

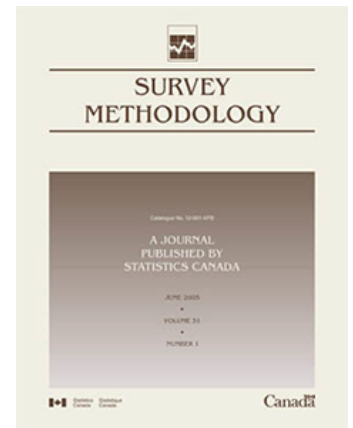
Catalogue no. 12-001-X  
ISSN 1492-0921

## Survey Methodology

# On combining independent probability samples

by Anton Grafström, Magnus Ekström, Bengt Gunnar Jonsson,  
Per-Anders Esseen and Göran Ståhl

Release date: June 27, 2019



Statistics  
Canada Statistique  
Canada

Canada

---

## How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, [www.statcan.gc.ca](http://www.statcan.gc.ca).

You can also contact us by

**Email at** [STATCAN.infostats-infostats.STATCAN@canada.ca](mailto:STATCAN.infostats-infostats.STATCAN@canada.ca)

**Telephone**, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- |   |                |
|---|----------------|
| • Statistical Information Service                             | 1-800-263-1136 |
| • National telecommunications device for the hearing impaired | 1-800-363-7629 |
| • Fax line  | 1-514-283-9350 |

### Depository Services Program

- |                  |                |
|------------------|----------------|
| • Inquiries line | 1-800-635-7943 |
| • Fax line       | 1-800-565-7757 |

## Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on [www.statcan.gc.ca](http://www.statcan.gc.ca) under “Contact us” > “[Standards of service to the public](#).”

## Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© Her Majesty the Queen in Right of Canada as represented by the Minister of Industry, 2019

All rights reserved. Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

**An HTML version is also available.**

*Cette publication est aussi disponible en français.*

---

# On combining independent probability samples

Anton Grafström, Magnus Ekström, Bengt Gunnar Jonsson,  
Per-Anders Esseen and Göran Ståhl<sup>1</sup>

## Abstract

Merging available sources of information is becoming increasingly important for improving estimates of population characteristics in a variety of fields. In presence of several independent probability samples from a finite population we investigate options for a combined estimator of the population total, based on either a linear combination of the separate estimators or on the combined sample approach. A linear combination estimator based on estimated variances can be biased as the separate estimators of the population total can be highly correlated to their respective variance estimators. We illustrate the possibility to use the combined sample to estimate the variances of the separate estimators, which results in general pooled variance estimators. These pooled variance estimators use all available information and have potential to significantly reduce bias of a linear combination of separate estimators.

**Key Words:** Horvitz-Thompson estimator; Inclusion probabilities; Linear combination estimator; Variance estimation.

## 1 Introduction

The idea of using all available information to produce better estimates is very appealing, but it is seldom clear how to proceed to achieve the best results. There is a vast literature on what has become known as meta-analysis, that builds on the idea of combining results of multiple studies. Cochran and Carroll (1953) and Cochran (1954) are two early papers that treat combination of estimates from different experiments. Koricheva, Gurevitch and Mengersen (2013) and Schmidt and Hunter (2014) are two books that provide an updated and more comprehensive treatment of meta-analysis. In this paper we do not treat combination of results from traditional experiments, but rather from multiple probability samples. We present all required design elements, such as inclusion probabilities of first and second order, for a general combination of multiple independent samples from different sampling designs. We also present new estimators for the variance of separate estimators based on the design of the combined samples. These suggested variance estimators can be thought of as general pooled variance estimators using all available information. In particular such pooled variance estimators can be used in a linear combination of separate estimators to reduce the mean square error (MSE) compared to using the separate, and thus independent, variance estimators.

