'}[E(\text{cov}(Z_{ijk}\widehat{\Theta}_{ijk}, Z_{i'jk'}\widehat{\Theta}_{i'jk'}|Z_{ijk}, Z_{i'jk'})) \\
&+ \text{cov}(E(Z_{ijk}\widehat{\Theta}_{ijk}|Z_{ijk}), E(Z_{i'jk'}\widehat{\Theta}_{i'jk'}|Z_{i'jk'}))] \\
&= \sum_{k,k'} y_{ijk}y_{i'jk'}\left[E(Z_{ijk}Z_{i'jk'})\text{cov}(\widehat{\Theta}_{ijk}, \widehat{\Theta}_{i'jk'}) + \text{cov}(Z_{ijk}\Theta_{ijk}, Z_{i'jk'}\Theta_{i'jk'})\right] \\
&= \sum_{k,k'=1} \frac{y_{ijk}y_{i'jk'}}{\Theta_{ijk}\Theta_{i'jk'}}\text{cov}(\widehat{\Theta}_{ijk}, \widehat{\Theta}_{i'jk'}),
\end{aligned}
$$

given that Z_{ijk} and $Z_{i'jk'}$ are independent. Therefore $\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j}) = 0$ when sighting probabilities are known. Second,

$$
\begin{aligned}
\text{cov}(I_j\overline{\hat{\tau}}_j, I_{j'}\overline{\hat{\tau}}_{j'}) &= \frac{1}{r^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}\text{cov}(I_j\hat{\tau}_{ij}, I_{j'}\hat{\tau}_{i'j'}) \\
&= \frac{1}{r^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}[E(\text{cov}(I_j\hat{\tau}_{ij}, I_{j'}\hat{\tau}_{i'j'}|I_jI_{j'})) + \text{cov}(E(I_j\hat{\tau}_{ij}|I_j), E(I_{j'}\hat{\tau}_{i'j'}|I_{j'}))] \\
&= \frac{1}{r^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}[E(I_jI_{j'})\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j'}) + \text{cov}(I_j\tau_{ij}, I_{j'}\tau_{i'j'})] \\
&= \frac{1}{r^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}[\pi_{jj'}\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j'}) + \tau_{ij}\tau_{i'j'}\text{cov}(I_j, I_{j'})] \\
&= \frac{1}{r^2}\sum_{i=1}^{r}\sum_{i'=1}^{r}[\pi_{jj'}\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j'}) + (\pi_{jj'} - \pi_j\pi_{j'})\tau_{ij}\tau_{i'j'}],
\end{aligned}
$$

where

$$
\text{cov}(\hat{\tau}_{ij}, \hat{\tau}_{i'j'}) = \sum_{k=1}^{M_{ij}}\sum_{k'=1}^{M_{i'j'}} \frac{y_{ijk}y_{i'j'k'}}{\Theta_{ijk}\Theta_{i'j'k'}}\text{cov}(\widehat{\Theta}_{ijk}, \widehat{\Theta}_{i'j'k'}),
$$

is zero when sighting probabilities are known.

3.4.1.1 When sighting probabilities are unknown

We have

$$
\hat{\tau}_{ij} = \sum_{k=1}^{M_{ij}} Z_{ijk}y_{ijk}\widehat{\Theta}_{ijk}.
$$

The first error component due to sampling of primary units is

$$
v_1 = \sum_{j=1}^{N}\frac{1-\pi_j}{\pi_j}\overline{\tau}_j^2 + \sum_{j\neq j'}^{N}\frac{\pi_{jj'} - \pi_j\pi_{j'}}{\pi_j\pi_{j'}}\overline{\tau}_j\overline{\tau}_{j'}.
$$

To find an unbiased estimator $\hat{v}_1$, consider

$$
\sum_{j=1}^{N}\frac{1-\pi_j}{\pi_j^2}I_j\overline{\hat{\tau}}_j^2 + \sum_{j\neq j'}^{N}\frac{\pi_{jj'} - \pi_j\pi_{j'}}{\pi_{jj'}\pi_j\pi_{j'}}I_jI_{j'}\overline{\hat{\tau}}_j\overline{\hat{\tau}}_{j'}.
$$

Expectation of the above expression is

$$v_1 + \frac{1}{r^2}\sum_{j=1}^{N}\left(\frac{1-\pi_j}{\pi_j}\left[\sum_{i=1}^{r}\text{var}(\hat{\tau}_{ij}) + \sum_{i\neq i'}^{r}\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j})\right] + \sum_{j\neq j'}^{N}\frac{\pi_{jj'}-\pi_j\pi_{j'}}{\pi_j\pi_{j'}}\sum_{i,i'=1}^{r}\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j'})\right).$$

Thus, an unbiased estimator for the first error component v_1 is

$$\begin{aligned}\hat{v}_1 &= \sum_{j=1}^{N}\frac{1-\pi_j}{\pi_j^2}I_j\bar{\hat{\tau}}_j^2 + \sum_{j\neq j'}^{N}\frac{\pi_{jj'}-\pi_j\pi_{j'}}{\pi_{jj'}\pi_j\pi_{j'}}I_jI_{j'}\bar{\hat{\tau}}_j\bar{\hat{\tau}}_{j'} \\ &- \frac{1}{r^2}\sum_{j=1}^{N}\left(\frac{1-\pi_j}{\pi_j^2}I_j\left[\sum_{i=1}^{r}\widehat{\text{var}}(\hat{\tau}_{ij}) + \sum_{i\neq i'}^{r}\widehat{\text{cov}}(\hat{\tau}_{ij},\hat{\tau}_{i'j})\right] + \sum_{j\neq j'}^{N}\frac{\pi_{jj'}-\pi_j\pi_{j'}}{\pi_{jj'}\pi_j\pi_{j'}}I_jI_{j'}\sum_{i,i'=1}^{r}\widehat{\text{cov}}(\hat{\tau}_{ij},\hat{\tau}_{i'j'})\right),\end{aligned}$$

where

$$\begin{aligned}\widehat{\text{var}}(\hat{\tau}_{ij}) &= \sum_{k=1}^{M_{ij}} Z_{ijk}y_{ijk}^2[\widehat{\Theta}_{ijk}^2 - \widehat{\Theta}_{ijk} - \widehat{\text{var}}(\widehat{\Theta}_{ijk})] \\ &+ \sum_{k=1}^{M_{ij}} Z_{ijk}y_{ijk}^2\widehat{\text{var}}(\widehat{\Theta}_{ijk}) + \sum_{k,k'=1}^{M_{ij}} Z_{ijk}Z_{ijk'}y_{ijk}y_{ijk'}\widehat{\text{cov}}(\widehat{\Theta}_{ijk},\widehat{\Theta}_{ijk'}),\end{aligned}$$

$$\widehat{\text{cov}}(\hat{\tau}_{ij},\hat{\tau}_{i'j}) = \sum_{k,k'=1}^{M_{ij}} Z_{ijk}Z_{i'jk'}y_{ijk}y_{i'jk'}\widehat{\text{cov}}(\widehat{\Theta}_{ijk},\widehat{\Theta}_{i'jk'})$$

and

$$\widehat{\text{cov}}(\hat{\tau}_{ij},\hat{\tau}_{i'j'}) = \sum_{k,k'=1}^{M_{ij}} Z_{ijk}Z_{i'j'k'}y_{ijk}y_{i'j'k'}\widehat{\text{cov}}(\widehat{\Theta}_{ijk},\widehat{\Theta}_{i'j'k'})$$

are unbiased estimators of $\text{var}(\hat{\tau}_{ij})$ (see section 3.2.2), $\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j})$ and $\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j'})$, respectively. The second error component due to sampling of groups is

$$v_2 = \sum_{j=1}^{N}\frac{1}{r^2\pi_j}\sum_{i=1}^{r}\sum_{k=1}^{M_{ij}} y_{ijk}^2(\Theta_{ijk}-1).$$

An unbiased estimator for v_2 is

$$\hat{v}_2 = \sum_{j=1}^{N}\frac{I_j}{r^2\pi_j^2}\sum_{i=1}^{r}\sum_{k=1}^{M_{ij}} Z_{ijk}y_{ijk}^2[\widehat{\Theta}_{ijk}^2 - \widehat{\Theta}_{ijk} - \widehat{\text{var}}(\widehat{\Theta}_{ijk})].$$

The third error component due to estimation of Θ_{ijk} is

$$\begin{aligned}v_3 &= \sum_{j=1}^{N}\frac{1}{r^2\pi_j}\sum_{i=1}^{r}\left[\sum_{k=1}^{M_{ij}}\frac{y_{ijk}^2}{\Theta_{ijk}}\widehat{\text{var}}(\widehat{\Theta}_{ijk}) + \sum_{k\neq k'}^{M_{ij}} g_{ijk}g_{ijk'}\frac{y_{ijk}y_{ijk'}}{\Theta_{ijk}\Theta_{ijk'}}\widehat{\text{cov}}(\widehat{\Theta}_{ijk},\widehat{\Theta}_{ijk'})\right] \\ &+ \sum_{j=1}^{N}\frac{1}{r^2\pi_j}\sum_{i\neq i'}^{r}\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j}) + \sum_{j\neq j'}^{N}\frac{1}{r^2}\sum_{i,i'=1}^{r}\frac{\pi_{jj'}}{\pi_j\pi_{j'}}\text{cov}(\hat{\tau}_{ij},\hat{\tau}_{i'j'}).\end{aligned}$$

An unbiased estimator for v_3 is

$$\begin{aligned}\hat{v}_3 &= \sum_{j=1}^{N} \frac{I_j}{r^2\pi_j^2}\left[\sum_{i=1}^{r}\sum_{k=1}^{M_{ij}} Z_{ijk} y_{ijk}^2 \widehat{\text{var}}(\widehat{\Theta}_{ijk}) + \sum_{k,k'=1}^{M_{ij}} Z_{ijk} Z_{ijk'} y_{ijk} y_{ijk'} \widehat{\text{cov}}(\widehat{\Theta}_{ijk}, \widehat{\Theta}_{ijk'})\right] \\ &+ \sum_{j=1}^{N} \frac{I_j}{r^2\pi_j^2} \sum_{i\neq i'}^{r} \widehat{\text{cov}}(\hat{\tau}_{ij}, \hat{\tau}_{i'j}) + \sum_{j\neq j'}^{N} \frac{I_j I_{j'}}{r^2 \pi_j \pi_{j'}} \sum_{i,i'=1}^{r} \widehat{\text{cov}}(\hat{\tau}_{ij}, \hat{\tau}_{i'j'}).\end{aligned}$$

3.4.1.2 When sighting probabilities are known

We have

$$\hat{\tau}_{ij} = \sum_{k=1}^{M_{ij}} \frac{Z_{ijk} y_{ijk}}{g_{ijk}}.$$

The first error component due to sampling of primary units is

$$v_1 = \sum_{j=1}^{N} \frac{1-\pi_j}{\pi_j} \bar{\tau}_j^2 + \sum_{j\neq j'}^{N} \frac{\pi_{jj'} - \pi_j \pi_{j'}}{\pi_j \pi_{j'}} \bar{\tau}_j \bar{\tau}_{j'}.$$

To find an unbiased estimator $\hat{v}_1$, consider

$$\sum_{j=1}^{N} \frac{1-\pi_j}{\pi_j^2} I_j \hat{\bar{\tau}}_j^2 + \sum_{j\neq j'}^{N} \frac{\pi_{jj'} - \pi_j \pi_{j'}}{\pi_{jj'} \pi_j \pi_{j'}} I_j I_{j'} \hat{\bar{\tau}}_j \hat{\bar{\tau}}_{j'}.$$

Expectation of the above expression is

$$v_1 + \frac{1}{r^2} \sum_{j=1}^{N} \frac{1-\pi_j}{\pi_j} \sum_{i=1}^{r} \text{var}(\hat{\tau}_{ij}).$$

Thus, an unbiased estimator for the first error component v_1 is

$$\hat{v}_1 = \sum_{j=1}^{N} \frac{1-\pi_j}{\pi_j^2} I_j \hat{\bar{\tau}}_j^2 + \sum_{j\neq j'}^{N} \frac{\pi_{jj'} - \pi_j \pi_{j'}}{\pi_{jj'} \pi_j \pi_{j'}} I_j I_{j'} \hat{\bar{\tau}}_j \hat{\bar{\tau}}_{j'} - \frac{1}{r^2} \sum_{j=1}^{N} \frac{1-\pi_j}{\pi_j^2} I_j \sum_{i=1}^{r} \widehat{\text{var}}(\hat{\tau}_{ij}),$$

where $\widehat{\text{var}}(\hat{\tau}_{ij}) = \sum_{k=1}^{M_{ij}} \frac{1-g_{ijk}}{g_{ijk}^2} Z_{ijk} y_{ijk}^2$ is an unbiased estimator of $\text{var}(\hat{\tau}_{ij})$ (see section 3.2.1). The second error component due to sampling of groups is

$$v_2 = \sum_{j=1}^{N} \frac{1}{r^2 \pi_j} \sum_{i=1}^{r} \sum_{k=1}^{M_{ij}} \frac{1-g_{ijk}}{g_{ijk}} y_{ijk}^2.$$

An unbiased estimator for v_2 is

$$\hat{v}_2 = \sum_{j=1}^{N} \frac{I_j}{r^2 \pi_j^2} \sum_{k=1}^{M_{ij}} \frac{1-g_{ijk}}{g_{ijk}^2} Z_{ijk} y_{ijk}^2.$$

3.4.2 Stratified random sampling of primary Units

The population size estimator in this case is

$$\hat{\tau} = \sum_{h=1}^{L} N_h \sum_{j=1}^{N_h} I_{hj} \frac{\bar{\hat{\tau}}_{hj}}{n_h}, \quad \bar{\hat{\tau}}_{hj} = \sum_{i=1}^{r} \frac{\hat{\tau}_{hji}}{r}$$

where $\hat{\tau}_{hji}$ is the size estimator of j^{th} primary unit in the h^{th} strata based on the i^{th} population survey. Then

$$\begin{aligned}
\text{var}(\hat{\tau}) &= \text{var}(\sum_{h=1}^{L} N_h \sum_{j=1}^{N_h} I_{hj} \frac{\bar{\hat{\tau}}_{hj}}{n_h}) \\
&= \sum_{h=1}^{L} N_h^2 \text{var}(\sum_{j=1}^{N_h} I_{hj} \frac{\bar{\hat{\tau}}_{hj}}{n_h}) + \sum_{h \neq h'}^{L} \frac{N_h}{n_h} \frac{N_{h'}}{n_{h'}} \sum_{j=1}^{N} \sum_{j'=1}^{N_{h'}} \text{cov}(I_{hj} \bar{\hat{\tau}}_{hj}, I_{h'j'} \bar{\hat{\tau}}_{h'j'}).
\end{aligned}$$

Now,

$$\begin{aligned}
\text{var}(\sum_{j=1}^{N_h} I_{hj} \frac{\bar{\hat{\tau}}_{hj}}{n_h}) = \text{var}(\sum_{j=1}^{n_h} \frac{\bar{\hat{\tau}}_{hj}}{n_h}) &= \text{var}(\sum_{j=1}^{n_h} E(\frac{\bar{\hat{\tau}}_{hj}}{n_h}|I)) + E(\text{var}(\sum_{j=1}^{N_h} \frac{I_{hj}\bar{\hat{\tau}}_{hj}}{n_h}|I)) \\
&= \text{var}(\sum_{j=1}^{n_h} \frac{\bar{\tau}_{hj}}{n_h}) + \frac{1}{n_h^2} \left[\sum_{j=1}^{N_h} E(I_{hj}) \text{var}(\bar{\hat{\tau}}_{hj}) + \sum_{j \neq j'}^{N_h} E(I_{hj} I_{hj'}) \text{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{hj'}) \right] \\
&= (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{j=1}^{N_h} \frac{\text{var}(\bar{\hat{\tau}}_{hj})}{N_h} + \frac{n_h - 1}{n_h} \sum_{j \neq j'}^{N_h} \frac{\text{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{hj'})}{N_h(N_h - 1)},
\end{aligned}$$

where

$$S_h^2 = \sum_{j=1}^{N_h} \frac{(\bar{\tau}_{hj} - \bar{\bar{\tau}}_h)^2}{N_h - 1}, \quad \bar{\tau}_{hj} = \sum_{i=1}^{r} \frac{\tau_{hji}}{r}, \quad \bar{\bar{\tau}}_h = \sum_{j=1}^{N_h} \frac{\bar{\tau}_{hj}}{N_h},$$

and

$$\begin{aligned}
\text{var}(\bar{\hat{\tau}}_{hj}) &= \text{var}(\sum_{i=1}^{r} \frac{\hat{\tau}_{hji}}{r}) \\
&= \frac{1}{r^2} [\sum_{i=1}^{r} \text{var}(\hat{\tau}_{hji}) + \sum_{i \neq i'}^{r} \text{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'})]
\end{aligned}$$

such that

$$\begin{aligned}
\text{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) &= \text{cov}(\sum_{k=1}^{M_{hji}} Z_{hjik} y_{hjik} \widehat{\Theta}_{hjik}, \sum_{k'=1}^{M_{hji'}} Z_{hjik'} y_{hjik'} \widehat{\Theta}_{hjik'}) \\
&= \sum_{k,k'=1} y_{hjik} y_{hjik'} \text{cov}(Z_{hjik} \widehat{\Theta}_{hjik}, Z_{hjik'} \widehat{\Theta}_{hjik'})
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k,k'=1} y_{hjik} y_{hjik'} [E(\mathrm{cov}(Z_{hjik}\widehat{\Theta}_{hjik}, Z_{hjik'}\widehat{\Theta}_{hjik'} | Z_{hjik}, Z_{hjik'})) \\
&+ \mathrm{cov}(E(Z_{hjik}\widehat{\Theta}_{hjik}|Z_{hjik}), E(Z_{hjik'}\widehat{\Theta}_{hjik'}|Z_{hjik'}))] \\
&= \sum_{k,k'=1} y_{hjik} y_{hjik'} [E(Z_{hjik}Z_{hjik'})\mathrm{cov}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}) \\
&+ \mathrm{cov}(Z_{hjik}\Theta_{hjik}, Z_{hjik'}\Theta_{hjik'})] \\
&= \sum_{k,k'=1} \frac{y_{hjik} y_{hjik'}}{\Theta_{hjik}\Theta_{hjik'}} \mathrm{cov}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}).
\end{aligned} \tag{3.4.1}
$$

Therefore $\mathrm{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) = 0$ when sighting probabilities are known. Also,

$$
\mathrm{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{hj'}) = \frac{1}{r^2} \sum_{i=1}^{r} \sum_{i'=1}^{r} \sum_{k=1}^{M_{hji}} \sum_{k'=1}^{M_{hj'i'}} \frac{y_{hjik} y_{hj'i'k'}}{\Theta_{hjik}\Theta_{hj'i'k'}} \mathrm{cov}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hj'i'k'}) \tag{3.4.2}
$$

is zero when sighting probabilities are known. Similarly,

$$
\begin{aligned}
\mathrm{cov}(I_{hj}\bar{\hat{\tau}}_{hj}, I_{h'j'}\bar{\hat{\tau}}_{h'j'}) &= E(\mathrm{cov}(I_{hj}\bar{\hat{\tau}}_{hj}, I_{h'j'}\bar{\hat{\tau}}_{h'j'} | I_{hj}, I_{h'j'}) \\
&+ \mathrm{cov}(E(I_{hj}\bar{\hat{\tau}}_{hj}|I_{hj}), E(I_{h'j'}\bar{\hat{\tau}}_{h'j'}|I_{h'j'})) \\
&= E(I_{hj}I_{h'j'})\mathrm{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{h'j'}) + \bar{\tau}_{hj}\bar{\tau}_{h'j'}\mathrm{cov}(I_{hj}, I_{h'j'}) \\
&= \frac{n_h}{N_h}\frac{n_{h'}}{N_{h'}} \mathrm{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{h'j'}),
\end{aligned}
$$

where

$$
\mathrm{cov}(\bar{\hat{\tau}}_{hj}, \bar{\hat{\tau}}_{h'j'}) = \frac{1}{r^2} \sum_{i=1}^{r} \sum_{i'=1}^{r} \sum_{k=1}^{M_{hji}} \sum_{k'=1}^{M_{h'j'i'}} \frac{y_{hjik} y_{h'j'i'k'}}{\Theta_{hjik}\Theta_{h'j'i'k'}} \mathrm{cov}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{h'j'i'k'}), \tag{3.4.3}
$$

is zero when sighting probabilities are known.

3.4.2.1 When sighting probabilities are unknown

We have

$$
\hat{\tau}_{hji} = \sum_{k=1}^{M_{hj}} Z_{hjik} y_{hjik} \widehat{\Theta}_{hjik}
$$

and

$$
\begin{aligned}
\mathrm{var}(\hat{\tau}) &= \sum_{h=1}^{L} N_h^2 \left[(\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{j=1}^{N_h} \frac{\mathrm{var}(\overline{\hat{\tau}}_{hj})}{N_h} + \frac{n_h - 1}{n_h} \sum_{j \neq j'}^{N_h} \frac{\mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{N_h(N_h-1)} \right] + \sum_{h \neq h'}^{L} \sum_{j=1}^{N_h} \sum_{j'=1}^{N_{h'}} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'}) \\
&= \sum_{h=1}^{L} \frac{N_h^2}{n_h} \left[S_h^2 + \sum_{j=1}^{N_h} \frac{\mathrm{var}(\overline{\hat{\tau}}_{hj})}{N_h} - \sum_{j \neq j'}^{N_h} \frac{\mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{N_h(N_h-1)} \right] - \sum_{h=1}^{L} \frac{N_h^2}{N_h} \left[S_h^2 - N_h \sum_{j \neq j'}^{N_h} \frac{\mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{N_h(N_h-1)} \right] \\
&+ \sum_{h \neq h'}^{L} \sum_{j=1}^{N_h} \sum_{j'=1}^{N_{h'}} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'}) \\
&= \left[\sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 \right] + \left[\sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{1}{r^2} \sum_{i=1}^{r} \sum_{k=1}^{M_{hj}} y_{hjik}^2 (\Theta_{hjik} - 1) \right] \\
&+ [\sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{I_{hj}}{r^2} \left[\sum_{i=1}^{r} \left(\sum_{k=1}^{M_{hji}} \frac{y_{hjik}^2}{\Theta_{hjik}} \widehat{\mathrm{var}}(\widehat{\Theta}_{hjik}) + \sum_{k \neq k'}^{M_{hji}} \frac{y_{hjik} y_{hjik'}}{\Theta_{hjik} \Theta_{hjik'}} \widehat{\mathrm{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}) \right) \right] \\
&+ \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{1}{r^2} \sum_{i \neq i'}^{r} \mathrm{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) + \sum_{h=1}^{L} \frac{N_h}{n_h} \frac{n_h - 1}{N_h - 1} \sum_{j \neq j'}^{N_h} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'}) + \sum_{h \neq h'}^{L} \sum_{j=1}^{N_h} \sum_{j'=1}^{N_{h'}} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'})] \\
&= v_1 + v_2 + v_3
\end{aligned}
$$

The first error component due to sampling of primary units is

$$v_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2$$

which is exactly the same as the first error component when sighting probabilities are known. To find an unbiased estimator of S_h^2, consider

$$s_h^2 = \sum_{j=1}^{n_h} \frac{(\overline{\hat{\tau}}_{hj} - \overline{\overline{\hat{\tau}}}_h)^2}{n_h - 1}, \text{where } \overline{\overline{\hat{\tau}}}_h = \sum_{j=1}^{n_h} \frac{\overline{\hat{\tau}}_{hj}}{n_h}.$$

Then,

$$
\begin{aligned}
E(s_h^2) &= \frac{1}{n_h - 1} E\left(\sum_{j=1}^{n_h} \overline{\hat{\tau}}_{hj}^2 - n_h \overline{\overline{\hat{\tau}}}_h^2 \right) \\
&= \frac{1}{n_h - 1} E\left[(\sum_{j=1}^{n_h} (\overline{\hat{\tau}}_{hj}^2 - \mathrm{var}(\overline{\hat{\tau}}_{hj} | I_{hj})) - n_h (\overline{\overline{\hat{\tau}}}_h^2 - \frac{1}{n_h} \sum_{j=1}^{n_h} \mathrm{var}(\overline{\hat{\tau}}_{hj} | I_{hj})) \right],
\end{aligned}
$$

such that

$$E(\overline{\overline{\hat{\tau}}}_h^2) = E(\sum_{j=1}^{N_h} \frac{I_{hj} \overline{\hat{\tau}}_{hj}}{n_h})^2 = \sum_{j=1}^{N_h} \frac{1}{n_h^2} E(I_{hj} \overline{\hat{\tau}}_{hj}^2) + \sum_{j \neq j'}^{N_h} \frac{1}{n_h^2} E(I_{hj} I_{hj'} \overline{\hat{\tau}}_{hj} \overline{\hat{\tau}}_{hj'})$$

$$
\begin{aligned}
&= \sum_{j=1}^{N_h} \frac{1}{n_h^2} E(E(I_{hj}\overline{\hat{\tau}}_{hj}^2 | I_{hj}) + \sum_{j \neq j'}^{N_h} \frac{1}{n_h^2} E(E(I_{hj} I_{hj'} \overline{\hat{\tau}}_{hj} \overline{\hat{\tau}}_{hj'} | I_{hj}, I_{hj'})) \\
&= \sum_{j=1}^{N_h} \frac{1}{n_h^2} E(I_{hj}) E(\overline{\hat{\tau}}_{hj}^2) + \sum_{j \neq j'}^{N_h} \frac{1}{n_h^2} E(I_{hj} I_{hj'}) E(\overline{\hat{\tau}}_{hj} \overline{\hat{\tau}}_{hj'}) \\
&= \sum_{j=1}^{N_h} \frac{n_h}{N_h} [\text{var}(\overline{\hat{\tau}}) + (E(\overline{\hat{\tau}}_{hj}^2))^2] + \sum_{j \neq j'}^{N_h} \frac{n_h - 1}{n_h N_h (N_h - 1)} [\text{cov}(\overline{\hat{\tau}}_{hj} \overline{\hat{\tau}}_{hj'}) + \overline{\tau}_{hj} \overline{\tau}_{hj'}].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(s_h^2) &= \sum_{j=1}^{N_h} \frac{1}{N_h} \overline{\tau}_{hj}^2 - \sum_{j \neq j'}^{N_h} \frac{\overline{\tau}_{hj} \overline{\tau}_{hj'}}{N_h(N_h - 1)} + \sum_{j=1}^{N_h} \frac{1}{N_h} \text{var}(\overline{\hat{\tau}}_{hj}) - \sum_{j \neq j'}^{N_h} \frac{\text{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{N_h(N_h - 1)} \\
&= S_h^2 + \sum_{j=1}^{N_h} \frac{1}{N_h} \sum_{i=1}^{r} \left(\frac{\text{var}(\hat{\tau}_{hji})}{r^2} + \sum_{i \neq i'}^{r} \frac{\text{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'})}{r^2} \right) - \sum_{j \neq j'}^{N_h} \frac{\text{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{N_h(N_h - 1)} \\
&= S_h^2 + E \left[\frac{1}{n_h} \sum_{j=1}^{N_h} I_{hj} \left(\sum_{i=1}^{r} \frac{\widehat{\text{var}}(\hat{\tau}_{hji})}{r^2} + \sum_{i \neq i'}^{r} \frac{\widehat{\text{cov}}(\hat{\tau}_{hji}, \hat{\tau}_{hji'})}{r^2} \right) \right] - E \left[\sum_{j \neq j'}^{N_h} I_{hj} I_{hj'} \frac{\widehat{\text{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{n_h(n_h - 1)} \right]
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{\text{var}}(\hat{\tau}_{hji}) &= \sum_{k=1}^{M_{hji}} Z_{hjik} y_{hjik}^2 [\widehat{\Theta}_{hjik}^2 - \widehat{\Theta}_{hjik} - \widehat{\text{var}}(\widehat{\Theta}_{hjik})] \\
&+ \sum_{k=1}^{M_{hji}} Z_{hjik} y_{hjik}^2 \widehat{\text{var}}(\widehat{\Theta}_{hjik}) + \sum_{k \neq k'}^{M_{hji}} Z_{hjik} Z_{hjik'} y_{hjik} y_{hjik'} \widehat{\text{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}),
\end{aligned}
$$

$$
\widehat{\text{cov}}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) = \sum_{k=1}^{M_{hji}} \sum_{k'=1}^{M_{hji'}} Z_{hjik} Z_{hji'k'} y_{hjik} y_{hji'k'} \widehat{\text{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hji'k'}),
$$

and

$$
\widehat{\text{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'}) = \frac{1}{r^2} \sum_{i,i'=1}^{r} \sum_{k=1}^{M_{hji}} \sum_{k'=1}^{M_{hj'i'}} Z_{hjik} Z_{hj'i'k'} y_{hjik} y_{hj'i'k'} \widehat{\text{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hj'i'k'}),
$$

are unbiased estimators of $\text{var}(\hat{\tau}_{hji})$ (see section 3.2.2), and of $\text{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'})$ and $\text{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})$ previously derived in (3.4.1) and (3.4.2). Hence, an unbiased estimator for S_h^2 is

$$
\widehat{S}_h^2 = s_h^2 - \frac{1}{n_h} \sum_{j=1}^{N_h} I_{hj} \left(\sum_{i=1}^{r} \frac{\widehat{\text{var}}(\hat{\tau}_{hji})}{r^2} + \sum_{i \neq i'}^{r} \frac{\widehat{\text{cov}}(\hat{\tau}_{hji}, \hat{\tau}_{hji'})}{r^2} \right) + \sum_{j \neq j'}^{N_h} I_{hj} I_{hj'} \frac{\widehat{\text{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'})}{n_h(n_h - 1)},
$$

and thus, an unbiased estimator for the first error component v_1 is

$$\hat{v}_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) \widehat{S}_h^2.$$

The second error component due to sampling of groups is

$$v_2 = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{1}{r^2} \sum_{i=1}^{r} \sum_{k=1}^{M_{hj}} y_{hjik}^2 (\Theta_{hjik} - 1)$$

and an unbiased estimator for this component is

$$\hat{v}_2 = \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{j=1}^{N_h} \frac{I_{hj}}{r^2} \sum_{i=1}^{r} \sum_{k=1}^{M_{hj}} Z_{hjik} y_{hjik}^2 [\widehat{\Theta}_{hjik}^2 - \widehat{\Theta}_{hjik} - \widehat{\mathrm{var}}(\widehat{\Theta}_{hjik})].$$

The third error component due to estimation of sighting probabilities is

$$\begin{aligned} v_3 \;=\; & \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{1}{r^2} \left[\sum_{i=1}^{r} \left(\sum_{k=1}^{M_{hji}} \frac{y_{hjik}^2}{\Theta_{hjik}} \widehat{\mathrm{var}}(\widehat{\Theta}_{hjik}) + \sum_{k \neq k'}^{M_{hji}} \frac{y_{hjik} y_{hjik'}}{\Theta_{hjik} \Theta_{hjik'}} \widehat{\mathrm{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}) \right) \right] \\ + & \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \frac{1}{r^2} \sum_{i \neq i'}^{r} \mathrm{cov}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) + \sum_{h=1}^{L} \frac{N_h}{n_h} \frac{n_h - 1}{N_h - 1} \sum_{j \neq j'}^{N_h} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'}) + \sum_{h \neq h'}^{L} \sum_{j=1}^{N_h} \sum_{j'=1}^{N_{h'}} \mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'}) \end{aligned}$$

and an unbiased estimator for this error component is

$$\begin{aligned} \hat{v}_3 \;=\; & \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{j=1}^{N_h} \frac{I_{hj}}{r^2} \left[\sum_{i=1}^{r} \left(\sum_{k=1}^{M_{hji}} Z_{hjik} y_{hjik}^2 \widehat{\mathrm{var}}(\widehat{\Theta}_{hjik}) + \sum_{k \neq k'}^{M_{hji}} Z_{hjik} Z_{hjik'} y_{hjik} y_{hjik'} \widehat{\mathrm{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{hjik'}) \right) \right] \\ + & \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{j=1}^{N_h} \frac{I_{hj}}{r^2} \sum_{i \neq i'}^{r} \widehat{\mathrm{cov}}(\hat{\tau}_{hji}, \hat{\tau}_{hji'}) + \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{j \neq j'}^{N_h} I_{hj} I_{hj'} \widehat{\mathrm{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{hj'}) \\ + & \sum_{h \neq h'}^{L} \frac{N_h}{n_h} \frac{N_{h'}}{n_{h'}} \sum_{j=1}^{N_h} \sum_{j'=1}^{N_{h'}} I_{hj} I_{h'j'} \widehat{\mathrm{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'}) \end{aligned}$$

where

$$\widehat{\mathrm{cov}}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'}) = \frac{1}{r^2} \sum_{i,i'=1}^{r} \sum_{k=1}^{M_{hji}} \sum_{k'=1}^{M_{h'j'i'}} Z_{hjik} Z_{h'j'i'k'} y_{hjik} y_{h'j'i'k'} \widehat{\mathrm{cov}}(\widehat{\Theta}_{hjik}, \widehat{\Theta}_{h'j'i'k'})$$

is an unbiased estimator of $\mathrm{cov}(\overline{\hat{\tau}}_{hj}, \overline{\hat{\tau}}_{h'j'})$ previously derived in (3.5.3).

3.4.2.2 When sighting probabilities are known

We have

$$\hat{\tau}_{hji} = \sum_{k=1}^{M_{hj}} Z_{hjik} y_{hjik} \Theta_{hjik}$$

and

$$\begin{aligned}
\text{var}(\hat{\tau}) &= \sum_{h=1}^{L} N_h^2 \left[(\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \frac{1}{n_h} \sum_{i=1}^{r} \frac{\sum_{j=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hji})}{N_h}}{r^2} \right] \\
&= \sum_{h=1}^{L} \frac{N_h^2}{n_h} \left[S_h^2 + \sum_{i=1}^{r} \frac{\sum_{j=1}^{N_h} \frac{\text{var}(\hat{\tau}_{hji})}{N_h}}{r^2} \right] - \sum_{h=1}^{L} \frac{N_h^2}{N_h} S_h^2 \\
&= \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2 + \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \sum_{i=1}^{r} \frac{\text{var}(\hat{\tau}_{hji})}{r^2} \\
&= v_1 + v_2.
\end{aligned}$$

The first error component due to sampling of primary units is

$$v_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) S_h^2.$$

To find an unbiased estimator of S_h^2, consider s_h^2 as before. We have

$$\begin{aligned}
E(s_h^2) &= S_h^2 + \sum_{j=1}^{N_h} \frac{1}{N_h} \text{var}(\hat{\tau}_{hj}) \\
&= S_h^2 + \sum_{j=1}^{N_h} \frac{1}{N_h} \sum_{i=1}^{r} \frac{\text{var}(\hat{\tau}_{hji})}{r^2} \\
&= S_h^2 + E \left[\frac{1}{n_h} \sum_{j=1}^{N_h} I_{hj} \sum_{i=1}^{r} \frac{\widehat{var}(\hat{\tau}_{hji})}{r^2} \right].
\end{aligned}$$

Hence an unbiased estimator for S_h^2 is

$$\widehat{S}_h^2 = s_h^2 - \frac{1}{n_h} \sum_{j=1}^{n_h} \sum_{i=1}^{r} \frac{\widehat{var}(\hat{\tau}_{hji})}{r^2}.$$

Then an unbiased estimator for v_1 is

$$\hat{v}_1 = \sum_{h=1}^{L} N_h^2 (\frac{1}{n_h} - \frac{1}{N_h}) \widehat{S}_h^2.$$

The second error component due to sampling of groups is

$$v_2 = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{j=1}^{N_h} \sum_{i=1}^{r} \frac{\mathrm{var}(\hat{\tau}_{hji})}{r^2}.$$

An unbiased estimator for v_2 is thus,

$$\hat{v}_2 = \sum_{h=1}^{L} \frac{N_h^2}{n_h^2} \sum_{j=1}^{N_h} I_{hj} \sum_{i=1}^{r} \frac{\widehat{\mathrm{var}}(\hat{\tau}_{hji})}{r^2},$$

where $\widehat{\mathrm{var}}(\hat{\tau}_{hji}) = \sum_{k=1}^{M_{hji}} Z_{hjik} y_{hjik}^2 (\Theta_{hjik}^2 - \Theta_{hjik})$ (see (3.2.3)) is an unbiased estimator of $\mathrm{var}(\hat{\tau}_{hji}) = \sum_{k=1}^{M_{hji}} y_{hjik}^2 (\Theta_{hjik} - 1)$ (see (3.2.2)).

3.4.3 Optimum allocations

Optimum allocation of sampling effort assuming that survey costs are a linear function of stratum sample sizes (n_h), satisfies (Cochran, 1977)

$$\frac{n_h}{n} = \frac{N_h S_h^* / \sqrt{c_h}}{\sum_{h=1}^{L} N_h S_h^* / \sqrt{c_h}},$$

where $n_1 + n_2 + \ldots + n_h = n$ and c_h is the cost of sampling a primary unit, which may vary from stratum to stratum, and S_h^* depends on whether sighting probabilities are known or estimated via a logistic model.

3.4.3.1 When sighting probabilities are known

We have

$$S_h^{*2} = S_h^2 + \sum_{i=1}^{r} \frac{\sum_{j=1}^{N_h} \frac{\mathrm{var}(\hat{\tau}_{hji})}{N_h}}{r^2},$$

where $\mathrm{var}(\hat{\tau}_{hji}) = \sum_{k=1}^{M_{hj}} y_{hjik}^2 (\Theta_{hjik} - 1)$.

3.4.3.2 When sighting probabilities are unknown

When sighting probabilities are unknown and are assumed to satisfy a logistic model defined in (2.1.3) with all the assumptions described in section 3.3.2, we have

$$S_h^{*2} = S_h^2 + \sum_{j=1}^{N_h} \frac{\mathrm{var}(\overline{\tau}_{hj})}{N_h} - \sum_{j \neq j'}^{N_h} \frac{\mathrm{cov}(\overline{\tau}_{hj}, \overline{\tau}_{hj'})}{N_h(N_h - 1)}.$$

4. THIRD CENTRAL MOMENT OF THE MODIFIED HORVITZ-THOMPSON ESTIMATOR FOR POPULATION SIZE

4.1 Third central moment in terms of conditional moments

In this chapter, we derive the third central moment of the modified Horvitz-Thompson Estimators. The motivation of this derivation is to explore the possibility of improving confidence interval estimates of the population size by incorporating the third central moment. The following derivations are done analytically by hand.

Recall that in section 2.2, we obtained explicit expressions and developed unbiased estimators for variances of the population size estimators via conditional distributions of indicator functions I_i and Z_{ij}. We will use the same method to find the third central moment. We shall begin with deriving the third central moment of an arbitrary random variable X in terms of its conditional moments conditioned on another arbitrary variable Y.

Now

$$\begin{aligned} M_3(X) &= E[(X-E(X))^3] = E(X^3) - 3E(X)\text{var}(X) - (E(X))^3 \\ &= E(X^3) - 3E(X)E(X^2) + 2(E(X))^3. \end{aligned} \tag{4.1.1}$$

Then the third central moment of X conditioned on $Y = y$ is

$$\begin{aligned} M_3(X|Y) &= E[(X-E(X|Y))^3|Y] \\ &= E(X^3|Y) - 3E(X|Y)E(X^2|Y) + 2(E(X|Y))^3 \\ &= E(X^3|Y) - 3E(X|Y)\text{var}(X|Y) - (E(X|Y))^3, \end{aligned}$$

where $\mathrm{var}(X|Y) = E(X^2|Y) - (E(X|Y))^2$. So

$$E(M_3(X|Y)) = E(X^3) - 3E[E(X|Y)\mathrm{var}(X|Y)] - E[(E(X|Y))^3]. \qquad (4.1.2)$$

On the other hand,

$$\begin{aligned} M_3[E(X|Y)] &= E[E(X|Y) - E(X)]^3 \\ &= E[E(X|Y)]^3 - 3E(X)E[E(X|Y)]^2 + 2[E(X)]^3 \\ &= E[E(X|Y)]^3 - 3E(X)E(X^2) + 3E[E(X)\mathrm{var}(X|Y)] + 2[E(X)]^3, \end{aligned}$$

$$(4.1.3)$$

since in line 2, $[E(X|Y)]^2 = E(X^2|Y) - \mathrm{var}(X|Y)$. Therefore, the sum of equations (4.1.2) and (4.1.3) is

$$E(M_3(X|Y)) + M_3(E(X|Y)) = M_3(X) - 3\mathrm{cov}(\mathrm{var}(X|Y), E(X|Y)).$$

Hence,

$$M_3(X) = E(M_3(X|Y)) + M_3(E(X|Y)) + 3E[\mathrm{var}(X|Y)(E(X|Y) - E(X))]. \quad (4.1.4)$$

4.2 Third central moment of the modified Horvitz-Thompson estimator of population size

We will now proceed to find the third central moment of the modified Horvitz-Thompson estimators of the population size defined in (2.1.2) when sighting probabilities are known and in (2.1.4) when sighting probabilities are unknown.