A restriction is that we only treat combination of independent probability samples selected from the same population at the same point in time, or under the assumption that there has been a non-significant change in the target variable. Further, we assume that each sampling design is known to the extent that inclusion probabilities of first and second order are known for all units. In general we will also need to be able to

---

1. Anton Grafström, Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-90183 Umeå, Sweden. E-mail: anton.grafstrom@slu.se; Magnus Ekström, Department of Statistics, USBE, Umeå University, SE-90187 Umeå, Sweden, and Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-90183 Umeå, Sweden. E-mail: magnus.ekstrom@umu.se; Bengt Gunnar Jonsson, Department of Natural Sciences, Mid Sweden University, SE-85170 Sundsvall, Sweden. E-mail: bengt-gunnar.jonsson@miun.se; Per-Anders Esseen, Department of Ecology and Environmental Science, Umeå University, SE-90187 Umeå, Sweden. E-mail: per-anders.esseen@umu.se; Göran Ståhl, Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-90183 Umeå, Sweden. E-mail: goran.stahl@slu.se.

uniquely identify each unit so that we can detect if the same unit is selected in more than one sample, or multiple times in the same sample. At least some of these assumptions may be quite restrictive as they may not hold in some practical circumstances.

Let  $U = \{1, 2, \dots, N\}$  be the set of labels of the  $N$  units in the population. Our objective is to estimate the total of a target variable  $y$ , that takes value  $y_i$  for unit  $i \in U$ . Thus we wish to estimate  $Y = \sum_{i=1}^N y_i$ . We assume access to  $k$  independent probability samples  $S^{(\ell)}$ ,  $\ell = 1, \dots, k$ , from  $U$ , where the samples may be from different sampling designs. Under these assumptions, we investigate different options for estimating the population total by use of all available information. Knowledge of what units have been included in multiple different samples is required in some cases. Such knowledge is more readily available today in environmental monitoring and natural resource surveys, following the widespread use of accurate satellite-based positioning systems (Næsset and Gjevestad, 2008). In environmental studies the units can often be considered as locations with given coordinates, so the situation is different from surveys of e.g., people that may be anonymous or unidentifiable. Further, in several countries landscape and forest monitoring programmes are performed (Tomppo, Gschwantner, Lawrence and McRoberts, 2009; Ståhl, Allard, Esseen, Glimskår, Ringvall, Svensson, Sundquist, Christensen, Gallegos Torell, Högström, Lagerqvist, Marklund, Nilsson and Inghe, 2011; Fridman, Holm, Nilsson, Nilsson, Ringvall and Ståhl, 2014) which sometimes need to be augmented by special sampling programmes in order to reach specific accuracy targets for certain regions or years (Christensen and Ringvall, 2013).

In Section 2 we first recall the theory for an optimal linear combination of separate independent estimators. Then, in Section 3, we present the theory for combining independent samples. As a unit may be included in more than one sample or multiple times in the same sample we need to choose between using single or multiple count of inclusion. By using single count the resulting design becomes a without replacement design and multiple count results in a form of with replacement design. Two examples comparing different alternatives for estimation are presented in Section 4. We end with a discussion in Section 5.

## 2 Combining separate estimates

We assume that we have  $k \geq 2$  estimators,  $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$  of a population total  $Y$ , resulting from  $k$  independent samples from the same population. Our options greatly depend on what information is available. If we have estimates and corresponding variance estimates, then a linear combination based on weights calculated from estimated variances may be an interesting option. We could also weight the estimators with respect to sample size, if available, but that is known to be far from optimal in some situations. We recall the theory for an optimal linear combination of independent unbiased estimators. The linear combination of  $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$  with the smallest variance is

$$\hat{Y}_L = \alpha_1 \hat{Y}_1 + \alpha_2 \hat{Y}_2 + \dots + \alpha_k \hat{Y}_k,$$

where

$$\alpha_i = \frac{1/V_i(\hat{Y}_i)}{\sum_{j=1}^k 1/V_j(\hat{Y}_j)}$$

are positive weights that sum to 1. The variance of  $\hat{Y}_L$  is

$$V(\hat{Y}_L) = \frac{1}{\sum_{j=1}^k 1/V_j(\hat{Y}_j)}.$$

It is common that variance estimates are used in place of the unknown variances when calculating the  $\alpha$ -weights, see Cochran and Carroll (1953) and Cochran (1954). If the variance estimators are consistent, that approach will asymptotically provide the optimal weighting. Moreover, under the assumption that the variance estimators are independent of the estimators  $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$ , the resulting estimator

$$\hat{Y}_L^* = \hat{\alpha}_1 \hat{Y}_1 + \hat{\alpha}_2 \hat{Y}_2 + \dots + \hat{\alpha}_k \hat{Y}_k,$$

is unbiased and its variance depends only on the variance of  $\hat{Y}_L$  and the MSEs of the  $\hat{\alpha}_i$ 's, see Rubin and Weisberg (1974). However, as we soon will illustrate, the assumption of independence is likely to be violated in many sampling applications. In case of positive correlations between the estimators and their variance estimators, we will on average put more weight on small estimates because they tend to have smaller estimated variances. Thus the combined estimator (using weights based on estimated variances) will be negatively biased and the negative bias can increase as the number of independent surveys we combine increases, see Example 1. The opposite holds as well, in case of negative correlation, but that is likely a rarer situation in sampling applications.

**Example 1:** *A very simplistic example that illustrates that the bias can increase as the number of independent surveys we combine increase. Let the unbiased estimator  $\hat{Y}$  for one sample take the values 1 or 2 with equal probabilities and let the variance estimator take values  $c$  times the estimator (perfectly correlated) and let it be unbiased ( $c = 1/6$ ). Clearly the expected value of  $\hat{Y}$  is 1.5. Next, we consider the linear combination of two independent estimators  $(\hat{Y}_1, \hat{Y}_2)$  of the same type as  $\hat{Y}$  using estimated variances. The pair  $(\hat{Y}_1, \hat{Y}_2)$  has the following four possible outcomes (1,1), (1,2), (2,1), (2,2), each with probability 1/4. The corresponding outcomes for the linear combination  $\hat{Y}_L^*$  with estimated variances are 1, 4/3, 4/3, 2 with expectation  $17/12 \approx 1.4167$ . It is negatively biased. If a third independent estimator of the same type is added we have the eight outcomes (1,1,1), (1,1,2), (1,2,1), (1,2,2), (2,1,1), (2,1,2), (2,2,1), (2,2,2), each with equal probability 1/8. The corresponding outcomes for  $\hat{Y}_L^*$  are 1, 6/5, 6/5, 3/2, 6/5, 3/2, 3/2, 2, with expectation  $111/80 = 1.3875$ . It is even more negatively biased, and the bias continues to grow as more independent estimators of the same type are added in the combination.*

## 2.1 Why positive correlation between estimator and variance estimator is common in sampling applications

The issue of positive correlation between the estimator of a total and its variance estimator has previously been noticed by e.g., Gregoire and Schabenberger (1999) when sampling skewed biological populations, but we show that a high correlation may appear in more general sampling applications. Assume that the

target variable is non-negative and that  $y_i > 0$  for exactly  $N'$  units. The proportion of non-zero (positive)  $y_i$ 's is denoted by  $p = N'/N$ . This is a very common situation in sampling and we get such a target variable if we estimate a domain total ( $y_i = 0$  outside of the domain) or if only a subset of the population has the property of interest.

The design-based unbiased Horvitz-Thompson (HT) estimator is given by

$$\hat{Y} = \sum_{i \in S} \frac{y_i}{\pi_i},$$

where  $S$  denotes the random set of sampled units and  $\pi_i = \Pr(i \in S)$ . Under fixed size designs the variance of  $\hat{Y}$  is

$$V(\hat{Y}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij} - \pi_i \pi_j) \left( \frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2,$$

where  $\pi_{ij} = \Pr(i \in S, j \in S)$  is the second order inclusion probability. The corresponding variance estimator is

$$\hat{V}(\hat{Y}) = -\frac{1}{2} \sum_{i \in S} \sum_{j \in S} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \left( \frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2.$$

Provided that all  $\pi_{ij}$  are strictly positive, it follows that the variance estimator is an unbiased estimator of  $V(\hat{Y})$ .