4.2.1 When sighting probabilities are known

When sighting probabilities g_{ij}'s (see definition of notation in (2.1.1)), recall that the unbiased population estimator (see (2.1.2)) is

$$\hat{\tau}_\pi = \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} \frac{Z_{ij} y_{ij}}{g_{ij}}.$$

In addition to the set of notation introduced in (2.1.1), we define the joint probability

$$\pi_{i_1 i_2 i_3} = P(I_{i_1} = 1, I_{i_2} = 1, I_{i_3} = 1). \qquad (4.2.1)$$

to be the probability that primary units I_1, I_2 and I_3 are all selected in the first stage sampling. Following (4.1.4), we will have, firstly,

$$\begin{aligned}
E & \quad (M_3(\hat{\tau}_\pi|I)) \\
= & \quad E[E[(\hat{\tau}_\pi - E(\hat{\tau}_\pi|I))^3|I]] \\
= & \quad E[E[(\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} \frac{Z_{ij} y_{ij}}{g_{ij}} - \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} y_{ij})^3|I)] \\
= & \quad E[E[(\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} (\frac{Z_{ij}}{g_{ij}} - 1) y_{ij})^3|I]] \\
= & \quad E[E[\sum_{i=1}^{N} \frac{I_i}{\pi_i^3} (\sum_{j=1}^{M_i} (\frac{Z_{ij}}{g_{ij}} - 1) y_{ij})^3 \\
+ & \quad 3 \sum_{i_1 \neq i_2} \frac{I_{i_1} I_{i_2}}{\pi_{i_1} \pi_{i_2}^2} (\sum_{j=1}^{M_{i_1}} (\frac{Z_{i_1 j}}{g_{i_1 j}} - 1) y_{i_1 j})(\sum_{j=1}^{M_{i_2}} (\frac{Z_{i_2 j}}{g_{i_2 j}} - 1) y_{i_2 j})^2 \\
+ & \quad \sum_{i_1 \neq i_2 \neq i_3} \frac{I_{i_1} I_{i_2} I_{i_3}}{\pi_{i_1} \pi_{i_2} \pi_{i_3}} (\sum_{j=1}^{M_{i_1}} (\frac{Z_{i_1 j}}{g_{i_1 j}} - 1) y_{i_1 j})(\sum_{j=1}^{M_{i_2}} (\frac{Z_{i_2 j}}{g_{i_2 j}} - 1) y_{i_2 j})(\sum_{j=1}^{M_{i_3}} (\frac{Z_{i_3 j}}{g_{i_3 j}} - 1) y_{i_3 j})|I]].
\end{aligned}$$

(4.2.2)

Since the Z_{ij}'s are independent and $E(Z_{ij}|I_i) = g_{ij}$, the above expression is simply

$$\begin{aligned}
E(M_3(\hat{\tau}_\pi|I)) &= E[E[\sum_{i=1}^{N} \frac{I_i}{\pi_i^3} (\sum_{j=1}^{M_i} (\frac{Z_{ij}}{g_{ij}} - 1) y_{ij})^3|I]] \\
&= E[E[\sum_{i=1}^{N} \frac{I_i}{\pi_i^3} (\sum_{j=1}^{M_i} (\frac{Z_{ij}}{g_{ij}} - 1)^3 y_{ij}^3 + 3 \sum_{j_1 \neq j_2} (\frac{Z_{ij_1}}{g_{ij_1}} - 1)(\frac{Z_{ij_2}}{g_{ij_2}} - 1)^2 y_{ij_1} y_{ij_2}^2 \\
&\qquad + \textstyle\sum_{j_1 \neq j_2 \neq j_3} (\frac{Z_{ij_1}}{g_{ij_1}} - 1)(\frac{Z_{ij_2}}{g_{ij_2}} - 1)(\frac{Z_{ij_3}}{g_{ij_3}} - 1) y_{ij_1} y_{ij_2} y_{ij_3})|I]] \\
&= \textstyle\sum_{i=1}^{N} \frac{1}{\pi_i^2} \sum_{j=1}^{M_i} \frac{1}{g_{ij}^3} E[(Z_{ij} - g_{ij})^3|I] y_{ij}^3 = \sum_{i=1}^{N} \frac{1}{\pi_i^2} \sum_{j=1}^{M_i} \frac{1 - 3g_{ij} + 2g_{ij}^2}{g_{ij}^2} y_{ij}^3.
\end{aligned}$$

(4.2.3)

Secondly,

$$
\begin{aligned}
M_3(E(\hat{\tau}_\pi|I)) &= M_3(\sum_{i=1}^{N}\frac{I_i}{\pi_i}\tau_i)\\
&= E((\sum_{i=1}^{N}\frac{I_i}{\pi_i}\tau_i)^3) - 3\tau\mathrm{var}(\sum_{i=1}^{N}\frac{I_i}{\pi_i}\tau_i) - \tau^3\\
&= E(\sum_{i=1}^{N}\frac{I_i}{\pi_i^3}\tau_i^3 + 3\sum_{i_1\neq i_2}\frac{I_{i_1}I_{i_2}}{\pi_{i_1}\pi_{i_2}^2}\tau_{i_1}\tau_{i_2}^2 + \sum_{i_1\neq i_2\neq i_3}\frac{I_{i_1}I_{i_2}I_{i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}}\tau_{i_1}\tau_{i_2}\tau_{i_3})\\
&- 3(\sum_{i=1}^{N}\tau_i)(\sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i}\tau_i^2 + \sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}-\pi_{i_1}\pi_{i_2}}{\pi_{i_1}\pi_{i_2}}\tau_{i_1}\tau_{i_2}) - (\sum_{i=1}^{N}\tau_i)^3\\
&= \sum_{i=1}^{N}\frac{1}{\pi_i^2}\tau_i^3 + 3\sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}}{\pi_{i_1}\pi_{i_2}^2}\tau_{i_1}\tau_{i_2}^2 + \sum_{i_1\neq i_2\neq i_3}\frac{\pi_{i_1i_2i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}}\tau_{i_1}\tau_{i_2}\tau_{i_3}\\
&- 3(\sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i}\tau_i^3 + \sum_{i_1\neq i_2}^{N}\frac{1-\pi_{i_2}}{\pi_{i_2}}\tau_{i_1}\tau_{i_2}^2 + 2\sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}-\pi_{i_1}\pi_{i_2}}{\pi_{i_1}\pi_{i_2}}\tau_{i_1}\tau_{i_2}^2\\
&+ \sum_{i_1\neq i_2\neq i_3}\frac{\pi_{i_2i_3}-\pi_{i_2}\pi_{i_3}}{\pi_{i_2}\pi_{i_3}}\tau_{i_1}\tau_{i_2}\tau_{i_3}) - \sum_{i=1}^{N}\tau_i^3 - 3\sum_{i_1\neq i_2}\tau_{i_1}\tau_{i_2}^2 - \sum_{i_1\neq i_2\neq i_3}\tau_{i_1}\tau_{i_2}\tau_{i_3}\\
&= \sum_{i=1}^{N}\frac{1-3\pi_i+2\pi_i^2}{\pi_i^2}\tau_i^3 + \sum_{i_1\neq i_2}\frac{3(\pi_{i_1i_2}-\pi_{i_1}\pi_{i_2})(1-2\pi_{i_2})}{\pi_{i_1}\pi_{i_2}^2}\tau_{i_1}\tau_{i_2}^2\\
&+ \sum_{i_1\neq i_2\neq i_3}\frac{\pi_{i_1i_2i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}} - 3\frac{\pi_{i_2i_3}-\pi_{i_2}\pi_{i_3}}{\pi_{i_1}\pi_{i_2}^2} - 1.
\end{aligned}
\tag{4.2.4}
$$

Finally, the last term of (4.1.4) is three times

$$
E[\mathrm{var}(\hat{\tau}_\pi|I)(E(\hat{\tau}_\pi|I) - E(\tau_\pi))] = E[(\sum_{i=1}^{N}\frac{I_i}{\pi_i^2}\sum_{j=1}^{M_i}\frac{1-g_{ij}}{g_{ij}}y_{ij}^2)(\sum_{i=1}^{N}\frac{I_i}{\pi_i}\tau_i - \tau)]
$$

$$
\begin{aligned}
&= E[\sum_{i=1}^{N}\frac{I_i}{\pi_i^3}\tau_i\sum_{j=1}^{M_i}\frac{1-g_{ij}}{g_{ij}}y_{ij}^2 + \sum_{i_1\neq i_2}\frac{I_{i_1}I_{i_2}}{\pi_{i_1}^2\pi_{i_2}}\tau_{i_2}\sum_{j=1}^{M_{i_1}}\frac{1-g_{i_1j}}{g_{i_1j}}y_{i_1j}^2 - \tau(\sum_{i=1}^{N}\frac{I_i}{\pi_i^2}\sum_{j=1}^{M_i}\frac{1-g_{ij}}{g_{ij}}y_{ij}^2)]\\
&= \sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i^2}\tau_i\sum_{j=1}^{M_i}\frac{1-g_{ij}}{g_{ij}}y_{ij}^2 + \sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}-\pi_{i_1}\pi_{i_2}}{\pi_{i_1}\pi_{i_2}^2}\tau_{i_1}\sum_{j=1}^{M_{i_2}}\frac{1-g_{i_2j}}{g_{i_2j}}y_{i_2j}^2\\
&= \sum_{i=1}^{N}\frac{1-\pi_i}{\pi_i^2}[\sum_{j=1}^{M_i}\frac{1-g_{ij}}{g_{ij}}y_{ij}^3 + \sum_{j_1\neq j_2}\frac{1-g_{ij_1}}{g_{ij_1}}y_{ij_1}^2y_{ij_2}] + \sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}-\pi_{i_1}\pi_{i_2}}{\pi_{i_1}\pi_{i_2}^2}\sum_{j_1=1}^{M_{i_1}}\sum_{j_2=1}^{M_{i_2}}\frac{1-g_{i_2j_2}}{g_{i_2j_2}}y_{i_1j_1}y_{i_2j_2}^2.
\end{aligned}
\tag{4.2.5}
$$

Hence the third central moment of the population size estimator is

$$M_3(\hat{\tau}_\pi) = (4.2.3) + (4.2.4) + 3 \times (4.2.5).$$

4.2.2 When sighting probabilities are unknown

When sighting probabilities are unknown and are assumed to satisfy a logistic model defined in (2.1.3), the population size estimator (see (2.1.4)) is

$$\hat{\tau}_{LR} = \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} y_{ij} \widehat{\Theta}_{ij}.$$

Following (4.1.4), the third central moment of $\hat{\tau}_{LR}$ is, firstly,

$$\begin{aligned}
E(M_3(\hat{\tau}_{LR}|I,Z)) &= E[E[(\hat{\tau}_{LR} - E(\hat{\tau}_{LR}|I,Z))^3|I,Z]] \\
&= E[E[(\textstyle\sum_{i=1}^N \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij}y_{ij}(\widehat{\Theta}_{ij} - \Theta_{ij}))^3|I]] \\
&= E[E[\textstyle\sum_{i=1}^N \frac{I_i}{\pi_i^3}(\sum_{j=1}^{M_i} Z_{ij}y_{ij}(\widehat{\Theta}_{ij} - \Theta_{ij}))^3 \\
&\quad +3\textstyle\sum_{i_1\neq i_2} \frac{I_{i_1}I_{i_2}}{\pi_{i_1}\pi_{i_2}^2}(\sum_{j=1}^{M_{i_1}} Z_{i_1j}y_{i_1j}(\widehat{\Theta}_{i_1j} - \Theta_{i_1j})^2(\sum_{j=1}^{M_{i_2}} Z_{i_2j}y_{i_2j}(\widehat{\Theta}_{i_2j} - \Theta_{i_2j})) \\
&\quad +\textstyle\sum_{i_1\neq i_2\neq i_3} \frac{I_{i_1}I_{i_2}I_{i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}}(\sum_{j=1}^{M_{i_1}} Z_{i_1j}y_{i_1j}(\widehat{\Theta}_{i_1j} - \Theta_{i_1j})(\sum_{j=1}^{M_{i_2}} Z_{i_2j}y_{i_2j}(\widehat{\Theta}_{i_2j} - \Theta_{i_2j}) \\
&\qquad (\textstyle\sum_{j=1}^{M_{i_3}} Z_{i_3j}y_{i_3j}(\widehat{\Theta}_{i_3j} - \Theta_{i_3j})|I,Z]],
\end{aligned} \tag{4.2.6}$$

With the definitions in section 2.1.2, we note that

$$\widehat{\Theta}_{ij} - \Theta_{ij} = e^{-\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{-\boldsymbol{x}'_{ij}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}/2} - 1)$$

and that $E(e^{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})}) = e^{\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}/2}$. Therefore

$$\begin{aligned}
E\ \ (\widehat{\Theta}_{ij} - \Theta_{ij})^3 &= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}E(e^{-\boldsymbol{x}'_{ij}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}/2} - 1)^3 \\
&= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}E(e^{-3\boldsymbol{x}'_{ij}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{3}{2}\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 3e^{-2\boldsymbol{x}_{ij}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} \\
&+ 3e^{-\boldsymbol{x}_{ij}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}/2} - 1) \\
&= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{3\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 3e^{\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} + 2) \\
&= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{3\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 1) + e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(-3e^{\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} + 3) \\
&= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{3\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 1) + 3(-e^{-\boldsymbol{x}'_{ij}\boldsymbol{\beta}})e^{-2\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 1) \\
&= e^{-3\boldsymbol{x}'_{ij}\boldsymbol{\beta}}(e^{3\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}} - 1) + 3(1 - \Theta_{ij})\mathrm{var}(\widehat{\Theta}_{ij}).
\end{aligned} \tag{4.2.7}$$

Next

$$E(\widehat{\Theta}_{ij_1} - \Theta_{ij_1})^2(\widehat{\Theta}_{ij_2} - \Theta_{ij_2})$$

$$= e^{-2\boldsymbol{x}'_{ij_1}\boldsymbol{\beta}-\boldsymbol{x}'_{ij_2}\boldsymbol{\beta}} E[(e^{-\boldsymbol{x}'_{ij_1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}/2} - 1)^2(e^{-\boldsymbol{x}_{ij_2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}/2} - 1)]$$

$$= e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}} E[e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}-\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}/2} - e^{-2\boldsymbol{x}'_{ij_1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}$$

$$-2e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}/2-\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}/2} + 2e^{-\boldsymbol{x}'_{ij_1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}/2}$$

$$+e^{-\boldsymbol{x}'_{ij_2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}/2} - 1]$$

$$= e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}+2\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}} - 2e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} + 2)$$

$$= e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}+2\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - 1) - e^{-\boldsymbol{x}'_{ij_2}\boldsymbol{\beta}}e^{-2\boldsymbol{x}'_{ij_1}\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}} - 1)$$

$$+2(-e^{-\boldsymbol{x}'_{ij_1}\boldsymbol{\beta}})e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - 1)$$

$$= e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}+2\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - 1) + (1 - \Theta_{ij_2})\mathrm{var}(\widehat{\Theta}_{ij_1})$$

$$+2(1 - \Theta_{ij_1})\mathrm{cov}(\widehat{\Theta}_{ij_1}, \widehat{\Theta}_{ij_2}),$$

Finally

$$E(\widehat{\Theta}_{ij_1} - \Theta_{ij_1})(\widehat{\Theta}_{ij_2} - \Theta_{ij_2})(\widehat{\Theta}_{ij_3} - \Theta_{ij_3})$$

$$= e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}} E[(e^{-\boldsymbol{x}'_{ij_1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}{2}} - 1)(e^{-\boldsymbol{x}'_{ij_2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}}{2}} - 1)$$

$$*(e^{-\boldsymbol{x}'_{ij_3}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_3}\Sigma\boldsymbol{x}_{ij_3}}{2}} - 1)]$$

$$= e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}} E[e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}{2}-\frac{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}}{2}-\frac{\boldsymbol{x}'_{ij_3}\Sigma\boldsymbol{x}_{ij_3}}{2}}$$

$$-e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}{2}-\frac{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}}{2}} - e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_3})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}{2}-\frac{\boldsymbol{x}'_{ij_3}\Sigma\boldsymbol{x}_{ij_3}}{2}}$$

$$-e^{-(\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}}{2}-\frac{\boldsymbol{x}'_{ij_3}\Sigma\boldsymbol{x}_{ij_3}}{2}} + e^{-\boldsymbol{x}'_{ij_1}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}{2}}$$

$$+e^{-\boldsymbol{x}'_{ij_2}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_2}}{2}} + e^{-\boldsymbol{x}'_{ij_3}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})-\frac{\boldsymbol{x}'_{ij_3}\Sigma\boldsymbol{x}_{ij_3}}{2}} - 1]$$

$$= e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}[e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}+\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}+\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}} - e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}} - e^{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}} + 1 + 1 + 1 - 1]$$

$$= e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}+\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}+\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}} + 2 - 3) + (-e^{-\boldsymbol{x}'_{ij_1}\boldsymbol{\beta}})e^{-(\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}} - 1)$$

$$+(-e^{-\boldsymbol{x}'_{ij_2}\boldsymbol{\beta}})e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}} - 1) + (-e^{-\boldsymbol{x}'_{ij_3}\boldsymbol{\beta}})e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}} - 1)$$

$$= e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}+\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}+\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}} + 1)$$

$$+(1-\Theta_{ij_1})\mathrm{cov}(\widehat{\Theta}_{ij_2},\widehat{\Theta}_{ij_3})+(1-\Theta_{ij_2})\mathrm{cov}(\widehat{\Theta}_{ij_1},\widehat{\Theta}_{ij_3})+(1-\Theta_{ij_3})\mathrm{cov}(\widehat{\Theta}_{ij_1},\widehat{\Theta}_{ij_2}).$$

Hence, in (4.2.6) which is the first term of the third central moment, we have

$$E[E[\sum_{i=1}^{N}\frac{I_i}{\pi_i^3}(\sum_{j=1}^{M_i}Z_{ij}(\widehat{\Theta}_{ij}-\Theta_{ij})y_{ij})^3|I,Z]]$$

$$=\sum_{i=1}^{N}\frac{1}{\pi_i^2}(\sum_{j=1}^{M_i}\frac{y_{ij}^3}{\Theta_{ij}}[3(1-\Theta_{ij})\mathrm{var}(\widehat{\Theta}_{ij})+e^{-3\boldsymbol{x}'\boldsymbol{\beta}}(e^{3\boldsymbol{x}'_{ij}\Sigma\boldsymbol{x}_{ij}}-1)]$$

$$+3\sum_{j_1\neq j_2}\frac{y_{ij_1}^2y_{ij_2}}{\Theta_{ij_1}\Theta_{ij_2}}[2(1-\Theta_{ij_1})\mathrm{cov}(\widehat{\Theta}_{ij_1},\widehat{\Theta}_{ij_2})+(1-\Theta_{ij_2})\mathrm{var}(\widehat{\Theta}_{ij_1})$$
$$+e^{-(2\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2})'\boldsymbol{\beta}}(e^{2\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}+\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_1}}-1)]$$

$$+\sum_{j_1\neq j_2\neq j_3}\frac{y_{ij_1}y_{ij_2}y_{ij_3}}{\Theta_{ij_1}\Theta_{ij_2}\Theta_{ij_3}}[(1-\Theta_{ij_1})\mathrm{cov}(\widehat{\Theta}_{ij_2},\widehat{\Theta}_{ij_3})+(1-\Theta_{ij_2})\mathrm{cov}(\widehat{\Theta}_{ij_1},\widehat{\Theta}_{ij_3})+(1-\Theta_{ij_3})\mathrm{cov}(\widehat{\Theta}_{ij_1},\widehat{\Theta}_{ij_2})$$

$$+e^{-(\boldsymbol{x}_{ij_1}+\boldsymbol{x}_{ij_2}+\boldsymbol{x}_{ij_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_2}+\boldsymbol{x}'_{ij_1}\Sigma\boldsymbol{x}_{ij_3}+\boldsymbol{x}'_{ij_2}\Sigma\boldsymbol{x}_{ij_3}}-1)]),\tag{4.2.8}$$

Similarly,

$$E[E[\sum_{i_1\neq i_2}\frac{I_{i_1}I_{i_2}}{\pi_{i_1}\pi_{i_2}^2}(\sum_{j=1}^{M_{i_1}}Z_{i_1j}y_{i_1j}(\widehat{\Theta}_{i_1j}-\Theta_{i_1j}))^2(\sum_{j=1}^{M_{i_2}}Z_{i_2j}y_{i_2j}(\widehat{\Theta}_{i_2j}-\Theta_{i_2j}))|I,Z]]$$

$$=\sum_{i_1\neq i_2}\frac{\pi_{i_1i_2}}{\pi_{i_1}\pi_{i_2}}(\sum_{j_1=1}^{M_{i_1}}\sum_{j_2=1}^{M_{i_2}}\frac{y_{i_1j_1}^2y_{i_2j_2}}{\Theta_{i_1j_1}\Theta_{i_2j_2}}[2(1-\Theta_{i_1j_1})\mathrm{cov}(\widehat{\Theta}_{i_1j_1},\widehat{\Theta}_{i_2j_2})+(1-\Theta_{i_2j_2})\mathrm{var}(\widehat{\Theta}_{i_1j_1})$$
$$+e^{-(2\boldsymbol{x}_{i_1j_1}+\boldsymbol{x}_{i_2j_2})'\boldsymbol{\beta}}(e^{2\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_2j_2}+\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_1j_1}}-1)]$$
$$+\sum_{j_1=1}^{M_{i_1}}\sum_{j_2\neq j_3}^{M_{i_2}}\frac{y_{i_1j_1}y_{i_2j_2}y_{i_2j_3}}{\Theta_{i_1j_1}\Theta_{i_2j_2}\Theta_{i_2j_3}}[(1-\Theta_{i_1j_1})\mathrm{cov}(\widehat{\Theta}_{i_2j_2},\widehat{\Theta}_{i_2j_3})+(1-\Theta_{i_2j_2})\mathrm{cov}(\widehat{\Theta}_{i1j_1},\widehat{\Theta}_{i_2j_3})$$
$$+(1-\Theta_{i_2j_3})\mathrm{cov}(\widehat{\Theta}_{i_1j_1},\widehat{\Theta}_{i_2j_2})$$
$$+e^{-(\boldsymbol{x}_{i_1j_1}+\boldsymbol{x}_{i_2j_2}+\boldsymbol{x}_{i_2j_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_2j_2}+\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_2j_3}+\boldsymbol{x}'_{i_2j_2}\Sigma\boldsymbol{x}_{i_2j_3}}-1)]).\tag{4.2.9}$$

Also,

$$E[E[\sum_{i_1\neq i_2\neq i_3}\frac{I_{i_1}I_{i_2}I_{i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}}(\sum_{j=1}^{M_{i_1}}Z_{i_1j}(\widehat{\Theta}_{i_1j}-\Theta_{i_1j})y_{i_1j})(\sum_{j=1}^{M_{i_2}}Z_{i_2j}(\widehat{\Theta}_{i_2j}-\Theta_{i_2j})y_{i_2j})(\sum_{j=1}^{M_{i_3}}Z_{i_3j}(\widehat{\Theta}_{i_3j}-\Theta_{i_3j})y_{i_3j})|I,Z]]$$

$$=\sum_{i_1\neq i_2\neq i_3}\frac{\pi_{i_1i_2i_3}}{\pi_{i_1}\pi_{i_2}\pi_{i_3}}\sum_{j_1=1}^{M_{i_1}}\sum_{j_2=1}^{M_{i_2}}\sum_{j_3=1}^{M_{i_3}}\frac{y_{i_1j_1}y_{i_2j_2}y_{i_3j_3}}{\Theta_{i_1j_1}\Theta_{i_2j_2}\Theta_{i_3j_3}}[(1-\Theta_{i_1j_1})\mathrm{cov}(\widehat{\Theta}_{i_2j_2},\widehat{\Theta}_{i_3j_3})+(1-\Theta_{i_2j_2})\mathrm{cov}(\widehat{\Theta}_{i_1j_1},\widehat{\Theta}_{i_3j_3})$$
$$+(1-\Theta_{i_3j_3})\mathrm{cov}(\widehat{\Theta}_{i_1j_1},\widehat{\Theta}_{i_2j_2})$$
$$+e^{-(\boldsymbol{x}_{i_1j_1}+\boldsymbol{x}_{i_2j_2}+\boldsymbol{x}_{i_3j_3})'\boldsymbol{\beta}}(e^{\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_2j_2}+\boldsymbol{x}'_{i_1j_1}\Sigma\boldsymbol{x}_{i_3j_3}+\boldsymbol{x}'_{i_2j_2}\Sigma\boldsymbol{x}_{i_3j_3}}-1)].\tag{4.2.10}$$

The second term of the third central moment (see (4.1.4)) of $\hat{\tau}_{LR}$ is

$$M_3(E(\hat{\tau}_{LR}|I,Z)) = M_3(\hat{\tau}_\pi).$$

Finally, the last term of the third central moment of $\hat{\tau}_{LR}$ is three times

$$E[\text{var}(\hat{\tau}_{LR}|I,Z)(E(\hat{\tau}_{LR}|I,Z) - E(\hat{\tau}_{LR}))] = E[\text{var}(\hat{\tau}_{LR}|I,Z)(\hat{\tau}_\pi - \tau)], \quad (4.2.11)$$

where

$$\text{var}(\hat{\tau}_{LR}|I,Z) = \sum_{i=1}^{N} \frac{I_i}{\pi_i^2}[\sum_{j=1}^{M_i} Z_{ij} y_{ij}^2 \text{var}(\Theta_{ij}) + \sum_{j_1 \neq j_2} Z_{ij_1} Z_{ij_2} y_{ij_1} y_{ij_2} \text{cov}(\widehat{\Theta}_{ij_1}, \widehat{\Theta}_{ij_2})]$$
$$+ \textstyle\sum_{i_1 \neq i_2} \frac{I_{i_1} I_{i_2}}{\pi_{i_1} \pi_{i_2}} \sum_{j_1=1}^{M_{i_1}} \sum_{j_2=1}^{M_{i_2}} Z_{i_1 j_1} Z_{i_2 j_2} y_{i_1 j_1} y_{i_2 j_2} \text{cov}(\widehat{\Theta}_{i_1 j_1}, \widehat{\Theta}_{i_2 j_2}).$$

Hence, the third central moment of the population size estimator is

$$M_3(\hat{\tau}_{LR}) = [(4.2.8) + 3 \times (4.2.9) + (4.2.10)] + M_3(\hat{\tau}_\pi) + 3 \times (4.2.11).$$

The complexity of those expressions of the third central moment, especially with estimated sighting probabilities, makes it difficult to find an unbiased (asymptotically) estimator for the third central moment. One possible bias estimator of the third central moment is obtained by simply replacing true values of τ_i, $\boldsymbol{\beta}$, Σ, Θ, variances and covariances of Θ's by their (asymptotically) unbiased estimators (see sections 2.1, 2.2 and 3.2) in expressions of $M_3(\hat{\tau}_\pi)$ and $M_3(\hat{\tau}_{LR})$. If it is possible to improve the confidence interval estimates for the population size, it would be worthwhile to attempt to find an (asymptotically) unbiased estimator for the third central moment of the population size estimator.

5. SIMULATION STUDY OF THE POPULATION SIZE ESTIMATOR AND THE UNBIASED ESTIMATOR OF THE VARIANCE OF THE POPULATION SIZE ESTIMATOR

5.1 Introduction

In this chapter, we want to evaluate small sample size performance of the modified Horvitz-Thompson estimator of the population size and small sample size performance of asymptotically unbiased estimators for the variance components and the variance of the population size estimator. As error (or variation) due to sampling of primary units can be controlled by the choice of sampling units and sampling design (Samuel and Steinhorst, 1989, Caughley, 1977), we consider the special case of a complete census of primary units (see section 3.2) to get a further insight of the control of error due to sampling of groups and error due to estimation of sighting probabilities. We will assume, for our simulated population, that the sighting probabilities $\{g_j\}$'s satisfy the logistic model defined in (3.2.4). The population size estimators in this special case are stated in (3.2.1) when sighting probabilities are known and in (3.2.5) when unknown sighting probabilities are estimated via logistic regression. Variance of the population size estimator is (3.2.2) when sighting probabilities are known and when unknown sighting probabilities are estimated via logistic regression, variance of the population size estimator is the sum of the first error component due to sampling of groups expressed in (3.2.6) and the second error component due to estimating the sighting probabilities expressed in (3.2.8). An (asymptotically) unbiased estimator for the variance is (3.2.3) when sighting probabilities are known and is the sum of (3.2.7) and (3.2.9), respectively, when unknown sighting probabilities are estimated via logistic regression. Expressions given in (3.2.7) and (3.2.9) are (asymptotically) unbiased estimators for the first and second error components respectively.

We will only simulate the scenario where the needed data are collected in 2 phases (see subsection 2.3.1), i.e. conducting sightability and population surveys separately. Phase I data collection or sightability survey is for estimating the sightability model and phase II data collection or population survey is for estimating the population size via the fitted model obtained in phase I. So a single population size estimate requires two independenet samples - a phase I sample and a phase II sample. We will use the fitted model of the data collected from five aerial surveys of elk in Pennsylvania as the true sightability model of our simulated population (see section 5.2). A sighting trial, in both phases, is the process through which a sample is collected. During a simulated sighting trial, $U(0,1)$ random numbers $\{u_j\}$ are generated, jointly independent, for the groups. Group j is sighted if u_j is less than its true sighting probability g_j.

We use SAS for programming the simulation. A sample of the program is listed in the Appendix. Models are fitted using procedure LOGISTIC. This procedure gives results of measures including the likelihood ratio test and Akaike Information Criteria (AIC), estimates of parameters $\widehat{\boldsymbol{\beta}}$ of the fitted model and estimated covariance matrix $\widehat{\Sigma}$ of the estimated parameters.

5.1.1 Measure performances of estimators by bias

The performance of the population size estimator will be measured by bias of the mean of a set of simulated population size estimates relative to the true population size of the simulated population. To measure the performances of asymptotically unbiased estimators of the error components and the variance by bias, we need to approximate the true covariance matrix Σ of the estimated parameters $\widehat{\boldsymbol{\beta}}$ of the sightability model. Two approximations were used. First, we let Σ be $\widehat{\Sigma}_1$, the sample covariance matrix of a set of estimated parameters $\widehat{\boldsymbol{\beta}}$ and second, we let Σ be $\widehat{\boldsymbol{\Sigma}}_2$, the mean of the set of estimated covariance matrices $\widehat{\Sigma}$ of those estimated parameters. Sample covariance matrix $\widehat{\Sigma}_1$ is an unbiased and a consistent estimator of the true covariance matrix Σ. By the weak law of large numbers, the mean of a set of covariance matrix estimates, $\widehat{\Sigma_2}$, converges in probability to the expectation of the covariance matrix estimator. The covariance matrix estimator in this case is a maximum likelihood estimator and thus, is a consistent estimator of the true covariance matrix Σ.

Cochran (1977) stated that "the effect of bias on the accuracy of an estimate is negligible if the bias is less than one tenth of the standard deviation of the estimate. If we have a biased method of estimation for which $B/\sigma < 0.1$, where B is the absolute value of the bias, it can be claimed that the bias is not an appreciabble disadvantage of the method. Even with $B/\sigma = 0.2$, the disturbance in the probability of the total error is modest." We will use this working rule to evaluate the significance of bias in estimating the population size.

Cochran (1977) also stated that we usually measure the precision instead of the accuracy of estimates. Precision refers to the standard deviation of the observed (or sampling) variance of the estimates. We will report the observed variance of simulated population size estimates and compute bias of the mean of a set of simulated variance estimates of the population size estimator relative to the sample variance of the corresponding set of population size estimates. Another way of evaluating $\widehat{var}(\hat{\tau}_{LR})$ is to measure bias of $\widehat{var}(\hat{\tau}_{LR})$ relative to the sample (or observed) variance of the corresponding set of population size estimates. By the weak law of large numbers, as stated before, the sample variance is expected to be close (asymptotically) to the true variance when the number of population estimation trials, denoted by k, is large.

5.1.2 Construct confidence interval estimate for population size

We further enhance our evaluation of the estimators by constructing 90% and a 95% confidence interval estimates for population size and study the coverage rate of these confidence interval estimates.

We will use the large sample theory (normality) to construct confidence interval estimates for the true population size τ. Furthermore, as indicated in section 1.4, in an earlier trial run of our simulation, we observed that the sample distribution of population size estimates is positively skewed. We have therefore considered four transformations of the population size estimate, namely, the natural log $\ln\hat{\tau}_{LR}$, square root of the reciprocal $\hat{\tau}^{-1/2}$, the reciprocal $\hat{\tau}^{-1}$ and power of three halves of the reciprocal $\hat{\tau}^{-3/2}$. We construct the confidence interval estimates by assuming large sample theory for sample distributions of transformed population size estimates. To obtain an approximate variance of the transformed population size estimator, we apply the delta method. For example,

$$\ln \hat{\tau}_{LR} \approx \ln \tau + \frac{1}{\tau}(\hat{\tau}_{LR} - \tau).$$

or

$$\text{var}(\ln \hat{\tau}_{LR}) \approx \frac{1}{\tau^2}\text{var}(\hat{\tau}_{LR}).$$

The interval $\ln \hat{\tau}_{LR} \pm z_{\alpha/2}\sqrt{\frac{\widehat{\text{var}}(\hat{\tau}_{LR})}{\hat{\tau}^4_{LR}}}$ is then back-transformed. Hence, we obtain an approximate $100(1-\alpha)\%$ confidence interval estimate $(\hat{\tau}e^{-z_{\alpha/2}cv}, \hat{\tau}e^{z_{\alpha/2}cv})$ for the true population size τ, where $\hat{\tau} = \hat{\tau}_{LR}$, $cv^2 = \widehat{\text{var}}(\hat{\tau})/\hat{\tau}^2$ and $z_{\alpha/2}$ is the $100(1-\frac{\alpha}{2})^{th}$ percentile of the standard normal distribution. Likewise, we obtain three other approximate $100(1-\alpha)\%$ confidence interval estimates for τ from the other three transformations of the population size estimator. They are back-transformed as

$(\frac{1}{\hat\tau} - \frac{cv}{\hat\tau} z_{\alpha/2}, \frac{1}{\hat\tau} + \frac{cv}{\hat\tau} z_{\alpha/2})$ from the reciprocal transformation, $(\frac{1}{\sqrt{\hat\tau}} - \frac{cv}{2\sqrt{\hat\tau}} z_{\alpha/2}, \frac{1}{\sqrt{\hat\tau}} + \frac{cv}{2\sqrt{\hat\tau}} z_{\alpha/2})$ from the square root of the reciprocal transformation and $(\frac{1}{\hat\tau^{3/2}} - \frac{3cv}{2\hat\tau^{3/2}} z_{\alpha/2}, \frac{1}{\hat\tau^{3/2}} + \frac{3cv}{2\hat\tau^{3/2}} z_{\alpha/2})$ from the three halves of the reciprocal transformation of population size estimates respectively.

In addition (also stated in section 1.4), we have assumed that the sampling distribution of $(\hat\tau_{LR} - T)$, where T is the total number of sighted animals in a phase II sample, is lognormally distributed. This was suggested by Dr. Kenneth Burnham of the Colorado Cooperative Fish & Wildlife Research Unit, Colorado State University. In this case, exact expressions of expectation and variance of $\ln(\hat\tau_{LR} - T)$ can be worked out as follows.

We know that if a random variable, say X, is lognormally distributed or $\ln X$ is normally distributed as $\ln X \sim N(\mu, \sigma^2)$, its expectation and variance are

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}, \quad \text{var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).$$

With $X = \hat\tau_{LR} - T$, we therefore obtain the following equations,

$$E(\hat\tau_{LR} - T) = \tau - T = e^{\mu + \frac{\sigma^2}{2}},$$

and

$$\text{var}(\hat\tau_{LR} - T) = \text{var}(\hat\tau_{LR}) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).$$

Solving this pair of equations for μ and σ^2, we have

$$\sigma^2 = \ln\left(1 + \frac{\text{var}(\hat\tau_{LR})}{(\tau - T)^2}\right) = \ln(1 + cv^2),$$

where $cv^2 = \frac{\text{var}(\hat\tau)}{(\tau - T)^2}$, and

$$\begin{aligned} \mu &= \ln(\tau - T) - \frac{\sigma^2}{2} \\ &= \ln(\tau - T) - \frac{1}{2}\ln(1 + cv^2) \\ &= \ln\frac{\tau - T}{\sqrt{1 + cv^2}}. \end{aligned}$$

Thus,

$$\frac{\ln rm(\hat{\tau} - T) - \ln \frac{\tau - T}{\sqrt{1+cv^2}}}{\sqrt{\ln(1 + cv^2)}} = z_{\alpha/2},$$

where $z_{\alpha/2}$ is the $100(1 - \frac{\alpha}{2})^{th}$ percentile of the standard normal distribution. If we estimate cv^2 by replacing $\text{var}(\hat{\tau}_{LR})$ with $\widehat{\text{var}}(\hat{\tau}_{LR})$ and τ with $\hat{\tau}_{LR}$, we obtain a $100(1-\alpha)\%$ confidence interval estimate of τ as $(T + [(\hat{\tau}_{LR} - T)/C]\sqrt{1 + cv^2}, (T + [(\hat{\tau}_{LR} - T)C]\sqrt{1 + cv^2})$ where $C = exp[z_{\alpha/2}\sqrt{\ln(1 + cv^2)}]$. We will use the label $\hat{\tau}_{lg2}$ to denote this confidence interval estimator. A simpler but possibly lower in coverage $100(1 - \alpha)\%$ confidence interval estimate of τ is $(T + (\hat{\tau}_{LR} - T)/C, T + (\hat{\tau}_{LR} - T)C)$ obtained by the delta method on $\ln(\hat{\tau} - T)$. We will use the label $\hat{\tau}_{log}$ to denote this confidence interval estimator. Note that confidence interval $\hat{\tau}_{ln2}$ is shifted and expanded, compared to confidence interval $\hat{\tau}_{log}$. Both confidence interval estimates will have a lower bound of at least T, the number of seen animals in a phase II sample. Confidence interval estimates constructed from the sampling distribution of $\hat{\tau}_{LR}$ and from the reciprocal transformations might have a lower bound less than T.

To test if the coverage rate of confidence interval estimates is significantly different from the nominal level, we apply the hypothesis test of the expected value of $100(\widehat{1 - \alpha}))$ is equal to a given value. Standard error for the z-statistics is $100\sqrt{\alpha(1 - \alpha)/k}\%$, where k is the number of population estimation trials or equivalently, number of phase I samples (or phase II samples), for a $100(1 - \alpha)\%$ nominal level. For example, let $\alpha = 0.1$ and $k = 500$. Coverage rate $\widehat{1 - \alpha}$ is not different from the 90% nominal level at 0.05 level of significance if $(\widehat{1 - \alpha} - 0.9)/\sqrt{0.9(1 - 0.9)/500}$ is between -1.96 and 1.96. For our simulation studies in this chapter, we will apply the hypothesis test at 0.05 level of significance. Then, for $k = 500$, coverage rate is not significantly different from the nominal level of 90% if it is within $90\% \pm 2.68\%$ and from 95% if it is within $95\% \pm 2\%$, both inclusive. For $k = 1000$, the bounds are $90\% \pm 2\%$ and $95\% \pm 1.4\%$, both inclusive, for nominal levels of 90% and 95% respectively.

By comparing coverage of various (approximate) confidence interval estimates of the population size τ, we will be able to determine if we should transform the population size estimates and if so, which transformation(s) is(are) better.

5.2 Simulated populations

We assume that there are M groups in the population. We also assume that the sighting probabilities g_j for $j = 1, 2, ..., M$ satisfy the following model,

$$\log \frac{g_j}{1 - g_j} = \beta_0 + x_{j,1}\beta_1 + x_{j,2}\beta_2 + x_{j,3}\beta_3, \tag{5.2.1}$$

where parameters and independent variables are

$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3) = (2.4132, 0.1322, -1.8224, -0.0423)$,

$x_{j,1}$ = group size,

$x_{j,2}$ = behavior of the group, 0 if resting and 1 if active,

$x_{j,3}$ = degree of vegetation cover (in percent from 0 to 100, in increments of 5 or 10).

This is a special case of (3.2.4). The data set listed in Table 5.1 was used as a basis for the generated population used in the simulation study. This data set is a composite of five aerial surveys of elk in Pennsylvania conducted by biogolists of the Pennsylvania Game Commission. The model defined in (5.2.1) is the fitted model of this composite data.

We randomly generated a number (= 116) to use as the number of groups in a simulated population. Nine groups were added to the 107 listed groups of Table 5.1 to obtain a simulated population of 116 groups, listed in Table 5.2. The population size is $\tau = \sum_1^{116} y_j = 866$ animals. This simulated population has a structure different from that of the simulated population of composite (raw) data of 107 groups. The independent variables were randomly ordered to form the population. Specifically, the groups are indexed from 1 to 116. For each independent variable, we randomly assign a rank (1-116) to each generated value so that it can attached to a group whose index matched its rank. We have thus constructed a simulated population of $M = 116$ groups. The three covariates are hence, mutually independent. The frequency distribution of the simulated population is listed in Table 5.2. The stem-and-leaf graphs of group size and vegetation cover of the simulated data are displayed in figure 5.1.

We use the fitted model defined in (5.2.1) of the composite (raw) data ($M =$ 107 groups) as the true sightability model for determining the sighting probabilities for our simulated population. For the composite data, the three covariates have low correlations between them (see footnote under Table 5.1). The covariates are pretty much mutually independent. The sighting probabilities $\{g_j\}$ of these 116 groups ranged from 0.0357 to 0.9971 and the first, second (Median) and third quantiles are 0.3150, 0.6280 and 0.8236 respectively. The average sighting probability is 0.5770.

We have also simulated two subpopulations from this simulated population.

- The first one is obtained by excluding groups in high vegetation cover, ie. vegetation cover more than 50 percent. This subpopulation, which will be referred to and labeled as *veg*, has 82 groups and size 573. With low vegetation cover, the groups in this population are expected to have higher sighting probabilities. The sighting probabilities ranged from 0.2210 to 0.9971, and the first, second and third quantiles are 0.4726, 0.6905 and 0.8934 respectively. We see that the sighting probabilities are at least 0.2. The average sighting probability for this population is 0.6771. The stem and leaf graphs of group size and vegetation cover is displayed in figure 5.2.

Table 5.1 Frequency distribution of the composite data from Pennsylvania [1]

Variable	No. of groups Missed	Seen	Total	Percent (Total/107)	Visibility (Seen/Total)
Group size					
1	14	4	18	16.8	0.22
2	5	5	10	9.3	0.50
3	9	7	16	15.0	0.44
4	4	5	9	8.4	0.55
5	1	6	7	6.5	0.85
6	2	3	5	4.7	0.60
7	2	2	4	3.7	0.50
8	3	3	6	5.6	0.50
9	1	3	4	3.7	0.75
10	2	2	4	3.7	0.50
11	2	2	4	3.7	0.50
12	0	3	3	2.8	1.00
13	0	1	1	0.9	1.00
16	0	1	1	0.9	1.00
17	0	1	1	0.9	1.00
19	1	0	1	0.9	0.00
20	0	1	1	0.9	1.00
21	0	1	1	0.9	1.00
22	0	2	2	1.9	1.00
24	0	1	1	0.9	1.00
25	1	1	2	1.9	0.50
27	0	2	2	1.9	1.00
29	0	1	1	0.9	1.00
31	0	1	1	0.9	1.00
34	0	1	1	0.9	1.00
43	0	1	1	0.9	1.00
Behavior					
0	6	28	34	31.8	0.44
1	41	32	73	68.2	0.82
Vegetation cover					
0	0	7	7	6.5	1.00
5	0	1	1	0.9	1.00
10	3	12	15	14.0	0.80
20	4	7	11	10.3	0.63
25	2	5	7	6.5	0.71
30	3	10	13	12.1	0.77
35	2	0	2	1.9	0.00
40	5	3	8	7.5	0.37
50	3	6	9	8.4	0.66
60	4	6	10	9.3	0.60
70	9	0	9	8.4	0.00
80	3	3	6	5.6	0.50
85	2	0	2	1.9	0.00
90	5	0	5	4.7	0.00
95	2	0	2	1.9	0.00

[1]There is a total of 107 marked groups, of which, 60 groups (56 percent) are sighted. cor(gp size,behav)=-0.024, cor(gp size,veg cover)=-0.146, cor(behav,veg cover)=0.153. Fit full model, with gp size, $\widehat{\boldsymbol{\beta}} = (-2.4132, -0.1322, 1.8224, 0.0423)$, std err= $(0.7, 0.04, 0.6, 0.01)$. Fit full model, with log(gp size), $\widehat{\boldsymbol{\beta}} = (-1.8828, -0.9694, 1.7704, 0.043)$, std err= $(0.7, 0.3, 0.6, 0.01)$.

Table 5.2 Frequency distribution of simulated population

Variable	No. of groups	Percent (out of 116)
Group size		
1	20	17.2
2	11	9.5
3	18	15.5
4	10	8.6
5	8	6.9
6	6	5.2
7	5	4.3
8	6	5.2
9	4	3.4
10	4	3.4
11	4	3.4
12	3	2.6
13	1	0.9
16	1	0.9
17	1	0.9
19	1	0.9
20	1	0.9
21	1	0.9
22	2	1.7
24	1	0.9
25	2	1.7
27	2	1.7
29	1	0.9
31	1	0.9
34	1	0.9
43	1	0.9
Behavior		
0	38	32.8
1	78	67.2
Vegetation cover		
0	8	6.9
5	2	1.7
10	17	14.7
20	12	10.3
25	8	6.9
30	15	12.9
35	3	2.6
40	8	6.9
50	9	7.8
60	10	8.6
70	9	7.8
80	6	5.2
85	2	1.7
90	5	4.3
95	2	1.7

```
                Group Size

Stem Leaf                           #
  43 0                              1
  42
  41
  40
  39
  38
  37
  36
  35
  34 0                              1
  33
  32
  31 0                              1
  29 0                              1
  28
  27 00                             2
  26
  25 00                             2
  24 0                              1
  23
  22 00                             2
  21 0                              1
  20 0                              1
  19 0                              1
  18
  17 0                              1
  16 0                              1
  15
  14
  13 0                              1
  12 000                            3
  11 0000                           4
  10 0000                           4
   9 0000                           4
   8 000000                         6
   7 00000                          5
   6 000000                         6
   5 00000000                       8
   4 0000000000                    10
   3 000000000000000000            18
   2 00000000000                   11
   1 00000000000000000000          20
     ----+----+----+----+----+---
```

```
            Vegetation Cover

Stem Leaf                         #
  95 00                           2
  90 00000                        5
  85 00                           2
  80 000000                       6
  75
  70 000000000                    9
  65
  60 0000000000                  10
  55
  50 000000000                    9
  45
  40 00000000                     8
  35 000                          3
  30 000000000000000             15
  25 00000000                     8
  20 000000000000                12
  15
  10 00000000000000000           17
   5 00                           2
   0 00000000                     8
     ----+----+----+----+
```

Figure 5.1 Stem-and-leaf graphs of group size and degree of vegetation cover for the simulated population of $M = 116$ groups

- The second subpopulation is obtained by excluding large groups, ie. group size of 10 or more. This subpopulation, which will be referred to and labeled as *small*, has 88 groups and size 331. With small group sizes, most groups, especially those with high vegetation cover, in this population are expected to have smaller sighting probabilities. The sighting probabilities ranged from 0.0357 to 0.9674 and the first, second and third quantiles are 0.2738, 0.5285 and 0.7462 respectively. We observed that 17% and 37.5% of the groups have sighting probabilities less than 0.2 and 0.4 respectively. The average sighting probability for this population is 0.5133. The stem and leaf graphs of group size and vegetation cover is displayed in figure 5.3.