The number of non-zero  $y_i$ 's in  $S$  (and hence in  $\hat{Y}$ ) is here denoted by  $n'$  and it will usually be a random number. It can be shown that the number of non-zero elements in  $\hat{V}(\hat{Y})$  is approximately proportional to  $n'$  if  $p$  is small, which indicates that there might be a strong correlation between  $\hat{Y}$  and  $\hat{V}(\hat{Y})$  in general if  $p$  is small. To show that the number of non-zero terms in  $\hat{V}(\hat{Y})$  is approximately proportional to  $n'$  we look at three cases, where the third case is the most general.

**Case 1:** Assume that all the non-zero  $y_i/\pi_i$ 's are different, i.e.,  $y_i/\pi_i \neq y_j/\pi_j$  for  $i \neq j$ , and  $\pi_{ij} \neq \pi_i \pi_j$  for all  $i, j$ . The double sum in  $\hat{V}(\hat{Y})$  then contains  $2n'(n - n')$  non-zero terms of the form

$$\frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \left( \frac{y_k}{\pi_k} \right)^2,$$

where  $k$  is equal to  $i$  or  $j$  and  $i \neq j$ . There are  $n'(n' - 1)$  non-zero terms of the form

$$\frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \left( \frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2,$$

where  $i \neq j$ . In total the number of non-zero terms is  $n'(2n - n' - 1)$ . If  $n$  is fairly large and  $p$  is small, then  $n' \ll n$  and roughly we have  $n'(2n - n' - 1) \approx 2n'n$ . The number of non-zero terms is approximately proportional to  $n'$ .

**Case 2:** Assume that all the non-zero  $y_i/\pi_i$ 's are equal, e.g.,  $y_i$  is an indicator variable and  $\pi_i = n/N$ , and  $\pi_{ij} \neq \pi_i\pi_j$  for all  $i, j$ . Then the double sum in  $\hat{V}(\hat{Y})$  contain  $2n'(n - n')$  non-zero terms of the form

$$\frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}} \left( \frac{y_k}{\pi_k} \right)^2,$$

where  $k$  is equal to  $i$  or  $j$  and  $i \neq j$ . If  $n$  is fairly large and  $p$  is small, then  $n' \ll n$  and roughly we have  $2n'(n - n') \approx 2n'n$ . Thus, the number of non-zero terms is still approximately proportional to  $n'$ .

**Case 3:** If some of the non-zero  $y_i/\pi_i$ 's are equal and the rest are different, then the number of non-zero terms will be between  $2n'(n - n')$  (case 2) and  $n'(2n - n' - 1)$  (case 1). Thus, the number of non-zero terms in  $\hat{V}(\hat{Y})$  is always approximately proportional to  $n'$  if  $p$  is small.