In the second paragraph of subsection 2.3.3, it was stated that from the past experience of some elk sightability surveys (Samuel et. al., 1987), the natural log transformation of group size might give a better fit in some cases. In addition to the model defined in (5.2.1), we fitted, to the composite data from Pennsylvania, a similar model to (5.2.1) but with the group size natural-log transformed. This model is defined as follows.

$$\log \frac{g_j}{1-g_j} = \beta_0 + x_{j,1}\beta_1 + x_{j,2}\beta_2 + x_{j,3}\beta_3, \tag{5.2.2}$$

where parameters and independent variables are

$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3) = (1.8828, 0.9694, -1.7704, -0.043),$

$x_{j,1}$ = natural log of group size, ie. log(group size),

$x_{j,2}$ = behavior of the group, 0 if resting and 1 if active,

$x_{j,3}$ = degree of vegetation cover (in percent from 0 to 100, in increments of 5 or 10).

```
          Group size

Stem Leaf                              #
  34 0                                 1
  33                                   1
  32
  31 0                                 1
  30
  29 0                                 1
  28
  27
  26
  25 0                                 1
  24 0                                 1
  23
  22 00                                2
  21 0                                 1
  20 0                                 1
  19 0                                 1
  18
  17 0                                 1
  16 0                                 1
  15
  14
  13
  12 00                                2
  11 000                               3
  10 0                                 1
   9 000                               3
   8 00                                2
   7 000                               3
   6 00000                             5
   5 00000                             5
   4 000000                            6
   3 00000000000000000                17
   2 000000000                         9
   1 00000000000000                   16
     ----+----+----+----+----+---
```

```
          Vegetation Cover

Stem Leaf                      #
  50 000000000                 9
  45
  40 00000000                  8
  35 000                       3
  30 000000000000000          15
  25 00000000                  8
  20 000000000000             12
  15
  10 00000000000000000        17
   5 00                        2
   0 00000000                  8
     ----+----+----+----+
```

Figure 5.2 Stem-and-leaf graphs of group size and degree of vegetation cover for the subpopulation *veg* of $M = 82$ groups

```
          Group size                                    Vegetation Cover

Stem Leaf                    #                Stem Leaf                    #
   9 0000                    4                  95 00                      2
   8 000000                  6                  90 00                      2
   7 00000                   5                  85 00                      1
   6 000000                  6                  80 00000                   5
   5 00000000                8                  75
   4 0000000000             10                  70 00000                   5
   3 000000000000000000     18                  65
   2 00000000000            11                  60 000000000               9
   1 00000000000000000000   20                  55
     ----+----+----+----+----+---               50 00000000                8
                                                45
                                                40 00000                   5
                                                35 000                     3
                                                30 00000000000            11
                                                25 000000                  6
                                                20 000000000               9
                                                15
                                                10 000000000000           12
                                                 5 00                      2
                                                 0 00000000                8
                                                   ----+----+----+----+
```

Figure 5.3 Stem-and-leaf graphs of group size and degree of vegetation cover for the subpopulation *small* of $M = 88$ groups

Model (5.2.2) will also be used to determined the sighting probabilities of those $M = 116$ groups in the simulated population to obtain a second simulated population and hence, a second set of subpopulations *veg* and *small*. We will also use models (5.2.1) and (5.2.2) to determine the sighting probabilities of the composite data from Pennsylvania to obtain a third and a fourth simulated populations. Note that the AIC of fitting model (5.2.2) to the composite data is a little lower ($101.274 < 103.113$) than fitting model (5.2.1) to the composite data, indicating that model (5.2.2) might be better.

5.3 Simulation

A single population size estimate in the simulation study requires 2 independent samples from the simulated population. The phase I sample is obtained by making one or more complete passes through the population. In a single complete pass through the population or a sighting trial, a group is determined to be sighted or not by comparing a randomly generated uniform random number, u_j (generated jointly independent for the groups), with the corresponding sighting probability g_j. If $u_j \leq g_j$, the j^{th} group for $j = 1, 2, ..., M$ is said to be sighted. A single pass in phase I generates 116 observations for estimating the parameters of the sightability model. We enlarge the size of a phase I sample by increasing the number of passes in phase I.

For phase II, a sample is obtained by making one or more complete pass through the population as in phase I. A phase II sample only consisted of sighted groups. Sighting probabilities of these sighted groups are estimated using the fitted model obtained in phase I. With group sizes and estimated sighting probabilities of these sighted groups, estimates of the population size (see (3.2.5)) and the error components (see (3.2.7) and (3.2.9)) whose sum is the variance of the population size estimator, are obtained. We enlarge the size of a phase II sample by increasing the number of passes in phase II.

Theoretically, if a phase I sample is enlarged by, say, making two complete passes through the population, the estimated covariance matrix $\widehat{\Sigma}$ of estimated parameters $\widehat{\boldsymbol{\beta}}$ would be halved as explained as follows.

Let $\boldsymbol{X} = [x_{j,i}]_{M \times 4}$ be the matrix of covariates such that $x_{j,0} = 1$ (see (5.2.1) or (5.2.2)). Maximizing the likelihood function, the estimating equation of the parameters $\boldsymbol{\beta}$ is

$$\mathbf{X}'(\mathbf{Z} - \mathbf{g}) = 0$$

where $\mathbf{Z}$ is the matrix of indicator functions Z_j (see section 3.2) and $\mathbf{g}$ is the matrix of g_j. The estimated covariance matrix of the estimated parameters $\widehat{\beta}$ based on a sample obtained from one complete pass through the population, is then

$$\widehat{\Sigma}_{(1,1)} = r\widehat{mcov}(\widehat{\beta}) = (\widehat{\mathbf{X}}'\widehat{\mathbf{V}}^{-}\widehat{\mathbf{X}})^{-1},$$

where $\widehat{\mathbf{V}}^{-}$ is the diagonal matrix with $\hat{g}_j(1-\hat{g}_j)$ on the diagonal. If we make two complete passes through the population, we will have

$$[\mathbf{X}\ \mathbf{X}]' \begin{bmatrix} \widehat{\mathbf{V}}^{-} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{V}}^{-} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{X} \end{bmatrix} = 2\mathbf{X}'\widehat{\mathbf{V}}^{-}\mathbf{X}$$

Hence, the estimated covariance matrix of the estimated parameters based on this larger sample is,

$$\widehat{\Sigma}_{(2,1)} = (2\widehat{\mathbf{X}}'\widehat{\mathbf{V}}^{-}\widehat{\mathbf{X}})^{-1} = \frac{1}{2}\widehat{\Sigma}_{(1,1)}$$

Similarly, if the phase I sample is enlarged by, in general, making r complete passes through the population, the estimated covariance matrix of estimated parameters would be divided by r.

A phase I sample obtained from a sightability survey is for estimating the sightability model. Therefore, enlarging the size of a phase I sample will not affect the first error component e_1, the error due to sampling of groups, but will affect the second error component e_2, the error due to estimating sighting probabilities. Indeed, we see that the derived expression (3.2.6) for e_1 does not involve estimated sighting probabilities (or involve Σ). The derived expression (3.2.8) for e_2 consists of variances and covariances of $\widehat{\Theta}_j$'s), the reciprocal of estimated sighting probabilities g_j, where

$$\mathrm{var}(\widehat{\Theta}_j) = e^{-2\boldsymbol{x}_j'\boldsymbol{\beta}}(e^{\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j} - 1),$$

and

$$\mathrm{cov}(\widehat{\Theta}_{j_1}, \widehat{\Theta}_{j_2}) = e^{-(\boldsymbol{x}_{j1}+\boldsymbol{x}_{j2})'\boldsymbol{\beta}}(e^{\boldsymbol{x}_{j1}'\Sigma\boldsymbol{x}_{j2}} - 1).$$

To compute two approximate 'true' second error components, we replace Σ in each of the above variance and covariance terms by $\widehat{\Sigma_1}$ and by $\widehat{\Sigma_2}$ (see section 5.1.2).

To see how the variance and the covariance terms are affected when a phase I sample is enlarged, we use Taylor's expansion to the first order. We have

$$e^{\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j} - 1 \approx \boldsymbol{x}_j'\Sigma\boldsymbol{x}_j.$$

If we make two passes through the population in phase I, we will reduce, as shown before, $\widehat{\Sigma}$ by half or $\widehat{\Sigma}/2$. So

$$e^{\boldsymbol{x}_j'\Sigma/2\boldsymbol{x}_j} - 1 \approx \frac{\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j}{2}.$$

Thus, both variance and covariance terms are approximately halved. This approximation has an error of about 0.04 when $\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j$ is 0.5, i.e., $(e^{0.5}-1)/2-e^{0.25} \approx 0.04$. The smaller the term $\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j$, the smaller is the error of the approximation. From (3.2.8), we see that the second error component is the expectation of the total of two sums. One sum sums the product of the variance term and the square of the respective group size while the other sum sums the product of the covariance term and the respective group sizes. Hence, the approximate 'true' second error component is reduced by half approximately if we make two complete passes through the population to obtain a phase I sample. Similarly, in general, if we make r complete passes through the population in phase I, the approximate 'true' second error component will be reduced by r times (or divided by r) compared to the approximate 'true' second error component obtained based on 1 complete pass through the population in phase I.

On the other hand, a phase II sample obtained from a population survey is for estimating the population size based on the estimated parameters and the estimated covariance matrix already obtained in phase I. If we enlarge the phase II sample by making, say, 2 complete passes through the population, the population size estimator in (3.2.5) now sums from $j = 1$ to $j = 2M$. Therefore, to obtain the required estimates, we need to divide the population size estimate by 2 and estimates $\hat{e}_1$, $\hat{e}_2$ (see (3.2.7) and (3.2.9)) of both error component by $2^2 = 4$. To obtain the required true error components, we also need to divide both error components e_1, e_2 by $2^2 = 4$. This is an example of the first scenario of the third special case described in section 3.4.

However, we notice that in fact, the expression (3.2.6) for e_1, now sums from $j = 1$ to $j = 2M$, is just twice the sum from $j = 1$ to $j = M$. So to obtain the required e_1, dividing the new e_1 (summing from $j = 1$ to $j = 2M$) by 4 is exactly the same as taking half of the old e_1 (summing from $j = 1$ to $j = M$). In other words, making 2 complete passes through the population in phase II, the required e_1 is just half of the old e_1 based on making 1 complete pass through the population. As for expression (3.2.8) for e_2, we first notice that the first sum now sums the product of the variance term and the square of the respective group size from $j = 1$ to $j = 2M$, which is twice the sum from $j = 1$ to $j = M$. For the second sum that sums the product of the covariance term and the respective group sizes, since the M groups are repeated, the second sum is now consisted of twice the old first sum (summing from $j = 1$ to $j = M$) and four times the old second sum (summing from $j = 1$ to $j = M$). Hence, the new e_2 (not the required e_2) is now four times the old e_2. In other words, the required e_2, which is the new e_2 divided by 4, is exactly equal to the old e_2.

Similarly, in general, if we make r complete passes through the population in phase II, to obtain the required estimates, we need to divide the population size estimate by r and estimates of the error components by r^2. To obtain the required true error components, we only need to divide the true first error component computed based on 1 complete pass through the population by r while the second true error component remains unchanged.

For one iteration of the simulation to obtain one population size estimate and one set of the error component estimates, we denote the number of complete passes through the population in phases I and II by r_1 and r_2 respectively. We will refer to and label this as (r_1, r_2). For example, $(1,1)$ denotes $r_1 = 1$ complete pass through the population in phase I and $r_2 = 1$ complete pass through the population in phase II, while $(2,1)$ denotes $r_1 = 2$ complete passes through the population in phase I and $r_2 = 1$ complete pass through the population in phase II, etc. We use the notation (∞, r_2) to denote the case of known sighting probabilities, indicating that sampling is not needed in phase I because there is no need to estimate the sightability model and r_2 complete passes through the population in phase II.

We consider the following cases for our simulation.

- (1) The simulated population of $M = 116$ groups, with sightability models (5.2.1) and (5.2.2) as true models, with phase I and phase II sampling specified as $(\infty, 1)$, $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$ and $(4,1)$,

- (2) Subpopulation *veg* of $M = 82$ groups, with sightability models (5.2.1) and (5.2.2.) as true models, with phase I and phase II sampling specified as $(1,1)$, $(2,1)$, $(2,2)$, $(4,1)$ and $(4,2)$,

- (3) Subpopulation *small* of $M = 88$ groups, with sightability models (5.2.1) and (5.2.2) as true models, with phase I and phase II sampling specified as $(1,1)$, $(2,1)$, $(2,2)$, $(4,1)$ and $(4,2)$,

- (4) The population of $M = 107$ groups and size $\tau = 834$ (the composite data from Pennsylvania listed in Table 5.1), with sightability model (5.2.2) as the true model, with phase I and phase II sampling specified as $(\infty, 1)$, $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$ and $(4,1)$,

- (5) The population of $M = 107$ groups (the composite data from Pennsylvania), with sightability model (5.2.1) as the true model, with phase I and phase II sampling specified as $(\infty, 1)$, $(\infty, 2)$, $(\infty, 3)$, $(1,1)$ and $(4,2)$.

We will refer to the population of $M = 116$ groups and the subpopulations as the simulated population, subpopulation *veg* and subpopulation *small*. We will label the population of $M = 107$ groups as the population *Penn*.

We want to evaluate small sample size performances of our estimators for estimating the population size, the error components and hence, the variances, which depends on the asymptotic properties of the maximum likelihood estimators of the parameters $\boldsymbol{\beta}$ of the sightability model. We know that by enlarging the sample size, we will reduce noise during simulation. With a larger phase I sample, we will obtain better parameter estimates, $\widehat{\beta}$, whose performances could also be measured in terms of bias relative to their true values, which in turn could improve estimation of population size, $\hat{\tau}_{LR}$ and estimation of variance $\text{var}(\hat{\tau})$ of the population size estimator, including confidence interval estimates of population size (see subsections 5.1.1, 5.1.2 on performance measures). In some cases, increasing the size of the phase I sample may not be sufficient to provide desirable estimates for the population size, the variance of the population size estimator and the confidence interval estimates of population size. In these cases, increasing the sample size for phase II would also be necessary.

We would also like to know how those point and confidence interval estimators perform for some particular types of population such as subpopulation *veg* of groups with low vegetation cover and higher sighting probabilities on average, and subpopulation *small* of small groups with lower sighting probabilities on average, compared to a more general population such as our simulated population of $M = 116$ groups. By looking at various combinations (r_1, r_2) for different types of small population, we hope to give some general suggestion on the sample sizes for both phase 1 and phase II.

5.4 Results of the simulation

We generated $k = 500$ or 1000 population estimation trials for each listed phase I-phase II sample sizes (r_1, r_2). In Tables 5.3A, 5.3B, 5.3C, 5.7A, 5.7B, 5.7C, 5.10 and 5.13, we have used, symbols e_1 to denote true error component 1. In addition, we used e_{21}, var_1 and e_{22}, var_2 to denote true error components 2 and true variances computed using $\widehat{\Sigma}_1$ and $\widehat{\Sigma}_2$ respectively. We used $\hat{e}_1(\hat{\tau}_{LR})$, $\hat{e}_2(\hat{\tau}_{LR})$ and $\widehat{\text{var}}(\hat{\tau}_{LR})$ denotes mean of 500 or 1000 estimated first, second error components and variance of $\hat{\tau}_{LR}$. Notation bias(x,y) denotes bias of x relative to y, i.e., $\text{bias}(x, y) = \left(\frac{x-y}{y}\right) 100$. We have labelled the observed (or sample) variance of 500 or 1000 population size estimates as obs $\text{var}(\hat{\tau}_{LR})$. The estimated expected interval width in parentheses next to coverage rates of the true population size in Tables 5.4, 5.5, 5.8, 5.9, 5.11, 5.12, 5.14 and 5.15 is calculated using the average population size estimate, the average estimate of the variance of population size estimator and the average number of animals seen of 500 or 1000 estimation trials.

SAS programs for obtaining true error components and variances are also attached in the Appendix.

5.4.1 Simulation results of the simulated population of 116 groups and the subpopulations with sightability model (5.2.1) as the true model

Assuming model (5.2.1) as the true sightability model, results of all listed (r_1, r_2) of the simulated population of $M = 116$ groups are presented in Table 5.3A, results of the subpopulation *veg* are presented in Table 5.3B, and results of the subpopulation *small* are presented in Table 5.3C. Coverage of 500 90% and 95% confidence interval estimates of the population size τ for all listed (r_1, r_2) of the simulated population and the subpopulations are presented in Tables 5.4 and 5.5. In the column labeled $\hat{\tau}_{log}$, we listed the coverage of the confidence interval estimates $(T+(\hat{\tau}_{LR}-T)/C, T+(\hat{\tau}_{LR}-T)C)$ and in the column labeled $\hat{\tau}_{lg2}$, we listed the coverage of the confidence interval estimates $(T+[(\hat{\tau}_{LR}-T)/C]\sqrt{1+cv^2}, (T+[(\hat{\tau}_{LR}-T)C]\sqrt{1+cv^2})$. These confidence interval estimates were not computed when sighting probabilities were known. In Tables 5.3B, 5.3C, 5.4 and 5.5, we have used prefixes *veg* and *small* on the notation (r_1, r_2) to indicate that subpopulations *veg* and *small* were used. Note that we only listed coverage of confidence interval estimates $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ for $veg(4,1)$, $veg(4,2)$, $small(4,1)$ and $small(4,2)$. We will not list the bias of estimated parameters $\widehat{\boldsymbol{\beta}}$ relative to the true parameters $\boldsymbol{\beta}$. We are more interested in bias of population size estimates, estimates of error components and variance of the population size estimator.

Averages of 500, original and transformed population size estimates, and of 500 corresponding estimates of the error components and variance are obtained by utilizing procedure UNIVARIATE of SAS. This procedure also gives the (observed) sample variance of those 500 population size estimates. Standard errors of these averages were also listed in the output of procedure UNIVARIATE.

We have observed that for one complete pass through the population of 116 groups, the average sample size was 67 groups or about 58% and average number of seen animals was 629 for the simulated population, the average sample size was 56 groups or about 68% and average number of seen animals was 465 for subpopulation *veg*, and, the average sample size was 46 or about 52% and average number of seen animals was 180 for subpopulation *small*. As expected, subpopulation *veg* ($M = 82$) generates a larger sample on average than subpopulation *small* ($M = 88$) for one complete pass through the population as the groups in subpopulation *veg* have higher sighting probabilities on average ($0.6771 > 0.5133$).

We immediately see, in Tables 5.3A, 5.3B and 5.3C, that the (asymptotic) population size estimator performed, on average, extremely well in all cases. The largest bias observed was 1.7% for $small(1,1)$ in Table 5.3C.

Table 5.3A Simulation results of 500 population estimation trials for listed (r_1, r_2) of the simulated population of $M = 116$ groups and $\tau = 866$, with sightability model (5.2.1) as the true model[2]

Statistic	$(\infty,1)$	$(\infty,2)$	$(\infty,3)$	$(1,1)$	$(1,2)$
$\hat{\tau}_{LR}$	867.1(3.2)	867.7(2.3)	866.3(1.9)	864.2(5.3)	864.3(4.8)
bias$(\hat{\tau}_{LR})$	0.1	0.2	0.1	−0.2	−0.2
e_1	5344.1	2672.1	1781.4	5344.1	2672.1
$\hat{e}_1(\hat{\tau}_{LR})$	5347.6(79.3)	2686.5(26.1)	1765.6(14.6)	5536.25(276.6)	2734.62(146.6)
bias$(\hat{e}_1)$	0.1	0.5	−0.9	3.6	2.3
e_{21}	–	–	–	9317.2	8437.8
e_{22}	–	–	–	8129.5	7975.7
$\hat{e}_2(\hat{\tau}_{LR})$	–	–	–	8214.5(1026.2)	8457.7(2000.5)
bias$(\hat{e}_2, e_{21})$	–	–	–	−11.8	0.2
bias$(\hat{e}_2, e_{22})$	–	–	–	−1.1	6.1
var_1	5344.1	2672.1	1781.4	14661.3	11109.9
var_2	5344.1	2672.1	1781.4	13473.6	10647.8
$\widehat{\text{var}}(\hat{\tau}_{LR})$	5347.6(79.3)	2686.5(26.1)	1765.6(14.6)	13750.7(1279.6)	11210.3(2140.2)
bias$(\widehat{\text{var}}, \text{var}_1)$	0.1	0.5	−0.9	−6.2	0.9
bias$(\widehat{\text{var}}, \text{var}_2)$	0.1	0.5	−0.9	2.1	5.3
obs var$(\hat{\tau}_{LR})$	5236.6	2537.9	1810.5	13925.8	11310.9
bias$(\text{var}_1$, obs var)	2.1	5.3	−1.6	5.4	−1.8
bias$(\text{var}_2$, obs var)	2.1	5.3	−1.6	−3.3	−5.9
bias$(\widehat{\text{var}}$, obs var)	2.1	5.9	−2.5	−1.3	−0.9

Statistic	$(2,1)$	$(2,2)$	$(4,1)$	$(4,2)$
$\hat{\tau}_{LR}$	866.9(4.3)	862.3(3.2)	867.9(3.6)	863.8(2.9)
bias$(\hat{\tau}_{LR})$	0.1	−0.4	0.2	−0.3
e_1	5344.1	2672.1	5344.1	2672.1
$\hat{e}_1(\hat{\tau}_{LR})$	5614.2(180.9)	2689.2(67.9)	5398.9(120.9)	2671.4(49.8)
bias$(\hat{e}_1)$	5.1	0.6	1.0	−0.1
e_{21}	3151.1	3151.1	1660.9	1609.2
e_{22}	3183.3	3183.3	1452.3	1481.7
$\hat{e}_2(\hat{\tau}_{LR})$	3222.3(166.6)	2813.8(112.3)	1421.9(49.1)	1322.1(35.9)
bias$(\hat{e}_2, e_{21})$	2.3	−10.7	−14.4	−17.8
bias$(\hat{e}_2, e_{22})$	1.2	−11.6	−2.1	−10.8
var_1	8495.1	5964.8	7005.1	4281.2
var_2	8527.4	5997.1	6796.3	4153.7
$\widehat{\text{var}}(\hat{\tau}_{LR})$	8836.4(343.1)	5502.9(179.0)	6820.9(168.5)	3993.4(85.1)
bias$(\widehat{\text{var}}, \text{var}_1)$	4.0	−7.7	−2.6	−6.7
bias$(\widehat{\text{var}}, \text{var}_2)$	3.6	−8.2	0.4	−3.8
obs var$(\hat{\tau}_{LR})$	9177.5	5223.3	6551.3	4061.7
bias$(\text{var}_1$, obs var)	−7.4	14.2	6.9	5.4
bias$(\text{var}_2$, obs var)	−7.1	14.8	3.7	2.2
bias$(\widehat{\text{var}}$, obs var)	−3.7	5.3	4.1	−1.6

[2]All estimates shown here are averages of 500 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 500 simulated estimates. Average sample size is 67 groups and average number of seen animals is 629 in a (1,1).

Table 5.3B Simulation results of 500 population estimation trials for listed (r_1, r_2) of the subpopulation *veg* of $M = 82$ groups and $\tau = 573$, with sightability model (5.2.1) as the true model [3]

Statistic	veg(1,1)	veg(2,1)	Case veg(2,2)	veg(4,1)	veg(4,2)
$\hat{\tau}_{LR}$	567.5(2.5)	572.3(2.1)	569.8(1.7)	567.5(1.9)	572.9(1.4)
bias($\hat{\tau}_{LR}$)	0.9	−0.1	−0.5	−0.9	−0.01
e_1	1428.7	1428.7	714.3	1428.7	714.3
$\hat{e}_1(\hat{\tau}_{LR})$	1309.6(45.6)	1405.7(39.1)	687.5(16.9)	1353.7(23.8)	713.2(11.4)
bias($\hat{e}_1$)	−8.3	−1.6	−3.7	−5.2	−0.2
e_{21}	2880.0	895.1	895.1	444.6	413.6
e_{22}	2731.5	928.5	928.5	414.2	405.7
$\hat{e}_2(\hat{\tau}_{LR})$	1488.7(96.6)	738.8(32.6)	690.7(26.3)	334.8(9.2)	351.3(8.0)
bias($\hat{e}_2, e_{21}$)	−48.3	−17.4	−22.8	−24.7	−15.1
bias($\hat{e}_2, e_{22}$)	−45.5	−20.4	−25.6	−19.1	−13.4
var_1	4308.7	2323.8	1609.5	1850.3	1127.9
var_2	4160.2	2357.2	1642.8	1819.8	1120.1
$\widehat{\text{var}}(\hat{\tau}_{LR})$	2798.3(139.3)	2144.5(71.0)	1378.3(42.8)	1688.53(32.7)	1064.5(19.3)
bias($\widehat{\text{var}}, \text{var}_1$)	−35.0	−7.7	−14.3	−8.7	−5.6
bias($\widehat{\text{var}}, \text{var}_2$)	−32.7	−9.0	−16.1	−7.2	−4.9
obs var($\hat{\tau}_{LR}$)	3128.8	2354.9	1481.3	1861.3	1015.3
bias(var_1, obs var)	37.7	−1.3	8.6	−0.6	11.1
bias(var_2, obs var)	32.9	0.1	10.9	−2.2	10.3
bias($\widehat{\text{var}}$, obs var)	−10.5	−8.9	−6.9	−9.3	4.8

[3]All estimates except those of *veg*(1, 1) shown in above tables are averages of the 500 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 500 simulated estimates. For subpopulation *veg*, average sample size is 56 groups and average number of seen animals is 465. For subpopulation *small*, average sample size is 46 groups and average number of seen animals is 180.

Table 5.3C Simulation results of 500 population estimation trials for listed (r_1, r_2) of the subpopulation *small* of $M = 88$ groups and $\tau = 331$, with sightability model(5.2.1) as the true model [4]

			Case		
Statistic	small(1, 1)	small(2, 1)	small(2, 2)	small(4, 1)	small(4, 2)
$\hat{\tau}_{LR}$	336.5(4.1)	333.7(3.2)	335.4(2.6)	332.6(2.8)	332.8(2.1)
bias($\hat{\tau}_{LR}$)	1.7	0.8	1.3	0.5	0.5
e_1	2870.8	2870.8	1435.4	2870.8	1435.4
$\hat{e}_1(\hat{\tau}_{LR})$	2978.5(185.5)	3057.8(153.2)	1581.9(67.9)	2982.0(110.8)	1492.9(42.8)
bias($\hat{e}_1$)	3.7	6.5	10.2	3.8	4.0
e_{21}	3939.4	1804.0	1804.0	801.3	871.7
e_{22}	4127.2	1771.2	1771.2	824.6	818.1
$\hat{e}_2(\hat{\tau}_{LR})$	5996.0(801.4)	2235.5(194.8)	2102.0(183.4)	924.0(59.3)	820.7(35.0)
bias($\hat{e}_2, e_{21}$)	52.2	23.9	16.5	15.3	−5.8
bias($\hat{e}_2, e_{22}$)	45.2	26.2	18.6	12.0	0.3
var_1	6810.3	4674.9	3239.4	3779.9	2307.2
var_2	6998.1	4642.0	3206.6	3803.2	2253.5
$\widehat{\text{var}}(\hat{\tau}_{LR})$	8974.6(976.4)	5293.4(343.9)	3683.9(247.6)	3906.0(167.5)	2313.6(77.3)
bias($\widehat{\text{var}}$, var_1)	31.8	13.2	13.7	3.3	0.3
bias($\widehat{\text{var}}$, var_2)	28.2	14.0	14.8	2.7	2.6
obs var($\hat{\tau}_{LR}$)	8511.6	5410.1	3484.7	3915.1	2300.5
bias(var_1, obs var)	−19.9	−13.6	−7.0	−3.4	0.3
bias(var_2, obs var)	−17.8	−14.2	−7.9	−2.8	−2.0
bias($\widehat{\text{var}}$, obs var)	5.4	−2.1	5.7	0.2	−0.5

[4]All estimates except those of *veg*(1, 1) shown in above tables are averages of the 500 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 500 simulated estimates. For subpopulation *veg*, average sample size is 56 groups and average number of seen animals is 465. For subpopulation *small*, average sample size is 46 groups and average number of seen animals is 180.

Table 5.4 Coverage of 500 simulated 90% confidence interval estimates that contained τ for all (r_1, r_2) listed in Tables 5.3A,5.3B, and 5.3C[4] [5]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	89(241)	88(242)	89(243)	89(245)	89(248)		
$(\infty, 2)$	90(171)	91(171)	91(171)	90(172)	91(174)		
$(\infty, 3)$	90(138)	90(139)	90(138)	91(139)	92(140)		
$(1, 1)$	85⋆(386)	85⋆(389)	86⋆(395)	86⋆(406)	86⋆(422)		
$(1, 2)$	84⋆(348)	85⋆(351)	85⋆(355)	85⋆(363)	85⋆(374)	86⋆(360)	87⋆(395)
$(2, 1)$	86⋆(309)	87⋆(311)	87⋆(315)	87⋆(320)	87⋆(326)		
$(2, 2)$	85⋆(244)	86⋆(245)	87⋆(246)	87⋆(249)	87⋆(253)		
$(4, 1)$	85⋆(272)	86⋆(273)	88(275)	88(279)	88(283)		
$(4, 2)$	88(208)	88(209)	89(209)	89(210)	90(213)	90(211)	90(218)
veg$(1, 1)$	82⋆(174)	82⋆(175)	83⋆(176)	83⋆(178)	83⋆(181)		
veg$(2, 1)$	87⋆(152)	87⋆(153)	87⋆(153)	87⋆(155)	88(157)		
veg$(2, 2)$	87⋆(122)	86⋆(123)	87⋆(123)	87⋆(124)	87⋆(125)		
veg$(4, 1)$	87⋆(135)	87⋆(136)	86⋆(136)	87⋆(137)	88(139)	87⋆(139)	88(150)
veg$(4, 2)$	91(107)	90(108)	90(108)	90(109)	90(109)	89(109)	90(114)
small$(1, 1)$	84⋆(312)	84⋆(323)	86⋆(348)	85⋆(396)	85⋆(506)		
small$(2, 1)$	85⋆(239)	84⋆(244)	85⋆(256)	85⋆(275)	85⋆(308)		
small$(2, 2)$	88(200)	89(202)	89(208)	89(219)	88(235)		
small$(4, 1)$	86⋆(206)	87⋆(209)	86⋆(216)	87⋆(227)	87⋆(245)	86⋆(212)	88(229)
small$(4, 2)$	90(158)	89(160)	89(163)	90(168)	90(175)	89(161)	89(169)

[5] Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ (see subsection 5.1.2) in both Tables 5.4 and 5.5 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. Coverage of these confidence interval estimates were not listed in some cases because simulation for these cases were done before we implemented the computation of these confidence interval estimates in our SAS program. Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2.68\%$ and $95\% \pm 2\%$, respectively. Numbers in parentheses are estimated expected interval widths.

Table 5.5 Coverage of 500 simulated 95% confidence interval estimates that contained τ for all (r_1, r_2) listed in Tables 5.3A,5.3B,5.3C [6]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	93(287)	94(288)	94(291)	95(294)	95(300)		
$(\infty, 2)$	94(203)	95(204)	95(204)	95(206)	95(208)		
$(\infty, 3)$	94(165)	94(165)	95(165)	95(166)	95(167)		
$(1, 1)$	89*(460)	90*(465)	91*(476)	91*(495)	92*(523)		
$(1, 2)$	88*(415)	89*(419)	90*(428)	90*(441)	91*(459)	93(444)	92*(487)
$(2, 1)$	90*(368)	91*(372)	91*(377)	91*(385)	92*(399)		
$(2, 2)$	90*(291)	92*(292)	92*(295)	92*(300)	92*(306)		
$(4, 1)$	91*(324)	92*(325)	94(329)	94(335)	95(344)		
$(4, 2)$	93(248)	93(248)	94(250)	94(253)	94(257)	95(255)	95(264)
veg$(1, 1)$	86*(207)	87*(209)	87*(211)	88*(215)	88*(220)		
veg$(2, 1)$	92*(182)	92*(182)	92*(184)	92*(187)	92*(189)		
veg$(2, 2)$	91*(146)	91*(146)	91*(147)	91*(148)	91*(150)		
veg$(4, 1)$	91*(161)	91*(162)	92*(163)	92*(164)	92*(166)	93(170)	93(184)
veg$(4, 2)$	94(128)	94(129)	94(129)	94(129)	95(131)	95(132)	95(138)
small$(1, 1)$	88*(371)	89*(391)	89*(435)	92*(534)	92*(861)		
small$(2, 1)$	88*(285)	90*(294)	91*(313)	92*(349)	92*(421)		
small$(2, 2)$	93(238)	94(243)	94(254)	95(272)	94(304)		
small$(4, 1)$	91*(245)	91*(250)	93(262)	93(283)	93(320)	93(260)	93(280)
small$(4, 2)$	93(189)	95(191)	95(196)	95(205)	95(219)	95(196)	95(205)

In Table 5.3A containing simulation results of the simulated population, estimators of the first error component, which is the error due to sampling of groups, also performed well on average. The largest bias observed was 5.1% of $(2, 1)$. In columns $(\infty, 1)$, $(\infty, 2)$ and $(\infty, 3)$ when sighting probabilities are known, the first error component is the variance. In this case, since there is no need to estimate the sighting probabilities, the second error component which is the error due to estimating the sighting probabilities is zero. From Tables 5.4 and 5.5, we see that when sighting probabilities are known, we need a phase II sample size of at least $200(> 116 \times 1)$ for coverage rates of confidence interval estimates and levels to be not significantly different.

[6] Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ (see subsection 5.1.2) in both Tables 5.4 and 5.5 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. Coverage of these confidence interval estimates were not listed in some cases because simulation for these cases were done before we implemented the computation of these confidence interval estimates in our SAS program. Coverage rates with $\star$ are different from levels at 0.05 level of significance as they are not within $90\% \pm 2.68\%$ and $95\% \pm 2\%$, respectively. Numbers in parentheses are estimated expected interval widths.

In Table 5.3A, for $(1,1)$, we observed from some basic statistics that simulated values of both estimated error components, and of the estimated variance, fluctuated in a very wide range (see large standard errors under (1,1)). These basic statistics, which will not be listed, include the average, the three quantiles and some percentiles at both ends of the frequency plot of the simulated values. These basic statistics are available from the output of procedure UNIVARIATE in SAS. We observed that maximum simulated values were more than 10 times and minimum simulated values were only about 10 percent of the average values. More than 50 percent of the simulated values were less than the average values. This could be due to a high correlation between the population size estimate and the variance estimate of the population size estimator. Increasing the sample size has shortened the range of the simulated estimates of both error components and the variance. The average values were quite close to the true values (Table 5.3A: $\text{bias}(\hat{e}_1) = 3.6\%$, $\text{bias}(\hat{e}_2, e_{21}) = -11.8\%$, $\text{bias}(\hat{e}_2, e_{22}) = -1.0\%$, $\text{bias}(\widehat{\text{var}}(\hat{\tau}_{LR}), \text{var}_1) = -6.2$, $\text{bias}(\widehat{\text{var}}(\hat{\tau}_{LR}), \text{var}_2) = 2.0$).

We have computed the bias of the approximate variances relative to observed variance of 500 population estimates, i.e. $\text{bias}(\text{var}_1, \text{obsvar})$ and $\text{bias}(\text{var}_2, \text{obsvar})$. The largest of these bias observed in Table 5.3A were 14.2% and 14.8% for $(2,2)$. In terms of bias, our approximate variances were quite close to the observed variance observedvar for the simulated population. We have also tabulated the bias of $\widehat{\text{var}}(\hat{\tau}_{LR})$ relative to $\text{obsvar}(\hat{\tau}_{LR})$ in the last row. The largest of all entries of $\text{bias}(\widehat{\text{var}}(\hat{\tau}_{LR}), \text{obsvar})$ was only 5.3% of $(2,2)$. Largest bias of $\widehat{\text{var}}(\hat{\tau}_{LR})$ relative to approximate 'true' variances $\text{var}_1, \text{var}_2$ was -8.2%. In terms of bias, the estimated variance $\widehat{\text{var}}(\hat{\tau}_{LR})$ seems to perform well whether it is compared to $\text{obsvar}(\hat{\tau}_{LR})$, the observed variance of population size estimates or to $\text{var}_1, \text{var}_2$, the approximate 'true' variances. If we further look at the coverage of the confidence interval estimates in Tables 5.4 and 5.5 for the simulated population, we see that we need a phase I sample size of at least $500(> 116 \times 4)$ and a phase II sample size of at least $200(> 116 \times 1)$ for coverage rates of confidence interval estimates to be not differ significantly from nominal levels.

The confidence interval $\hat{\tau}_{log}$ performs uniformly better and has the shortest expected interval width compared to those constructed using reciprocal transformations and of $\hat{\tau}_{lg2}$.

In Tables 5.3B and 5.3C for subpopulation *veg* and subpopulation *small* respectively, the first error components was unbiasedly estimated (largest bias was 10.2% of $small(2,2)$). On the other hand, for all cases except $veg(4,2)$, $small(4,1)$ and $small(4,2)$, there was a severe underestimate and overestimate of the approximate 'true' second error components e_{21}, e_{22}. For example, $\text{bias}(\hat{e}_2, e_{21})$ and $\text{bias}(\hat{e}_2, e_{22})$ under $veg(1,1)$ were -48.3% and -45.5% and under $small(1,1)$ were 52.2% and 45.2%. This resulted in severe underestimation and overestimation of the approximate 'true' variances $\text{var}_1, \text{var}_2$. For example, $\text{bias}(\widehat{\text{var}}(\hat{\tau}_{LR}), \text{var}_1)$ and $\text{bias}(\widehat{\text{var}}(\hat{\tau}_{LR}), \text{var}_2)$ under $veg(1,1)$ were -35.0% and -32.7% and under $small(1,1)$ were 31.8% and 28.2%. We also observed that there was significant bias when the approximate variances were compared to the observed variance $\text{obsvar}(\hat{\tau}_{LR})$ of the simulated population size estimates. The average estimated variance of the population size estimator $\widehat{\text{var}}(\hat{\tau}_{LR})$ had much smaller bias when compared to the observed variance $\text{obsvar}(\hat{\tau}_{LR})$ of the simulated population size estimates. Normally, since the approximate 'true' variances were computed using the estimated covariance of the estimated parameters, we would expect the estimated variance of the population size estimator, $\widehat{\text{var}}(\hat{\tau}_{LR})$, to be closer to the approximate 'true' variances than to the observed variance $\text{obsvar}(\hat{\tau}_{LR})$ of the simulated population size estimates. However, in our simulation, the results went the other way except for $veg(4,1)$, $veg(4,2)$ and $small(4,2)$. This phenomenon was also observed in Table 5.3A for the simulated population except for $(\infty, 1)$ and

(4, 1). Note that in $(\infty, 1)$, when sighting probabilities are known, there was no need to approximate the 'true' variance since there was no second error component involved.

To verify that e_2 was poorly estimated due to noise, i.e. probably due to small sample size, we repeated the $k = 500$ simulation runs of $veg(1,1)$, $veg(2,1)$, $veg(2,2)$ and $small(1,1)$, $small(2,1)$, $small(2,2)$ for another 3 times. Results for all 3 simulations were similar to what was tabulated in Tables 5.3B and 5.3C. We will not list the individual biases in percentage as before. Instead, we list the standard errors of the average of 500 simulated estimates of population total, error components 1 and 2, and of $\widehat{\text{var}}(\hat{\tau}_{LR})$ in the following table.

Table 5.6 Standard errors of 500 simulated estimates of some cases listed in Tables 5.3B and 5.3C

Cases	$\hat{\tau}_{LR}$	$\hat{e}_1(\hat{\tau}_{LR})$	$\hat{e}_2(\hat{\tau}_{LR})$	$\widehat{\text{var}}(\hat{\tau}_{LR})$
	4.4	226.9	1526.3	1733.8
$small(1,1)$	4.7	219.8	1675.6	1877.0
	4.3	292.0	1372.1	1656.4
	3.4	175.2	211.7	384.5
$small(2,1)$	3.2	126.2	147.9	270.2
	3.3	130.2	147.4	275.4
	2.5	56.8	115.7	170.9
$small(2,2)$	2.7	83.1	168.8	250.3
	2.5	55.4	131.8	185.4
	2.3	47.1	94.8	139.4
$veg(1,1)$	2.4	47.7	112.1	156.6
	2.5	46.5	111.3	153.9
	2.2	32.0	24.1	55.6
$veg(2,1)$	2.1	34.3	29.8	63.1
	2.0	39.7	37.2	76.0
	1.7	15.4	25.5	40.5
$veg(2,2)$	1.7	15.5	22.9	38.0
	1.7	16.6	25.5	41.7

Notice that for $small(1,1)$, the standard errors were very large in estimating the second error component and thus the variance, almost as large as that of (1, 1) in Table 5.3A of the simulated population. These standard errors of $small(1,1)$ were also large compared to those of $small(1,1)$ tabulated in Table 5.3C (801.41 for $\hat{e}_2$ and 976.4 for $\widehat{\text{var}}(\hat{\tau}_{LR})$). These much smaller standard errors could be outliers or could be due to small sample size. The standard error of other cases in Table 5.6 was pretty much the same as what was tabulated in Tables 5.3B and 5.3C under the same case. Standard errors were reduced considerably after increasing the sample size (also see $(4,1)$ and $(4,2)$ in Tables 5.3B and 5.3C). So the poor bias we obtained was probably mainly due to noise during simulation and we need to increase the sample size to control it.

We also see that, in Tables 5.4 and 5.5, coverage of the confidence interval estimates was poor for $veg(1,1)$ due to the severe underestimation of the variance. On the other hand, coverage rates for the $small(1,1)$ were better than $(1,1)$ of the simulated population because of the severe overestimation of the variance. The estimated variances were so large that the confidence interval estimates covered the true population size close to the nominal level. As we increase the sample size, coverage improves for both subpopulations. There were satisfactory results for $(4,2)$ of both subpopulations where coverage rates of at least one confidence interval estimates were not significantly different from nominal levels. In other words, for both subpopulations, we need a phase I sample size of at least $400(> 88 \times 4)$ and a phase II sample size of at least $200(> 88 \times 2)$.

In Tables 5.4 and 5.5, we noticed that overall, approximate confidence interval estimates obtained from the three reciprocal transformations of population size estimates had better coverage for the simulated population and for both subpopulations. Both confidence interval estimates $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ also gave good coverage. The confidence interval $\hat{\tau}_{log}$ performs uniformly better and has the shortest expected interval width compared to those constructed using reciprocal transformations and of $\hat{\tau}_{lg2}$.

For the simulated population and the subpopulations, we have also observed that approximate 'true' second error components e_{21},e_{22} of $(2, 1)$ were less than half (about a third) of that of $(1, 1)$ (see theoretical explanation in third paragraph of section 5.3). The approximate 'true' second error components for $(4, 1)$ were about half of that of $(2, 1)$, agreeing with the theory. This implies that the larger the sample size, the smaller the term $\boldsymbol{x}_j'\Sigma\boldsymbol{x}_j$ (and less noise in the simulation) and the better is the approximation of halving the 'true' second error component.

5.4.2 Simulation results of the simulated population of 116 groups and the subpopulations with sightability model (5.2.2) as the true model

In this subsection, we assume sightability model (5.2.2), taking log of group size, as the true model for the simulated population of $M = 116$ groups and the subpopulations *veg*, *small*. Simulation results of all listed (r_1, r_2) are tabulated, using the same notation as before, in Tables 5.7A, 5.7B and 5.7C. Coverage of 90% and 95% confidence interval estimate of population size are tabulated in Tables 5.8 and 5.9.

For this simulated population, sighting probabilities ranged from 0.0185 to 0.9884 and the three quantiles are 0.3332, 0.6440 and 0.8309. The average sighting probability is 0.5814. One complete pass through the population generates an average sample size of 68 groups or about 59%) and the average number of seen animals was 616. For subpopulation *veg*, sighting probabilities ranged from 0.1669 to 0.9884 and the three quantiles are 0.4898, 0.7326 and 0.8730. The average sighting probability is 0.679. One complete pass through subpopulation *veg* generates an average sample size of 56 groups or about 68% and the average number of seen animals was 470. For subpopulation *small*, sighting probabilities ranged from 0.0185 to 0.9781 and the three quantiles are $0.2804, 0.5268 and 0.7674$. The average sighting probability is 0.5229. One complete pass through subpopulation

Table 5.7A Simulation results of $k = 1000$ population estimation trials for listed (r_1, r_2) of the simulated population of $M = 116$ groups and size $\tau = 866$, with sightability model (5.2.2) as the true model[7]

Statistic	Case $(\infty,1)$	$(\infty,2)$	$(\infty,3)$	$(1,1)$	$(1,2)$
$\hat{\tau}_{LR}$	867.8(2.6)	864.1(1.8)	866.1(1.5)	869.8(4.1)	870.5(3.7)
$\text{bias}(\hat{\tau}_{LR})$	0.2	−0.2	0.1	0.4	0.5
e_1	7309.8	3654.9	2436.6	7309.8	3654.9
$\hat{e}_1(\hat{\tau}_{LR})$	7424.0(92.6)	3604.1(32.5)	2432.5(18.4)	7666.1(412.5)	3888.8(154.7)
$\text{bias}(\hat{e}_1)$	1.5	−1.3	−0.1	4.8	6.4
e_{21}	–	–	–	8802.7	9777.4
e_{22}	–	–	–	8606.0	8625.3
$\hat{e}_2(\hat{\tau}_{LR})$	–	–	–	10681.1(1553.6)	9973.6(829.8)
$\text{bias}(\hat{e}_2, e_{21})$	–	–	–	21.3	2.0
$\text{bias}(\hat{e}_2, e_{22})$	–	–	–	24.1	15.6
var_1	7309.8	3654.9	2436.6	16112.5	13432.3
var_2	7309.8	3654.9	2436.6	15915.9	12280.2
$\widehat{\text{var}}(\hat{\tau}_{LR})$	7424.0(92.6)	3604.1(32.5)	2432.5(18.4)	18347.2(1948.9)	13862.4(972.5)
$\text{bias}(\widehat{\text{var}}, \text{var}_1)$	1.5	−1.3	−0.1	13.8	3.2
$\text{bias}(\widehat{\text{var}}, \text{var}_2)$	1.5	−1.3	−0.1	15.2	12.8
$\text{obs var}(\hat{\tau}_{LR})$	7055.8	3534.0	2485.2	17085.8	14226.3
$\text{bias}(\text{var}_1, \text{obs var})$	3.6	3.4	−1.9	−5.7	−5.5
$\text{bias}(\text{var}_2, \text{obs var})$	3.6	3.4	−1.9	−6.8	−13.6
$\text{bias}(\widehat{\text{var}}, \text{obs var})$	5.2	1.9	−2.1	7.3	−2.5

Statistic	Case $(2,1)$	$(2,2)$	$(4,1)$	$(4,2)$
$\hat{\tau}_{LR}$	868.7(3.5)	868.4(2.7)	870.3(3.1)	864.6(2.2)
$\text{bias}(\hat{\tau}_{LR})$	0.3	0.2	0.5	−0.1
e_1	7309.8	3654.9	7309.8	3654.9
$\hat{e}_1(\hat{\tau}_{LR})$	7505.9(210.8)	3671.8(77.7)	7621.6(164.0)	3628.4(56.3)
$\text{bias}(\hat{e}_1)$	2.6	0.4	4.2	−0.7
e_{21}	3821.1	4056.5	1949.7	1748.0
e_{22}	3815.0	3809.5	1824.5	1805.8
$\hat{e}_2(\hat{\tau}_{LR})$	4147.3(195.0)	3683.7(107.6)	1945.1(59.0)	1697.3(36.7)
$\text{bias}(\hat{e}_2, e_{21})$	8.5	−9.1	−0.2	−2.9
$\text{bias}(\hat{e}_2, e_{22})$	8.4	−3.3	6.6	−6.0
var_1	11131.0	7711.5	9259.5	5402.9
var_2	11124.9	7464.5	9134.3	5460.7
$\widehat{\text{var}}(\hat{\tau}_{LR})$	11653.3(399.7)	7355.6(183.9)	9566.8(221.5)	5325.7(92.3)
$\text{bias}(\widehat{\text{var}}, \text{var}_1)$	4.6	−4.6	3.3	−1.4
$\text{bias}(\widehat{\text{var}}, \text{var}_2)$	4.7	−1.4	4.7	−2.4
$\text{obs var}(\hat{\tau}_{LR})$	12567.1	7298.7	10075.9	5081.7
$\text{bias}(\text{var}_1, \text{obs var})$	−11.4	5.6	−9.1	6.3
$\text{bias}(\text{var}_2, \text{obs var})$	−11.4	2.2	−9.3	7.4
$\text{bias}(\widehat{\text{var}}, \text{obs var})$	−7.2	0.7	−5.0	4.8

[7] All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 1000 simulated estimates. Average sample size is 68 groups and average number of seen animals is 616 in a (1,1).

Table 5.7B Simulation results of 1000 population estimation trials for listed cases of the population *veg* of $M = 82$ groups and $\tau = 573$, with sightability model (5.2.2) as the true model [8]

	Case				
Statistic	veg(1, 1)	veg(2, 1)	veg(2, 2)	veg(4, 1)	veg(4, 2)
$\hat{\tau}_{LR}$	569.4(1.7)	571.0(1.4)	572.1(1.2)	573.4(1.3)	573.8(1.0)
bias($\hat{\tau}_{LR}$)	−0.6	−0.3	−0.1	0.1	0.1
e_1	1389.6	1389.6	694.8	1389.6	694.8
$\hat{e}_1(\hat{\tau}_{LR})$	1294.9(30.5)	1354.0(21.7)	673.1(10.5)	1383.9(14.3)	689.8(7.3)
bias($\hat{e}_1$)	−6.8	−2.5	−3.1	−0.4	−0.7
e_{21}	2046.4	834.1	818.3	342.5	395.6
e_{22}	1827.4	777.4	769.8	355.5	357.9
$\hat{e}_2(\hat{\tau}_{LR})$	1477.8(75.1)	694.1(19.8)	667.3(17.0)	336.7(5.1)	331.2(4.9)
bias($\hat{e}_2, e_{21}$)	−27.8	−16.8	−18.5	−1.6	−16.2
bias($\hat{e}_2, e_{22}$)	−19.1	−10.7	−13.3	−5.2	−7.4
var_1	3435.8	2223.7	1513.1	1732.1	1090.4
var_2	3217.0	2167.0	1464.6	1745.1	1052.7
$\widehat{\text{var}}(\hat{\tau}_{LR})$	2772.7(102.1)	2048.2(40.5)	1340.5(27.2)	1720.7(19.3)	1021.1(13.1)
bias($\widehat{\text{var}}, \text{var}_1$)	−19.3	−7.9	−11.4	−0.6	−6.3
bias($\widehat{\text{var}}, \text{var}_2$)	−13.8	−5.5	−8.5	−1.4	−3.0
obs var($\hat{\tau}_{LR}$)	3040.1	2175.6	1483.3	1703.4	1036.7
bias(obs var, var_1)	−11.5	−2.1	−1.9	−1.6	−4.9
bias(obs var, var_2)	−5.5	0.4	1.3	−2.4	−1.5
bias($\widehat{\text{var}}$, obs var)	−8.8	−5.1	−9.6	1.0	−1.5

[8]All estimates except those of *veg*(1, 1) shown in above tables are averages of the 1000 simulated estimates. estimation trials. Bias is in %. Numbers in parentheses are the standard errors of the average of 1000 simulated estimates. For subpopulation *veg*, average sample size is 56 groups and average number of seen animals is 470. For subpopulation *small*, average sample size is 46 groups and average number of seen animals is 194.

Table 5.7C Simulation results of 1000 population estimation trials for listed (r_1, r_2) of the population *small* of $M = 88$ groups and $\tau = 331$, with sightability model (5.2.2) as the true model [9]

	Case				
Statistic	small(1, 1)	small(2, 1)	small(2, 2)	small(4, 1)	small(4, 2)
$\hat{\tau}_{LR}$	336.5(2.4)	333.4(2.2)	332.4(1.5)	327.8(1.6)	328.8(1.2)
bias($\hat{\tau}_{LR}$)	1.6	0.7	0.5	−0.9	−0.6
e_1	2168.3	2168.3	1084.1	2168.3	1084.1
$\hat{e}_1(\hat{\tau}_{LR})$	2198.1(81.8)	2434.1(190.1)	1119.2(28.0)	2135.9(50.4)	1056.5(19.6)
bias($\hat{e}_1$)	0.9	12.2	3.2	−1.5	−2.5
e_{21}	3806.5	1367.3	1426.6	650.1	613.1
e_{22}	3315.9	1370.2	1365.9	638.7	637.2
$\hat{e}_2(\hat{\tau}_{LR})$	4995.8(593.9)	2246.3(596.8)	1430.4(60.7)	653.8(28.9)	580.1(16.7)
bias($\hat{e}_2, e_{21}$)	31.2	64.3	0.3	0.6	−5.3
bias($\hat{e}_2, e_{22}$)	50.6	63.9	4.7	2.4	−9.9
var_1	5974.8	3535.7	2510.8	2818.4	1697.2
var_2	5484.3	3538.5	2450.1	2807.0	1721.3
$\widehat{\text{var}}(\hat{\tau}_{LR})$	7194.0(659.2)	4680.6(782.3)	2549.6(88.0)	2789.8(78.2)	1636.7(36.1)
bias($\widehat{\text{var}}, \text{var}_1$)	20.4	32.4	1.5	−1.0	−3.6
bias($\widehat{\text{var}}, \text{var}_2$)	31.17	32.27	4.06	−0.61	−4.92
obs var($\hat{\tau}_{LR}$)	6137.6	5226.0	2500.8	2856.95	1610.47
bias(obs var, var_1)	2.72	47.81	−0.40	1.37	−5.11
bias(obs var, var_2)	11.91	47.69	2.07	1.78	−0.06
bias($\widehat{\text{var}}$, obs var)	17.21	−10.44	1.95	−10.76	1.63

[9] All estimates except those of *veg*(1, 1) shown in above tables are averages of the 1000 simulated estimates. estimation trials. Bias is in %. Numbers in parentheses are the standard errors of the average of 1000 simulated estimates. For subpopulation *veg*, average sample size is 56 groups and average number of seen animals is 470. For subpopulation *small*, average sample size is 46 groups and average number of seen animals is 194.

Table 5.8 Coverage of 1000 simulated 90% confidence interval estimates that contained τ for all (r_1, r_2) listed in Tables 5.7A, 5.7B and 5.7C [10]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	88(283)	88(284)	88(288)	88(291)	88(297)		
$(\infty, 2)$	89(198)	89(198)	89(199)	90(200)	90(202)		
$(\infty, 3)$	89(162)	89(163)	90(163)	91(164)	91(165)		
(1, 1)	84*(446)	84*(450)	85*(461)	85*(477)	85*(501)	85*(467)	84*(529)
(1, 2)	84*(387)	84*(391)	84*(398)	84*(407)	83*(423)	84*(402)	85*(443)
(2, 1)	85*(355)	87*(357)	88(362)	87*(370)	88(382)	88(367)	88(398)
(2, 2)	88(282)	87*(283)	88(286)	88(289)	88(296)	88(288)	88(305)
(4, 1)	87*(322)	87*(324)	87*(327)	87*(333)	87*(341)	85*(330)	85*(354)
(4, 2)	89(240)	90(241)	90(242)	89(245)	89(248)	89(245)	89(254)
veg(1, 1)	85*(173)	84*(174)	84*(176)	84*(177)	83*(180)	85*(182)	86*(205)
veg(2, 1)	88(149)	87*(149)	88(150)	88(151)	87*(154)	87*(155)	87*(169)
veg(2, 2)	86*(120)	86*(120)	86*(121)	86*(122)	87*(123)	88(123)	88(131)
veg(4, 1)	90(136)	90(136)	91(138)	91(138)	91(140)	89(141)	88(151)
veg(4, 2)	88(105)	88(105)	89(106)	89(106)	89(107)	89(107)	88(112)
small(1, 1)	84*(279)	86*(287)	86*(305)	86*(337)	86*(400)	86*(295)	85*(343)
small(2, 1)	86*(225)	87*(230)	86*(238)	87*(254)	87*(280)	87*(235)	87*(261)
small(2, 2)	87*(166)	87*(168)	88(172)	89(177)	88(186)	88(170)	88(181)
small(4, 1)	85*(174)	85*(176)	85*(180)	85*(187)	86*(197)	85*(179)	87*(193)
small(4, 2)	87*(133)	88(134)	89(136)	89(139)	89(143)	89(135)	88(142)

[10]Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ in both Tables 5.8 and 5.9 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. These CI were not computed when sighting probabilities are known. Coverage with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

Table 5.9 Coverage of 1000 simulated 95% confidence interval estimates that contained τ for all (r_1, r_2) listed in Tables 5.7A,5.7B,5.7C [11]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	92*(338)	93*(340)	93*(345)	93*(351)	93*(361)		
$(\infty, 2)$	94(235)	94(236)	94(238)	94(240)	94(243)		
$(\infty, 3)$	94(193)	95(194)	94(195)	95(195)	95(197)		
$(1,1)$	90*(531)	90*(539)	90*(557)	91*(585)	91*(632)	92*(582)	92*(660)
$(1,2)$	87*(462)	89*(467)	89*(478)	90*(496)	90*(524)	91*(496)	92*(546)
$(2,1)$	90*(423)	91*(428)	91*(436)	91*(450)	92*(471)	93*(450)	94(490)
$(2,2)$	92*(336)	93*(338)	93*(343)	93*(349)	93*(359)	93*(350)	93*(370)
$(4,1)$	91*(383)	92*(386)	92*(393)	93*(403)	93*(418)	92*(404)	91*(433)
$(4,2)$	93*(286)	93*(288)	93*(290)	93*(295)	94(300)	94(295)	94(308)
veg$(1,1)$	87*(206)	88*(207)	89*(210)	89*(213)	89*(218)	91*(226)	92*(256)
veg$(2,1)$	91*(177)	92*(178)	92*(180)	92*(182)	93*(185)	93*(190)	93*(208)
veg$(2,2)$	91*(144)	92*(144)	92*(145)	92*(146)	92*(147)	93*(150)	94(160)
veg$(4,1)$	94(163)	95(163)	95(164)	95(166)	95(168)	93*(172)	93*(185)
veg$(4,2)$	94(125)	94(126)	95(126)	95(127)	94(128)	94(130)	94(135)
small$(1,1)$	87*(332)	90*(346)	91*(377)	91*(440)	92*(596)	93*(371)	93*(432)
small$(2,1)$	90*(268)	92*(275)	92*(291)	93*(320)	93*(374)	94(290)	94(323)
small$(2,2)$	91*(198)	93*(201)	93*(207)	92*(217)	93*(234)	93*(207)	94(221)
small$(4,1)$	89*(207)	90*(210)	92*(218)	93*(230)	93*(249)	92*(219)	93*(235)
small$(4,2)$	92*(159)	93*(160)	93*(164)	94(169)	94(176)	94(164)	94(171)

small generates an average sample size of 46 groups or about 52% and the average number of seen animals was 180. In this subsection, we simulated $k = 1000$ population estimation trials. The simulation results showed all the phenomena we observed in the simulation results of the simulated population and subpopulations with sightability model (5.2.1) as the true model. This time, however, when sighting probabilities are known, we need a larger phase II sample size of at least $300(> 116 \times 2)$ for the coverage of confidence interval estimates of population size of the simulated population for coverage rates to be not significantly different from nominal levels (see coverage under $(\infty, 1)$ and $(\infty, 2)$ in Tables 5.8 and 5.9). When sighting probabilities are estimated, we also need a larger phase II sample

[11] Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ in both Tables 5.8 and 5.9 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. These CI were not computed when sighting probabilities are known. Coverage with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

than when we assumed sightability model (5.2.1) for coverage rates of at least one confidence interval estimates to be not significantly different from nominal levels. The size of phase I sample needed to be at least $500(> 116\times4)$ and the size of phase II sample needed to be at least $300(> 116 \times 2)$. In Tables 5.8 and 5.9, we noticed, again, that approximate confidence interval estimates obtained from the three reciprocal transformation of population size estimates had better coverage for the simulated population and for both subpopulations. Both confidence interval estimates $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ also gave good coverage. Again, the confidence interval $\hat{\tau}_{log}$ performs better uniformly and has the shortest expected interval width when compared to that of $\hat{\tau}_{lg2}$ and of confidence intervals constructed from reciprocal transformations.

For the simulated population, the standard errors of all except population size estimates and $\hat{e}_2$ of $(2,2)$ in Table 5.7A were larger than the respective standard errors in Table 5.3A. For example, the standard error of $\hat{e}_1$ was 276.6 in Table 5.3A and was 412.57 in Table 5.7A. Also note that the error components and hence the variance in Table 5.7A were larger. Recall that when sighting probabilities are known, the variance of population size estimator (see (3.2.1)) and the asymptotically unbiased estimator of the variance (see (3.2.2)) are sums of product of the square of the group size and the reciprocal of the sighting probability. Therefore, both the variance and its estimator will be dominated by larger group sizes. If sighting probabilities of larger group sizes are proportionally larger and the majority of the groups are large groups, both the variance and its estimator will be smaller. We have observed for the simulated population that for group sizes of at least 20, the sighting probabilities are larger when we assumed sightability model (5.2.1). Although there are only 13 out of 116 groups ($13/116 \approx 11\%$) with group size of at least 20, the variance and its estimate under $(\infty, 1)$ and the true first error component e_1 under other listed (r_1, r_2) were smaller in Table 5.3A compared to their respective values in Table 5.7A.

When sighting probabilities are estimated, the true second error component e_2 is also dominated by large group sizes and consisted of $\text{var}(\widehat{\Theta}_j)$ and $\text{cov}(\widehat{\Theta}_{j1}, \widehat{\Theta}_{j2})$. If we compare simulation results obtained using models (5.2.1) and (5.2.2) as true models, it is not obvious when the variance of the population size estimator will be smaller. For the simulated population of $M = 116$ groups when sighting probabilities are estimated, the variance and its estimate were smaller for all listed (r_1, r_2) in Table 5.3A than their respective values in Table 5.7A. Later in the following subsection, when we look at the simulation results of population *Penn* of $M = 107$ groups, we will see that this is not true.

In Tables 5.7B and 5.7C, standard errors of all except $\hat{e}_2$ and $\widehat{\text{var}}(\hat{\tau}_{LR})$ of $small(2,1)$ were smaller than those in Tables 5.3B and 5.3C. Another two simulation runs of $k = 1000$ for $small(2,1)$ gave standard errors 100.56 and 119.57 for $\hat{e}_2$ and 168.98 and 190.85 for $\widehat{\text{var}}(\hat{\tau}_{LR})$. These standard errors were smaller than or close to those tabulated in Table 5.3C and Table 5.6. So the standard errors we obtained from the first simulation run of $small(2,1)$, tabulated in Table 5.7C, could be outliers.

5.4.3 Simulation results of population *Penn* - the composite data of 107 groups from Pennsylvania, with sightability models (5.2.1) and (5.2.2) as true models

In this subsection, we perform simulation based on population *Penn*, the composite data of $M = 107$ groups from Pennsylvania. We first assumed sightability model (5.2.2). The simulation results are presented in Table 5.10 and coverage of confidence interval estimates for population size are tabulated in Tables 5.11, 5.12 as follows.

Table 5.10 Simulation results of 1000 population estimation trials for listed (r_1, r_2) of population *Penn* of $M = 107$ groups and size $\tau = 834$, with sightability model (5.2.2) as the true model[12]

Statistic	Case $(\infty,1)$	$(\infty,2)$	$(\infty,3)$	$(1,1)$	$(1,2)$
$\hat{\tau}_{LR}$	836.0(2.5)	835.7(1.8)	838.2(1.4)	840.8(4.2)	836.3(3.6)
$\text{bias}(\hat{\tau}_{LR})$	0.2	0.2	0.5	0.8	0.2
e_1	6707.0	3353.5	2235.6	6707.0	3353.5
$\hat{e}_1(\hat{\tau}_{LR})$	6777.0(90.8)	3372.2(33.4)	2268.5(18.0)	7099.7(333.3)	3327.6(152.1)
$\text{bias}(\hat{e}_1)$	1.0	0.5	1.4	5.8	−0.7
e_{21}	–	–	–	10232.0	8947.8
e_{22}	–	–	–	9053.0	8911.4
$\hat{e}_2(\hat{\tau}_{LR})$	–	–	–	12577.3(1403.1)	11139.9(1213.4)
$\text{bias}(\hat{e}_2, e_{21})$	–	–	–	22.9	24.5
$\text{bias}(\hat{e}_2, e_{22})$	–	–	–	38.9	25.0
var_1	6707.0	3353.5	2235.6	16939.0	12301.3
var_2	6707.0	3353.5	2235.6	15760.0	12264.9
$\widehat{\text{var}}(\hat{\tau}_{LR})$	6777.0(90.8)	3372.2(33.4)	2268.5(18.0)	19677.1(1719.2)	14467.5(1359.2)
$\text{bias}(\widehat{\text{var}}, \text{var}_1)$	1.0	0.5	1.4	16.1	17.6
$\text{bias}(\widehat{\text{var}}, \text{var}_2)$	1.0	0.5	1.4	24.8	17.9
obs $\text{var}(\hat{\tau}_{LR})$	6708.1	3301.5	2197.2	18116.3	13571.8
$\text{bias}(\text{var}_1, \text{obs var})$	−0.2	1.5	1.7	−6.5	−9.3
$\text{bias}(\text{var}_2, \text{obs var})$	−0.2	1.5	1.7	−13.0	−9.6
$\text{bias}(\widehat{\text{var}}, \text{obs var})$	1.0	2.1	3.2	8.6	6.6

Statistic	Case $(2,1)$	$(2,2)$	$(4,1)$	$(4,2)$
$\hat{\tau}_{LR}$	841.3(3.4)	830.9(2.6)	838.3(3.0)	833.8(2.2)
$\text{bias}(\hat{\tau}_{LR})$	0.8	−0.3	0.5	−0.1
e_1	6707.0	3353.5	6707.0	3353.5
$\hat{e}_1(\hat{\tau}_{LR})$	7247.6(245.0)	3373.3(86.4)	7074.0(147.8)	3366.1(56.2)
$\text{bias}(\hat{e}_1)$	8.0	0.5	5.4	0.3
e_{21}	3892.2	4018.9	1805.4	1706.0
e_{22}	3904.3	3883.5	1828.3	1828.6
$\hat{e}_2(\hat{\tau}_{LR})$	4825.4(355.4)	3817.5(157.9)	2001.8(62.7)	1722.7(44.3)
$\text{bias}(\hat{e}_2, e_{21})$	23.9	−5.0	10.8	0.9
$\text{bias}(\hat{e}_2, e_{22})$	23.5	−1.7	9.4	−5.7
var_1	10599.2	7372.4	8512.4	5059.5
var_2	10611.3	7237.0	8535.3	5182.1
$\widehat{\text{var}}(\hat{\tau}_{LR})$	12073.1(590.5)	7190.9(242.4)	9075.9(209.0)	5088.8(99.8)
$\text{bias}(\widehat{\text{var}}, \text{var}_1)$	13.9	−2.4	6.6	0.5
$\text{bias}(\widehat{\text{var}}, \text{var}_2)$	13.7	−0.6	6.3	−1.8
obs $\text{var}(\hat{\tau}_{LR})$	11597.3	7208.6	9156.1	4847.5
$\text{bias}(\text{var}_1, \text{obs var})$	−8.6	2.2	−7.0	4.3
$\text{bias}(\text{var}_2, \text{obs var})$	−8.5	0.3	−6.7	6.9
$\text{bias}(\widehat{\text{var}}, \text{obs var})$	4.1	−0.2	−0.8	4.9

[12]All estimates shown here are averages of the 1000 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 1000 estimates. Average sample size is 61 groups and the number of seen animals is 614 in a (1,1).

Table 5.11 Coverage of 1000 simulated 90% confidence interval estimates that contained τ for all (r_1, r_2) listed in Table 5.10[13]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	88(271)	88(272)	89(274)	88(278)	88(283)		
$(\infty, 2)$	90(191)	90(191)	90(193)	90(193)	90(195)		
$(\infty, 3)$	90(157)	90(157)	90(157)	90(158)	90(160)		
$(1, 1)$	82*(462)	83*(467)	84*(479)	84*(499)	83*(529)	85*(490)	86*(576)
$(1, 2)$	83*(396)	84*(399)	84*(407)	85*(419)	84*(437)	86*(415)	87*(472)
$(2, 1)$	87*(361)	88(364)	87*(369)	87*(379)	87*(392)	87*(377)	86*(418)
$(2, 2)$	85*(279)	86*(280)	86*(283)	86*(287)	86*(293)	87*(287)	87*(308)
$(4, 1)$	86*(313)	86*(315)	85*(319)	85*(325)	85*(333)	85*(324)	85*(352)
$(4, 2)$	89(235)	89(235)	89(237)	89(240)	89(243)	90(239)	90(251)

Table 5.12 Coverage of 1000 simulated 95% confidence interval estimates that contained τ for all (r_1, r_2) listed in Table 5.10 [13]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	92*(323)	93*(325)	93*(329)	94(335)	94(344)		
$(\infty, 2)$	94(228)	95(228)	95(230)	95(232)	95(235)		
$(\infty, 3)$	95(187)	95(187)	95(187)	95(189)	95(190)		
$(1, 1)$	87*(550)	89*(560)	90*(580)	90*(616)	91*(674)	92*(618)	92*(726)
$(1, 2)$	87*(472)	89*(478)	89*(491)	89*(512)	90*(545)	93*(518)	93*(589)
$(2, 1)$	90*(431)	91*(435)	91*(445)	92*(461)	91*(485)	93*(466)	93*(517)
$(2, 2)$	89*(332)	91*(334)	91*(339)	91*(346)	92*(357)	93*(351)	94(376)
$(4, 1)$	90*(373)	91*(377)	91*(383)	91*(393)	92*(408)	92*(397)	91*(432)
$(4, 2)$	92*(280)	93*(281)	93*(284)	94(287)	94(293)	95(290)	94(305)

Sighting probabilities ranged from 0.0518 to 0.9884 and the three quantiles are $0.2217, 0.5712 and 0.8947$. The average sighting probability is 0.5610. One complete pass through the population generates an average sample size of 61 or about 57% and the average number of seen animals was 619.

Everything we observed from the simulation results of the simulated population of $M = 116$ groups, presented in Table 5.3A and Table 5.7A, were again observed in the simulation results in Table 5.10 of this population *Penn*. For instance, it took a phase I sample size of at least $500(> 107 \times 4)$ or $(4, 1)$ to half the

[13]Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ in both Tables 5.11 and 5.12 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. These CI were not computed when sighting probabilities are known. Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths.

'true' second error component e_2 of $(2,1)$ and the 'true' second error component of $(2,1)$ was also about a third of that of $(1,1)$, contradicting the approximation theory. Except for $(\infty,1)$ and $(4,2)$, the estimated variance of the population size estimator was closer to the observed variance of population size estimates than to the approximate 'true' variance, as observed in Table 5.3A and Table 5.7A.

In Table 5.10, as before, the asymptotically unbiased estimator $\hat{e}_1$ of the first error component did well with the largest bias of 8.0% under $(2,1)$. Notice however the significant bias of $\hat{e}_2$, the asymptotically unbiased estimator of the second error component, under (1,1),(1,2) and (2,1) (see bias$(\hat{e}_2, e_{21})$ and bias$(\hat{e}_2, e_{22})$). This significant bias affected the performance of the asymptotically unbiased estimator of the variance (see bias$(\widehat{\mathrm{var}}(\hat{\tau}_{LR}), \mathrm{var}_1)$ and bias$(\widehat{\mathrm{var}}(\hat{\tau}_{LR}), \mathrm{var}_2)$). This situation has improved considerably when we tried $(2,2)$. Under $(2,2)$, $(4,1)$ and $(4,2)$, bias of estimates of second error components relative to the approximate 'true' second error components, and bias of estimates of the variance relative to the approximate 'true' variances were much smaller (largest was bias$(\hat{e}_2, e_{21}) = 10.8$ under $(4,1)$).

From Tables 5.11 and 5.12, when sighting probabilities are known, we need a phase II sample size of at least $200(> 107 \times 1)$ for coverage rates of at least one of the confidence interval estimates to not be significantly different from nominal levels. When sighting probabilities are unknown and are estimated, we need a phase I sample size of at least $500(> 107 \times 4)$ and a phase II sample size of at least $300(> 107 \times 2)$ for coverage rates of at least one confidence interval estimates to be not significantly different from nominal levels. Once again, the approximate confidence interval estimates obtained from the three reciprocal transformations of population size estimates and both confidence interval estimates constructed with the assumption that $(\hat{\tau} - T)$ is lognormally distributed gave better coverage. The confidence interval $\hat{\tau}_{log}$ performs better uniformly and has the shortest

Table 5.13 Simulation results of 1000 population estimation trials of population *Penn* of $M = 107$ groups and size $\tau = 834$, with sightability model (5.2.1) as the true model[14]

Statistic	Case $(\infty, 1)$	$(\infty, 2)$	$(\infty, 3)$	$(1, 1)$	$(4, 4)$
$\hat{\tau}_{LR}$	836.8(2.5)	834.4(1.8)	835.89(1.4)	838.2(5.1)	833.4(1.7)
bias($\hat{\tau}_{LR}$)	0.3	0.1	0.2	0.5	−0.1
e_1	6604.5	3302.2	2201.5	6604.5	1651.1
$\hat{e}_1(\hat{\tau}_{LR})$	6585.7(123.7)	3329.9(44.4)	2219.5(24.2)	7331.6(480.1)	1666.3(29.1)
bias($\hat{e}_1$)	−0.2	0.8	0.8	11.0	0.92
e_{21}	–	–	–	12079.9	1634.2
e_{22}	–	–	–	9983.9	1822.7
$\hat{e}_2(\hat{\tau}_{LR})$	–	–	–	16675.2(2820.1)	1527.2(35.9)
bias($\hat{e}_2, e_{21}$)	–	–	–	38.0	−6.5
bias($\hat{e}_2, e_{22}$)	–	–	–	67.0	−16.2
var_1	6604.5	3302.2	2201.5	18684.4	3285.3
var_2	6604.5	3302.2	2201.5	16588.4	3473.9
$\widehat{\text{var}}(\hat{\tau}_{LR})$	6585.7(123.7)	3329.9(44.4)	2219.5(24.2)	24006.9(3246.5)	3193.5(64.5)
bias($\widehat{\text{var}}, \text{var}_1$)	−0.2	0.8	0.8	28.4	−2.8
bias($\widehat{\text{var}}, \text{var}_2$)	−0.2	0.8	0.8	44.7	−8.1
obs var($\hat{\tau}_{LR}$)	6695.6	3451.7	2162.1	26329	2989.3
bias(var_1, obs var)	−1.3	−4.3	1.8	−29.0	9.9
bias(var_2, obs var)	−1.3	−4.3	1.8	−37.0	16.2
bias($\widehat{\text{var}}$, obs var)	−1.6	−3.5	2.6	−8.8	6.83

Table 5.14 Coverage of 1000 simulated 90% confidence interval estimates that contained τ for (r_1, r_2) listed in Table 5.13 [15]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	86*(267)	87*(268)	86*(271)	87*(274)	85*(279)		
$(\infty, 2)$	87*(190)	87*(191)	87*(191)	87*(192)	87*(194)		
$(\infty, 3)$	89(155)	89(156)	89(156)	89(157)	89(157)		
$(1, 1)$	80*(510)	80*(518)	79*(534)	79*(562)	78*(605)	80*(548)	83*(670)
$(4, 4)$	89(186)	88(186)	88(187)	88(189)	89(190)	90(188)	90(195)

[14]All estimates shown here are averages of the 1000 simulated estimates. Bias is in %. Numbers in parentheses are the standard errors of the average of 1000 respective simulated estimates. Average sample size is 61 groups and the number of seen animals is 619 in a (1,1).

[15]Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ in both Tables 5.14 and 5.15 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. These CI were not computed when sighting probabilities are known. Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths.

Table 5.15 Coverage of 1000 simulated 95% confidence interval estimates that contained τ for (r_1, r_2) listed in Table 5.13 [16]

Case	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
$(\infty, 1)$	91*(318)	91*(321)	91*(324)	92*(330)	92*(339)		
$(\infty, 2)$	92*(226)	92*(227)	92*(229)	93*(231)	92*(233)		
$(\infty, 3)$	94(185)	94(185)	95(186)	95(187)	95(189)		
$(1, 1)$	82*(607)	84*(621)	84*(650)	84*(699)	85*(786)	89*(701)	90*(857)
$(4, 4)$	93*(222)	93*(222)	94(224)	94(226)	94(228)	95(227)	95(235)

expected interval width when compared to that of $\hat{\tau}_{lg2}$ and of confidence intervals constructed from reciprocal transformations.

Tables 5.13, 5.14 and 5.15 give simulation results based on population *Penn* with sightability model (5.2.1) as the true model. Sighting probabilities ranged from 0.0460 to 0.9977 and the three quantiles are 0.2699,0.5545 and 0.8754. The average sighting probability is 0.5545. One complete pass through the population generates an average sample size of 61 groups or about 57% and the average number of seen animals was 619.

The asymptotically unbiased estimator of the first error component also performed well for this fourth simulated population. When sighting probabilities are estimated, under (1,1), bias of the asymptotically unbiased estimator of the second error component relative to approximate 'true' second error components were significant (38.04% and 67.02%) as observed in Table 5.10 for our third simulated population. The estimated variance had, thus, significant bias relative to approximate 'true' variances (28.49%, 44.72%). Coverage of all confidence interval estimates were very low (see Tables 5.14,5.15). When we tried $(4, 4)$, all bias were no longer significant. Also, coverage rates of confidence interval estimates

[16] Notation $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ in both Tables 5.14 and 5.15 refer to the confidence interval estimates derived from the assumption that $\hat{\tau}_{LR} - T$ has a log-normal distribution, where T is total count of animals in the sample. These CI were not computed when sighting probabilities are known. Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths.

$\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ were acceptable when compared to nominal levels. This implies we need a phase I sample size of at least $428(> 107 \times 4)$ and a phase II sample size of at least $428(> 107 \times 4)$. When sighting probabilities are known, it needed a phase II sample size of at least $321(107 \times 3)$. For both sighting probabilities known and estimated case, as before, confidence interval estimates obtained from the three reciprocal transformation of population size estimates and both confidence interval estimates $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ gave better coverage. The approximate confidence interval estimates obtained from the three reciprocal transformations of population size estimates and both confidence interval estimates constructed with the assumption that $(\hat{\tau} - T)$ is lognormally distributed gave better coverage. The confidence interval $\hat{\tau}_{log}$ performs better uniformly and has the shortest expected interval width when compared to that of $\hat{\tau}_{lg2}$ and of confidence intervals constructed from reciprocal transformations.

We have noticed, as in our simulated population of $M = 116$ groups, that for population *Penn*, sighting probabilities of groups with sizes of at least 20 are a little larger (the difference is not as much as those of the simulated population of $M = 116$ groups) when we assumed sightability model (5.2.1) than when we assumed sightability model (5.2.2). There are 13 such groups ($13/107 \approx 12\%$). As we compare the 'true' first error component e_1 of Tables 5.10 and 5.13, we see that e_1 in Table 5.13 is slightly smaller ($6604.52 < 6707.01$). However, if we compare the observed variance observed var($\hat{\tau}_{LR}$) of population size estimates under $(1, 1)$ of both tables, we see that observed var($\hat{\tau}_{LR}$) in Table 5.13 is a lot larger ($26329 > 18116.3$). This result differs from the result we observed from simulation results of the simulated population of $M = 116$ groups earlier. We obtained similar results for another 2 simulations of $(1, 1)$ of this simulated population.

5.5 Dominant factor in the variance estimator

From (3.2.2) and in our simulation results, we have observed, for a complete census of primary units, that large group sizes dominate the variance of the population size estimator when sighting probabilities are known. When sighting probabilities are estimated, from (3.2.7) and (3.2.9), it is unclear what is dominant in the variance estimator. To get a better understanding of how e_2 is dominated by large group sizes, we assumed, for the composite data of $M = 107$ groups, the sightability model defined in (5.2.1) but with the signs of the parameters switched, i.e. $\boldsymbol{\beta} = (-2.4132, -0.1322, 1.8224, 0.0423)$. We refer to this sightability model as model (5.2.3). With sightability model (5.2.3) as the true model, sighting probabilities were totally backwards. Large groups have extremely small sighting probabilities (< 0.1) even when they were moving and were under low vegetation cover (eg. 10) during the sighting trial. On the other hand, small groups have extremely high sighting probabilities (close to or more than a 0.9) even when they were stationary and were under high vegetation cover (eg. 80). In this 'abnormal' case, small groups were sighted and large groups were missed most of the time. Thus, the population size estimates were too low (average of 1000 estimates was about 700) most of the time and the estimates of both error components were extremely large (both average over 20000). Although the estimated variance of the population size estimator was large (average over 40000), coverage rates of approximate and confidence interval estimates $\hat{\tau}_{log}$ and $\hat{\tau}_{lg2}$ were still very low. With a sample size of $(10, 10)$, coverage for 90% and 95% confidence levels were only about 75% and 80%. We saw that the second error component could get very large when sighting probabilities of large groups are extremely small although, when we studied the estimated sighting probabilities of large groups of 20 population estimation trials, we observed that sighting probabilities of most large groups were overestimated. We also saw that the second error component could be blown up with large $\text{var}(\widehat{\Theta}_j)$ and $\text{cov}(\widehat{\Theta}_{j1}, \widehat{\Theta}_{j2})$ and we did observed that for some large groups, $\text{var}(\widehat{\Theta}_j)$ and $\text{cov}(\widehat{\Theta}_{j1}, \widehat{\Theta}_{j2})$ were large ($>> 500$). When we assumed sightability models (5.2.1) and (5.2.2) as true models for the composite data of $M = 107$ groups, we found that for large groups (size > 20), $\text{var}(\widehat{\Theta}_j)$ is small (< 10). Also, when at least one of the two groups is a large group, the majority of $\text{cov}(\widehat{\Theta}_{j1}, \widehat{\Theta}_{j2})$ were close to 0 and none of them is > 10 or < -10. These observations indicate that when unknown sighting probabilities are estimated via logistic regression, for the variance of the population size estimator (or both error components) not to blow up, it is important for large groups to have large sighting probabilities when they are expected to have large sighting probabilities.

5.6 Estimation of elk population size in Pennsylvania

The composite data of $M = 107$ groups from Pennsylvania is a phase I sample obtained from 5 sightability surveys conducted on the elk population. Also, 5 replicated population surveys were conducted on the elk population to obtain an independent phase II sample. A total of 82 groups were sighted. This is an example of a complete census of primary units with replicated primary units (see section 3.4). The phase I sample and the phase II sample would be a $(r_1, r_2) = (5, 5)$ using our notation defined for the simulation study. The following table shows the population size estimates, estimates of both error components and variance of the population size estimator for using both sightability models (5.2.1) and (5.2.2) as fitted models. We see that both fitted models gave the same population size estimates of 189. There was little difference between the estimated variance components and the estimated variance of the population size estimator. The variance of population size estimator estimated based on fitted model (5.2.2) with natural log of the group size was 453.94, about 5% less than 478.28, the

Table 5.16 Population size estimates, estimates of error components and variance of the population size estimator for the composite data of 82 sighted groups of 5 independent elk population surveys from Pennsylvania

Estimates	(5.2.1) as fitted model	(5.2.2) as fitted model
Population size, $\hat{\tau}_{LR}$	189	189
Error due to sampling of groups, $\hat{e}_1$	217.53	198.82
Error due to estimation of g_j, $\hat{e}_2$	260.75	255.13
Variance of population size estimator, $\widehat{\text{var}}(\hat{\tau}_{LR})$	478.28	453.94

variance estimated based on fitted model (5.2.1) without transformation of group size. Confidence intervals of 90% and 95% are constructed using reciprocal transformations of the population size estimate and from the assumption that $(\hat{\tau} - T)$, where $T = 834/5 \approx 167$ is the average total elk count of the 5 population surveys, has a lognormal distribution is are presented in the following tables.

Table 5.17 A 90% confidence interval of the population size of elk in Pennsylvania constructed from the population size estimates, with reciprocal transformations and from assumption that $(\hat{\tau} - T)$ is lognormally distributed, obtained based on the composite data with fitted models (5.2.1), (5.2.2)[17]

Sightability model	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
(5.2.1)	(158, 231), 73	(159, 234), 75	(160, 237), 77	(173, 254), 81	(175, 289), 114
(5.2.2)	(159, 230), 71	(160, 233), 73	(161, 235), 74	(173, 251), 78	(176, 284), 108

Table 5.18 A 95% confidence interval of the population size of elk in Pennsylvania constructed from the population size estimates, with reciprocal transformations and from assumption that $(\hat{\tau} - T)$ is lognormally distributed, obtained based on the composite data with fitted models (5.2.1), (5.2.2)[17]

Sightability model	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
(5.2.1)	(153, 241), 88	(155, 245), 90	(156, 250), 94	(172, 279), 107	(174, 325), 151
(5.2.2)	(154, 234), 80	(155, 243), 88	(157, 248), 91	(172, 276), 104	(174, 318), 144

Intervals constructed based on fitted model (5.2.2) is shorter. From our simulation results, we have observed that coverage rates of confidence interval estimates were poor (too low) when both phase I and phase II samples have less than 200

[17]The number next to the interval is the interval width

groups. Therefore, these confidence interval estimates in Tables 5.17 and 5.18 constructed based on a phase I sample of 107 groups and a phase II sample of 82 groups might not include the true population size. Our simulation results suggested that it requires a phase I sample of at least 464 and a phase II sample of at least 464 groups to obtain a confidence interval estimate with reasonable coverage rate.

5.7 Discussion

From the simulation results, we have made the following conclusions.

- The largest bias of population size estimates obtained was 1.7% for the case of subpopulation *small*, $M = 88$ groups and size $\tau = 331$ with model (5.2.1) as the true sightability model. When sighting probabilities are known, all ratios $B/\sigma = |(\hat{\tau} - \tau)/\sqrt{\text{observedvar}(\hat{\tau}_{LR})}|$ were less than 0.1. According to Cochran's (1977) working rule, the bias were negligible. When sighting probabilities were unknown and estimated via logistic regression, all ratios B/σ except for $veg(4,2)$ were less than 0.1. For $veg(4,2)$, the ratio was 0.13. Therefore, all bias of population size estimates were negligible and modest for $veg(4,2)$.

- As expected, increasing the sample size decreases the variance of the population size estimator or increases the precision of estimating the population size. Increasing the sample size also gave us a better estimate of the variance, reducing noise and standard errors of estimates of both error components. Coverage rates of confidence interval estimates of the population size are not different from nominal levels at 0.05 level of significance when phase I and II sample sizes are large.

- The sampling distribution of the population size estimator was positively skewed. Approximate confidence interval estimates obtained from transformations $\hat{\tau}^{-1}$, $\hat{\tau}^{-1/2}$ and $\hat{\tau}^{-3/2}$ of population size estimates gave better coverage rates. Confidence interval estimates constructed from the assumption that $(\hat{\tau} - T)$ has a lognormal distribution also gave better results than $\hat{\tau} \pm z_{\alpha/2}\sqrt{\widehat{\text{var}}(\hat{\tau}_{LR})}$. The confidence interval $\hat{\tau}_{log}$ performs better uniformly and has the shortest estimated expected interval width when compared to that of $\hat{\tau}_{lg2}$ and of confidence intervals constructed from reciprocal transformations.

- By excluding groups with smaller sighting probabilities, estimation of e_2, the error component due to estimation of sighting probabilities, given subpopulation *veg* did not perform well. Approximate 'true' second error component e_2 was severely underestimated and coverage rates of confidence interval estimates for population size were low for $veg(1,1), veg(2,1), veg(1,2)$. When we increased both phase I and phase II sample sizes to $veg(4,2)$, bias of estimating the second error component were reduced considerably (under or close to 10%) and coverage rates of confidence interval estimates were not significantly different from nominal levels.

- By only including groups with smaller sighting probabilities, estimation of e_2 given subpopulation *small* did not perform well. Approximate 'true' second error component was severely overestimated with high standard errors for $small(1,1), small(2,1), small(2,2)$. As a result, confidence interval estimates of population size covered the true population size most of the time and thus, nominal levels were close to nominal levels. When we increased both phase I and phase II sample sizes to $small(4,2)$, bias of estimating the second error component were reduced considerably (under or close to 10%)

and coverage rates of confidence interval estimates were not significantly different from nominal levels.

- We have shown (in section 5.3) that in phase I, if we make 2 complete passes through the population, the approximate 'true' second error component will be half of that of making 1 complete pass through the population. A larger sample size was needed to simulate this theoretical approximate 50% reduction in the second error component. It requires a phase I sample size of at least 464 (or $(4, 1)$) to simulate an approximate 'true' second error component that is half of the approximate 'true' second error component simulated based on half the phase I sample size (or $(2, 1)$). To test this conclusion, we have tried to simulate the theoretical 50% reduction with another two subpopulations, groups with sighting probabilities of at least 0.4 and 0.6 respectively. Our simulation results agreed with this conclusion.

- We have also shown (in section 5.3) that in phase II, if we make $r_2 > 1$ complete passes through the population, to obtain the first true error component e_1, we divide the first true error component computed based on making $r_2 = 1$ complete pass through the population by r_2. The number of complete passes through the population in phase II has no effect on the second error component e_2, i.e. e_2 is the same for $r_2 = 1, 2, \ldots$.

- Based on the simulation results of all 4 simulated populations ($M = 116$ and $M = 107$ groups, with (5.2.1) and (5.2.2) as true models) and both subpopulations *veg* and *small* of the simulated population of $M = 116$ groups, we found for the choice of sample size in both phase I and phase II, coverage rates of at least one of the confidence interval estimates we constructed to not be significantly different from nominal levels at 0.05 level of significance. When sighting probabilities are known, it requires a phase

II sample size of at least 348. When sighting probabilities are estimated via logistic regression, it requires a phase I sample size of at least 464 and a phase II sample size of at least 464.

- When the sighting probabilities are known, the first error component due to sampling of groups is the variance of the population size estimator since there is no second error component due to estimating sighting probabilities. From expressions (3.2.2), we see that the variance of the population size estimator is dominated by large groups. The variance will be smaller if the majority of the population is large groups with larger sighting probabilities. Similarly, when sighting probabilities are estimated via logistic regression, the first error component e_1 in (3.2.6) will be smaller if the majority of the population is large groups with larger sighting probabilities. We observed this result when we compared e_1 of $(\infty, 1)$ and $(1, 1)$ in Tables 5.3A, 5.7A and when we compared e_1 of $(\infty, 1)$ and $(1, 1)$ in Tables 5.10 and 5.13. When sighting probabilities are estimated, it is not obvious when the second error component e_2, which is also dominated by large group sizes (see (3.2.8)), will be smaller because terms such as $\text{var}(\widehat{\Theta}_j)$ and $\text{cov}(\widehat{\Theta}_{j1}, \widehat{\Theta}_{j2})$ are involved. From simulation results obtained based on the 'abnormal' simulated population, We have observed that when unknown sighting probabilities are estimated via logistic regression, it is important for large groups to have large sighting probabilities when they are expected to have large sighting probabilities so that the estimate of the variance of the population size estimator (or both error components) will not get too large. This also implies that if you don't see most of the animals, estimators will do poorly.