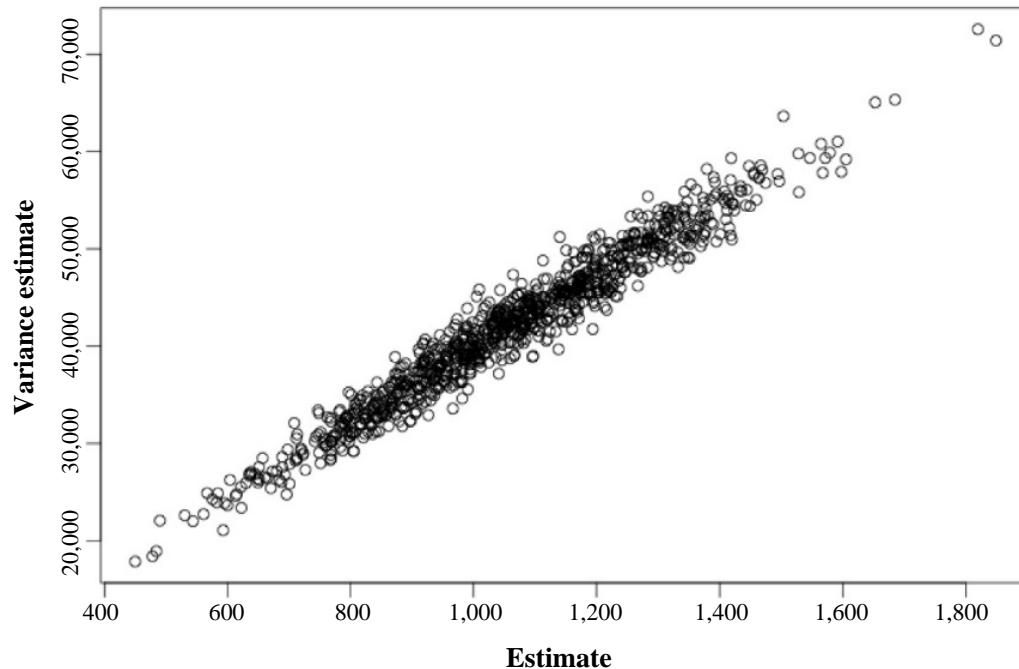
If  $\pi_{ij} \leq \pi_i\pi_j$  for all  $i \neq j$ , then all non-zero terms are positive. This condition holds e.g., for simple random sampling (SRS) and high entropy unequal probability designs such as Conditional Poisson, Sampford and Pareto. More discussion about entropy of sampling designs can be found in e.g., Grafström (2010). The average size of the positive terms in  $\hat{V}(\hat{Y})$ , or  $\hat{Y}$ , is not likely to depend much on  $n'$ . Thus, if  $\hat{Y}$  contains  $n'$  positive terms, and  $\hat{V}(\hat{Y})$  contains a number of positive terms that is proportional to  $n'$ , their sizes are mainly determined by  $n'$ . A high relative variance in  $n'$  can cause a high correlation between  $\hat{Y}$  and  $\hat{V}(\hat{Y})$ , see Example 2.

Commonly used designs can produce a high relative variance for  $n'$ . If we do simple random sampling without replacement we get  $n' \sim \text{Hyp}(N, N', n)$  and  $V(n')/E(n') = (1 - p)(N - n)/(N - 1) \approx (1 - p)(1 - n/N)$ , which means that we need a large  $p$  or a large sample fraction  $n/N$  in order to achieve a small relative variance for  $n'$ . In many applications we will have a rather small  $p$  and a small sampling fraction  $n/N$  and, thus, for many designs (that do not use prior information which can explain to some extent if  $y_i \neq 0$  or not) there will be a high relative variance for  $n'$ . To illustrate the magnitude of the resulting correlation between the estimator and its variance estimator an example for simple random sampling without replacement follows.

**Example 2:** For this example we first simulate a population of size  $N = 1,000$  where  $N' = 100$ , i.e.,  $p = 0.1$ . The 100 non-zero  $y$ -values are simulated from  $N(\mu, \sigma^2)$  with  $\mu = 10$  and  $\sigma = 2$ . We select samples of size  $n = 200$  with simple random sampling, so  $\pi_i = n/N$  and  $\pi_{ij} = n(n - 1)/(N(N - 1))$  for  $i \neq j$ . The observed correlation between  $\hat{Y}$  and  $\hat{V}(\hat{Y})$  was 0.974 for  $10^6$  samples, see Figure 2.1 for the first 1,000 observations of  $(\hat{Y}, \hat{V}(\hat{Y}))$ . If we increase  $p$  to 0.3, the correlation is still above 0.9. The results remain unchanged if the ratio  $\sigma/\mu$  remains unchanged, e.g., we get the same correlations if  $\mu = 100$  and  $\sigma = 20$ .