6. BOOTSTRAP CONFIDENCE INTERVAL FOR THE POPULATION SIZE

6.1 Introduction

In chapter 5, we reported on the performance of confidence intervals for the population size constructed using the large sample theory and various transformations of the population estimate or by assuming that $(\hat{\tau}_{LR} - T)$ has a lognormal distribution. In this chapter, we report on the performance of confidence intervals for population size constructed using bootstrap methods. We were unable to find an existing bootstrapping procedure which could be directly applied to our problem, namely, a finite population of animal groups whose unknown, unequal sighting probabilities satisfy a logistic model (2.1.3). It is not appropriate, in our case, to apply the usual bootstrap simple random sample with replacement given our phase I and a phase II sampling schemes. Recently, Booth et. al. (1994) and Sitter (1992) have proposed sampling without replacement procedures to obtain bootstrap samples for stratified random sampling and the Rao-Hartley-Cochran method of unequal probability sampling (PPS sampling) of a finite population. Neither of these sampling schemes is applicable to our problem. We have therefore proposed some procedures for obtaining bootstrap samples for our problem.

We shall begin with the case of a complete census of primary units.

6.2 Case 1: A complete census of primary units

We assume that the needed data to construct a single population size estimate are collected in 2 phases. Phase I data collection is for estimating the sightability model and phase II data collection is for estimating the population size via the fitted model obtained in phase I. A single population size estimate thus requires 2 samples, assumed to be independent, from the population.

We proposed independent bootstrapping procedures in each phase of sampling. Only bootstrapping in phase II is needed when sighting probabilities are known.

6.2.1 Bootstrapping in phase I

When we refer to the 'true' sighting trial or 'true' sample, we mean the sighting trial is conducted on a real population or on our simulated population. When we refer to an empirical sighting trial or empirical sample, we mean sampling from an empirical population we generated from the 'true' sample.

Suppose there are M groups in the 'true' phase I sample. In the case of a complete census of primary units, these M groups would form the entire population. We proposed two bootstrapping procedures in phase I :

- (1) After the 'true' sighting trial, add to the 'true' phase I sample the values of the indicator function Z_j (1 if group j is sighted and 0 otherwise), for $j = 1, 2, ..., M$, as a characteristic of the groups. Thereafter, to obtain a bootstrap phase I sample, select a simple random sample with replacement of size M from the 'true' phase I sample. A new sighting trial is not conducted on this empirical population of M groups, rather, the attached indicator variable values are used in fitting a sightability model.

- (2) After the 'true' sighting trial, fit a sightability model based on a 'true' phase I sample and then use it to estimate sighting probabilities $\{g_j^\star\}$ of all groups in the sample, i.e. for $j = 1, 2, ..., M$. These estimated sighting probabilities $\{g_j^\star\}$ are then attached to the groups, as though they were the 'true' sighting probabilities. Then to obtain a bootstrap phase I sample, an empirical sighting trial is conducted with new, independently generated uniform random variable values $\{u_j\}$ and the estimated sighting probabilities $\{g_j^\star\}$. In an empirical sighting trial, a group is determined to be sighted or not by comparing a randomly generated uniform random number, u_j (generated jointly independent for the groups), with the corresponding estimated sighting probability $g_j^\star$. If $u_j \leq g_j$, the j^{th} group for $j = 1, 2, ..., M$ is said to be sighted.

The first procedure is an imitation of the usual bootstrap procedure, i.e. simple random sampling with replacement. The second procedure, on other hand, is somewhat different from the usual bootstrap method. It uses the fitted model of the 'true' phase I sample to obtain the bootstrap phase I sample. We shall refer to procedures 1 and 2 as the nonparametric and the parametric procedures, respectively. The estimated parameters obtained from bootstrap iterations will be refered to as the bootstrap parameter estimates.

6.2.2 Bootstrapping in phase II

Let $\tilde{\boldsymbol{\beta}}_0$ (see (3.2.4)) be the estimated parameters obtained from the 'true' sighting trial of phase I. Also let $\tilde{\boldsymbol{\beta}}_i$ for $i = 1, 2, ..., b$ be the bootstrap parameter estimates of b bootstrap iterations obtained in phase I from either nonparametric or parametric procedure. The 'true' phase II sample is obtained in a second 'true' sighting trial conducted independently from the phase I 'true' sighting trial.

We also proposed two methods to obtain a bootstrap phase II sample.

- (1) - A single empirical population is constructed and independent empirical sighting trials conducted using the phase 1 'true' sighting parameter estimates to estimate sighting probabilities for each group of the phase II 'true' sample. An empirical population is constructed for phase II sampling using these estimated probabilities and for selecting the bootstrap phase II samples. Then a bootstrap phase I sample and a bootstrap phase II sample are used to construct a population size estimate.

To construct the empirical population, we used the $\widetilde{\boldsymbol{\beta}}_0$ obtained in phase I to get $\widehat{\Theta}_{0j} = 1/\hat{g}_{0j}$ where $\hat{g}_{0j}$ is the estimated sighting probability of the j^{th} sighted group (using $\widetilde{\boldsymbol{\beta}}_0$ for model (3.2.4)) in the 'true' phase II sighting trial. Then we replicated the j^{th} sighted group by $[\widehat{\Theta}_{0j} + 0.5]$ times. This replication was repeated for all sighted groups in the 'true' phase II sample to create an empirical population. The notation $[x]$ where x is a positive real number is the largest integer less than x. Next, to obtain a bootstrap phase II sample, we used $\{\hat{g}_{0j}\}$ as 'true' sighting probabilities for the corresponding group in the empirical population. An independent empirical sighting trial was then conducted for each bootstrap iteration i for $i = 1, 2, ..., b$ in phase II. For each group in the i^{th} bootstrap phase II sample, we used the i^{th} phase I bootstrap sample estimated parameters $\widetilde{\beta}_i$ to construct its sighting probability, and hence, to obtain $\hat{\tau}_{LR}$ (see (3.2.5)) and $\widehat{\text{var}}(\hat{\tau}_{LR})$ (sum of (3.2.7) and (3.2.9)).

- (2) - Procedure (2) differs from procedure (1) only in how the empirical population was generated. For each bootstrap phase II sample, a potentially different empirical population was generated. For this procedure, the j^{th} sighted group in the 'true' phase II sighting trial was replicated in the following way.

  First we generated a $U(0, 1)$ random variate p'_j. The j^{th} sighted group was replicated $[\widehat{\Theta}_{0j}]$ times if $p'_j < 1 - dec(\widehat{\Theta}_{0j})$ and $([\widehat{\Theta}_{0j}] + 1)$ times otherwise, where $dec(\widehat{\Theta}_{0j})$ is the decimal part of $\widehat{\Theta}_{0j}$.

Since p'_j for group j in procedure 2 is a random variate generated independently for each phase 2 bootstrap sample, different empirical populations could have been generated for each bootstrap iteration whereas in procedure 1, there is only one empirical population for all the bootstrap iterations.

We shall refer to the first procedure as procedure (1) and the second procedure as the selection rule. A population size estimate obtained from a bootstrap iteration will be referred to as a bootstrap population size estimate.

6.3 Case 2: Simple random sampling without replacement of primary units

Suppose the survey area is divided into N primary units and a sample of n units is selected using simple random sampling. We want to generate an empirical population of primary units from the sample. We will employ one of the methods proposed by Booth et. al. (1994) to generate the empirical population of primary units. The method is stated as follows:

- If $\frac{N}{n} = m$ is an integer, replicate each selected primary unit m times to get an empirical population of $m \times n = N$ primary units.
- If $\frac{N}{n}$ is not an integer, replicate each selected primary unit $[\frac{N}{n}] = m$ times. Since this will not give us an empirical population of N primary units, we have to fill up the remaining $(N - m \times n)$ units. We do that by selecting without replacement $(N - m \times n)$ units from the sample of n units.

For each bootstrap iteration, we take a simple random sample without replacement of the n primary units from the empirical population to get a sample of n primary units. To get an empirical population for the next bootstrap iteration, we simply repeat the step(s) described above. For each bootstrap iteration, use the empirical sample of n units and the 'true' sample of secondary units for the population size estimator defined in (2.1.2) or (2.1.4) and the respective unbiased estimator of the variance of the population size estimator derived in (2.2.2) or (2.2.12).

The method we just described for bootstrapping the first-stage sampling could be easily extended to the case of stratified random sampling. As stated in Booth et. al. (1994), simply use this method to generate empirical strata or subpopulations.

6.4 Simulation study

We want to evaluate, using the Monte Carlo method, small sample performance of our proposed bootstrapping procedures. In order to study our ideas more closely, we examine the case of a complete census of primary units and use a smaller simulated population with $M = 54$ groups.

In our sampling scheme, both phase I and phase II samples are obtained by making one or more complete passes (denoted by r_1 in phase I and r_2 in phase II) through the population. We will use the same notation (r_1, r_2) defined in section 5.3 to denote the number of complete passes made in phase I and phase II, respectively, to construct a single population size estimate. Computer programs for this simulation study are listed in the Appendix.

6.4.1 Simulated population I

We randomly selected $M = 54$ groups from the composite data of $M = 107$ groups of the Pennsylvania elk data. This simulated population has a population size of 422 ($\sum_{j=1}^{M=54} y_j = 422$). To make things even simplier, we considered the following sightability model :

$$\log \frac{g_j}{1-g_j} = \beta_0 + x_j\beta_1. \tag{6.4.1}$$

where x_j is the group size of group j. This is a special case of the model defined in (3.2.4) where there is only one independent variable involved.

6.4.2 Phase I

In phase I, we evaluate the credibility of our proposed procedures by measuring the bias of the average of a set of b bootstrap parameter estimates $\tilde{\boldsymbol{\beta}}_i$ for $i = 1, 2, ..., b$ relative to $\tilde{\boldsymbol{\beta}}_0$, the parameter estimates obtained from the 'true' sighting trial of phase I.

One hundred 'true' phase I sighting trials ($S = 100$) with $r_1 = 1$ were conducted on the simulated population and these $S = 100$ 'true' phase I samples were used to evaluate both nonparametric and parametric procedures. Based on each 'true' phase I sample, we obtained $b = 1000$ bootstrap parameter estimates $\tilde{\boldsymbol{\beta}}_i$ for $i = 1, 2, ..., 1000$. The average of these 1000 bootstrap parameter estimates is then compared with $\tilde{\boldsymbol{\beta}}_0$, the parameter estimate obtained from fitting the 'true' phase I sample to get a bias in %.

Table 6.1 shows, for both procedures, the mean and the standard errors of 100 estimates of bias in %, i.e. $\left(\frac{\tilde{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_0}{\tilde{\boldsymbol{\beta}}_0} 100\right)$.

Table 6.1 Mean and standard error of 100 biases on $S = 100$ 'true' phase I samples with $r_1 = 1$, using nonparametric and parametric procedures of phase I

Nonparametric

				H0: bias=0	
Variable	S	Mean	Std Err	z-stat	p-value
BETA_0	100	-7.79	0.49	-15.90*	0.00
BETA_1	100	13.60	1.22	11.15*	0.00

Parametric

				H0: bias=0	
Variable	S	Mean	Std Err	z-stat	p-value
BETA_0	100	-7.24	1.02	-7.10*	0.00
BETA_1	100	23.72	3.14	7.55*	0.00

We compute a $z-$statistic to test the equality of average bias between two procedures for both parameters. The $z-$statistics are -0.486 for β_0 and -3.001

for β_1. Therefore, the difference between average bias of bootstrap estimates of β_0 obtained using nonparametric and parametric procedures is not significant at $\alpha = 0.05$ level However, the average bias of bootstrap estimates of β_1 obtained using parametric procedure is significantly larger than that obtained using nonparametric procedure at $\alpha = 0.05$ level. We repeated this run with various sets of 100 'true' phase I samples and we obtained similar results. According to Cochran's (1977) working rule (see subsection 5.1.1), since the ratio B/σ = bias/std err is more than 0.7 for simulation results in Table 6.1, bias of bootstrap parameter estimates obtained using either of the procedures is not negligible.

We further evaluate these two procedures by considering $r_1 = 2$, making two complete passes through the population. By doing so, we have increased the sample size. The simulation results for $S = 100$ 'true' phase I samples are as given in Table 6.2.

Table 6.2 Mean and standard error of 100 biases based on $S = 100$ 'true' phase I sighting trials with $r_1 = 2$, using nonparametric and parametric procedures of phase I

Nonparametric

Variable	S	Mean	Std Err	HO: bias=0 z-stat	p-value
BETA_0	100	-3.81	0.17	-22.41*	0.00
BETA_1	100	6.43	0.32	20.09*	0.00

Parametric

Variable	S	Mean	Std Err	HO: bias=0 z-stat	p-value
BETA_0	100	-3.74	0.09	-41.56*	0.00
BETA_1	100	12.56	0.15	83.73*	0.00

Average bias is about half of that in Table 6.1 when $r_1 = 1$. The $z-$statistics for testing equality of average bias between the two procedures are -0.341 for β_0 and -17.433 for β_1. Therefore, the difference between the average bias of bootstrap estimates of β_0 obtained using nonparametric and parametric procedures is not significant at $\alpha = 0.05$ level. However, the average bias of bootstrap estimates of β_1 obtained using parametric procedure is significantly larger than that obtained using nonparametric procedure at $\alpha = 0.05$ level.

The ratio $B/\sigma > 2$ (Cochran, 1977) indicates that bias of bootstrap parameter estimates obtained using either of the procedures is still significant. This observation, however, does not explicitly indicate that these procedures will be unsuccessful in estimating sighting probabilities. We now proceed to test the two procedures we proposed for phase II.

6.4.3 Phase II

To evaluate the procedures we proposed for phase II without phase I, we assume that the sighting probabilities are known. We conducted 1000 'true' sighting ($r_2 = 1$) trials on our simulated population of $M = 54$ groups. In other words, we make $r_2 = 1$ complete pass (or $(\infty, 1)$ using the notation we defined in chapter 5) of the simulated population to obtain one phase II sample. For each sample, we ran $b = 500$ bootstrap iterations to obtain 500 bootstrap $\hat{\tau}_\pi$ (see (3.2.1)). We also evaluated $\widehat{\text{var}}(\hat{\tau}_{LR})$ (see (3.2.3)). With the (5^{th}, 95^{th}) percentiles and (2.5^{th}, 97.5^{th}) percentiles of the sampling distribution of 500 bootstrap $\hat{\tau}_\pi$, we form a 2-sided 90% and a 2-sided 95% bootstrap confidence intervals for τ, the true population size. This will be referred to as the *percentile method* (Efron, 1981a, 1982). We also obtained bootstrap $t-$ statistics (Booth et. al., 1994) for those bootstrap $\hat{\tau}_\pi$'s as follows:

$$t = \frac{\hat{\tau}_\pi - \hat{\tau}_0}{\sqrt{\widehat{\text{var}}(\hat{\tau}_\pi)}}, \qquad (6.4.3)$$

where $\hat{\tau}_0$ is the population size estimate obtained using the 'true' sample. Note that the bootstrap t-statistic does not necessarily has the distribution of the *student* t-statistic. To get confidence intervals for τ, we use the 5^{th}, 95^{th}, 2.5^{th},

97.5^{th} percentiles of the sampling distribution of 500 t-statistics, with $\hat{\tau}_0$ and $\widehat{\text{var}}(\hat{\tau}_0)$ to compute the bounds. To obtain another set of confidence intervals, we replace $\widehat{\text{var}}(\hat{\tau}_0)$ with the observed standard variation $\hat{\sigma}_\pi$ of those 500 bootstrap $\hat{\tau}_\pi$.

If our proposed bootstrap procedures are appropriate, coverage rates of these confidence intervals for the true population size (=422) should be close to their respective actual levels. Simulation results are tabulated in Table 6.3.

Table 6.3 True population size coverage rates of 1000 bootstrap confidence intervals constructed based on 1000 'true' phase II samples, using 500 bootstrap iterations, when sighting probabilities are known[1]

				Using bootstrap$-t$ statistic with $\sqrt{\widehat{\text{var}}(\hat{\tau}_0)}$		with $\hat{\sigma}_\pi$	
	Confidence level	90	95	90	95	90	95
Method	Procedure 1	89.4	93.0*	86.5*	91*	86.3*	90.9*
	Selection Rule	92.5*	96.7*	91.4	94.0	92.7*	94.9

Recall, from subsection 5.1.2, that for $k = 1000$ population estimation trials, a coverage rate is not different from the actual level at 0.05 level of significance if it is within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, both inclusive, for 90% and 95% actual levels respectively. The results in Table 6.3 show that coverage rates of bootstrap confidence intervals constructed using the selection rule and the bootstrap t-statistic with $\sqrt{\widehat{\text{var}}(\hat{\tau}_0)}$, the estimated variance of the population size estimate obtained based on the 'true' phase II sample, were reasonable. Coverage rates of other bootstrap confidence intervals were different from respective nominal levels at 0.05 level of significance. To obtain the results in Table 6.3, we have used a small phase II sample of 54 groups and population size 422.

[1]Coverage rates with $\star$ are different from actual levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively

6.4.4 Combining phase I and phase II

We combine each bootstrap procedure of phase I with each bootstrap procedure of phase II, thereby giving us 4 possible methods to obtain a bootstrap confidence interval for population size when unknown sighting probabilities are estimated using logistic regression. The simulated population is the composite raw data from Pennsylvania with (5.2.2) as the true sightability model. The sightability model has three predictors : the natural log of group size, the behavorial code and percent vegetation cover. There are $M = 107$ groups in the population and the population size is $\tau = 834$. We consider $(1, 1)$, i.e. we make 1 complete pass through the population in phase I and 1 complete pass through of the population in phase II.

We simulate 1000 pairs of 'true' phase I and phase II samples. Based on each 'true' phase I sample, one set of 500 bootstrap estimated parameters are obtained using the nonparametric procedure or the parametric procedure. This set of bootstrap estimated parameters are matched with 500 bootstrap phase II samples generated using by procedure 1 or the selection rule from the paired 'true' phase II sample to construct 500 bootstrap population size estimates. A bootstrap confidence interval is then constructed using the *percentile method.* This is repeated for 1000 pairs of 'true' phase I and phase II samples to give us 1000 bootstrap confidence intervals. Coverage rates for the true population size of these 1000 bootstrap confidence intervals are tabulated in Table 6.4.

In addition, for each bootstrap iteration in phase II, we constructed $log(\hat{\tau}_b - T_b)$, where $\hat{\tau}_b$ is the bootstrap population size estimate and T_b is the number of animals seen in the bootstrap phase II sample. We then constructed a bootstrap interval using percentiles (5^{th}, 90^{th} percentiles for a 90% confidence interval and 2.5^{th}, 97.5^{th} percentiles for a 95% confidence interval) from the sampling distribution of 500 $\log(\hat{\tau}_b - T_b)$ with T_0, the number of animals seen in the 'true' phase II sample. For example, a 90% bootstrap confidence interval would be $(\exp(\log(\hat{\tau}_b - T_b)_{0.05}) + T_0$, $(\exp(\log(\hat{\tau}_b - T_b)_{0.95}) + T_0)$. The reason for considering this method is we have observed from our simulation results in chapter 5 that interval estimates constructed by assuming $\log(\hat{\tau} - T)$ is normally distributed had reasonable coverage rates for true population size and shorter estimated interval widths. Coverage rates of these bootstrap confidence intervals are also reported in Table 6.4.

Table 6.4 True population size coverage rates of 1000 bootstrap confidence intervals constructed based on 1000 pairs of (1,1) 'true' phase I and phase II samples for the composite data of $M = 107$ groups and size $\tau = 834$ from Pennsylvania with (5.2.2) as the true model, using 500 bootstrap iterations.[2]

Nonparametric procedure of Phase I with

	Confidence Level		Using $\log(\hat{\tau} - T)$	
Phase II method	90	95	90	95
Procedure 1	87.0⋆(11.7)	93.0⋆(6.7)	84.7⋆(9.5)	91.7⋆(5.5)
Average interval width	597	811	549	754
Selection Rule	90.7(6.4)	95.8(3.3)	85.1⋆(8.9)	90.8⋆(5.6)
Average interval width	606	815	547	748

Parametric procedure of Phase I with

	Confidence Level		Using $\log(\hat{\tau} - T)$	
Phase II method	90	95	90	95
Procedure 1	81.7⋆(17.9)	87.7⋆(12.0)	82.5⋆(15.5)	88.9⋆(10.1)
Average interval width	475	634	424	575
Selection Rule	89.6(9.7)	93.8(6.1)	84.6⋆(13.2)	90.2⋆(9.0)
Average interval width	494	661	433	590

We see that for both nonparametric and parametric procedures in phase I when combined with the selection rule in phase II have observed coverage levels that are not significantly different from nominal levels. All other bootstrap confidence

[2]Coverage rates with $\star$ are different from actual levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are percentages of intervals with an upper bound lower than the true population size.

intervals have coverage rates $5 - 10\%$ lower ($p < 0.05$) than nominal levels. We also see that intervals constructed using the nonparametric procedure in phase I are longer on average and coverage rates have improved, though not very much improvement for those using $\log(\hat{\tau} - T)$. We noticed that on average, lower and upper bounds of intervals constructed using the nonparametric procedure were about 20 and 100 more than lower and upper bounds of intervals constructed using the parametric procedure. So, for using nonparametric procedure, the percentage of confidence intervals not covering the true population size because the upper bound is too low (compare parenthesized numbers) has decreased, quite significantly in some cases. For using the nonparametric procedure, since the average lower bound is also higher, the percentage of confidence intervals not covering the true population size because the lower bound is too high has also increased a little.

We have observed that at least 90% of the time, the number of groups and animals seen in a bootstrap phase II sample generated using procedure I in phase II is less than those in a phase II sample generated using the selection rule. We also noticed that for using the parametric procedure in phase I, among those intervals that did not cover the true population size, majority of them have an upper bound lower than the true population size (see parenthesized numbers next to coverage rates).

We have also constructed bootstrap t confidence intervals as described in subsection 6.4.3. For this method, we need to estimate variance of the population size estimator for each bootstrap phase II sample. Using SAS coding on UNIX System V Release 4.0 SunOS 5.5 machines, the CPU time required to obtain one bootstrap$-t$ confidence interval is about 15 minutes. This is at least 8 times the CPU time required to obtain one bootstrap interval using the *percentile method.* The longer CPU time is due to the computation of estimated variance of the population size estimator of a bootstrap phase II sample. For this method, since it is time consuming, we generated only 300 pairs of (1,1) phase I and phase II samples. Table 6.5 contains the preliminary simulation result of using the parametric procedure in phase I with 500 bootstrap iterations.

Table 6.5 True population size coverage rates of 300 bootstrap t confidence intervals constructed based on 300 pairs of (1,1) 'true' phase I and phase II samples for the composite data of $M = 107$ groups and size $\tau = 834$ from Pennsylvania with (5.2.2) as the true model, using 500 bootstrap iterations.[3]

Parametric procedure of Phase I with

	Confidence Level with $\sqrt{\widehat{\text{var}}(\hat{\tau}_0)}$		with $\hat{\sigma}_{LR}$	
Phase II method	90	95	90	95
Procedure 1	68.0*(32.0)	71.3*(28.7)	73.3*(26.7)	76.0*(24.0)
Average interval width	538	669	955	1191
Selection Rule	75.7*(24.3)	77.7*(22.3)	82.7*(17.3)	87.3*(12.7)
Average interval width	584	726	1004	1253

The notation $\hat{\sigma}_{LR}$ denotes the standard deviation of the 500 population size estimates constructed based on 500 bootstrap phase I and II samples. Notice how poor these coverage rates are although average interval widths are longer than those in Table 6.4. These bootstrap t confidence intervals are shifted to the left. lower (much lower than the number of animals seen in the 'true' phase II sample and are negative when we use the bootstrap $t-$statistic with $\hat{\sigma}_{LR}$) than those of confidence intervals constructed using the *percentile method* and by assuming $\log(\hat{\tau} - T)$ is normally distributed whose coverage rates are in Table 6.4. Average upper bounds of these bootstrap t confidence intervals are also about $300 - 400$ lower than those of confidence intervals whose coverage rates are in Table 6.4. These bootstrap t confidence intervals are so shifted to the left that all intervals not covering the true population size have upper bounds less than the true population size (see parenthesized numbers in Table 6.5). We think that this shift is

[3]Coverage rates with $\star$ are different from actual levels at 0.05 level of significance as they are not within $90\% \pm 3.5\%$ and $95\% \pm 2.5\%$, respectively. Numbers in parentheses are percentages of intervals with an upper bound lower than the true population size.

due to the inconsistency of the estimated variance of the population size estimator. We observed in chapter 5 that one needs a phase I and phase II sample of at least size 464 each to get a confidence interval of reasonable coverage rate. For the simulation result in Table 6.5, we used a phase I and phase II sample of 107 each. We will probably have to increase the sample size to improve the bootstrap t intervals. That will require a longer CPU time to compute estimated variances of the population size estimator and thus, to obtain bootstrap t confidence intervals. Using a lower level computer language such as Fortran or C+ could speed up simulations of bootstrap t confidence intervals. We also think that using the nonparametric instead of the parametric procedure in phase I, bootstrap t confidence intervals will probably give similar coverage rates. This conclusion is based on the observation from Table 6.4 that coverage rates of confidence intervals constructed using both parametric and nonparametric procedures in phase I are similar.

6.5 Discussion

Results from the simulation study of our proposed bootstrap procedures in phase I and phase II were encouraging. We summarized our simulation results as follows.

- Obtained using both nonparametric and parametric procedures in phase I, biases of bootstrapped estimated parameters of fitted logistic models are not negligible according to Cochran's (1977) working rule. This simulation result is obtained based on 1 and 2 complete passes through a simulated population of 54 groups and size 422 elk with a true one predictor, which is the group size, sightability model. However, this result does not explicitly imply that these procedures will be unsuccessful in estimating sighting probabilities. In fact, when we combined these phase I bootstrap procedures with those phase II bootstrap procedures, based on another simulated population, we obtained bootstrap confidence intervals with reasonable coverage rates of the true population size.

- Based on 1000 complete passes for population size estimation only (or $(\infty, 1)$) of the same simulated population of 54 groups of elk (sighting probabilities are known), coverage rates of confidence intervals constructed using the selection rule in phase II and the bootstrap $t-$statistic with $\sqrt{\widehat{\text{var}}(\hat{\tau}_0)}$ is not different from nominal levels at 0.05 level of significance. Note that $\widehat{\text{var}}(\hat{\tau}_0)$ is the estimated variance of population size estimator obtained based on the 'true' phase II sample. Coverage rates of confidence intervals constructed using the *percentile method* were too high for the selection rule. Coverage rate of the 95% confidence interval constructed using procedure I and the *percentile method* was low. We saw later when we combined bootstrap procedures in phase I and phase II, based on another simulated population, coverage rates of confidence intervals constructed using the *percentile method* were reasonable for the selection rule and much lower for procedure I.

- When sighting probabilities are not known and estimated via logistic regression, we used the simulated population (the composite raw data from Pennsylvania) of 107 groups and size 834 elk with (5.2.2) as the true sightability model. In addition to using the *percentile method*, we constructed a bootstrap confidence interval based on the assumption that $\log(\hat{\tau}_b - T_b)$ is normally distributed, where $\hat{\tau}_b$ and T_b are the population size estimate and the number of animals seen of a bootstrap phase II sample. We saw that when both nonparametric and parametric procedures in phase I are combined with the selection rule in phase II, observed coverage levels are not significantly different from nominal levels. All other bootstrap confidence intervals have coverage rates $5 - 10\%$ lower than nominal levels. We also see that intervals constructed using the nonparametric procedure in phase I were longer on average and coverage rates have improved but are still low when confidence intervals are not constructed using the selection rule with the *percentile method*. We observed that for using the parametric procedure in phase I, among those intervals that did not cover the true population size, majority of them have an upper bound lower than the true population size (see parenthesized numbers next to coverage rates).

- We observed that possibly larger phase I and II samples are needed to obtain a bootstrap t confidence interval of reasonable coverage. With smaller phase I and II samples, the bias correction method of Efron (1987) with an *acceleration constant* may provide improved bootstrap intervals constructed by the *percentile method* and by assuming that $\log(\hat{\tau} - T)$ is normally distributed using a smaller sample. This improved bootstrap bias-correction method is called the BC_a method. Efron (1987) has proved and demonstrated that the BC_a method gives second-order correct (asymptotically in Edgeworth expansion) intervals under reasonable conditions. He also gave examples where the BC_a endpoints match those of the corresponding exact intervals. However, he also showed that if this second-order accuracy is desired, it requires a lot more bootstrap replications, on the order of 1000.

For a preliminary result of Efron's bias correction method, we simulated 100 pairs of (1,1) phase I and II 'true' samples and for each pair of phase I and II samples, we ran $b = 3000$ bootstrap iterations with the parametric method in phase I and both bootstrap . We estimated the *acceleration constant* a by taking one sixth of the estimated skewness of the 3000 population size estimates constructed based on bootstrap phase I and II samples. We restricted the *acceleration con-*

stant and the bias correction constant z_0 to between -0.2 and 0.2 (Efron, 1987). The simulation results are in Table 6.6.

Table 6.6 True population size coverage rates of 100 bias-corrected BC_a confidence intervals constructed based on 100 pairs of (1,1) 'true' phase I and phase II samples for the composite data of $M = 107$ groups and size $\tau = 834$ from Pennsylvania with (5.2.2) as the true model, using 500 bootstrap iterations.[4]

Parametric procedure of Phase I with

	Confidence Level		Using $\log(\hat{\tau} - T)$	
Phase II method	90	95	90	95
Procedure 1	91.0(3.0)	96.0(2.0)	80.0*(8.0)	90.0*(5.0)
Average interval width	2563	3290	834	1316
Selection Rule	92.0(2.0)	96.0(2.0)	80.0*(8.0)	90.0*(5.0)
Average interval width	2604	3332	822	1280

Coverage rates of confidence intervals constructed using the *percentile method* using procedure I in phase II have improved and are now not significantly different from nominal levels. For confidence intervals constructed with the assumption that $\log(\hat{\tau} - T)$ is normally distributed, coverage rates did not improve. For all confidence intervals, average lower bounds are much larger than those without bias correction. At least half of those confidence intervals that did not cover the true population size have lower bounds larger than the true population size. Average upper bound and average interval width of confidence intervals constructed using the *percentile method* are much larger (up to a couple of thousands) than those without bias correction. It looks like upper bounds constructed using this bias correction are blown up too much. For those constructed with the assumption that $\log(\hat{\tau} - T)$ is normally distributed, average upper bound and average interval width are a few hundreds larger than those without bias correction. We noticed that average lower bound constructed with the normal distribution assumption

[4]Coverage rates with $\star$ are different from actual levels at 0.05 level of significance as they are not within $90\% \pm 6\%$ and $95\% \pm 4.4\%$, respectively. Numbers in parentheses are percentages of intervals with an upper bound lower than the true population size.

are higher (about 40 more) and average upper bound are much lower than (about half) those constructed without the normal distribution assumption. Therefore, in this case, lower bounds need to be adjusted lower and upper bounds need to be blown up a little bit more to have a better coverage rate of the true population size.

7. LOGISTIC MODEL SELECTION FOR UNKNOWN SIGHTING PROBABILITIES

7.1 Introduction

When sighting probabilities are unknown, we assume that a useful sightability model is a logistic model defined in (2.1.3). Although in reality the true model is much more complex than any other fitted model can be, we want to fit a logistic model that is not only useful and suitable but also parsimonious. This involves choosing appropriate independent variables and in some cases, transforming some or all of those independent variables. Some conventional measures for logistic model selection are the AIC (Akaike Information Criteria), Likelihood Ratio test and p-values for hypothesis testing specified parameters are zero.

Group size and percent vegetation cover were found to be influential independent variables in 3 elk sightability data sets (Colorado Department of Wildlife, Pennsylvania Game Commission and Idaho Department of Fish & Game). However, use of the natural log of group size might give a better fit (Samuel et. al., 1987). In view of this, for the simulation study in chapter 5, we have used sightability model (5.2.1) with group size and model (5.2.2) with natural log of group size for both the simulated population of $M = 116$ groups and size $\tau = 866$, and the composite data of $M = 107$ groups and size $\tau = 834$ which is labeled as population *Penn* (see section 5.2). We use these simulated populations again for our simulation study of logistic model selection.

During estimation of sightability model in our simulation study, we consider the full model (with all 3 independent variables) and 2 one-predictor models with group size (or natural log of group size) and percent vegetation cover as the only independent variable. This is because it was found that group size and percent vegetation cover are two influential independent variables. We will refer to both one-predictor models as the short models. In addition to comparing AIC, we want to diagnose the adequacy of all three models by evaluating the performances of asymptotically unbiased estimators of the population size and the variance of the population size estimator in terms of bias relative to the expected value and variance of the population size estimator. We will also compute the bias of estimated variance of the population size estimator relative to the observed or sample variance of simulated population size estimates. Confidence intervals for population size will also be constructed as described in section 5.1.2.

7.1.1 Expectation and covariance matrix of estimated parameters when the fitted model is simpler than the full model

When sighting probabilities are unknown and are estimated using logistic regression, error components of variance of the population size estimator presented in chapters 2 and 3 were derived under the assumption that the fitted model is the full model and $E(\widehat{\Theta}) = \Theta$ (asymptotically). Under this assumption, the population size estimator and the estimator of each error component of variance of the population size estimator stated in chapters 2 and 3 are (asymptotically) unbiased. Fitting a simpler model than the full model, such as the short model, violates this assumption because we may not have $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$, where $\boldsymbol{\beta}$ are true the parameters. When $\widehat{\boldsymbol{\beta}}$ is biased, our population size estimator will no longer be unbiased. Also, our expressions for the error components of variance of the population size estimator may be incorrect and hence, the estimator for each error component may no longer be unbiased.

When, say, a short model is fitted to a phase I sample, to derive theoretically the bias and the mean-squared-error in estimating the population size relative to fitting a full model, we need to know $E(\widehat{\boldsymbol{\beta}})$ and $\mathrm{cov}(\widehat{\boldsymbol{\beta}})$. If we have expressions for expectation and covariance matrix of the estimated parameters, we will be able to derive approximate expectation and variance of the population size estimator using the delta method. Derivation of $E(\widehat{\boldsymbol{\beta}})$ and $\mathrm{cov}(\widehat{\boldsymbol{\beta}})$ will not be as direct as that for linear models because for logistic models, estimated parameters obtained by maximizing the likelihood function are *iterative reweighted least square estimates.* So, $E(\widehat{\boldsymbol{\beta}})$ will also be obtained iteratively.

From (4.32) in Agresti (1990), the iterative expression for obtaining $\boldsymbol{\beta}^{(t+1)}$ of the $(t+1)^{th}$ iteration is

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} + (X'V^{-1(t)}X)^{-1}X'(y - m^{(t)}), \tag{7.1.1}$$

where X is the matrix of covariates, $V^{-1(t)}$ is the diagonal matrix with $\{g_j^{(t)}(1 - g_j^{(t)})\}$ on the diagonal, y is the vector of indicator functions Z_j of the j^{th} group in a phase I sample and $m^{(t)}$ is the vector of $\{g_j^{(t)}\}$. Recall that g_j is the sighting probability of the j^{th} group in a phase I sample and Z_j is 1 if the j^{th} group is sighted and 0 otherwise. So $g_j^{(t)}$ is the estimated sighting probability of the j^{th} group obtained based on $\boldsymbol{\beta}^{(t)}$ of the t^{th} iteration. In our simulation study, the matrix of covariates for fitting the full model is $X = (1, x_{j,1}, x_{j,2}, x_{j,3})_{j=1}^{M}$ where $x_{j,1}$=group size or natural log of group size, $x_{j,2}$=behavioral code and $x_{j,3}$=percent vegetation cover (see (5.2.1), (5.2.2)) of the j^{th} group while the matrix of covariates for the short models are $X = (1, x_{j,1})$ and $X = (1, x_{j,3})$. The expectation of (7.1.1) is

$$E(\boldsymbol{\beta})^{(t+1)} = E(\boldsymbol{\beta})^{(t)} + (X'V^{-1(t)}X)^{-1}X'(E(y) - m^{(t)}), \tag{7.1.2}$$

where $E(y)$ is the vector of $E(Z_j) = g_j$. With the same initial values and convergence criteria procedure LOGISTIC in SAS used for obtaining $\widehat{\boldsymbol{\beta}}$ by convergence of (7.1.1), we obtain $E(\widehat{\boldsymbol{\beta}})$ by convergence of (7.1.2). The estimated covariance matrix $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ of estimated parameters of a fitted model is

$$\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}}) = (X'V^{-1}X)^{-1}X'V^{-1\star}X(X'V^{-1}X)^{-1} \qquad (7.1.3)$$

where V^{-1} is the diagonal matrix with $\{\hat{g}_j(1-\hat{g}_j)\}$ on the diagonal such that $\hat{g}_j$ is the estimated sighting probability obtained based on $E(\widehat{\boldsymbol{\beta}})$ and $V^{-1\star}$ is the diagonal matrix with $\{g_j(1-g_j)\}$ as diagonal elements. If we fit the full model, (7.1.2) will converge to the true parameters. Also, $V^{-1(t)}$ will converge to $V^{-1\star}$ and thus, (7.1.3) will reduce to $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}}) = (X'V^{-1\star}X)^{-1}$, the estimated covariance matrix of estimated parameters of a fitted full model. We have therefore derived expressions to obtain $E(\widehat{\boldsymbol{\beta}})$ iteratively and $\widehat{\text{cov}}(\widehat{\boldsymbol{beta}})$ for estimated parameters $\widehat{\boldsymbol{\beta}}$ of a fitted model simpler than the full model.

1.1.1.1 Simulation study

We evaluated, for the short models, the performance of $E(\widehat{\boldsymbol{\beta}})$ obtained from (7.1.2) and $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ obtained from (7.1.3) by comparing their values to the average of 1000 estimated parameters and the sample covariance matrix of these estimated parameters obtained based on $k = 1000$ simulated phase I samples. We considered the four simulated populations we used in chapter 5. They are :

- (1)The simulated population of $M = 116$ groups and size $\tau = 866$ with sightability model (5.2.1) as the true model,
- (2)The simulated population of $M = 116$ groups and size $\tau = 866$ with sightability model (5.2.2) as the true model,
- (3)The simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.1) as the true model,
- (4)The simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.2) as the true model.

As in chapter 5, we considered the case of a complete census of primary units (see section 3.2.2). Each phase I sample is obtained with 1 complete pass through the simulated population. For fitting the short models, the following tables show the bias (%) of $E(\widehat{\boldsymbol{\beta}})$ relative to $(\overline{\widehat{\beta}_0}, \overline{\widehat{\beta}_1})$, the average of 1000 $(\widehat{\beta}_0, \widehat{\beta}_1)$, for all 4 simulated populations. In other words, we compute $\left(\frac{E(\widehat{\boldsymbol{\beta}})-\overline{\widehat{\boldsymbol{\beta}}}}{\overline{\widehat{\boldsymbol{\beta}}}}\right)100$. Results in Table 7.0A are for fitting the group size (or log(group size)) short model and and results in Table 7.0B are for fitting the vegetation cover short model.

Table 7.0A Considering the short model of taking the group size or the natural log of group size as the only predictor, bias (%) of $E(\widehat{\boldsymbol{\beta}})$, obtained from computing (7.1.2) iteratively, relative to the average of 1000 estimate parameters obtained based on 1000 phase I samples

Simulated population	$E(\widehat{\boldsymbol{\beta}})$	$(\overline{\widehat{\beta}_0}, \overline{\widehat{\beta}_1})$ Average of 1000 $(\widehat{\beta}_0, \widehat{\beta}_1)$	Bias$(E(\widehat{\boldsymbol{\beta}}), \overline{\widehat{\boldsymbol{\beta}}})$
1	$(-0.3254, 0.0963)$	$(-0.3512, 0.1018)$	$(7.35, -5.40)$
2	$(-0.6256, 0.6532)$	$(-0.6321, 0.6644)$	$(1.03, -1.69)$
3	$(-0.5831, 0.1232)$	$(-0.6148, 0.1318)$	$(5.16, -6.53)$
4	$(-1.0593, 0.8709)$	$(-1.0791, 0.8852)$	$(1.83, -1.62)$

Table 7.0B Considering the short model of taking the percent vegetation cover as the independent variable, bias (%) of $E(\widehat{\boldsymbol{\beta}})$, obtained from computing (7.1.2) iteratively, relative to the average of 1000 estimate parameters obtained based on 1000 phase I samples

Simulated population	$E(\widehat{\boldsymbol{\beta}})$	$(\overline{\widehat{\beta}_0}, \overline{\widehat{\beta}_1})$ Average of 1000 $(\widehat{\beta}_0, \widehat{\beta}_1)$	Bias$(E(\widehat{\boldsymbol{\beta}}), \overline{\widehat{\boldsymbol{\beta}}})$
1	$(1.4849, -0.0297)$	$(1.5187, -0.0306)$	$(-2.23, 2.94)$
2	$(1.4988, -0.0296)$	$(1.5286, -0.0301)$	$(-1.95, 1.66)$
3	$(1.9999, -0.0428)$	$(2.0526, -0.0439)$	$(-2.57, 2.51)$
4	$(1.9994, -0.0427)$	$(2.0477, -0.0440)$	$(-2.36, 2.96)$

In Table 7.0A, standard errors of the average of 1000 estimated parameters are $0.0077, 0.0095, 0.0088, 0.0106$ for $\widehat{\beta}_0$ and $0.0010, 0.0056, 0.0013, 0.0063$ for $\widehat{\beta}_1$ for the 4 simulated populations respectively. In Table 7.0B, standard errors are 0.0110519, $0.0110807, 0.0131172, 0.0125918$ for $\widehat{\beta}_0$ and $0.00027701, 0.00021242, 0.00023082$, 0.000284954 for $\widehat{\beta}_1$ respectively. We can see from above tables that $E(\widehat{\boldsymbol{\beta}})$ was very close to the average of 1000 estimated parameters. The largest bias was 7.35% of $\widehat{\beta}_0$

in Table 7.0A. We will not list the bias of $\widehat{\text{cov}}(\boldsymbol{\beta})$ obtained from (7.1.3) with respect to the sampling covariance matrix of 1000 $(\widehat{\beta}_0, \widehat{\beta}_1)$. We observed that $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ was fairly close to the sampling covariance matrix in terms of bias. In Table 7.0A, bias is less than a two tenth of the (observed) standard deviation of the estimate, i.e. $|\overline{\widehat{\beta}_i} - E(\widehat{\beta}_i)|/\sigma(\widehat{\beta}_i) < 0.2$ for $i = 0, 1$. In Table 7.0B, $|\overline{\widehat{\beta}_i} - E(\widehat{\beta}_i)|/\sigma(\widehat{\beta}_i) < 0.15$ for $i = 0.1$. According to Cochran's (1977) working rule, the bias is at most modest.

We will use those values of $E(\widehat{\boldsymbol{\beta}})$ and $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ we obtained to compute approximate values for $E(\hat{\tau}_{LR})$, the expectation of population size estimator, and $\text{var}(\hat{\tau}_{LR})$, the variance of population size estimator, of fitting the short models. We could then compute approximate values of the bias and the mean-squared-error ($= \text{var}(\hat{\tau}_{LR}) + \text{bias}^2$) in estimating the population size of fitting a short model relative to a full model.

7.1.2 Expectation and variance of the modified Horvitz-Thompson estimator when the fitted model is simpler than the full model

Consider the special case of a complete census of primary units. When a short model is fitted based on a phase I sample, the expectation and variance of the modified Horvitz-Thompson population size estimator (see (3.2.5)) are

$$E(\hat{\tau}_{LR}) = \sum_{j=1}^{M} g_j y_j E(\widetilde{\Theta}_j), \tag{7.1.4}$$

and

$$\begin{aligned}\text{var}(\hat{\tau}_{LR}) &= \sum_{j=1}^{M} \text{var}(Z_j y_j \widetilde{\Theta}_j) + \sum_{j \neq j'}^{M} \text{cov}(Z_j y_j \widetilde{\Theta}_j, Z_{j'} y_{j'} \widetilde{\Theta}_{j'}) \\ &= \sum_{j=1}^{M} \left[g_j \text{var}(\widetilde{\Theta}_j) + g_j(1 - g_j)(E(\widetilde{\Theta}_j))^2 \right] + \sum_{j \neq j'}^{M} g_j g_{j'} \text{cov}(\widetilde{\Theta}_j, \widetilde{\Theta}_{j'}),\end{aligned} \tag{7.1.5}$$

where $\widetilde{\Theta}_j = 1/\widehat{g}_j = 1 + e^{-\mathbf{x}_j' \widehat{\boldsymbol{\beta}} - \mathbf{x}_j' \Sigma \mathbf{x}_j'/2}$ such that $\widehat{\boldsymbol{\beta}}$ are estimated parameters of a fitted short model and Σ is the covariance matrix of the estimated parameters. We will use the delta method to derive approximate expressions for $E(\widetilde{\Theta}_j)$, $\text{var}(\widetilde{\Theta}_j)$ and $\text{cov}(\widetilde{\Theta}_j, \widetilde{\Theta}_{j'})$ in terms of $E(\widehat{\boldsymbol{\beta}})$ and Σ.

To begin with, we assume that $\Sigma \approx \widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ and so, our task is reduced to deriving approximate expressions for $E(e^{-\mathbf{x}_j'\boldsymbol{\beta}})$, $\text{var}(e^{-\mathbf{x}_j'\boldsymbol{\beta}})$ and $\text{cov}(e^{-\mathbf{x}_j'\boldsymbol{\beta}}, e^{-\mathbf{x}_{j'}'\boldsymbol{\beta}})$. With Taylor's expansion to the first order, we obtain

$$E(e^{-\mathbf{x}_j'\boldsymbol{\beta}}) \approx e^{-\mathbf{x}_j'E(\widehat{\boldsymbol{\beta}})}(1 + \mathbf{x}_j'\Sigma\mathbf{x}_j/2),$$

$$\text{var}(e^{-\mathbf{x}_j'\boldsymbol{\beta}}) \approx e^{-2\mathbf{x}_j'E(\widehat{\boldsymbol{\beta}})}\mathbf{x}_j'\Sigma\mathbf{x}_j,$$

and

$$\text{cov}(e^{-\mathbf{x}_j'\boldsymbol{\beta}}, e^{-\mathbf{x}_{j'}'\boldsymbol{\beta}}) \approx e^{-(\mathbf{x}_j+\mathbf{x}_{j'})E(\widehat{\boldsymbol{\beta}})}\mathbf{x}_j'\Sigma\mathbf{x}_{j'}.$$

Thus, we have

$$E(\widetilde{\Theta}_j) \approx 1 + e^{-\mathbf{x}_j'E(\widehat{\boldsymbol{\beta}})-\mathbf{x}_j'\Sigma\mathbf{x}_j/2}(1 + \mathbf{x}_j'\Sigma\mathbf{x}_j/2),$$

$$\text{var}(\widetilde{\Theta}_j) \approx e^{-2\mathbf{x}_j'E(\widehat{\boldsymbol{\beta}})-\mathbf{x}_j'\Sigma\mathbf{x}_j}\mathbf{x}_j'\Sigma\mathbf{x}_j,$$

and

$$\text{cov}(\widetilde{\Theta}_j, \widetilde{\Theta}_{j'}) \approx e^{-(\mathbf{x}_j+\mathbf{x}_{j'})E(\widehat{\boldsymbol{\beta}})-\mathbf{x}_j'\Sigma\mathbf{x}_j/2-\mathbf{x}_{j'}'\Sigma\mathbf{x}_{j'}/2}\mathbf{x}_j'\Sigma\mathbf{x}_{j'}.$$

We used the value of $E(\widehat{\boldsymbol{\beta}})$ obtained from the convergence of (7.1.2) and let $\Sigma \approx \widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ obtained from (7.1.3) for above approximate expressions. We then used these approximate values for (7.1.4) and (7.1.5) to obtain approximate values of $E(\hat{\tau}_{LR})$ and $\text{var}(\hat{\tau}_{LR})$. These results can be extended easily to the case of simple random sampling of primary units in chapter 2 and the other 2 special cases in chapter 3. In expressions of $E(\widetilde{\Theta}_j)$, $\text{var}(\widetilde{\Theta}_j)$ and $\text{cov}(\widetilde{\Theta}_j, \widetilde{\Theta}_j')$, simply replace $\widetilde{\Theta}_j$ and x_j by $\widetilde{\Theta}_{ij}$ and x_{ij} where subscript i indices the primary units. Note that we did not find approximate expressions for individual error components of the variance of the population size estimator.

7.2 A simulation study

Consider the case of a complete census of primary units. For each simulated population, we generated $k = 1000$ $(r_1, r_2) = (1, 1)$ population estimation trials. In other words, for each population estimation trial, we make $r_1 = 1$ complete pass through the population to obtain a phase I sample and $r_2 = 1$ complete pass through the population to obtain a phase II sample. Based on the phase I sample, 3 estimated sightability models (fitting the full and both short models) are obtained. Then, with each fitted model and the phase II sample, we compute population size estimates (see (3.2.5)). Along with each population size estimate, we compute estimates of both error components (see (3.2.7) and (3.2.9)) and thus, estimated variance $\widehat{\text{var}}(\hat{\tau}_{LR})$ of the population size estimator. For each simulated population, we also computed approximate values of $E(\hat{\tau}_{LR})$ and $\text{var}(\hat{\tau}_{LR})$ for the short models. As we did not have explicit expressions for the two error components, we will not evaluate their performances. We evaluate the performance of $E(\hat{\tau}_{LR})$ by computing its biases (%) relative to the true population size τ and relative to the average of 1000 population size estimates. We evaluate performances of the estimated variance $\widehat{\text{var}}(\hat{\tau}_{LR})$ and the approximate variance $\text{var}(\hat{\tau}_{LR})$ by computing their bias relative to the sample variance of 1000 population size estimates.

7.3 Simulation result

Approximate values of $E(\hat{\tau}_{LR})$, $\text{var}(\hat{\tau}_{LR})$ and averages of 1000 simulated estimates of the population size, $\hat{\tau}_{LR}$, and 1000 estimated variances of the population size estimator, $\widehat{\text{var}}(\hat{\tau}_{LR})$, are tabulated in Tables 7.1A, 7.2A, 7.3A and 7.4A for 4 simulated populations respectively. As in chapter 5, notation $\text{bias}(x, y)$ denote the bias (%) of x relative to y, i.e. $\frac{x-y}{y}100\%$. Coverage rates of 90% and 95% nominal confidence interval estimates for population size are tabulated in

Tables (7.1B,7.1C), (7.2B,7.2C), (7.3B,7.3C) and (7.4B,7.4C) respectively. In all tables, we use labels $full$, $group$ and veg to refer to fitting the full and both short models. In Tables 7.1A, 7.2A, 7.3A and 7.4A, we listed the ratios $|\hat{\tau}_{LR} - \tau|/\sqrt{\text{observed var}(\hat{\tau}_{LR})}|$ and $|E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}|/\sqrt{\text{observed var}(\hat{\tau}_{LR})}|$. By Cochran's (1977) working rule, the first ratio is to evaluate the bias of the population size estimate relative to the true population size and the second ratio is to evaluate the bias of the approximate value of $E(\hat{\tau}_{LR})$ as an estimate of the expected population size estimate. We also listed the approximate estimated mean-squared-error ($\widehat{MSE} = \text{var}(\hat{\tau}_{LR}) + (\hat{\tau}_{LR} - \tau)^2$) of estimating the population size based on a fitted short model.

In Tables 7.1A, 7.2A, 7.3A and 7.4A, we see that fitting the short model $group$ with group size as the only predictor gave a population size estimate as good as fitting $full$, the full model. Largest bias was only -1.28% and the ratio $|\hat{\tau}_{LR} - \tau|/\sqrt{\text{observed var}}$ was 0.14 in Table 7.2A for estimating the population size $\tau = 866$ of the simulated population of $M = 116$ groups, with sightability model (5.2.2) as the true model. According to Cochran's (1977) working rule, the bias of $\hat{\tau}_{LR}$ relative to the true population size τ is at most modest. Standard errors (parenthesized numbers right next to the average population size estimate) of the mean of 1000 population size estimates were also smaller than (about half) their respective values of fitting the full model. The sampling or observed variance of 1000 simulated population size estimates were about a third of that fitting the full model in Tables 7.1A, 7.2A, about a sixth of that fitting the full model in Table 7.3A and about a fourth of that fitting the full model in Table 7.4A. These smaller sampling variances in Tables 7.1A, 7.3A and 7.4A were even smaller than those obtained from fitting a full model with a $(4, 2)$ or a $(4, 4)$ (see obs $\text{var}(\hat{\tau}_{LR})$) under $(4, 2)$ or $(4, 4)$ in Tables 5.3A, 5.10 and 5.13 respectively). The sampling variance 5997.14 under the short model $group$ in Table 7.2A was a little larger (about 15%) than the sampling variance 5081.78 in Table 5.7A obtained from fitting a full model with a $(4, 2)$. The approximate estimated mean-squared-error or $\widehat{MSE}$ of fitting the short model *group* were also much smaller than that of fitting the full model.

The other short model *veg* with percent vegetation cover as the only predictor did not perform well. Population size was severely overestimated. The least bias was 33.50% in Table 7.2A. Standard errors of the mean of 1000 population size estimates were larger than (about 1.5 times) their respective values of fitting the full model. All ratios $|\hat{\tau}_{LR} - \tau|/\sqrt{\text{observed var}}$ were above 1.0 indicating that the bias of estimating the true population size with $\hat{\tau}_{LR}$ is significant (Cochran, 1977). Cochran (1977) stated that when this ratio is more than 0.5, it is not strictly correct to compare the MSE. Instead, one should just compare the variances for evaluating the precision of the population size estimator. So, we will not compare the MSE of the *veg* model to that of the *full* model. Instead, we noticed that all $\text{var}(\hat{\tau}_{LR})$ and sampling variances for fitted short model *veg* were very much larger than those obtained for fitted full model.

We have explained in section 7.1.1 that expressions (3.2.6) and (3.2.8) might not be the expressions for true error components of fitting a short model and estimators stated in (3.2.7) and (3.2.9) may no longer be unbiased. When we measure the bias of the estimated variance $\widehat{\text{var}}(\hat{\tau}_{LR})$ relative to the sample or observed variance ovar, i.e. $\text{bias}(\widehat{\text{var}}, \text{observedvar})$, for both short models, we see that there are signs of overestimation. The overestimation of the sampling variance for fitting both short models indicates that our expression for $\widehat{\text{var}}(\hat{\tau}_{LR})$ as a sum of (3.2.7) and (3.2.9) could be biased and overestimated the true variance of population size estimator when fitting a short model.

For fitting the short model *veg*, $\widehat{\text{var}}(\hat{\tau}_{LR})$ severely overestimated the observed variance. Nominal levels of confidence interval estimates constructed based on the

Table 7.1A Simulation results of $k = 1000$ (1,1) population estimation trials of the simulated population of $M = 116$ groups and size $\tau = 866$ with sightability model (5.2.1) as the true model, fitting the full model and both short models to each phase I sample[1]

	Model		
Statistic	*full*	*group*	*veg*
$\hat{\tau}_{LR}$	874.4(4.2)	858.1(2.2)	1267.2(6.3)
bias($\hat{\tau}_{LR}, \tau$)	1.0	−0.9	46.3
$E(\hat{\tau}_{LR})$	866	862.6	1256.7
bias($E(\hat{\tau}_{LR}), \hat{\tau}_{LR}$)	−1.0	0.5	−0.8
bias($E(\hat{\tau}_{LR}), \tau$)	–	−0.4	45.1
$\widehat{\text{var}}(\hat{\tau}_{LR})$	15415.2(1288.7)	5609.2(80.2)	86143.1(2491.3)
observed var($\hat{\tau}_{LR}$)	17877.1	5015.4	40347.7
bias($\widehat{\text{var}}$, observed var)	−13.8	11.8	53.2
var($\hat{\tau}_{LR}$)	14644.7	4403.5	30568.9
bias(var($\hat{\tau}_{LR}$), observed var)	−18.1	−12.2	−24.2
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.06	0.11	1.99
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.06	0.06	0.05
$\widehat{MSE}$	14715.1	4466.1	191529.4

Table 7.1B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 866$ of the simulated population of $M = 116$ groups with sightability model (5.2.1) as the true model[2]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	86*(408)	85*(412)	85*(419)	84*(432)	83*(450)	84*(427)	83*(478)
group	89(246)	89(247)	89(249)	90(251)	90(255)	90(251)	91(265)
veg	76*(966)	43*(989)	33*(1039)	26*(1130)	21*(1292)	18*(1001)	13*(1102)

Table 7.1C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 866$ of the simulated population of $M = 116$ groups with sightability model (5.2.1) as the true model[2]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	89*(487)	90*(493)	91*(506)	90*(527)	90*(561)	90*(529)	90*(594)
group	93*(294)	94(295)	94(298)	95(303)	95(309)	95(305)	96(321)
veg	97*(1151)	68*(1191)	55*(1279)	44*(1449)	34*(1817)	32*(1235)	22*(1360)

[1] All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[2] Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

Table 7.2A Simulation results of $k = 1000$ (1,1) population estimation trials of the simulated population of $M = 116$ groups and size $\tau = 866$ with sightability model (5.2.2) as the true model, fitting the full model and both short models to each phase I sample[3]

Statistic	Model *full*	*group*	*veg*
$\hat{\tau}_{LR}$	866.3(4.2)	854.9(2.5)	1156.1(6.1)
$\text{bias}(\hat{\tau}_{LR}, \tau)$	0.1	−1.3	33.5
$E(\hat{\tau}_{LR})$	866	861.5	1162.5
$\text{bias}(E(\hat{\tau}_{LR}), \hat{\tau}_{LR})$	−0.1	0.8	0.6
$\text{bias}(E(\hat{\tau}_{LR}), \tau)$	–	−0.5	34.2
$\widehat{\text{var}}(\hat{\tau}_{LR})$	17946.1(1314.6)	7065.5(92.3)	55038.2(1876.0)
observed $\text{var}(\hat{\tau}_{LR})$	17952.9	5997.1	36780.9
$\text{bias}(\widehat{\text{var}}, \text{observed var})$	−0.1	17.8	49.6
$\text{var}(\hat{\tau}_{LR})$	16183.9	5606.3	31579.3
$\text{bias}(\text{var}(\hat{\tau}_{LR}), \text{observed var})$	−9.8	−6.5	−14.1
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.002	0.14	1.51
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.002	0.09	0.03
$\widehat{MSE}$	16184.0	5729.3	115748.9

Table 7.2B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 866$ of the simulated population of $M = 116$ groups with sightability model (5.2.2) as the true model [4]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	83*(441)	83*(446)	83*(456)	83*(471)	83*(495)	84*(462)	85*(524)
group	89(277)	90(277)	90(280)	91(284)	92(289)	92(283)	92(300)
veg	81*(772)	57*(786)	51*(816)	46*(869)	42*(954)	33*(798)	40*(869)

Table 7.2C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 866$ of the simulated population of $M = 116$ groups with sightability model (5.2.2) as the true model[4]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	89*(525)	89*(533)	90*(550)	90*(579)	90*(623)	91*(576)	91*(653)
group	94(330)	95(331)	95(335)	95(342)	95(352)	96(344)	96(365)
veg	96(920)	78*(944)	67*(997)	58*(1092)	53*(1271)	51*(981)	43*(1070)

[3]All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[4]Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

Table 7.3A Simulation results of $k = 1000$ (1,1) population estimation trials of the simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.1) as the true model, fitting the full model and both short models to each phase I sample[5]

	Model		
Statistic	*full*	*group*	*veg*
$\hat{\tau}_{LR}$	838.2(5.1)	825.9(2.1)	1169.8(6.9)
bias($\hat{\tau}_{LR}, \tau$)	0.5	−1.0	40.2
$E(\hat{\tau}_{LR})$	834	835.2	1167.9
bias($E(\hat{\tau}_{LR}), \hat{\tau}_{LR}$)	−0.5	1.1	−0.2
bias($E(\hat{\tau}_{LR}), \tau$)	–	0.2	40.1
$\widehat{\text{var}}(\hat{\tau}_{LR})$	24006.9(3246.5)	5001.2(67.6)	38716.1(2318.0)
observed var($\hat{\tau}_{LR}$)	26329.0	4169.9	48065.2
bias($\widehat{\text{var}}$, observed var)	−8.8	19.9	71.8
var($\hat{\tau}_{LR}$)	18684.5	3732.3	37821.3
bias(var($\hat{\tau}_{LR}$), observed var)	−29.0	−10.5	−21.3
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.03	0.12	1.53
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.03	0.14	0.01
$\widehat{MSE}$	18702.0	3797.2	150562.8

Table 7.3B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model [6]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	80*(510)	80*(518)	79*(534)	79*(562)	78*(605)	80*(548)	83*(670)
group	90(233)	91(233)	91(235)	91(238)	91(241)	92(238)	92(251)
veg	91(647)	59*(655)	50*(673)	42*(701)	38*(744)	26*(663)	34*(704)

Table 7.3C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model[6]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	83*(607)	84*(621)	84*(650)	84*(699)	85*(786)	89*(701)	90*(857)
group	93*(277)	95(278)	96(281)	96(285)	96(291)	96(289)	96(305)
veg	99*(771)	83*(785)	69*(815)	58*(866)	50*(949)	46*(807)	37*(857)

[5]All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[6]Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within 90% ± 2% and 95% ± 1.4%, respectively. Numbers in parentheses are estimated expected interval widths

Table 7.4A Simulation results of $k = 1000$ (1,1) population estimation trials of the simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.2) as the true model, fitting the full model and both short models to each phase I sample[7]

Statistic	Model *full*	*group*	*veg*
$\hat{\tau}_{LR}$	840.9(4.3)	830.6(2.1)	1128.5(6.8)
$\text{bias}(\hat{\tau}_{LR}, \tau)$	0.8	−0.4	35.3
$E(\hat{\tau}_{LR})$	834	835.6	1125.0
$\text{bias}(E(\hat{\tau}_{LR}), \hat{\tau}_{LR})$	−0.8	0.6	−0.3
$\text{bias}(E(\hat{\tau}_{LR}), \tau)$	–	3.5	30.3
$\widehat{\text{var}}(\hat{\tau}_{LR})$	19677.1(1719.2)	6044.8(75.9)	67114.7(3498.2)
observed $\text{var}(\hat{\tau}_{LR})$	18116.3	4347.4	46034.7
$\text{bias}(\widehat{\text{var}}, \text{observed var})$	8.6	39.1	45.8
$\text{var}(\hat{\tau}_{LR})$	16939.1	4085.8	36414.7
$\text{bias}(\text{var}(\hat{\tau}_{LR}), \text{observed var})$	−6.5	−6.0	−20.9
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.05	0.05	1.37
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.05	0.08	0.02
$\widehat{MSE}$	16986.5	4097.6	123133.2

Table 7.4B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model [8]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	82*(462)	83*(467)	84*(479)	84*(499)	83*(529)	85*(490)	86*(576)
group	93(256)	94(256)	95(258)	94(262)	94(266)	94(262)	94(279)
veg	89(852)	64*(873)	56*(916)	49*(994)	43*(1134)	39*(889)	33*(996)

Table 7.4C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model[8]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	87*(550)	89*(560)	90*(580)	90*(616)	90*(674)	92*(618)	92*(726)
group	96(305)	97*(306)	97*(310)	98*(316)	98*(323)	97*(319)	97*(339)
veg	99*(1016)	83*(1050)	72*(1126)	64*(1273)	57*(1586)	52*(1104)	43*(1236)

[7] All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[8] Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

sampling distribution of $\hat{\tau}_{LR}$ were too high. On the other hand, due to the severe overestimation of the population size, nominal levels of approximate confidence interval estimates constructed based on transformations of population size estimates were too low (see coverage under *veg* in Tables (7.1B,7.1C), (7.2B,7.2C), (7.3B,7.3C), (7.4B,7.4C)). Very often, lower bounds of confidence interval estimates constructed from natural log and reciprocal transformations of population size estimates, and assuming that $\ln(\hat{\tau} - T)$ is normally distributed were greater than the true population size. Since these confidence intervals were so much shifted to the left (to larger values), coverage rates for the true population size were poor even though these confidence intervals were wider than those of fitting the full and *group* models.

For fitting the short model *group*, nominal levels of confidence interval estimates were good compared to fitting the full model in Tables (7.1B,7.1C), (7.2B,7.2C), (7.3B,7.3C) when $\text{bias}(\widehat{\text{var}}, \text{observedvar})$ was no more than 20%. The nominal levels under *group* in Table (7.4B,7.4C) were a little too high when $\text{bias}(\widehat{\text{var}}, \text{ovar})$ was 39.05%. Under the fitted *group* model, average population size estimates were about the same and average estimated variances were $\leq 1/3$ of those obtained from fitting the full model. Confidence interval estimates were, therefore, shorter than those of fitting the full model and yet, coverage rates were so much closer to or even higher than nominal levels.

For both short models, the approximate value of $E(\hat{\tau}_{LR})$ computed from (7.1.4) was very close to the average of 1000 simulated population size estimates. The largest bias was 1.11% of *group* in Table 7.3A and the ratio $|E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}|/\sqrt{\text{observed var}}$ was about 0.14, less than a 0.2. According to Cochran's 1977) working rule, the bias is at most modest. This implies that the approximate value of $E(\hat{\tau}_{LR})$ could give us an idea of how close the population size estimate obtained based on a fitted short model is to the population size estimate obtained based on a fitted full model. For the short model *group*, the approximate value of $\text{var}(\hat{\tau}_{LR})$ was quite close to the observed variance. The largest bias was -12.2% in Table 7.1A. For the short model *veg*, the approximate value of $\text{var}(\hat{\tau}_{LR})$ was about 20% less than the observed variance. This large bias could mean that we need a higher order Taylor's series when applying the delta method to obtain $\text{var}(\widetilde{\Theta}_j)$ and $\text{cov}(\widetilde{\Theta}_j, \widetilde{\Theta}_{j'})$. The large bias could also indicate that the value of $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ computed from (7.1.3) using $E(\widehat{\boldsymbol{\beta}})$ was not close enough to the true covariance matrix. Further, note that our simulation study was based on $(1, 1)$ of the simulated population. Increasing the phase I and phase II sample size reduces the noise in simulation (as observed in chapter 5) and in turn, may reduce this bias. Still, the approximate value of $\text{var}(\hat{\tau}_{LR})$ could at least give us some idea of how much the variance of the population size estimator based on a fitted short model would be.

We also compared the AIC's of all 3 fitted models. We have randomly chosen 10 (out of $k = 1000$) phase I samples for each simulated population. Based on these 10 chosen phase I samples, the AIC's of all three estimated models are presented in Tables 7.5, 7.6, 7.7 and 7.8 for the 4 simulated populations respectively.

Table 7.5 AIC's of estimated full and short models based on 10 randomly chosen (out of $k = 1000$) phase I samples of the simulated population of $M = 116$ groups with sightability model (5.2.1) as the true model

Sample	*full*	*group*	*veg*
1	105.6	151.6	151.8
2	114.2	152.3	140.6
3	118.3	158.1	144.5
4	113.4	148.7	142.3
5	126.6	160.4	143.0
6	129.1	145.3	153.5
7	106.1	145.7	140.7
8	111.8	139.7	129.3
9	131.1	144.7	158.7
10	132.4	160.3	140.2

Table 7.6 AIC's of estimated full and short models based on 10 randomly chosen (out of $k = 1000$) phase I samples of the simulated population of $M = 116$ groups with sightability model (5.2.2) as the true model

Sample	*full*	*group*	*veg*
1	102.3	147.3	129.1
2	111.9	154.4	143.8
3	121.5	152.8	150.3
4	122.0	152.0	138.6
5	128.1	151.8	151.0
6	88.6	149.5	148.7
7	122.8	150.9	137.5
8	120.0	154.9	149.1
9	136.8	158.4	137.3
10	105.1	150.5	143.8

Table 7.7 AIC's of estimated full and short models based on 10 randomly chosen (out of $k = 1000$) phase I samples of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model

Sample	*full*	*group*	*veg*
1	104.3	135.7	121.1
2	90.5	125.4	120.3
3	82.0	122.9	117.4
4	95.0	127.6	122.1
5	106.0	133.3	124.1
6	124.2	145.2	133.6
7	91.0	140.0	110.6
8	106.9	137.5	113.1
9	109.7	141.3	128.2
10	80.3	144.8	113.2

Table 7.8 AIC's of estimated full and short models based on 10 randomly chosen (out of $k = 1000$) phase I samples of the simulated population of $M = 107$ groups with sightability model (5.2.2) as the true model

Sample	*full*	*group*	*veg*
1	91.7	134.9	121.1
2	87.7	136.6	101.5
3	91.2	131.9	123.7
4	85.7	113.2	126.7
5	110.9	133.6	129.5
6	123.4	140.1	129.0
7	111.3	136.1	124.0
8	103.6	132.9	131.5
9	101.4	134.2	121.4
10	94.4	136.5	111.2

As we can see in each table, for all 10 chosen phase I samples of the simulated population, the AIC of fitting the full model is much less than the AIC of fitting a short model. Also, for all except sample no. 9 in Table 7.6, p-values of hypothesis tests (Wald statistics) of estimated coefficients of group size and vegetation cover are significant at 0.05 level. Thus, both conventional measures suggested that we should use the full model. Between the two short models, all except samples no. 1, 6, 9 in Table 7.5 and sample no. 4 in Table 7.8 suggested that we should choose *veg* over *group*. Both AIC and the Wald-statistic had selected the full model (or the correct model) as the best (or the most parsimonious) model that explains the variation in sighting probabilities. On the contrary, our simulation results indicated that the short model (or the wrong model) *group* has done better than the full model with a much higher precision (a smaller observed variance of the population size estimator) in estimating the population size and with confidence interval estimates of coverage rates close to or even higher than nominal levels.

7.4 Discussion

To get a further insight of the model selection problem, we plotted, using frequencies, percent vegetation cover versus group size for the composite data of $M = 107$ groups in figure 7.1. The scatter plot for the simulated population of $M = 116$ groups is similar and so, it will not be displayed.

In Table 7.5, the data are first sorted by percent vegetation cover and then by group size within each category of percent vegetation cover. In Table 7.6, the data are first sorted by group size and then by percent vegetation cover within each category of group size. In both Tables 7.5 and 7.6, the column labeled PROBF refers to sighting probabilities determined by model (5.2.2). In Table 7.7, groups are listed in ascending order of group size. Note that values of behavorial code and percent vegetation cover were not listed in Table 7.7 and our variable of interest is group size. The first column, labeled OBS, numbers the groups. Group sizes are listed in the second column labeled GROUP, followed by PROBF, and PROBG which are sighting probability determined by the fitted short model *group*. In the fourth column, labeled PROBV, are sighting probabilities determined by the fitted short model *veg*. The last 2 columns, labeled RG and RF, are $\frac{GROUP}{PROBG}PROBF$ and $\frac{GROUP}{PROBV}PROBF$, respectively. Values of RG or RF should be approximately equal to the group size listed in column GROUP if PROBG or PROBV is approximately equal to PROBF. So, comparing values of RG and RF as estimates of group sizes to respective true group sizes in column GROUP would help us to evaluate the fitted short models.

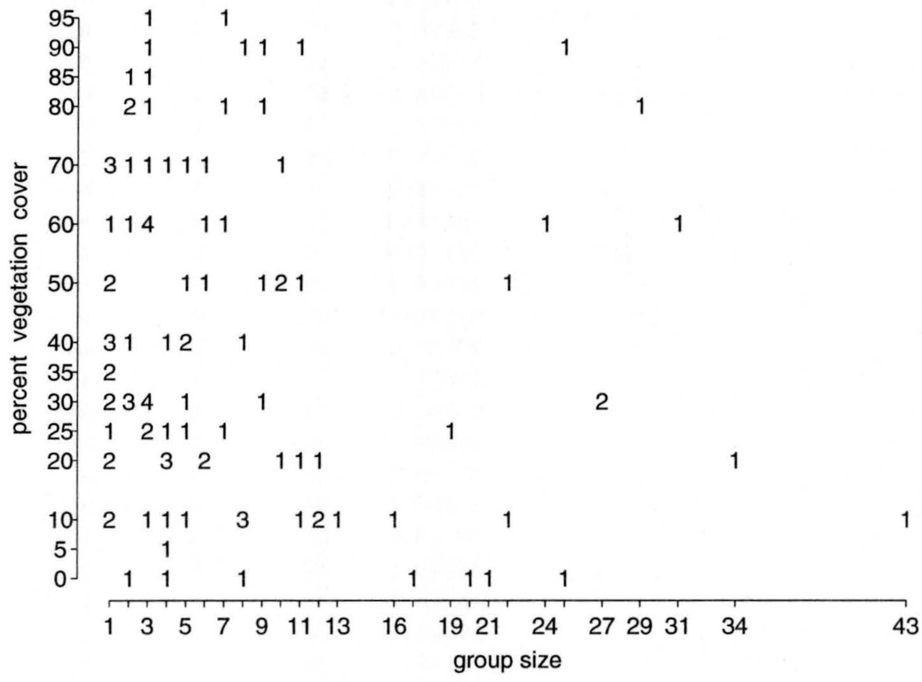

Fig. 7.1 Freq plot of group size versus veg cover for the composite raw data of 107 groups from Pennsylvania

Table 7.5 The composite elk data of 107 groups from Pennsylvania first sorted by ascending percent veg cover and then by ascending group size within each category of percent veg cover

OBS	GROUP	BEHAV	VEG	PROBF
1	2	1	0	0.68661
2	4	1	0	0.81096
3	8	1	0	0.89362
4	17	1	0	0.94578
5	20	1	0	0.95331
6	21	1	0	0.95537
7	25	1	0	0.96205
8	4	0	5	0.95310
9	1	1	10	0.42126
10	1	0	10	0.81043
11	3	1	10	0.67861
12	4	0	10	0.94250
13	5	1	10	0.77601
14	8	1	10	0.84530
15	8	0	10	0.96978
16	8	0	10	0.96978
17	11	1	10	0.88152
18	12	0	10	0.97940
19	12	1	10	0.89005
20	13	1	10	0.89742
21	16	1	10	0.91452
22	22	1	10	0.93577
23	43	1	10	0.96539
24	1	0	20	0.73552
25	1	1	20	0.32134
26	4	1	20	0.64480
27	4	1	20	0.64480
28	4	0	20	0.91425
29	6	0	20	0.94046
30	6	0	20	0.94046
31	10	0	20	0.96285
32	11	1	20	0.82877
33	12	0	20	0.96868
34	34	0	20	0.98836
35	1	0	25	0.69164
36	3	0	25	0.86678
37	3	1	25	0.52558
38	4	0	25	0.89583
39	5	0	25	0.91435
40	7	0	25	0.93668
41	19	1	25	0.86895
42	1	1	30	0.23548
43	1	1	30	0.23548
44	2	1	30	0.37621
45	2	1	30	0.37621
46	2	0	30	0.77984
47	3	1	30	0.47188
48	3	1	30	0.47188
49	3	1	30	0.47188
50	3	0	30	0.83994

51	5	0	30	0.89595
52	9	0	30	0.93836
53	27	1	30	0.88261
54	27	0	30	0.97786
55	1	1	35	0.19899
56	1	1	35	0.19899
57	1	1	40	0.16692
58	1	0	40	0.54061
59	1	1	40	0.16692
60	2	1	40	0.28178
61	4	1	40	0.43445
62	5	0	40	0.84851
63	5	1	40	0.48815
64	8	1	40	0.60066
65	1	1	50	0.11531
66	1	1	50	0.11531
67	5	0	50	0.78465
68	6	1	50	0.42539
69	9	1	50	0.52308
70	10	1	50	0.54848
71	10	0	50	0.87707
72	11	1	50	0.57124
73	22	1	50	0.72289
74	1	1	60	0.07816
75	2	0	60	0.49368
76	3	1	60	0.19740
77	3	1	60	0.19740
78	3	1	60	0.19740
79	3	1	60	0.19740
80	6	1	60	0.32505
81	7	1	60	0.35865
82	24	0	60	0.91557
83	31	0	60	0.93287
84	1	1	70	0.05227
85	1	1	70	0.05227
86	1	1	70	0.05227
87	2	1	70	0.09747
88	3	1	70	0.13793
89	4	1	70	0.17455
90	5	1	70	0.20794
91	6	1	70	0.23855
92	10	1	70	0.33951
93	2	1	80	0.06564
94	2	1	80	0.06564
95	3	1	80	0.09427
96	7	0	80	0.58156
97	9	0	80	0.63941
98	29	1	80	0.48417
99	2	0	85	0.24969
100	3	0	85	0.33022
101	3	1	90	0.06341
102	8	1	90	0.14909
103	9	1	90	0.16416
104	11	1	90	0.19262

105	25	1	90	0.34587
106	3	1	95	0.05178
107	7	1	95	0.11044

Table 7.6 The composite elk data of 107 groups from Pennsylvania first sorted by ascending group size and then by ascending percent veg cover within each category of group size

OBS	GROUP	BEHAV	VEG	PROBF
1	1	1	10	0.42126
2	1	0	10	0.81043
3	1	0	20	0.73552
4	1	1	20	0.32134
5	1	0	25	0.69164
6	1	1	30	0.23548
7	1	1	30	0.23548
8	1	1	35	0.19899
9	1	1	35	0.19899
10	1	1	40	0.16692
11	1	0	40	0.54061
12	1	1	40	0.16692
13	1	1	50	0.11531
14	1	1	50	0.11531
15	1	1	60	0.07816
16	1	1	70	0.05227
17	1	1	70	0.05227
18	1	1	70	0.05227
19	2	1	0	0.68661
20	2	1	30	0.37621
21	2	1	30	0.37621
22	2	0	30	0.77984
23	2	1	40	0.28178
24	2	0	60	0.49368
25	2	1	70	0.09747
26	2	1	80	0.06564
27	2	1	80	0.06564
28	2	0	85	0.24969
29	3	1	10	0.67861
30	3	0	25	0.86678
31	3	1	25	0.52558
32	3	1	30	0.47188
33	3	1	30	0.47188
34	3	1	30	0.47188
35	3	0	30	0.83994
36	3	1	60	0.19740
37	3	1	60	0.19740
38	3	1	60	0.19740
39	3	1	60	0.19740
40	3	1	70	0.13793
41	3	1	80	0.09427
42	3	0	85	0.33022
43	3	1	90	0.06341

44	3	1	95	0.05178
45	4	1	0	0.81096
46	4	0	5	0.95310
47	4	0	10	0.94250
48	4	1	20	0.64480
49	4	1	20	0.64480
50	4	0	20	0.91425
51	4	0	25	0.89583
52	4	1	40	0.43445
53	4	1	70	0.17455
54	5	1	10	0.77601
55	5	0	25	0.91435
56	5	0	30	0.89595
57	5	0	40	0.84851
58	5	1	40	0.48815
59	5	0	50	0.78465
60	5	1	70	0.20794
61	6	0	20	0.94046
62	6	0	20	0.94046
63	6	1	50	0.42539
64	6	1	60	0.32505
65	6	1	70	0.23855
66	7	0	25	0.93668
67	7	1	60	0.35865
68	7	0	80	0.58156
69	7	1	95	0.11044
70	8	1	0	0.89362
71	8	1	10	0.84530
72	8	0	10	0.96978
73	8	0	10	0.96978
74	8	1	40	0.60066
75	8	1	90	0.14909
76	9	0	30	0.93836
77	9	1	50	0.52308
78	9	0	80	0.63941
79	9	1	90	0.16416
80	10	0	20	0.96285
81	10	1	50	0.54848
82	10	0	50	0.87707
83	10	1	70	0.33951
84	11	1	10	0.88152
85	11	1	20	0.82877
86	11	1	50	0.57124
87	11	1	90	0.19262
88	12	0	10	0.97940
89	12	1	10	0.89005
90	12	0	20	0.96868
91	13	1	10	0.89742
92	16	1	10	0.91452
93	17	1	0	0.94578
94	19	1	25	0.86895
95	20	1	0	0.95331
96	21	1	0	0.95537
97	22	1	10	0.93577

98	22	1	50	0.72289
99	24	0	60	0.91557
100	25	1	0	0.96205
101	25	1	90	0.34587
102	27	1	30	0.88261
103	27	0	30	0.97786
104	29	1	80	0.48417
105	31	0	60	0.93287
106	34	0	20	0.98836
107	43	1	10	0.96539

Table 7.7 The composite elk data of 107 groups from Pennsylvania first sorted by ascending group size and then by ascending percent veg cover within each category of group size

OBS	GROUP	PROBF	PROBG	PROBV	RG	RV
1	1	0.73552	0.25723	0.75836	2.85934	0.96988
2	1	0.19899	0.25723	0.62286	0.77359	0.31948
3	1	0.32134	0.25723	0.75836	1.24923	0.42374
4	1	0.23548	0.25723	0.67166	0.91545	0.35060
5	1	0.23548	0.25723	0.67166	0.91545	0.35060
6	1	0.42126	0.25723	0.82803	1.63766	0.50875
7	1	0.16692	0.25723	0.57143	0.64891	0.29211
8	1	0.05227	0.25723	0.26967	0.20321	0.19384
9	1	0.81043	0.25723	0.82803	3.15056	0.97875
10	1	0.19899	0.25723	0.62286	0.77359	0.31948
11	1	0.07816	0.25723	0.36163	0.30385	0.21614
12	1	0.54061	0.25723	0.57143	2.10164	0.94606
13	1	0.05227	0.25723	0.26967	0.20321	0.19384
14	1	0.69164	0.25723	0.71701	2.68877	0.96461
15	1	0.11531	0.25723	0.46498	0.44828	0.24799
16	1	0.16692	0.25723	0.57143	0.64891	0.29211
17	1	0.11531	0.25723	0.46498	0.44828	0.24799
18	1	0.05227	0.25723	0.26967	0.20321	0.19384
19	2	0.37621	0.38776	0.67166	1.94043	1.12024
20	2	0.06564	0.38776	0.19399	0.33856	0.67675
21	2	0.06564	0.38776	0.19399	0.33856	0.67675
22	2	0.37621	0.38776	0.67166	1.94043	1.12024
23	2	0.68661	0.38776	0.88077	3.54144	1.55913
24	2	0.77984	0.38776	0.67166	4.02229	2.32213
25	2	0.09747	0.38776	0.26967	0.50273	0.72288
26	2	0.24969	0.38776	0.16270	1.28786	3.06940
27	2	0.28178	0.38776	0.57143	1.45336	0.98621
28	2	0.49368	0.38776	0.36163	2.54634	2.73034
29	3	0.86678	0.47412	0.71701	5.48459	3.62663
30	3	0.09427	0.47412	0.19399	0.59648	1.45784
31	3	0.05178	0.47412	0.11242	0.32763	1.38181
32	3	0.13793	0.47412	0.26967	0.87275	1.53441
33	3	0.19740	0.47412	0.36163	1.24907	1.63761
34	3	0.47188	0.47412	0.67166	2.98583	2.10767
35	3	0.19740	0.47412	0.36163	1.24907	1.63761
36	3	0.47188	0.47412	0.67166	2.98583	2.10767

37	3	0.47188	0.47412	0.67166	2.98583	2.10767
38	3	0.19740	0.47412	0.36163	1.24907	1.63761
39	3	0.19740	0.47412	0.36163	1.24907	1.63761
40	3	0.06341	0.47412	0.13560	0.40124	1.40287
41	3	0.83994	0.47412	0.67166	5.31476	3.75164
42	3	0.52558	0.47412	0.71701	3.32560	2.19902
43	3	0.33022	0.47412	0.16270	2.08946	6.08894
44	3	0.67861	0.47412	0.82803	4.29393	2.45866
45	4	0.64480	0.53667	0.75836	4.80600	3.40105
46	4	0.64480	0.53667	0.75836	4.80600	3.40105
47	4	0.43445	0.53667	0.57143	3.23813	3.04111
48	4	0.81096	0.53667	0.88077	6.04446	3.68299
49	4	0.91425	0.53667	0.75836	6.81431	4.82226
50	4	0.95310	0.53667	0.85640	7.10386	4.45166
51	4	0.89583	0.53667	0.71701	6.67698	4.99753
52	4	0.94250	0.53667	0.82803	7.02484	4.55298
53	4	0.17455	0.53667	0.26967	1.30099	2.58908
54	5	0.77601	0.58450	0.82803	6.63832	4.68593
55	5	0.20794	0.58450	0.26967	1.77876	3.85539
56	5	0.89595	0.58450	0.67166	7.66427	6.66966
57	5	0.84851	0.58450	0.57143	7.2585	7.4244
58	5	0.78465	0.58450	0.46498	6.7122	8.4374
59	5	0.91435	0.58450	0.71701	7.8217	6.3761
60	5	0.48815	0.58450	0.57143	4.1758	4.2713
61	6	0.42539	0.62247	0.46498	4.1004	5.4892
62	6	0.32505	0.62247	0.36163	3.1331	5.3931
63	6	0.94046	0.62247	0.75836	9.0651	7.4408
64	6	0.94046	0.62247	0.75836	9.0651	7.4408
65	6	0.23855	0.62247	0.26967	2.2993	5.3075
66	7	0.93668	0.65346	0.71701	10.0339	9.1445
67	7	0.58156	0.65346	0.19399	6.2297	20.9852
68	7	0.11044	0.65346	0.11242	1.1831	6.8771
69	7	0.35865	0.65346	0.36163	3.8419	6.9423
70	8	0.14909	0.67931	0.13560	1.7557	8.7954
71	8	0.84530	0.67931	0.82803	9.9549	8.1669
72	8	0.89362	0.67931	0.88077	10.5239	8.1167
73	8	0.96978	0.67931	0.82803	11.4208	9.3696
74	8	0.96978	0.67931	0.82803	11.4208	9.3696
75	8	0.60066	0.67931	0.57143	7.0738	8.4092
76	9	0.63941	0.70123	0.19399	8.2065	29.6650
77	9	0.52308	0.70123	0.46498	6.7135	10.1245
78	9	0.16416	0.70123	0.13560	2.1069	10.8952
79	9	0.93836	0.70123	0.67166	12.0434	12.5737
80	10	0.54848	0.72009	0.46498	7.6168	11.7957
81	10	0.87707	0.72009	0.46498	12.1799	18.8623
82	10	0.33951	0.72009	0.26967	4.7148	12.5899
83	10	0.96285	0.72009	0.75836	13.3712	12.6965
84	11	0.82877	0.73651	0.75836	12.3778	12.0213
85	11	0.57124	0.73651	0.46498	8.5316	13.5138
86	11	0.19262	0.73651	0.13560	2.8768	15.6251
87	11	0.88152	0.73651	0.82803	13.1657	11.7107
88	12	0.97940	0.75095	0.82803	15.6505	14.1938
89	12	0.89005	0.75095	0.82803	14.2227	12.8989
90	12	0.96868	0.75095	0.75836	15.4792	15.3281

91	13	0.89742	0.76376	0.82803	15.2750	14.0895
92	16	0.91452	0.79482	0.82803	18.4096	17.6714
93	17	0.94578	0.80330	0.88077	20.0152	18.2548
94	19	0.86895	0.81816	0.71701	20.1796	23.0262
95	20	0.95331	0.82471	0.88077	23.1188	21.6474
96	21	0.95537	0.83077	0.88077	24.1498	22.7789
97	22	0.93577	0.83639	0.82803	24.6141	24.8626
98	22	0.72289	0.83639	0.46498	19.0147	34.2027
99	24	0.91557	0.84649	0.36163	25.9584	60.7629
100	25	0.34587	0.85106	0.13560	10.1601	63.7660
101	25	0.96205	0.85106	0.88077	28.2604	27.3072
102	27	0.88261	0.85936	0.67166	27.7307	35.4800
103	27	0.97786	0.85936	0.67166	30.7232	39.3088
104	29	0.48417	0.86671	0.19399	16.2002	72.3802
105	31	0.93287	0.87328	0.36163	33.1156	79.9690
106	34	0.98836	0.88192	0.75836	38.1035	44.3116
107	43	0.96539	0.90161	0.82803	46.0421	50.1336

We first look at the scatter plot with group size on the x-axis and percent vegetation cover on the y-axis. We see that in each of the categories of 25, 30, 50, 60, 80 and 90 percent vegetation cover, group sizes are either small or large. For example, given 25 percent vegetation cover, group sizes are either ≤ 7 or ≥ 19. When percent vegetation cover is low, such as 25 and 30 percent, sighting probabilities (PROBF) of some small and large groups are not very different. For instance, in Table 7.5, two groups under 25 percent vegetation cover with sizes 7 and 19 have sighting probabilities 0.94 and 0.87. Also, in Table 7.5, three groups under 30 percent vegetation cover with sizes 9, 27 and 27 have sighting probabilities 0.94, 0.88 and 0.97. However, when percent vegetation cover is high, large groups have much larger sighting probabilities on average than small groups. For example, in Table 7.5, under 60 percent vegetation cover, 3 groups with sizes 7, 24 and 31 have sighting probabilities 0.36, 0.92 and 0.93. We know that if the fitted model has only 1 predictor, estimated sighting probabilities of all groups in one category are the same and is sort of an average of their true sighting probabilities. Then, sighting probabilities of large groups under high percent vegetation cover will be severely underestimated and, thus, will result in overestimating the number of animals in these large groups. In Table 7.7, in the column RV, we see that using the fitted short model *veg* expands the group with group size 24 under 60 percent vegetation cover (OBS no. 99) to 61 animals and the group with group size 31 under 60 percent vegetation cover (OBS no. 105) is expanded to 73. Since the population size estimate is dominated by large group sizes, the severe over-expansion from the large sized groups is probably the reason why the population size estimate obtained under the fitted short model *veg* was inflated (30%).

On the other hand, if we look at the scatter plot with group size on the y-axis and percent vegetation cover on the x-axis, we immediately see that for group sizes ≤ 13, except for size 11, group sizes are well distributed across the range of percent vegetation cover. In Table 7.6, with group size 11, the group with 50 percent and the group with 90 percent vegetation cover have sighting probabilities 0.57 and 0.19. If we cross reference these two groups in Table 7.7 (OBS no. 85 and 86), we see that the difference of about 0.4 in their true sighting probabilities results in a severe over-expansion of the sighting probability of the group with 90 percent vegetation cover when the fitted model is the short model *group* (estimated sighting probability is 0.73 under PROBG compared to the true sighting probability of 0.19). Thus, fitting the short model *group* underestimates the group size of this group. However, since group size 11 is not large, the bias (estimated group size is 3) contribution to the total population size estimate is small.