Now, assume we have access to more than one sample for the estimation of  $Y$ . As previously noted, with high positive correlations between the estimators and their corresponding variance estimators there is a risk of severe bias if we use a linear combination with estimated variances. The interest of using combined

information may be the largest for small domains or rare properties, in which case the problem of high correlation is the most likely. Next, we turn to alternative options for using combined information from multiple samples.



**Figure 2.1** Relationship between Horvitz-Thompson estimator and its variance estimator for a variable with 90% zeros.

### 3 Combining samples

Here we derive the design elements (e.g., inclusion probabilities of first and second order) for the combined sample. There are however different options to combine samples. We must e.g., choose between multiple or single count for the combined design. When combining independent samples selected from the same population we need to know the inclusion probabilities of all units in the samples, for all designs. Second order inclusion probabilities are needed for variance estimation. In some cases we also need to have unique identifiers (labels) for the units so they can be matched, e.g., when we use single count or when at least one separate design has unequal probabilities. Bankier (1986) considered the single count approach for the special case of combining two independently selected stratified simple random samples from the same frame. Roberts and Binder (2009) and O’Muircheartaigh and Pedlow (2002) discussed different options for combining independent samples from the same frame, but not with general sampling designs.

A somewhat similar problem is estimation based on samples from multiple overlapping frames, see e.g., the review articles by Lohr (2009, 2011) and the referenced articles therein. Even though having the same



frame can be considered as a special case of multiple frames, we have not found derivations of the design elements (in particular second order inclusion probabilities and second order of expected number of inclusions) for the combination of general sampling designs. Below we present, for general probability sampling designs, in detail two main ways to combine probability samples and derive corresponding design features needed for unbiased estimation and unbiased variance estimation.

### 3.1 Combining with single count

Here we first combine two independent samples  $S^{(1)}$  and  $S^{(2)}$  selected from the same population, and look at the union of the two samples as our combined sample. Thus, the inclusion of a unit is only counted once even if it is included in more than one sample. The first order inclusion probabilities are

$$\pi_i^{(1,2)} = \pi_i^{(1)} + \pi_i^{(2)} - \pi_i^{(1)}\pi_i^{(2)}, \tag{3.1}$$

where  $\pi_i^{(1,2)} = \Pr(i \in S^{(1)} \cup S^{(2)})$  and  $\pi_i^{(\ell)} = \Pr(i \in S^{(\ell)})$  for  $\ell = 1, 2$ . We let  $I_i^{(1)}$ ,  $I_i^{(2)}$  and  $I_i^{(1,2)}$  be the inclusion indicator for unit  $i$  in  $S^{(1)}$ ,  $S^{(2)}$  and  $S^{(1)} \cup S^{(2)}$  respectively. The resulting design is no longer a fixed size design (even if the separate designs are of fixed size). The expected size of the union  $S^{(1)} \cup S^{(2)}$  is given by  $E(n^{(1,2)}) = \sum_{i=1}^N \pi_i^{(1,2)}$ , where  $n^{(1,2)} = \sum_{i=1}^N I_i^{(1,2)}$  denotes the random size of the union. If we are interested in how much the samples will overlap on average, the expected size of the overlap is given by the sum  $\sum_{i=1}^N \pi_i^{(1)}\pi_i^{(2)}$ .