From Table 7.6, in categories of group sizes ≥ 13, except for group sizes 22, 25 and 27, there is only one group in each category. Both groups of size 27 are under the same percent vegetation cover and so their sighting probabilities do not differ much (0.88 and 0.98). The two groups of size 22 are under 10 and 50 percent vegetation cover. Since both groups do not have high percent vegetation cover and they have large group sizes, their sighting probabilities do not differ much (0.93 and 0.72). So the bias in estimating their group sizes are not large. From

Table 7.7, we see that (OBS 97 and 98) estimated group sizes are 25 and 20 for the groups of group size 22, and (OBS 102 and 103) estimated group sizes are 28 and 31 for the groups of group size 27.

In Table 7.6, we see that for group size 25, the two groups are under 0 and 90 percent vegetation cover. Since one group is under no vegetation cover whereas the other is under extremely high vegetation cover, their sighting probabilities differ very much (0.96 and 0.34) even though they were both large groups. In Table 7.7, we see that the sighting probability of the group under 90 percent vegetation cover (OBS no. 100) was severely overestimated and the group size is, thus, underestimated by a bias of 15. This is the largest bias so far for considering the fitted short model *group*. This bias is much smaller compared to those biases in the inflated estimated group sizes when we consider the fitted short model *veg*. The much smaller bias incurred from fitting the short model *group* is probably why model *group* does much better than model *veg*, and that the population size is so well estimated based on model *group*.

In reality, true models are much more complex than the one we considered for our simulation. The 'picture' will not be as clear as in the scatter plot and the tables we have just studied. In addition to using AIC to select the 'best' model, one could compute approximate values of $E(\hat{\tau}_{LR})$ and $\text{var}(\hat{\tau}_{LR})$ with the expressions we derived in (7.1.4) and (7.1.5) to evaluate fitted models simpler than the 'best' model selected by AIC.

When we compared coverage rates and estimated expected interval widths of confidence interval estimates of the fitted full and short models, one interesting observation was that confidence intervals constructed based on the short model *group* has shorter estimated expected interval width, about 50-60% of that of the full model, and yet, have higher coverage rates closer to nominal levels. We observed the same phenomenon when we increased the sample size from a (1,1) to a (4,1). Tables 7.8A, 7.8B, 7.8C, 7.9A, 7.9B and 7.9C show simulation results of 1000 (4,1) population estimation trials of the simulated population of $M = 107$ groups and size $\tau = 834$ with sightability models (5.2.1) and (5.2.2) as true models.

We saw, from Tables 7.8A and 7.9A, that both full and *group* model performed well with negligible bias (Cochran's rule : $|\hat{\tau} - \tau|/\sqrt{\text{observedvar}} = 0.02, 0.04 < 0.1$). The other short model *veg* severely overestimated ($|\hat{\tau} - \tau|/\sqrt{\text{observedvar}} = 2.18, 1.84 > 0.2$) the population size with 40.3% bias. Values of $E(\hat{tau}_{LR})$ obtained iteratively from (7.1.4) well approximated the simulated $\hat{\tau}$. In both Tables 7.8A and 7.9A, the largest $|E(\hat{\tau}_{LR}) - \hat{\tau}|/\sqrt{\text{observedvar}}$ is 0.07, which is < 0.1 and according to Cochran's working rule, this bias is negligible. Values of $var(\hat{\tau}_{LR})$ evaluated from (7.1.5) approximated the observed variance much better than in the (1,1) case. The largest bias in this case, a (4,1), was -6.7% of the fitted full model in Table 7.9A. The estimated MSE's of the fitted full and *group* models were a lot less different than in the (1,1) case. However, one can still see that between these 2 fitted models, the *group* model has a much less estimated MSE. Also, in Tables 7.8A and 7.9A, for the short models, our formula for an asymptotically unbiased estimated variance of the population size, $\widehat{var}(\hat{\tau}_{LR})$, overestimated the observed variance as much as 120% for the fitted *veg* model in Table 7.8A.

With a larger sample size, coverage rates of confidence intervals constructed based on the fitted short model *group* are now too high compared to nominal levels ($p < 0.05$). Still, we see that confidence intervals constructed based on the fitted short model *group* are shorter than those constructed based on the fitted full model and yet, have higher coverage rates. To get a further insight of this phenomenon, for the (1,1) simulation, we study a set of 1000 population size estimates, estimates of the variance of population size estimator and confidence interval estimates obtained based on models *full* and *group*. We found that confidence interval estimates based on the true model did not cover the true

Table 7.8A Simulation results of $k = 1000$ (4,1) population estimation trials of the simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.1) as the true model, fitting the full model and both short models to each phase I sample[9]

	Model		
Statistic	*full*	*group*	*veg*
$\hat{\tau}_{LR}$	832.0(2.9)	832.8(1.6)	1170.5(4.9)
$\text{bias}(\hat{\tau}_{LR}, \tau)$	−0.2	−0.1	40.3
$E(\hat{\tau}_{LR})$	834	836.1	1169.0
$\text{bias}(E(\hat{\tau}_{LR}), \hat{\tau}_{LR})$	0.2	0.4	−0.1
$\text{bias}(E(\hat{\tau}_{LR}), \tau)$	–	0.3	40.2
$\widehat{\text{var}}(\hat{\tau}_{LR})$	8249.7(228.1)	3460.1(22.7)	52296.9(891.8)
observed $\text{var}(\hat{\tau}_{LR})$	8384.1	2524.8	23760.5
$\text{bias}(\widehat{\text{var}}, \text{observed var})$	−1.6	0.4	120.1
$\text{var}(\hat{\tau}_{LR})$	8075.5	2494.8	23673.1
$\text{bias}(\text{var}(\hat{\tau}_{LR}), \text{observed var})$	−3.7	−1.2	−0.4
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.02	0.02	2.18
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.02	0.07	0.01
$\widehat{MSE}$	8079.5	2496.2	136905.4

Table 7.8B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model [10]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	85*(299)	85*(300)	86*(304)	86*(309)	86*(316)	86*(309)	87*(335)
group	94*(194)	94*(194)	94*(195)	94*(196)	94*(198)	94*(196)	93*(203)
veg	66*(752)	41*(766)	34*(793)	28*(839)	23*(914)	21*(775)	17*(839)

Table 7.8C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.1) as the true model[10]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	90*(356)	91*(359)	91*(365)	91*(373)	91*(385)	92*(379)	92*(412)
group	98*(231)	98*(231)	98*(233)	98*(235)	98*(238)	97*(238)	98*(246)
veg	93*(896)	65*(918)	52*(966)	43*(1050)	36*(1204)	33*(951)	25*(1030)

[9]All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[10]Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

Table 7.9A Simulation results of $k = 1000$ (4,1) population estimation trials of the simulated population of $M = 107$ groups and size $\tau = 834$ with sightability model (5.2.2) as the true model, fitting the full model and both short models to each phase I sample[11]

	Model		
Statistic	*full*	*group*	*veg*
$\hat{\tau}_{LR}$	832.2(3.0)	832.1(1.6)	1123.3(5.0)
bias($\hat{\tau}_{LR}, \tau$)	−0.2	−0.2	34.7
$E(\hat{\tau}_{LR})$	834	835.8	1126.0
bias($E(\hat{\tau}_{LR}), \hat{\tau}_{LR}$)	0.2	0.4	0.2
bias($E(\hat{\tau}_{LR}), \tau$)	–	0.2	35.0
$\widehat{\text{var}}(\hat{\tau}_{LR})$	8692.6(278.5)	4040.5(23.6)	41301.8(915.2)
observed var($\hat{\tau}_{LR}$)	8897.1	2715.0	24779.9
bias($\widehat{\text{var}}$, observed var)	−2.3	48.8	66.7
var($\hat{\tau}_{LR}$)	8277.1	2720.8	24507.5
bias(var($\hat{\tau}_{LR}$), observed var)	−6.7	0.2	−1.1
$\lvert\hat{\tau}_{LR} - \tau\rvert/\sqrt{\text{observed var}}$	0.02	0.04	1.84
$\lvert E(\hat{\tau}_{LR}) - \hat{\tau}_{LR}\rvert/\sqrt{\text{observed var}}$	0.02	0.07	0.02
$\widehat{MSE}$	8280.3	2724.4	108202

Table 7.9B True population size coverage rates of 1000 confidence interval estimates of 90% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.2) as the true model [12]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	84*(307)	85*(308)	85*(312)	85*(318)	85*(325)	87*(316)	88*(345)
group	94*(209)	95*(210)	96*(211)	96*(212)	96*(215)	95*(212)	95*(221)
veg	66*(669)	46*(678)	39*(699)	34*(734)	31*(788)	29*(496)	24*(535)

Table 7.9C True population size coverage rates of 1000 confidence interval estimates of 95% confidence level for population size $\tau = 834$ of the simulated population of $M = 107$ groups with sightability model (5.2.2) as the true model[12]

Model	$\hat{\tau}_{LR}$	$\ln\hat{\tau}_{LR}$	$\hat{\tau}^{-1/2}$	$\hat{\tau}^{-1}$	$\hat{\tau}^{-3/2}$	$\hat{\tau}_{log}$	$\hat{\tau}_{lg2}$
full	90*(365)	90*(368)	90*(375)	91*(384)	92*(398)	92*(389)	92*(423)
group	97*(249)	97*(250)	98*(252)	98*(255)	98*(259)	97*(257)	97*(268)
veg	91*(797)	68*(814)	57*(849)	49*(911)	41*(1018)	39*(600)	32*(646)

[11] All estimates shown here are averages of 1000 simulated estimates. Bias is in %. Numbers in parentheses are standard errors of the mean of 1000 simulated estimates.

[12] Coverage rates with $\star$ are different from nominal levels at 0.05 level of significance as they are not within $90\% \pm 2\%$ and $95\% \pm 1.4\%$, respectively. Numbers in parentheses are estimated expected interval widths

population size when both population size and variance of the population size are somewhat underestimated. When this occurred,the corresponding population size estimate based on the model *group* was larger and closer to the true value, and the estimate of the variance of population size estimator was also larger so that confidence intervals were a little longer and thus, covered the true population size. When both confidence intervals, one based on the full model and one based on the model *group*, covered the true population size, the confidence interval constructed based on the fitted full model was very much longer than that constructed based on model *group*. So confidence intervals constructed based on model *group* have better coverage rates and on average, these intervals have shorter estimated expected interval widths.

Although it is the correct model, one possible reason why the full model did not perform as well as the short model *group* is that relative to all 3 variables, we often find that the sample is not representative of the population. For instance,

small groups with $>$ 50 percent vegetation cover are almost always missed during the sighting trial. As a result, in this case, the fitted full model underexpands these groups relative to a fitted short model *group* which gives the same estimated sighting probability or approximately the average of true sighting probabilities to groups of the same group size. When the small group $>$ 50 percent vegetation cover does occur, the full model uses a smaller probability for expansion than the *group* model, contributing to the full model estimates larger variability in general than the short model.

8. SUMMARY, CONCLUSION AND TOPICS FOR FUTURE RESEARCH

In this final chapter, we give a summary and conclusion of our work, and suggest some topics for future research.

8.1 Summary and conclusion

We assume that data for population size estimation was collected in 2 phases. A phase I sample is the data set collected from a sightability survey conducted on the population for estimating the sightability model (or the sighting probabilities). A population survey is then conducted on sample primary units to obtain a phase II sample which is independent of the phase I sample.

When sighting probabilities are unknown and estimated via logistic regression, we have written the variance of the (asymptotically) unbiased population size estimator explicitly as a sum of three error components, namely, error due to sampling of primary units, error due to sampling of groups within sample primary units and error due to estimation of sighting probabilities via logistic regression. An unbiased estimator for each error component is derived.We also look at 3 special cases - (1) a complete census of primary units, (2) stratified random sampling of primary units, and (3) combining data of $r \geq 2$ replicated population surveys. For all three special cases we presented the unbiased population size estimator, error components of the variance of the population size estimator and an unbiased estimator for each error component, given known sighting probabilities or unknown sighting probabilities estimated via logistic regression. When the first stage sampling scheme of a population survey was stratified random sampling, optimum allocations of stratum sample sizes were given when sighting probabilities are known and when unknown sighting probabilities are estimated via logistic regression. Third central moments of the population size estimator when sighting probabilities are known and when unknown sighting probabilities are estimated via logistic regression were also derived.

We conducted a simulation study given a complete census of primary units, to evaluate small sample performance of the population size estimator and the unbiased estimators of error components. Data for each phase sample was obtained by making independent complete passes through the population. In a single complete pass through the population or a sighting trial, a group is determined to be sighted or not by comparing a randomly generated uniform random number, u_j (generated jointly independent for the groups), with the corresponding sighting probability g_j. If $u_j \leq g_j$, the j^{th} group for $j = 1, 2, ..., M$ is said to be sighted. We denote the number of complete passes through the population in phase I by r_1 and in phase II by r_2. Our simulation results show that the population size estimator (see (2.1.2), (2.1.4)) performed well (bias was negligible or at most modest according to Cochran's (1977) working rule). Furthermore, we observed that the sampling distribution of simulated population size estimates was positively skewed. When sighting probabilities are known, coverage rates of 90%, 95% confidence intervals for the true population size constructed using reciprocal transformations of population size estimates, $\hat{\tau}^{-1/2}$, $\hat{\tau}^{-1}$, $\hat{\tau}^{-3/2}$, where $\hat{\tau}$ is a population size estimate, are not different from nominal levels at 0.05 level of significance if the phase II sample size is at least 348. When sighting probabilities are unknown and estimated via logistic regression, we need a phase I and a phase II sample size of at least 464 each for coverage rates of all 90%, 95% confidence intervals constructed using reciprocal transformations to be not different from nominal levels at 0.05 level of significance. With these sample sizes, coverage rates of confidence intervals for the true population size constructed from the assumption that $(\hat{\tau} - T)$ is lognormal distribution, where T is the total count of animals in the phase II sample, were also reasonable. Furthermore, average interval width of approximate confidence intervals constructed for this method were shorter than average interval width of the other interval estimators.

We proposed two bootstrap procedures for phase I and two bootstrap procedures for phase II to construct a bootstrap confidence interval for the true population size. A simulation study was conducted to evaluate our procedures of each phase separately and for combining procedures of both phases. In phase I with a sample of 108 groups of elk, we observed that biases of bootstrapped estimated parameters of the sightability model are not negligible according to Cochran's (1977) working rule. This observation, however, does not explicitly imply that these procedures will be unsuccessful in estimating sighting probabilities. In phase II with a sample of 54 groups from a population survey and with 500 bootstrap iterations, we observed that coverage rates of 1000 confidence intervals constructed using percentiles of the bootstrap t-statistic (see (6.4.3)) and the square root of the estimated variance of the population size estimator, $\sqrt{\widehat{\text{var}}(\hat{\tau}_0)}$, obtained based on 'true' phase I and II samples are reasonable (see Table 6.3). Note that we refer to samples obtained from empirical populations constructed from our bootstrapping procedures as empirical samples or bootstrap samples. 'True' samples are samples simulated from the true population or our simulated populations.

For combining bootstrapping procedures of both phases, we conducted a simulation study based on the simulated population of $M = 107$ groups of elk and a population size of 834 from Pennsylvania state. We assumed that model (5.2.2) is the true sightability model. We generated 1000 pairs of 'true' phase I and II

$(r_1 = 1, r_2 = 1)$ samples. With 500 bootstrap iterations, we found that when both nonparametric and parametric procedures in phase I are combined with the selection rule in phase II, observed coverage levels are not significantly different from nominal levels. All other bootstrap confidence intervals have coverage rates $5 - 10\%$ lower than nominal levels. We also see that intervals constructed using the nonparametric procedure in phase I were longer on average and coverage rates have improved but are still low when confidence intervals are not constructed using the selection rule and percentiles of the sampling distribution of population size estimates obtained based on bootstrap samples. We observed that for using the parametric procedure in phase I, among those intervals that did not cover the true population size, majority of them have an upper bound lower than the true population size. Using the parametric procedure in phase I, we also constructed bootstrap t confidence intervals for 300 $(1, 1)$ 'true' phase I and II samples. Coverage rates of these confidence intervals are poor. We think that this is due to the inconsistency of estimated variances of the population size estimator. We will probably have to increase phase I and II sample sizes to obtain reasonable coverage rates. This conclusion is based on simulation results of confidence interval estimates constructed using the large sample theory and reciprocal transformations we obtained earlier.

Finally, we address the problem of logistic model selection. We have derived an iterative expression for obtaining $E(\widehat{\boldsymbol{\beta}})$ and an expression for obtaining $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ of estimated parameters $\widehat{\boldsymbol{\beta}}$ of a fitted model simpler than a given full model. We have also applied the delta method to derive approximate expressions of $E(\hat{\tau}_{LR})$ and $\text{var}(\hat{\tau}_{LR})$ when the fitted model is simpler than the full model. A simulation study was conducted, to evaluate the performance of our estimators when the fitted model exclude one of the predictors, or the short model, in the true sightability model. Thus we compare results based on a full versus a short model.

Our full model has group size and while our one-predictor short models use either group size (or the natural log of group size) or percent vegetation cover as the predictor. Simulation results in Tables 7.1A, 7.2A,, 7.3A and 7.4A of $(r_1 = 1, r_2 = 1)$ show that biases of population size estimates are negligible for the group size only model but the population size was severely overestimated for the vegetation cover only model. We also saw that approximate values of $E(\hat{\tau}_{LR})$ has negligible bias relative to the average of 1000 simulated population size estimates according to Cochran's (1977) working rule. Approximate values of $\text{var}(\hat{\tau}_{LR})$ underestimated (about 10% for the group size only model and 20% for the vegetation cover only model) the observed (or sampling) variance of those 1000 simulated population size estimates. Approximate values of both $E(\hat{\tau}_{LR})$ and $\text{var}(\hat{\tau}_{LR})$, especially the expectation, apparently give us an idea how different the expected population size and the variance of the population size estimator would be from those obtained based on a fitted full model.

We also found when less than the full model is fitted (Tables 7.1A, 7.2A, 7.3A and 7.4A) that the unbiased estimator of the variance of the population size estimator we derived (see section 2.2) under the assumption that the fitted model is the full model severely overestimated the observed variance of those simulated population size estimates. This indicates that when the fitted model is not the full model, the estimator of the variance may be biased and overestimate the true variance.

We have diagnosed the adequacy of all three fitted models using conventional measures such as the AIC and p values of hypothesis test of a parameter set to zero (or Wald statistics). The simulation results of 10 randomly selected phase I samples show that a fitted full model should be chosen over the short models and between the two short model, the AIC of the vegetation cover short model is lower most of the time. However, our simulation results indicate that the group size short model perform better than the full model with negligible bias and very much smaller mean-squared-error (reduced at least a third). Coverage rates of 90%, 95% confidence interval estimates under the group size model were close to nominal. Further, these confidence intervals are shorter than those constructed under the fitted full model. The vegetation cover short model on the other hand, severely overestimated (about 38% bias) the population size and the observed variance of population size estimates was much larger (at least twice) than those of the full model. The population size was so severely overestimated that even with the severely overestimated variance of the population size estimator, coverage rates of confidence interval estimates constructed with reciprocal transformations and from the assumption that $(\hat{\tau} - T)$ is lognormally distributed were too low. An explanation as to why the group size short model did well and why the vegetation cover short model did badly was given in section 7.4.

At this point, we are unable to make any substantial suggestion for logistic model selection. Currently, it is still unclear what exactly one should do to find a useful and simple model because in reality, true models are much more complex than the ones we considered. In addition to using conventional measures for selecting a model, the approximate expressions of $E(\hat{\tau}_{LR})$ and $\mathrm{var}(\hat{\tau}_{LR})$ for fitting short models rather than the full model apparently give us an idea, especially for the expected value of the population size estimator, how different the expected value of population size and the variance of the population size estimator obtained under various fitted models are.

8.2 Suggested topics for future research

We suggest using Efron's (1987) bias correction with an *acceleration constant* to improve the bootstrap intervals in Table 6.4. This improved bootstrap bias-correction method is called the BC_a method. Efron (1987) has proved and demonstrated that the BC_a method gives second-order correct (asymptotically in Edgeworth expansion) intervals under reasonable conditions. He also gave examples where the BC_a endpoints match those of the exact intervals. However, he also showed that if this second-order accuracy is desired, it requires a lot more bootstrap replications, on the order of 1000. Our preliminary simulation result of 100 pairs of phase I and II samples using the parametric procedure with 3000 bootstrap iterations in each phase showed that it is possible to improve the coverage rate by making bias correction. We have used estimated skewness of population size estimates constructed based on bootstrap samples to estimate the *acceleration constant* for the BC_a method. One other potential topic for future research is to see if the interval can be improved by incorporating the third central moment of the population we derived in chapter 4. We could replace population size and variance of population size estimator in expressions of the third central moment by their asymptotically unbiased estimates to obtain an (bias) estimate of the third central moment. An estimate of the third central moment can then be used to obtain the *acceleration constant* for the BC_a method. If the interval estimation of population size can be improved by incorporating the third central moment, it would be worthwhile to find an unbiased estimator for the third central moment.

We have conducted a simulation study of complete census of primary units using multiple phase II samples. Combining phase II samples is more difficult when the phase II sample is not a census. It would be useful to report an evaluation of performances of the point and interval population size estimator and our unbiased estimator of error components of the variance of the population size estimator using a composite data of such replicated surveys. In this case, the bootstrapping procedure of sample primary units proposed in section 6.3 for constructing bootstrap confidence intervals could also be evaluated.

When sighting probabilities are not known, Thompson and Seber (1994) have written the population size estimator as a function of the estimated sighting probability instead of the reciprocal of the estimated sighting probability. Using the delta method, they have presented an approximate third error component of the variance of the population size estimator. Their estimated sighting probability was not explicitly stated and like Samuel and Steinhorst (1989), they did not provide an unbiased estimator for each error component of their approximate variance, and for the approximate variance. We suggest, for future research on our problem, a derivation of an unbiased estimator for each error component of their approximate variance as a function of estimated sighting probabilities estimated using logistic regression. A simulation study using our simulated populations could then be conducted.

In addressing the problem of logistic model selection, our approximate variance of the population size estimator, $\text{var}(\hat{\tau}_{LR})$ (see subsection 7.1.2) based on a fitted short model has a significant bias of about -20%. It may be because we need to use more terms in the Taylor's expansion or it could be that the estimated covariance matrix of estimated parameters, $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$, was only approximating the observed covariance matrix. Also, increasing the phase I and phase II sample size reduces the noise in simulation (as observed in chapter 5) and in turn, may reduce this bias. Improving the value of $\widehat{\text{cov}}(\widehat{\boldsymbol{\beta}})$ and/or $\text{var}(\hat{\tau}_{LR})$ is a potential topic for future research as model selection is important in many applications of statistical modeling.

The following is another suggestion for future work. We have assumed that group size is known. In reality, when a group of elk flee from the helicopter, the observer might not have counted all of them. It is still unclear what impact the measurement error in group size as a covariate has on the population size estimation. As in linear regression, this is addressing the problem of having measurement error in covariates.

REFERENCES

1. Abromovitch, L. and Singh, K. (1985), Edgeworth Corrected Pivotal Statistics and the Bootstrap, *Ann. of Stat.*, **13**, 116-132.

2. Anderson, C. C. (1958), The elk of Jackson Hole, *Wyo. Game and Fish Comm. Bull*, **10**, 184pp.

3. Babu, G. J. and Singh, K. (1983), Inference on means using the bootstrap, *Ann. of Stat.*, **11**, 999-1003.

4. Babu, G. J. and Singh, K. (1985), Edgeworth Expansions for Sampling Without Replacement From Finite Populations, *J. Mult. Ana.*, **17**, 261-278.

5. Bailey, N. T. J. (1951), On estimating the size of mobile population sfrom recapture data, *Biometrika*, **38**, 293-306.

6. Beran, R. J. (1982), Estimated sampling distributions: the bootstrap and competitions, *Ann. of Stat.*, **10**, 212-225.

7. Beran, R. J. (1984), Jackknife approximation to bootstrap estimates, *Ann. of Stat.*, **12**, 101-118.

8. Beran, R. (1987), Prepivoting to Reduce Level Error of Confidence Sets, *Biometrika*, **74**, 457-468.

9. Bhattacharya, R. N. and Ghosh, J. K. (1978), On the Validity of Formal Edgeworth Expansions, *Ann. of Stat.*, **6**, 434-451.

10. Bhattacharya, R. N. and Qumsiyeh, M. (1989), Second-Order L^p comparisons Between the Bootstrap and Empirical Edgeworth Expansions Methodologies, *Ann. of Stat*, **6**, 434-451

11. Bhattacharya, R. N. and Rango Rao, R. (1976), *Normal Approximation and Asymptotic expansion*, John Wiley, NY.

12. Bickel, P. J. and Freedman, D. A. (1981), Some asymptotics on the bootstrap, *Ann. of Stat.*, **9**, 1196-1217.

13. Bickel, P. J. and Freedman, D. A. (1984), Asymptotic Normality and the Bootstrap in Stratified Random Sampling, *Ann. of Stat.*, **12**, 470-482.

14. Biggins, D. E. and Jackson, M. R. (1984), Biases in aerial surveys of mule deer, *Thorne Ecol. Inst. Tech. Publ.*, **14**, 60-65.

15. Booth, J. G., Butler, R. W. and Peter Hall (1994), Bootstrap Methods for Finite Populations, *J. Amer. Stat. Assoc.*, **89**, 1282-1289.

16. Bowden, D. C., Anderson, A. E. and Medin, D. E. (1984), Sampling plans for mule deer sex and age ratios, *J. Wildl. Manage.*, **48**, 500-509.

17. Buckland, S. T. (1980), A modified analysis of the Jolly-Seber capture-recapture model, *Biometrics*, **36**, 419-435.

18. Buckland, S. T. (1983), Monte Carlo methods for confidence interval construction using the bootstrap technique, *Bias*, **10**, 194-212.

19. Buechner, H. K., Buss, I. O. and Bryan, H. F. (1951), Censusing elk by airplane in the Blue Mountains fo Washington, *J. Wildl. Manage.*, **15**, 81-87.

20. Caughley, G. (1974), Bias in aerial survey, *J. Wildl. Manage*, **38**, 921-933.

21. Caughley, G. (1977a), Sampling in aerial survey, *J. Wildl. Manage*, **41**, 605-615.

22. Caughley, G. (1977b), *Analysis of vertebrate populations*, John Wiley, NY.

23. Caughley, G., Sinclair, R. and Scott-Kemmis, D. (1976), Experiments in aerial survey, *J. Wildl Manage*, **40**, 290-300.

24. Chen, A. and Sitter, R. R. (1991), Edgeworth Expansions and the Bootstrap for Stratified Random Sampling Without Replacement From a Finite Population, Technical Report 163, Carleton University, Laboratory for Research in Statistics and Probability.

25. Chao, M. T. and Lo, S. H. (1985), A Bootstrap Method for Finite Populations, *Sankhya*, **Ser. A, 47**, 399-405.

26. Cochran, W. G. (1977), *Sampling Technique* (3rd ed.), John Wiley, NY.

27. Connolly, G. E. (1981), Assessing populations, Mule and black tailed deer of North America, Wallmo, O. C. (ed.), **pp.** 287-345, Univ. Nebraska Press.

28. Cook, R. D. and Jacobson, J. O. (1979), A design for estimating visibility bias in aerial surveys, *Biometrics*, **35**, 735-742.

29. Cook, R. D. and Martin, F. B. (1974), A model for quadrat sampling with "visibility bias", *J. Amer. Stat. Assoc.*, **69**, 345-349.

30. Cox, D. R. (1980), Local ancillarity, *Biometrika*, **67**, 279-286.

31. Cox, D. R. and Hinkley, D. V. (1974), *Theoretical Statistics*, Chapman and Hall, London.

32. Czaplewski, R. L., Crowe, D. M. and McDonald, L. L. (1983), Sample sizes and confidence intervals for wildlife population ratios, *Wildl. Soc. Bull.*, **11**, 121-128.

33. Dalke, P. D., Beeman, F. J., Kindel, F. J., Robel, R. J. and Williams, T. R. (1965), Seasonal movements of elk in the Selway river drainage, Idaho, *J. Wildl. Manage.*, **29**, 333-338.

34. DiCiccio, T. J., Martin, M. A. and Young, G. A. (1992), Fast and accurate approximate double bootstrap confidence intervals, *Biometrika*, **79**, 285-295.

35. Dixon, W. J., et. al. (1981), *BMDP statistical software 1981*, Univ. California Press, Berkeley, 726pp.

36. Downing, R. L., Michael, E. D. and Poux, R. J. Jr. (1977), Accuracy of sex and age ratio counts of white-tailed deer, *J. Wildl. Manage.*, **41**,, 709-714.

37. Eberhardt, L. L. (1978), Appraising variability in population studies, *J. Wildl. Manage.*, **42**, 207-238.

38. Eberhardt, L. L. and Simmons, M. A. (1987), Calibrating population indices by double sampling, *J. Wildl. Manage.*, **50**, 665-675.

39. Efron, B. (1979), Bootstrap methods: another look at the jackknife, *Ann. of Stat.*, **7**, 1-26.

40. Efron, B. (1981a), Nonparametric standard errors and confidence intervals (with discussion), *Can. J. Stat.*, **9**, 139-172.

41. Efron, B. (1981b), Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods, *Biometrika*, **58**, 589-599.

42. Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia: Society for industrial and Applied Mathematics.

43. Efron, B. (1983), Estimating the error rate of a prediction rule: Improvement on cross-validation, *J. Amer. Stat. Assoc.*, **78**, 316-331.

44. Efron, B. (1985), Bootstrap confidence intervals for a class of parametric problems, *Biometrika*, **72**, 45-58.

45. Efron, B. (1987), Better Bootstrap Confidence Intervals (with discussion), *J. Amer. Stat. Assoc.*, **82**, 171-185

46. Faulkenberry, G. D. and Garoui, A. (1991), Estimating a population total using an area frame, *J. Amer. Stat. Assoc.*, **86**, 445-449.

47. Feller, W. (1966), *An Introduction to Probability Theory and Its Applications: Volume II* (2nd ed.), John Wiley, NY.

48. Floyd, T. D., Mech, L. D. and Nelson, M. E. (1979), An improved method of censusing deer in deciduous-coniferous forests, *J. Wildl. Manage.*, **43**, 258-261.

49. Francis, R. I. C. C. (1984), An adaptive strategy for stratified random trawl surveys, *NZ J. Marine and Freshwater Research*, **18**, 59-71.

50. Fuller, W. A. (1976), *Introduction to Statistical Time Series*, John Wiley, NY.

51. Gasaway, W. C., DuBois, S. D. and Harbo, S. J. (1985), Moose survey procedures development, Alaska Dept. of Fish and Game, project report, Federal Aid in Wildlife Restoration Project W-17-11.

52. Geary, R. C. (1947), Testing for normality, *Biometrika*,, **34**, 209-242.

53. Gilbert, J. R. (1978), Estimating population characteristics, Big game of North America, Schmidt, J. L. and Gilbert, D. L. (eds), **pp.** 297-304, Stackpole Books, Harisburg, PA.

54. Goodman, L. A. (1960), On the exact variance of products, *J. Amer. Stat. Assoc.*, **55**, 708-713.

55. Graham, A. and Bell, R. (1969), Factors influencing the countability of animals, *East Afr. Agric. For. J.*, **34**(special issue), 38-43.

56. Gross, S. (1980), Median Estimation in Sample Surveys, *Proceedings of the Section on Survey Research Methods*, Amer. Stat. Assoc., pp. 181-184.

57. Hall, P. (1983), Inverting an Edgeworth expansion, *Ann. of Stat.*, **11**, 569-576.

58. Hall, P. (1985), A tabular method for correcting skewness, *Math. Proc. Cambridge Phil. Soc.*, **97**, 525-540.

59. Hall, P. (1986), On the Boostrap and Confidence Intervals, *Ann. of Stat.*, **14**, 1431-1452.

60. Hall, P. (1986), On the Number of Bootstrap Simulations Required to Construct a Confidence Interval, *Ann. of Stat.*, **14**, 1453-1462.

61. Hall, P. (1987), On the bootstrap and likelihood-based confidence regions, *Biometrika*, **74**, 481-493.

62. Hall, P. (1988), Theoretical Comparison of Bootstrap Confidence Intervals, *Ann. of Stat.*, **16**, 927-985.

63. Hall, P. (1988), On bootstrap resampling and iteration, *Biometrika*, **75**, 661-671.

64. Hall, P. (1990), On the Relative Performance of Bootstrap and Edgeworth Approximation of a Distribution Function. *J. Mult. Ana.*, **35**, 108-129.

65. Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer, NY.

66. Hall, P. (1994), Monte Carlo approximation and the iterated bootstrap, *Biometrika*, **81**, 331-340.

67. Hall, P. and Martin, M. A. (1988), On Bootstrap Resampling and Iteration, *Biometrika*, **75**, 661-671.

68. Hansen, M. H., Hurwitz, W. N. and Madow, W. G. (1953), Sample survey methods and theory, *Methods and applications*, Volume I, John Wiley, NY., 638pp.

69. Hansen, M. H., Madow, W. G. and Tepping, B. J. (1983), An Evaluation of Model-Dependent and Probability-Sampling Inferences in Sample Survey, *J. Amer. Stat. Assoc.*, **78**, 776-807.

70. Hanson, W. R. (1967), Estimating the density of an animal population, *J. Res. Lepidoptera*, **6**, 203-247.

71. Harter, H. L. (1964), Criteria for best substitute interval estimator, with an application to the normal distribution, *J. Amer. Stat. Assoc.*, **59**, 1133-1140.

72. Hartigan, J. A. (1986), Discussion of "Bootstrap methods for standard errors, confidence intervals and other measures of statistical theory" by Efron and Tibshirani, *Stat Sci.*, **1**, 75-77.

73. Hinkley, D. V. and Shi, S. (1989), Importance resampling and the nested bootstrap, *Biometrika*, **76**, 435-446.

74. Hinkley, D. V. and Wei, B. C. (1984), Improvement of jackknife confidence limit methods, *Biometrika*, **71**, 331-339.

75. Horvitz, D. G. and Thompson, D. J. (1952), A generalization of sampling without replacement from a finite universe, *J. Amer. Stat. Assoc.*, **47**, 663-685.

76. Johnson, N. J. (1978), Modified $t-$tests and confidence intervals for asymmetrical populations, *J. Amer. Stat. Assoc.*, **73**, 536-544.

77. Jolly, G. M. (1969), Sampling methods for aerial censuses of wildlife populations, *East African Agriculture and Forestry Journal*, **34** (special issue), 46-49.

78. Kendall, M. G. and Stuart, A. (1979), *The Advanced Theory of Statistics*, **2**, Griffin, London.

79. Kovar, J. G., Rao, J. N. K. and Wu, C. F. J. (1988), Bootstrap and Other Methods to Measure Errors in Survey Estimates, *Canadian J. of Stat.*, **165**, 25-45.

80. Krewski, S. and Rao, J. N. K. (1981), Inference from Stratified Samples: Properties of Linearization, Jackknife and Balanced Repeated Replication Methods, *Ann. of Stat.*, **2**, 1010-1019.

81. Kufeld, R. C., Olterman, J. H. and Bowden, D. C. (1980), A helicopter quadrat census for mule deer on Uncompahgre Plateau, *Colorado J. Wildl. Manage.*, **44**, 632-639.

82. Lawless, J. F. (1981), Mean squared error properties of generalized ridge estimators, *J. Amer. Stat. Assoc.*, **76**, 462-466.

83. Lehmann, E. L. (1986), *Testing Statistical Hypotheses*, 2nd ed., John Wiley, NY.

84. LeResche, R. E. and Rausch, R. A. (1974), Accuracy and precision of aerial censusing, *J. Wildl. Manage.*, **38**, 175-182.

85. Little, R. J. A. (1986), Survey nonresponse adjustments for estimates of means, *International Statistical Review*, **54**, 139-157.

86. Little, R. J. A. (1988), Missing-data adjustments in large surveys, *J. Buss. Econ. Stat.*, **6**, 287-296.

87. Lo, A. Y. (1988), A Bayesian Bootstrap for a Finite Population, *Ann. of Stat*, **16**, 1684-1695.

88. Loh, W. Y. (1987), Calibrating confidence coefficients, *J. Amer. Stat. Assoc.*, **82**,

89. Lovaas, A. L., Egan, J. L. and Knight, R. R. (1966), Aerial counting of two Montana elk herds, *J. Wildl. Manage.*, **30**, 364-369.

90. Macarthy, P. J. and Snowden, C. B. (1985), The Bootstrap and Finite Population Sampling, *Vital and Health Statistics* (Ser. 2, No. 95), Public Health Service Publications, Washington, DC: U.S. government Printing Office, 85-1369.

91. Martin, M. A. (1990), On bootstrap iteration for coverage correction in confidence intervals, *J. Amer. Stat. Assoc.*, **85**, 1105-1118.

92. Mc Cullagh, P. (1984), Local sufficiency, *Biometrika*, **71**, 233-244.

93. Mehta, J. S. and Srinivasan, R. (1971), Estimation of the mean by shrinkage to a point, *J. Amer. Stat. Assoc.*, **66**, 86-90.

94. Oh, H. L. and Scheuren, F. J. (1983), Weighting adjustment for unit nonresponse, *Incomplete Data in Sample Surveys*, Madow, W. G., Olkin, I. and Rubin, D. B. (eds), 143-184, Academic Press, NY.

95. Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978), Statistical Inference from capture data on closed animal populations, *Wildl. Monogr.*, **62**, 135pp.

96. Parr, W. C. (1983), A note on the jackknife, the bootstrap and the delta method estimators of bias and variance, *Biometrika*, **70**, 719-722.

97. Petrov, V. V. (1975), *Sums of Independenet Random Variables*, Springer, Berlin.

98. Pfanzagl, J. (1979), Nonparametric minimum contrast estimators, *Selecta Statistica Canadiana*, **5**, 105-140. Dept. of Math., McMaster Unov., Hamilton, Ontario.

99. Pollock, K. H., Hines, J. E. and Nichols, J. D. (1984), The use of auxiliary variables in capture-recapture and removal experiments, *Biometrics*, **40**, 329-340.

100. Pollock, K. H. and Kendall, W. L. (1987), Visibility bias in aerial surveys: A review of estimation procedures, *J. Wildl. Manage.*, **51**, 502-510.

101. Pratt, J. W. (1961), Length of confidence intervals, *J. Amer. Stat. Assoc.*, **56**, 549-567.

102. Pratt, J. W. (1963),, Shorter confidence intervals for the mean of a normal distribution with known variance, *Ann. Math. Stat.*, **34**, 574-586.

103. Prentice, R. L. (1975), Discriminnation among some parametric models, *Biometrika*, **62**, 607-614.

104. Rao, C. R. (1973), *Linear Statistical Inference and Its Applications* (2nd ed.), John Wiley, NY.

105. Rao, J. N. K., Hartley, H. O. and Cochran, W. G. (1962), On a Simle Procedure of Unequal Probability Sampling Without Replacement, *J. Royal Stat. Soc.*, **Ser. B, 24**, 482-491.

106. Rao, J. N. K. (1975), Unbiased variance estimation for multistage designs, *Sankhya Ser. C*, **37**, 133-139.

107. Rao, J. N. K. and Wu, C. F. J. (1985), Inference From Stratified Samples: Second-Order Analysis of Three Methods for Nonlinear Statistics, *J. Amer. Stat. Assoc.*, **80**, 620-630.

108. Rao, J. N. K. and Wu, C. F. J. (1987), Methods for Standard Errors and Confidence Intervals From Sample Survey Data: Some Recent Work, *Bulletin of the International Statistical Institute*, Proceedings of the 46th Session, **Book 3**, pp. 5-19.

109. Rao, J. N. K. and Wu, C. F. J. (1988), Resampling Ingerence With Complex Survey Data, *J. Amer. Stat. Assoc.*, *83*, 231-241.

110. Rice, W. R. and Harder, J. D. (1977), Application of multiple aerial sampling to a mark-capture census of white-tailed deer, *J. Wildl. Manage.*, **41**, 197-206.

111. Robel, R. J. (1960), Determining elk movement through periodic aerial counts, *J. Wildl. Manage.*, **24**, 103-104.

112. Robinson, J. (1978), An Asymptotic Expansion for Samples From a Finite Population, *Ann. of Stat.*, **6**, 1005-1011.

113. Rolley, R. E. and Keith, L. B. (1980), Moose population dynamics and winter habitat use at Rochester, Alberta, 1965-1979, *Canadian Field-Naturalist*, **94**, 9-18.

114. Samuel, M. D. (1984), *An evluation of elk sightability in northcentral Idaho with application to aerial census and herd composition counts*, Ph.D. Thesis, Univ. Idaho, Moscow, 118pp.

115. Samuel, M. D. (1985), *Nonresponse Horvitz-Thompson methods for aerial surveys of wildlife populations*, M.S. Thesis, Univ. of Idaho, Moscow, 44pp.

116. Samuel, M. D., Garton, E. O., Schlegel, M. W. and Carson, R. G. (1987), Visibility bias during aerial surveys of elk in noncentral Idaho, *J. Wildl. Manage.*, **51**, 622-630.

117. Samuel, M. D. and Pollock, K. H. (1981), Correction of visibility bias in aerial surveys where animals occur in groups, *J. Wildl. Manage.*, **45**, 993-997.

118. Samuel, M. D., Steinhorst, R. K., Garton, E. O. and Unsworth, J. W. (1992), Estimation of wildlife population ratios incorporating survey design and visibility bias, *J. Wildl. Manage.*, **56**, 718-725.

119. Seber, G. A. F. (1982), *The Estimation of Animal Abundance* (2nd ed.), Griffin, London.

120. Siegmund, D. (1985), *Sequential Analysis; Tests and Confidence Intervals*, Springer, NY.

121. Singh, K. (1981), On the asymptotic accuracy of Efron's bootstrap, *Ann. of Stat.*, **9**, 1187-1195.

122. Sinha, B. K. (1991), Sequential Methods for Finite Populations, *Handbook of Sequential Analysis*, B.K. Ghosh and P. K. Sen (eds), 313-329, Marcel Dekker, NY.

123. Siniff, D. B. and Skoog, D. O. (1964), Aerial censusing of caribou using stratified random sampling, *J. Wildl. Manage.*, **28**, 391-401.

124. Sitter, R. R. (1992), A Resampling Procedure for Complex Survey Data, *J. Amer. Stat. Assoc.*, **87**, 755-765.

125. Steinhorst, R. K. and Samuel, M. D. (1989), Sightability adjustment methods for aerial surveys of wildlife populations, *Biometrics*, **45**, 415-425.

126. Sukhatme, P. V. and Sukhatme, B. W. (1970), *Sampling Theory of Surveys With Applications*, Ames IA: Iowa State University Press.

127. Thompson, J. R. (1968), Some shrinkage techniques for estimating the mean, *J. Amer. Stat. Assoc.*, **63**, 113-122.

128. Thompson, S. K. (1990), Adaptive cluster sampling, *J. Amer. Stat. Assoc.*, **85**, 1050-1059.

129. Thompson, S. K. (1991a), Adaptive cluster sampling: Designs and primary and secondary units, *Biometrics*, **47**, 1103-1115.

130. Thompson, S. K. (1991b), Stratified adaptive cluster sampling, *Biometrika*, **78**, 389-397.

131. Thompson, S. K. (1992), *Sampling*, John Wiley, NY, pp.343.

132. Thompson, S. K. and Ramsey, F. L. (1987), Detectability functions in observing spatial point processes, *Biometrics*, **43**, 355-362.

133. Thompson, S. K., Ramsey, F. L. and Seber, G. A. F. (1992), An adaptive procedure for sampling animal populations, *Biometrics*, **48**, 1195-1199.

134. Thompson, S. K. and Seber, G. A. F. (1994), Detectability in Conventional and Adaptive Sampling, *Biometrics*, **50**, 712-724.

135. Vadiveloo, J. (1977), *On the theory of modified randomization tests for nonparametric hypotheses*, Ph.D. thesis, Dept. of Stat., Univ. of California, Berkeley.

136. Vadiveloo, J. (1983), On the theory of modified randomization tests for nonparametric hypotheses, *Comm. Stat. A - Theory Methods*, **12**, 1581-1596.

137. Wilson, P. D. and Tonascia, J. (1971), Tables for shortest confidence ntervals on the standard deviation and variance ratio from normal distributions, *Ann. Math. Stat.*, **29**, 635-654.

138. Withers, C. S. (1983), Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals, *Ann. of Stat.*, **11**, 577-587.

139. Withers, C. S. (1984), Asymptotic expansions for distributions and quantiles with power series cumulants, *J. Roy. Stat. Soc. Ser. B*, **46**, 389-396.

140. Woodruff, R. S. (1952), Confidence Intervals for Medians and Other Position Measures, *J. Amer. Stat. Assoc.*, **47**, 635-646.

141. Wu, C. F. J. (1986), Jackknife, bootstrap and other resampling methods in regression analysis, *Ann. of Stat.*, **14**, 1261-1295.

10. APPENDIX: COMPUTER CODES

```
*** Computer program for simulation study in chap 5, 7: (1,1) of full,
group and veg models;

options nonotes nosource nosource2 ;
%macro part1;
data sample; set elk;
p=uniform(0); sight=0;
if p < prob then sight=1;
keep group behav veg sight;
%mend;

%macro part2;
proc iml; use beta; read all into beta;
use beta1; read all into beta1;
use beta2; read all into beta2;
nsim=nrow(beta)/5;
use elk; read all into pop; N=nrow(pop); pop=pop[1:N,]; pop=j(N,1,0)||pop;
/*pop_size=exp(pop[,2]); pop_size=pop_size[+,];*/ pop_size=pop[+,{2}];
print pop_size;

start sample;
      do i=1 to N;
         p=uniform(0);
         if p < pop[i,5] then pop[i,1]=1;
      end;
      num1=pop[+,{1}];
      x=j(num1,5,0); k=0;
      do i=1 to N;
         if pop[i,1]=1 then;
          do;
            k=k+1;
            x[k,]=pop[i,];
         end;
      end;
pop[,1]=j(N,1,0);
finish;

*****************    estimation;
create result var {l1 tau term1 term2 var};
*create result var {ind90 indL90 ind290 ind190 ind390 ind95 indL95 ind295 ind195 ind395};
l90=0; u90=0; low90=j(3,1,0); up90=j(3,1,0); count90=j(3,1,0); l95=0; u95=0; low95=j(3,1,0);
up95=j(3,1,0); count95=j(3,1,0);
lL90=0; uL90=0; lowL90=j(3,1,0); upL90=j(3,1,0); countL90=j(3,1,0); lL95=0; uL95=0; lowL95=j(3,1,0);
upL95=j(3,1,0); countL95=j(3,1,0);
l290=0; u290=0; low290=j(3,1,0); up290=j(3,1,0); count290=j(3,1,0); l295=0; u295=0; low295=j(3,1,0);
up295=j(3,1,0); count295=j(3,1,0);
l190=0; u190=0; low190=j(3,1,0); up190=j(3,1,0); count190=j(3,1,0); l195=0; u195=0; low195=j(3,1,0);
up195=j(3,1,0); count195=j(3,1,0);
l390=0; u390=0; low390=j(3,1,0); up390=j(3,1,0); count390=j(3,1,0); l395=0; u395=0; low395=j(3,1,0);
up395=j(3,1,0); count395=j(3,1,0);
le190=0; ue190=0; lowe190=j(3,1,0); upe190=j(3,1,0); coune190=j(3,1,0); le195=0; ue195=0; lowe195=j(3,1,0);
upe195=j(3,1,0); coune195=j(3,1,0);
le290=0; ue290=0; lowe290=j(3,1,0); upe290=j(3,1,0); coune290=j(3,1,0); le295=0; ue295=0; lowe295=j(3,1,0);
upe295=j(3,1,0); coune295=j(3,1,0);

b=j(1,4,0); esigma=j(4,4,0);
b1=j(1,2,0); esigma1=j(2,2,0); b2=b1; esigma2=esigma1;
```

```
do l=0 to nsim-1;
run sample;
x1=x[,1]||x[,4];
/*
ind90=0; indL90=0; ind290=0; ind190=0; ind390=0;
ind95=0; indL95=0; ind295=0; ind195=0; ind395=0;
*/
b=beta[l*5+1,]; esigma[1,]=beta[l*5+2,]; esigma[2,]=beta[l*5+3,];
esigma[3,]=beta[l*5+4,]; esigma[4,]=beta[l*5+5,];
b1=beta1[l*3+1,]; esigma1[1,]=beta1[l*3+2,]; esigma1[2,]=beta1[l*3+3,];
b2=beta2[l*3+1,]; esigma2[1,]=beta2[l*3+2,]; esigma2[2,]=beta2[l*3+3,];
one=j(num1,1,1);
e=j(num1,1,2.7182818); g=x[,2];
submat=x[,1:4]*esigma*x[,1:4]`#i(num1); submat=submat[+,]`;
exp1=-(x[,1:4]*b`)-(submat)/2;
theta=j(num1,3,0);
theta[,1]=one+e##exp1;
submat1=x[,1:2]*esigma1*x[,1:2]`#i(num1); submat1=submat1[+,]`;
exp1=-(x[,1:2]*b1`)-(submat1)/2;
theta[,2]=one+e##exp1;
submat2=x1[,1:2]*esigma2*x1[,1:2]`#i(num1); submat2=submat2[+,]`;
exp1=-(x1[,1:2]*b2`)-(submat2)/2;
theta[,3]=one+e##exp1;

*********** Calculating estimated Cov mx of theta hat;
evar=j(num1,3,0);
exp2=-2#(x[,1:4]*b`)-2#submat;
evar[,1]=e##exp2#(e##(submat)-one);
exp2=-2#(x[,1:2]*b1`)-2#submat1;
evar[,2]=e##exp2#(e##(submat1)-one);
exp2=-2#(x1[,1:2]*b2`)-2#submat2;
evar[,3]=e##exp2#(e##(submat2)-one);
ecov=shape({0},num1,num1); ecov1=ecov; ecov2=ecov;
 do i1=1 to num1;
   do j1=i1+1 to num1;
     z=x[i1,1:4]+x[j1,1:4];
     exp3=-z*b`-z*esigma*z`/2;
     ecvar=(2.718282##exp3)#(2.718282##(x[i1,1:4]*esigma*x[j1,1:4]`)-1);
     ecvar=g[i1]*g[j1]*ecvar;
     ecov[i1,j1]=ecvar; ecov[j1,i1]=ecvar;
     z=x[i1,1:2]+x[j1,1:2];
     exp3=-z*b1`-z*esigma1*z`/2;
     ecvar=(2.718282##exp3)#(2.718282##(x[i1,1:2]*esigma1*x[j1,1:2]`)-1);
     ecvar=g[i1]*g[j1]*ecvar;
     ecov1[i1,j1]=ecvar; ecov1[j1,i1]=ecvar;
     z=x1[i1,1:2]+x1[j1,1:2];
     exp3=-z*b2`-z*esigma2*z`/2;
     ecvar=(2.718282##exp3)#(2.718282##(x1[i1,1:2]*esigma2*x1[j1,1:2]`)-1);
     ecvar=g[i1]*g[j1]*ecvar;
     ecov2[i1,j1]=ecvar; ecov2[j1,i1]=ecvar;
   end;
 end;

term22=j(3,1,0);
term22[1]=ecov[+,+]; term22[2]=ecov1[+,+]; term22[3]=ecov2[+,+];
do l1=1 to 3;
term2=term22[l1];
extra=g#g#evar[,l1]; term2=term2+extra[+,]; *term2=term2/4;

********** Correction factor, variance, and covariates;
T=theta[,l1]#g; tau=T[+,];
TT=g#g#(theta[,l1]#theta[,l1]-theta[,l1]-evar[,l1]); term1=TT[+,];
var=term1+term2;

********** Calculate the interval estimates of population total;
se=sqrt(var); cv=se/tau;
tau1=1/tau; tau2=sqrt(tau1); tau3=tau1*tau2;
```

```
N1=1/pop_size; N2=sqrt(N1); N3=N1*N2;
l90=tau-1.645*se; u90=tau+1.645*se; l95=tau-1.96*se; u95=tau+1.96*se;
lL90=tau*exp(-1.645*cv); uL90=tau*exp(1.645*cv);
lL95=tau*exp(-1.96*cv); uL95=tau*exp(1.96*cv);
l290=tau2-1.645*cv*tau2/2; u290=tau2+1.645*cv*tau2/2;
l295=tau2-1.96*cv*tau2/2; u295=tau2+1.96*cv*tau2/2;
l190=tau1-1.645*cv*tau1; u190=tau1+1.645*cv*tau1;
l195=tau1-1.96*cv*tau1; u195=tau1+1.96*cv*tau1;
l390=tau3-1.645*cv*tau3*3/2; u390=tau3+1.645*cv*tau3*3/2;
l395=tau3-1.96*cv*tau3*3/2; u395=tau3+1.96*cv*tau3*3/2;
T=g[+,];cv2=sqrt(log(1+var/(tau-T)**2));
C=exp(1.645*cv2);
C1=sqrt(1+var/(tau-T)**2);
le190=T+(tau-T)/C; ue190=T+(tau-T)*C;
le290=T+(tau-T)/C*C1; ue290=T+(tau-T)*C*C1;
C=exp(1.96*cv2);
le195=T+(tau-T)/C; ue195=T+(tau-T)*C;
le295=T+(tau-T)/C*C1; ue295=T+(tau-T)*C*C1;

if l90 > pop_size then
    low90[l1]=low90[l1]+1;
  else if u90 < pop_size then
           up90[l1]=up90[l1]+1;
        else
  do;
          count90[l1]=count90[l1]+1;
    *ind90=1;
        end;
if l95 > pop_size then
    low95[l1]=low95[l1]+1;
  else if u95 < pop_size then
           up95[l1]=up95[l1]+1;
        else
  do;
           count95[l1]=count95[l1]+1;
    *ind95=1;
     end;
if lL90 > pop_size then
    lowL90[l1]=lowL90[l1]+1;
  else if uL90 < pop_size then
           upL90[l1]=upL90[l1]+1;
        else
  do;
          countL90[l1]=countL90[l1]+1;
    *indL90=1;
  end;
if lL95 > pop_size then
    lowL95[l1]=lowL95[l1]+1;
  else if uL95 < pop_size then
           upL95[l1]=upL95[l1]+1;
        else
  do;
           countL95[l1]=countL95[l1]+1;
    *indL95=1;
  end;
if l290 > N2 then
    low290[l1]=low290[l1]+1;
  else if u290 < N2 then
           up290[l1]=up290[l1]+1;
        else
  do;
          count290[l1]=count290[l1]+1;
    *ind290=1;
  end;
if l295 > N2 then
    low295[l1]=low295[l1]+1;
  else if u295 < N2 then
```

```
            up295[l1]=up295[l1]+1;
          else
   do;
            count295[l1]=count295[l1]+1;
     *ind295=1;
   end;
if l190 > N1 then
     low190[l1]=low190[l1]+1;
   else if u190 < N1 then
            up190[l1]=up190[l1]+1;
          else
   do;
           count190[l1]=count190[l1]+1;
    *ind190=1;
          end;
if l195 > N1 then
     low195[l1]=low195[l1]+1;
   else if u195 < N1 then
            up195[l1]=up195[l1]+1;
          else
   do;
            count195[l1]=count195[l1]+1;
     *ind195=1;
          end;
if l390 > N3 then
     low390[l1]=low390[l1]+1;
   else if u390 < N3 then
            up390[l1]=up390[l1]+1;
          else
   do;
            count390[l1]=count390[l1]+1;
     *ind390=1;
   end;
if l395 > N3 then
     low395[l1]=low395[l1]+1;
   else if u395 < N3 then
            up395[l1]=up395[l1]+1;
          else
   do;
            count395[l1]=count395[l1]+1;
     *ind395=1;
   end;
if le190 > pop_size then
     lowe190[l1]=lowe190[l1]+1;
   else if ue190 < pop_size then
            upe190[l1]=upe190[l1]+1;
          else
   do;
           coune190[l1]=coune190[l1]+1;
    *ind190=1;
          end;
if le195 > pop_size then
     lowe195[l1]=lowe195[l1]+1;
   else if ue195 < pop_size then
            upe195[l1]=upe195[l1]+1;
          else
   do;
            coune195[l1]=coune195[l1]+1;
    *ind195=1;
          end;
if le290 > pop_size then
     lowe290[l1]=lowe290[l1]+1;
   else if ue290 < pop_size then
            upe290[l1]=upe290[l1]+1;
          else
   do;
           coune290[l1]=coune290[l1]+1;
```

```
     *ind190=1;
           end;
if le295 > pop_size then
        lowe295[l1]=lowe295[l1]+1;
   else if ue295 < pop_size then
              upe295[l1]=upe295[l1]+1;
           else
   do;
              coune295[l1]=coune295[l1]+1;
      *ind195=1;
           end;
append;
end;
end;
close result;
/*
print low90 up90 count90 low95 up95 count95;
print lowL90 upL90 countL90 lowL95 upL95 countL95;
print low290 up290 count290 low295 up295 count295;
print low190 up190 count190 low195 up195 count195;
print low390 up390 count390 low395 up395 count395;
print lowe190 upe190 coune190 lowe195 upe195 coune195;
print lowe290 upe290 coune290 lowe295 upe295 coune295;
*/
do l1=1 to 3;
lper90=low90[l1]/nsim; uper90=up90[l1]/nsim; per90=count90[l1]/nsim;
lper95=low95[l1]/nsim; uper95=up95[l1]/nsim; per95=count95[l1]/nsim;
lperL90=lowL90[l1]/nsim; uperL90=upL90[l1]/nsim; perL90=countL90[l1]/nsim;
lperL95=lowL95[l1]/nsim; uperL95=upL95[l1]/nsim; perL95=countL95[l1]/nsim;
lper290=low290[l1]/nsim; uper290=up290[l1]/nsim; per290=count290[l1]/nsim;
lper295=low295[l1]/nsim; uper295=up295[l1]/nsim; per295=count295[l1]/nsim;
lper190=low190[l1]/nsim; uper190=up190[l1]/nsim; per190=count190[l1]/nsim;
lper195=low195[l1]/nsim; uper195=up195[l1]/nsim; per195=count195[l1]/nsim;
lper390=low390[l1]/nsim; uper390=up390[l1]/nsim; per390=count390[l1]/nsim;
lper395=low395[l1]/nsim; uper395=up395[l1]/nsim; per395=count395[l1]/nsim;
leper190=lowe190[l1]/nsim; ueper190=upe190[l1]/nsim; eper190=coune190[l1]/nsim;
leper195=lowe195[l1]/nsim; ueper195=upe195[l1]/nsim; eper195=coune195[l1]/nsim;
leper290=lowe290[l1]/nsim; ueper290=upe290[l1]/nsim; eper290=coune290[l1]/nsim;
leper295=lowe295[l1]/nsim; ueper295=upe295[l1]/nsim; eper295=coune295[l1]/nsim;
print lper90 uper90 per90 lper95 uper95 per95;
print lperL90 uperL90 perL90 lperL95 uperL95 perL95;
print lper290 uper290 per290 lper295 uper295 per295;
print lper190 uper190 per190 lper195 uper195 per195;
print lper390 uper390 per390 lper395 uper395 per395;
print leper290 ueper290 eper290 leper295 ueper295 eper295;
print leper190 ueper190 eper190 leper195 ueper195 eper195;
end;
quit;

/*
data result; set result; file 'veg1.ind';
put ind90 indL90 ind290 ind190 ind390 ind95 indL95 ind295 ind195 ind395;
*/
proc sort data=result; by l1;
proc univariate data=result normal plot;
by l1;
var tau term1 term2 var;
%mend;

%macro part0;
%do i=2 %to 1000;
    %part1;
    proc sort data=sample; by descending sight;
    data sample; set sample; set=&i;
    proc append base=sample1 data=sample force;
%end;
%mend;
```

```
data elk;
infile 'elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
u=2.4132+0.1322*group-1.8224*behav-0.0423*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;
/*
infile 'ssamuel.dat' lrecl=80;
input group behav veg prob;
group=log(group);
u=1.8828+0.9694*group-1.7704*behav-0.043*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;
*/

%part1;
data sample1; set sample; set=1;
proc sort data=sample1; by descending sight;
%part0;

proc logistic data=sample1 covout order=data outest=sub1 noprint;
 by set;
 model sight=group behav veg / covb;
proc logistic data=sample1 covout order=data outest=sub2 noprint;
 by set;
 model sight=group / covb;
proc logistic data=sample1 covout order=data outest=sub3 noprint;
 by set;
 model sight=veg / covb;

data beta; set sub1; keep intercep group behav veg;
data beta; set beta; if intercep=. then delete;
data beta1; set sub2; keep intercep group;
data beta1; set beta1; if intercep=. then delete;
data beta2; set sub3; keep intercep veg;
data beta2; set beta2; if intercep=. then delete;

%part2;
data beta; set beta;
file 'pori.cov';
put intercep group behav veg;
data beta1; set beta1;
file 'pgroup.cov';
put intercep group;
data beta2; set beta2;
file 'pveg.cov';
put intercep veg;
run;
```

```
*** Computer program for nonparametric bootstrap procedure in phase I;

options nonotes nosource nosource2;
%macro sample;
proc iml; use A_sample; read all into sample1; N=nrow(sample1);
sample2=j(N,4,0);
      do i=1 to N;
         m=int(uniform(0)*N+0.5); if m=0 then m=1; if m>N then m=N;
 sample2[i,]=sample1[m,];
      end;
create sample3 from sample2[colname={'group','behav','veg','sight'}];
append from sample2; close sample3;
quit;
%mend sample;

%macro more;
%sample;
/*
data sample4; set sample3; set=1;
proc sort data=sample4; by descending sight;
*/
%do i=2 %to 501;
    %sample;
    data sample3; set sample3; set=&i;
    proc sort data=sample3; by descending sight;
    proc append base=sample1 data=sample3 force;
%end;
%mend more;

%macro part0;
data sample1; set samuel;
p=uniform(0); sight=0; set=1;
if p < prob then sight=1;
keep sight group behav veg set;
data A_sample; set sample1;
keep sight group behav veg;
/*proc sort data=sample1; by descending sight;*/

%more;

proc logistic data=sample1 covout order=data outest=sub1 noprint;
by set;
 model sight= group behav veg/ covb;* influence iplots;
data beta; set sub1; keep intercep group behav veg; *if _N_=1;

/*
proc logistic data=sample4 covout order=data outest=sub1 noprint;
by set;
 model sight= group / covb;* influence iplots;
data bbeta; set sub1; keep intercep group; *if mod(_N_,3)=1;

proc means data=bbeta noprint;
var intercep group;
output out=result mean=mean1 mean2;
data result; merge result beta;
data result; set result; percent1=(mean1-intercep)/abs(intercep)*100;
percent2=(mean2-group)/abs(group)*100; keep percent1 percent2;
*/
%mend;

%macro more1;
%part0; /*data result1; set result;
%do i1=2 %to 50; %part0;
proc append base=result1 data=result force;
%end;
proc means data=result1 N mean std stderr;
var percent1 percent2;
```

```
*/
%mend;

data samuel;
/*
infile 'trial.dat';
input group prob;
data samuel; set samuel samuel;
*/
infile 'samuel/elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
group=log(group);
u=1.8828+0.9694*group-1.7704*behav-0.043*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;

%more1;
data beta; set beta;
file 'npboot.dat';
put intercep group behav veg;
run;
```

```
*** Computer program for combining the bootstrap parametric procedure in phase I
and bootstrap procedures in phase II without bootstrap t-statisic;

options nonotes nosource nosource2;
%macro part1;
proc iml; use beta; read all into beta; nsim=nrow(beta);
use samuel; read all into pop; N=nrow(pop); pop=j(N,1,1)||pop;

*****************   estimation;
b=j(1,4,0); esigma=j(4,4,0);

b=beta[1,]; esigma[1:4,]=beta[2:5,];
one=j(N,1,1);
e=j(N,1,2.7182818);
submat=pop[,1:4]*esigma*pop[,1:4]`#i(N); submat=submat[+,]`;
exp1=-(pop[,1:4]*b`)-(submat)/2;
ptheta=one+e##exp1; pop=pop[,2:4];
pi=1/ptheta; pop=pop||pi;
create ssamuel from pop[colname={'group','behav','veg','prob'}];
append from pop; close ssamuel;
quit;
%mend part1;

%macro part2;
data sample3; set ssamuel;
p=uniform(0); sight=0;
if p < prob then sight=1;
keep sight group behav veg;

proc sort data=sample3; by descending sight;
%mend part2;

%macro more;
%do i1=2 %to 601;
    %part2;
    data sample3; set sample3; set=&i1;
    proc append base=sample1 data=sample3 force;
%end;
%mend more;

%macro part3;
proc iml; use beta; read all into beta; nsim=nrow(beta);
use samuel; read all into pop; N=nrow(pop); pop=j(N,1,0)||pop;
pop_size=exp(pop[,2]); pop_size=pop_size[+,];
*pop_size=pop[{2},+];

          p=uniform(j(N,1,0));
          pop[loc(p<pop[,5]),1]=1;
       x=pop[loc(pop[,1]=1),];

*****************   estimation;
create result var {tau Tlog seen pop_size ytau yTlog yseen};
b=beta[1,]`; esigma=j(4,4,0); num1=nrow(x);
one=j(num1,1,1); e=j(num1,1,2.7182818); g=exp(x[,2]);
esigma[1:4,]=beta[2:5,];
submat=x[,1:4]*esigma*x[,1:4]`#i(num1); submat=submat[+,]`;
exp1=-(x[,1:4]*b)-(submat)/2;
theta=one+e##exp1; T=theta#g;
tau=T[+,]; seen=g[+,]; Tlog=log(tau-seen);
append;

nsim=nsim/5-1;
x[,5]=1/theta;
m=1-(1/x[,5]-int(1/x[,5]));

y1=j(num1,1,0); y=y1;
 do l=1 to 500;
```

```
************** create a population then sample;
  p=uniform(j(num1,1,0));
  y=ranbin(j(num1,1,0),int(1/x[,5]+0.5),x[,5]);
  y1[loc(p<m),]=ranbin(0,int(1/x[loc(p<m),5]),x[loc(p<m),5]);
  y1[loc(p>=m),]=ranbin(0,int(1/x[loc(p>=m),5])+1,x[loc(p>=m),5]);
b=beta[5*l+1,]`; esigma[1:4,]=beta[5*l+2:5*l+5,];
num2=y1[+,]; x2=j(num2,5,0); m2=0; num3=y[+,]; x1=j(num3,5,0); m1=0;
do k=1 to num1;
   if y1[k] > 0 then
         do;
         x2[m2+1:m2+y1[k],]=shape(x[k,],y1[k],5);
         m2=m2+y1[k];
         end;
   if y[k] > 0 then
         do;
         x1[m1+1:m1+y[k],]=shape(x[k,],y[k],5);
         m1=m1+y[k];
         end;
end;
one=j(num2,1,1);
e=j(num2,1,2.7182818); g2=exp(x2[,{2}]);
submat=x2[,1:4]*esigma*x2[,1:4]`#i(num2); submat=submat[+,]`;
exp1=-(x2[,1:4]*b)-(submat)/2;
theta=one+e##exp1; tau=theta#g2; tau=tau[+,]; seen=g2[+,];
Tlog=log(tau-seen);
one=j(num3,1,1);
e=j(num3,1,2.7182818); g2=exp(x1[,{2}]);
submat=x1[,1:4]*esigma*x1[,1:4]`#i(num3); submat=submat[+,]`;
exp1=-(x1[,1:4]*b)-(submat)/2;
theta=one+e##exp1; ytau=theta#g2; ytau=ytau[+,]; yseen=g2[+,];
yTlog=log(ytau-yseen);
append;
end;
************ cal Studentized statistic;
close result; quit;

/*
proc means data=result noprint;
var seen;
output out=burn mean=bseen;
data burn; set burn; keep bseen;

proc means noprint;
var tau seen ytau yseen;
output out=check mean=tave save ytave ysave;
*/

proc univariate data=result noprint;
var tau Tlog ytau yTlog;
output out=interv N=n bn mean=ave bave yave ybave p5=per5 bper5 yper5 ybper5
       p95=per95 bper95 yper95 ybper95 pctlpts=2.5 97.5 pctlpre=per bper yper ybp;
data result; set result; if _N_=1; keep seen pop_size;
data interv; merge interv result;
data interv; set interv;
burn5=exp(bper5)+seen; burn95=exp(bper95)+seen;
burn2_5=exp(bper2_5)+seen; burn97_5=exp(bper97_5)+seen;
yburn5=exp(ybper5)+seen; yburn95=exp(ybper95)+seen;
yburn2_5=exp(ybp2_5)+seen; yburn97=exp(ybp97_5)+seen;
data new; set interv;
keep n bn pop_size seen
     per5 per95 per2_5 per97_5
     burn5 burn95 burn2_5 burn97_5
     yper5 yper95 yper2_5 yper97_5
     yburn5 yburn95 yburn2_5 yburn97;
%mend part3;

%macro part0;
```

```
data sample1; set samuel;
p=uniform(0); sight=0; set=1;
if p < prob then sight=1;
keep sight group behav veg set;

proc sort data=sample1; by descending sight;
proc logistic covout order=data outest=beta noprint;
 model sight= group behav veg/ covb;* influence iplots;

data beta; set beta;
keep intercep group behav veg;

%part1;
%more;

proc logistic data=sample1 covout order=data outest=beta noprint;
 by set;
 model sight= group behav veg/ covb;* influence iplots;

data beta; set beta; keep intercep group behav veg;
data beta; set beta; if intercep='.' then delete;

%part3;
%mend part0;

%macro more1;
%part0; data old; set new; /*data check1; set check;*/
%do i2=2 %to 300; %part0; proc append base=old data=new force;
 *proc append base=check1 data=check force; %end;
%mend more1;

data samuel;
infile 'elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
group=log(group);
u=1.8828+0.9694*group-1.7704*behav-0.043*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;
%more1;

data old; set old;
if per5<pop_size and pop_size<per95 then
   do; p90+1; p95+1; end;
  else
   if per2_5<pop_size and pop_size<per97_5 then p95+1;
w90=per95-per5; w95=per97_5-per2_5;
if burn5<pop_size and pop_size<burn95 then
   do; bp90+1; bp95+1; end;
  else
   if burn2_5<pop_size and pop_size<burn97_5 then bp95+1;
bw90=burn95-burn5; bw95=burn97_5-burn2_5;
if yper5<pop_size and pop_size<yper95 then
   do; yp90+1; yp95+1; end;
  else
   if yper2_5<pop_size and pop_size<yper97_5 then yp95+1;
yw90=yper95-yper5; yw95=yper97_5-yper2_5;
if yburn5<pop_size and pop_size<yburn95 then
   do; ybp90+1; ybp95+1; end;
  else
   if yburn2_5<pop_size and pop_size<yburn97 then ybp95+1;
ybw90=yburn95-yburn5; ybw95=yburn97-yburn2_5;
proc print data=old;
*proc print data=check1;
run;
```

```
*** Computer program for combining the bootstrap parametric procedure in phase I
and bootstrap procedures in phase II for bootstrap t-statistic CI;

options nonotes nosource nosource2;
%macro part1;
proc iml; use beta; read all into beta; nsim=nrow(beta);
use samuel; read all into pop; N=nrow(pop); pop=j(N,1,1)||pop;

*****************   estimation;
b=j(1,4,0); esigma=j(4,4,0);

b=beta[1,]; esigma[1:4,]=beta[2:5,];
one=j(N,1,1);
e=j(N,1,2.7182818);
submat=pop[,1:4]*esigma*pop[,1:4]`#i(N); submat=submat[+,]`;
exp1=-(pop[,1:4]*b`)-(submat)/2;
ptheta=one+e##exp1; pop=pop[,2:4];
pi=1/ptheta; pop=pop||pi;
create ssamuel from pop[colname={'group','behav','veg','prob'}];
append from pop; close ssamuel;
quit;
%mend part1;

%macro part2;
data sample3; set ssamuel;
p=uniform(0); sight=0;
if p < prob then sight=1;
keep sight group behav veg;

proc sort data=sample3; by descending sight;
%mend part2;

%macro more;
%do i1=2 %to 601;
    %part2;
    data sample3; set sample3; set=&i1;
    proc append base=sample1 data=sample3 force;
%end;
%mend more;

%macro part3;
proc iml; use beta; read all into beta; nsim=nrow(beta);
use samuel; read all into pop; N=nrow(pop); pop=j(N,1,0)||pop;
pop_size=exp(pop[,2]); pop_size=pop_size[+,];

        p=uniform(j(N,1,0));
         pop[loc(p<pop[,5]),1]=1;
      x=pop[loc(pop[,1]=1),];

*****************   estimation;
b=j(1,4,0); esigma=j(4,4,0);

create result var {tau t_stat ytau yt_stat tau1 var1 pop_size};
b=beta[1,]; esigma[1:4,]=beta[2:5,]; num1=nrow(x);
one=j(num1,1,1);
e=j(num1,1,2.7182818); g=exp(x[,{2}]);
submat=x[,1:4]*esigma*x[,1:4]`#i(num1); submat=submat[+,]`;
exp1=-(x[,1:4]*b`)-(submat)/2;
theta=one+e##exp1;
*********** Calculating estimated Cov mx of theta hat;
exp2=-2#(x[,1:4]*b`)-2#submat;
evar=e##exp2#(e##(submat)-one);
ecov=shape({0},num1,num1);
 do i1=1 to num1;
   do j1=i1+1 to num1;
     z=x[i1,1:4]+x[j1,1:4];
     exp3=-z*b`-z*esigma*z`/2;
```

```
      ecvar=(2.718282##exp3)#(2.718282##(x[i1,1:4]*esigma*x[j1,1:4]`)-1);
      ecvar=g[i1]*g[j1]*ecvar;
      ecov[i1,j1]=ecvar; ecov[j1,i1]=ecvar;
    end;
  end;
 term2=ecov[+,+]; *g=exp(g); extra=g#g#evar; term2=term2+extra[+,];
 T=theta#g; tau=T[+,]; tau1=tau;
 TT=g#g#(theta#theta-theta-evar); term1=TT[+,];
 var1=term1+term2; append;
 x=x[,1:4]; x=x||theta; x[,5]=1/x[,5];

 nsim=nsim/5-1;
 m=1-(1/x[,5]-int(1/x[,5]));

 y=j(num1,1,0); y1=y;
 do l=0 to nsim;
 **************  create a population then sample;
   p=uniform(j(num1,1,0));
   y=ranbin(j(num1,1,0),int(1/x[,5]+0.5),x[,5]);
   y1[loc(p<m),]=ranbin(0,int(1/x[loc(p<m),5]),x[loc(p<m),5]);
   y1[loc(p>=m),]=ranbin(0,int(1/x[loc(p>=m),5])+1,x[loc(p>=m),5]);

 *********** Calculating estimated Cov mx of theta hat;
 b=beta[5*l+1,]; esigma[1:4,]=beta[5*l+2:5*l+5,];
 num2=y1[+,]; x2=j(num2,5,0); m2=0; num3=y[+,]; x1=j(num3,5,0); m1=0;
 do k=1 to num1;
    if y1[k] > 0 then
          do;
          x2[m2+1:m2+y1[k],]=shape(x[k,],y1[k],5);
          m2=m2+y1[k];
          end;
    if y[k] > 0 then
          do;
          x1[m1+1:m1+y[k],]=shape(x[k,],y[k],5);
          m1=m1+y[k];
          end;
 end;
 one=j(num2,1,1);
 e=j(num2,1,2.7182818); g2=exp(x2[,{2}]);
 submat=x2[,1:4]*esigma*x2[,1:4]`#i(num2); submat=submat[+,]`;
 exp1=-(x2[,1:4]*b`)-(submat)/2;
 theta=one+e##exp1;
 exp2=-2#(x2[,1:4]*b`)-2#submat;
 evar=e##exp2#(e##(submat)-one);
 ecov=shape({0},num2,num2);
  do i1=1 to num2;
    do j1=i1+1 to num2;
      z=x2[i1,1:4]+x2[j1,1:4];
      exp3=-z*b`-z*esigma*z`/2;
      ecvar=(2.718282##exp3)#(2.718282##(x2[i1,1:4]*esigma*x2[j1,1:4]`)-1);
      ecvar=g2[i1]*g2[j1]*ecvar;
      ecov[i1,j1]=ecvar; ecov[j1,i1]=ecvar;
    end;
  end;
 term2=ecov[+,+]; *g2=exp(g2); extra=g2#g2#evar; term2=term2+extra[+,];
 T=g2#theta; tau=T[+,];
 TT=g2#g2#(theta#theta-theta-evar); term1=TT[+,];
 var=term1+term2;
 t_stat=(tau-tau1)/sqrt(var);

 one=j(num3,1,1);
 e=j(num3,1,2.7182818); g2=exp(x1[,{2}]);
 submat=x1[,1:4]*esigma*x1[,1:4]`#i(num3); submat=submat[+,]`;
 exp1=-(x1[,1:4]*b`)-(submat)/2;
 theta=one+e##exp1;
 exp2=-2#(x1[,1:4]*b`)-2#submat;
 evar=e##exp2#(e##(submat)-one);
```

```
ecov=shape({0},num3,num3);
 do i1=1 to num3;
   do j1=i1+1 to num3;
     z=x1[i1,1:4]+x1[j1,1:4];
     exp3=-z*b'-z*esigma*z'/2;
     ecvar=(2.718282##exp3)#(2.718282##(x1[i1,1:4]*esigma*x1[j1,1:4]')-1);
     ecvar=g2[i1]*g2[j1]*ecvar;
     ecov[i1,j1]=ecvar; ecov[j1,i1]=ecvar;
   end;
 end;
term2=ecov[+,+]; *g2=exp(g2); extra=g2#g2#evar; term2=term2+extra[+,];
T=g2#theta; ytau=T[+,];
TT=g2#g2#(theta#theta-theta-evar); term1=TT[+,];
var=term1+term2;
yt_stat=(ytau-tau1)/sqrt(var);
append;
end;
close result; quit;

proc means data=result noprint;
var tau ytau;
output out=sigma std=stdtau ystdtau;
data sigma; set sigma; keep stdtau ystdtau;

proc univariate data=result noprint;
var tau t_stat ytau yt_stat;
output out=interv N=n tn yn ytn mean=ave tave yave ytave
       p5=per5 tper5 yper5 ytper5
       p95=per95 tper95 yper95 ytper95
       pctlpts=2.5 97.5 pctlpre=per tper yper ytp;
data result; set result; if _N_=1; keep tau1 var1 pop_size;
data interv; merge interv result sigma;
data interv; set interv;
Sper5=tper5*sqrt(var1)+tau1; Sper95=tper95*sqrt(var1)+tau1;
Sper2_5=tper2_5*sqrt(var1)+tau1; Sper97_5=tper97_5*sqrt(var1)+tau1;
Eper5=tper5*stdtau+tau1; Eper95=tper95*stdtau+tau1;
Eper2_5=tper2_5*stdtau+tau1; Eper97_5=tper97_5*stdtau+tau1;
ySper5=ytper5*sqrt(var1)+tau1; ySper95=ytper95*sqrt(var1)+tau1;
ySper2_5=ytp2_5*sqrt(var1)+tau1; ySp97_5=ytp97_5*sqrt(var1)+tau1;
yEper5=ytper5*ystdtau+tau1; yEper95=ytper95*ystdtau+tau1;
yEper2_5=ytp2_5*ystdtau+tau1; yEp97_5=ytp97_5*ystdtau+tau1;
data new; set interv;
keep n tn pop_size
     per5 per95 per2_5 per97_5 Sper5 Sper95 Sper2_5 Sper97_5
     Eper5 Eper95 Eper2_5 Eper97_5
     yper5 yper95 yper2_5 yper97_5 ySper5 ySper95 ySper2_5 ySp97_5
     yEper5 yEper95 yEper2_5 yEp97_5;
%mend;

%macro part0;
data sample1; set samuel;
p=uniform(0); sight=0; set=1;
if p < prob then sight=1;
keep sight group behav veg set;
data A_sample; set sample1;
keep sight group behav veg;

proc sort data=sample1; by descending sight;
proc logistic covout order=data outest=beta noprint;
model sight= group behav veg/ covb;* influence iplots;

data beta; set beta;
keep intercep group behav veg;

%part1;
%more;
```

```
proc logistic data=sample1 covout order=data outest=beta noprint;
 by set;
 model sight= group behav veg/ covb;* influence iplots;

data beta; set beta; keep intercep group behav veg;
data beta; set beta; if intercep='.' then delete;

%part3;
%mend part0;

%macro more1;
%part0; data old; set new;
%do i2=2 %to 300; %part0; proc append base=old data=new force;
%end;
%mend more1;

data samuel;
infile 'elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
group=log(group);
u=1.8828+0.9694*group-1.7704*behav-0.043*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;
%more1;

data old; set old;
if per5<pop_size and pop_size<per95 then
   do; p90+1; p95+1; end;
  else
   if per2_5<pop_size and pop_size<per97_5 then p95+1;
w90=per95-per5; w95=per97_5-per2_5;
if Eper5<pop_size and pop_size<Eper95 then
   do; Ep90+1; Ep95+1; end;
  else
   if Eper2_5<pop_size and pop_size<Eper97_5 then Ep95+1;
Ew90=Eper95-Eper5; Ew95=Eper97_5-Eper2_5;
if Sper5<pop_size and pop_size<Sper95 then
   do; Sp90+1; Sp95+1; end;
  else
   if Sper2_5<pop_size and pop_size<Sper97_5 then Sp95+1;
Sw90=Sper95-Sper5; Sw95=Sper97_5-Sper2_5;
if yper5<pop_size and pop_size<yper95 then
   do; yp90+1; yp95+1; end;
  else
   if yper2_5<pop_size and pop_size<yper97_5 then yp95+1;
yw90=yper95-yper5; yw95=yper97_5-yper2_5;
if yEper5<pop_size and pop_size<yEper95 then
   do; yEp90+1; yEp95+1; end;
  else
   if yEper2_5<pop_size and pop_size<yEp97_5 then yEp95+1;
yEw90=yEper95-yEper5; yEw95=yEp97_5-yEper2_5;
if ySper5<pop_size and pop_size<ySper95 then
   do; ySp90+1; ySp95+1; end;
  Else
   if ySper2_5<pop_size and pop_size<ySp97_5 then ySp95+1;
ySw90=ySper95-ySper5; ySw95=ySp97_5-ySper2_5;
proc print data=old;
run;
```

```
*** Compute expectation and variance of estimated parameters and
population size estimator when fit a simpler model (chap 7);

data samuel;
infile 'elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
group=log(group);
u=1.8828+0.9694*group-1.7704*behav-0.043*veg;
prob=1/(1+exp(-u));
keep group behav veg prob;
/*
infile 'ssamuel.dat';
input group behav veg prob;
u=2.4132+0.1322*group-1.8224*behav-0.0423*veg;
*/

proc iml;
use samuel; read all into x;
x=j(nrow(x),1,1)||x;  x1=x[,1]||x[,4]; *x1=x[,1:2];
pop_size=exp(x[,{2}]); pop_size=pop_size[+,];*pop_size=x[+,{2}];
print pop_size;
b1=j(4,1,0);
b1[1,1]=0.348307; b1[2,1]=0; b1[3,1]=0; b1[4,1]=0;
u=x1*b1[1:2]; prob=1/(1+exp(-u));
D=prob#(1-prob); D=diag(D);
z=inv(x1`*D*x1)*x1`*(x[,5]-prob);

beta=b1[1:2]; beta1=beta+z;
a=j(2,1,0.000001);
iter=1;
do while (abs(beta[1]-beta1[1])>a[1]|abs(beta[2]-beta1[2])>a[2]);
*do while (abs(beta-beta1)>a);
   iter=iter+1;
   u=x1*beta1; prob=1/(1+exp(-u));
   D=prob#(1-prob); D=diag(D);
   z=inv(x1`*D*x1)*x1`*(x[,5]-prob);
   beta=beta1; beta1=beta+z;
end;
sigma=inv(x1`*D*x1);
D1=x[,5]#(1-x[,5]); D1=diag(D1);
sigma1=sigma*x1`*D1*x1*sigma`;
print iter beta1 sigma1;
b=j(1,2,0); esigma=j(2,2,0);
b=beta1`; esigma=sigma1/4; num1=nrow(x);
one=j(num1,1,1);
e=j(num1,1,2.7182818); g=exp(x[,2]);
submat=x1*esigma*x1`#i(num1); submat=submat[+,]`;
exp1=-x1*b`-submat/2;
theta=1+e##exp1#(x1[,1]+x1[,1]/2#submat);

*********** Calculating estimated Cov mx of theta hat;
exp2=-2#(x1*b`)-submat;
evar=e##exp2#submat;
ecov=shape({0},num1,num1);
 do i1=1 to num1-1;
   do j1=i1+1 to num1;
     z=x1[i1,]+x1[j1,];
     exp3=-z*b`-submat[i1,]/2-submat[j1,]/2;
     ecvar=(2.718282**exp3)*(x1[i1,]*esigma*x1[j1,]`);
     ecvar=ecvar*g[i1]*g[j1]*x[i1,5]*x[j1,5];
     ecov[i1,j1]=ecvar; ecov[j1,i1]=ecvar;
   end;
 end;

expect=x[,5]#g#theta; expect=expect[+,]; extra=g#g#(x[,5]#evar+x[,5]#(1-x[,5])#theta#theta);
var=extra[+,]+ecov[+,+]; var=var; print expect var; run;
```

```
*** Compute mean and observed variance matrix of a set of estimated parameters.
Also compute average of estimated variance matrix.

data a1;
*infile 'h06101.dat';
*infile 'small42.cov';
*infile 'veg42.cov';
infile 'p4x4.cov';
input x1-x20;
*if _N_ <= 500;
if x1='.' then delete;
proc means noprint;
var x5-x20;
output out=result1 mean=m5-m20 N=n;

proc iml;
use a1; read all var {x1 x2 x3 x4} into beta;
use result1; read all var {m5 m6 m7 m8 m9 m10 m11 m12 m13 m14 m15
   m16 m17 m18 m19 m20} into m;
nsim=nrow(beta); nsim1=nsim;
print nsim;
sumsq=j(4,4,0);
do i=1 to nsim;
   sumsq=sumsq+beta[i,]`*beta[i,];
end;
sum=beta[+,]; cov1=(sumsq-sum`*sum/nsim)/(nsim-1);
var=j(4,4,0); var[1,]=m[,1:4]; var[2,]=m[,5:8];
var[3,]=m[,9:12]; var[4,]=m[,13:16];
ave1=sum/nsim; print ave1;
print cov1 var;
run;
```

```
*** Compute true error components 1 and 2 for a complete census;

data elk;
infile 'elk.dat';
input sight group hab behav veg;
if behav>2 then behav=2; behav=behav-1;
prob=2.4132+0.1322*group-1.8224*behav-0.0423*veg;
prob=1/(1+exp(-prob));
keep group behav veg prob;
/*
infile 'ssamuel.dat' lrecl=80;
input group behav veg prob;
group=log(group);
prob=1.8828+0.9694*group-1.7704*behav-0.043*veg;
if prob < 0.6 then delete;
if veg > 50 then delete;
if group > 9 then delete;
*/

data obs_cov;
infile 'cov.dat';
input x1-x4;

data covar;
infile 'var.dat';
input x1-x4;

proc iml;
use elk; read all into pop; N=nrow(pop); pop=j(N,1,1)||pop;
use obs_cov; read all into ocovar;
use covar; read all into covar;

***************  calculate true variance;
pi=pop[,5]; one=j(N,1,1); e=j(N,1,2.7182818); g=(pop[,2]);
pop_size=g[+,]; print pop_size;
/*
tbeta=j(1,4,0); tbeta[1,1]=1.8828; tbeta[1,2]=0.9694;
tbeta[1,3]=-1.7704; tbeta[1,4]=-0.043;
*/
tterm1=g#g#(1/pi-j(N,1,1)); tterm1=tterm1[+,]/4;
tbeta=j(1,4,0); tbeta[1,1]=2.4132; tbeta[1,2]=0.1322;
tbeta[1,3]=-1.8224; tbeta[1,4]=-0.0423;
submat1=pop[,1:4]*ocovar*pop[,1:4]`#i(N); submat1=submat1[+,]`;
submat2=pop[,1:4]*covar*pop[,1:4]`#i(N); submat2=submat2[+,]`;
texp2=-2#(pop[,1:4]*tbeta`);
evar1=e##texp2#(e##(submat1)-one); opart1=g#g#evar1#pi;
evar2=e##texp2#(e##(submat2)-one); part1=g#g#evar2#pi;
otterm2=0; tterm2=0;
do i2=1 to N-1;
   do j2=i2+1 to N;
      z=pop[i2,1:4]+pop[j2,1:4];
      texp3=-z*tbeta`;
      ecvar1=(2.718282##texp3)#
             (2.718282##(pop[i2,1:4]*ocovar*pop[j2,1:4]`)-1);
      ecvar2=(2.718282##texp3)#
             (2.718282##(pop[i2,1:4]*covar*pop[j2,1:4]`)-1);
     otterm2=otterm2+g[i2]*g[j2]*ecvar1*pi[i2]*pi[j2];
     tterm2=tterm2+g[i2]*g[j2]*ecvar2*pi[i2]*pi[j2];
   end;
end;
run;
t1=part1[+,]; ot1=opart1[+,]; t2=2*tterm2; ot2=2*otterm2;
*print t1 t2 ot1 ot2;
tterm2=(part1[+,]+2*tterm2); true_v=tterm1+tterm2;
otterm2=(opart1[+,]+2*otterm2); otrue_v=tterm1+otterm2;
total=g[+,]; print total;
print tterm1 otterm2 tterm2 otrue_v true_v; run;
```