The second order inclusion probabilities  $\pi_{ij}^{(1,2)}$  for the union  $S^{(1)} \cup S^{(2)}$  can be written in terms of first and second order inclusion probabilities of the two respective designs. Let  $B = (i \in S^{(1)} \cup S^{(2)}, j \in S^{(1)} \cup S^{(2)})$ , then  $\pi_{ij}^{(1,2)} = \Pr(B)$ . By conditioning on the outcomes for  $i$  and  $j$  in  $S^{(1)}$  we get the following four cases

$m$	$A_m$	$\Pr(A_m)$	$\Pr(B   A_m)$
1	$i \in S^{(1)}, j \in S^{(1)}$	$\pi_{ij}^{(1)}$	1
2	$i \in S^{(1)}, j \notin S^{(1)}$	$\pi_i^{(1)} - \pi_{ij}^{(1)}$	$\pi_j^{(2)}$
3	$i \notin S^{(1)}, j \in S^{(1)}$	$\pi_j^{(1)} - \pi_{ij}^{(1)}$	$\pi_i^{(2)}$
4	$i \notin S^{(1)}, j \notin S^{(1)}$	$1 - \pi_i^{(1)} - \pi_j^{(1)} + \pi_{ij}^{(1)}$	$\pi_{ij}^{(2)}$

where  $\pi_{ij}^{(\ell)} = \Pr(i \in S^{(\ell)}, j \in S^{(\ell)})$  for  $\ell = 1, 2$ . The events  $A_m, m = 1, 2, 3, 4$ , are disjoint and  $\sum_{m=1}^4 \Pr(A_m) = 1$ . Thus, by the law of total probability, we have  $\pi_{ij}^{(1,2)} = \Pr(B) = \sum_{m=1}^4 \Pr(B | A_m) \Pr(A_m)$ . This gives us

$$\pi_{ij}^{(1,2)} = \pi_{ij}^{(1)} + \pi_j^{(2)} (\pi_i^{(1)} - \pi_{ij}^{(1)}) + \pi_i^{(2)} (\pi_j^{(1)} - \pi_{ij}^{(1)}) + \pi_{ij}^{(2)} (1 - \pi_i^{(1)} - \pi_j^{(1)} + \pi_{ij}^{(1)}). \tag{3.2}$$

The equations (3.1) and (3.2) can be generalized to recursively obtain first and second order inclusion probabilities of the union of an arbitrary number  $k$  of independent samples. After having derived probabilities for the union of the first two samples, we can combine the result with the probabilities of the third design using the same formulas and so on. To exemplify, let  $\pi_i^{(1, \dots, \ell)}$  be the first order inclusion probability of unit  $i$  in the union of the first  $\ell$  samples. Then we have

$$\pi_i^{(1, \dots, \ell+1)} = \pi_i^{(1, \dots, \ell)} + \pi_i^{(\ell+1)} - \pi_i^{(1, \dots, \ell)} \pi_i^{(\ell+1)},$$

as the first order inclusion probability of unit  $i$  in the union of the first  $\ell + 1$  samples. Similarly, for the second order inclusion probabilities we get the recursive formula

$$\begin{aligned} \pi_{ij}^{(1, \dots, \ell+1)} &= \pi_{ij}^{(1, \dots, \ell)} + \pi_j^{(\ell+1)} (\pi_i^{(1, \dots, \ell)} - \pi_{ij}^{(1, \dots, \ell)}) + \pi_i^{(\ell+1)} (\pi_j^{(1, \dots, \ell)} - \pi_{ij}^{(1, \dots, \ell)}) \\ &\quad + \pi_{ij}^{(\ell+1)} (1 - \pi_i^{(1, \dots, \ell)} - \pi_j^{(1, \dots, \ell)} + \pi_{ij}^{(1, \dots, \ell)}). \end{aligned}$$

Henceforth, for the combination of  $k$  independent samples, we use the simplified notation  $\pi_i = \pi_i^{(1, \dots, k)}$ ,  $\pi_{ij} = \pi_{ij}^{(1, \dots, k)}$  and  $I_i = I_i^{(1, \dots, k)}$ . Since the individual samples may overlap, the resulting design is not of fixed size. The unbiased combined single count (SC) estimator, which has Horvitz-Thompson form, is given by

$$\hat{Y}_{\text{SC}} = \sum_{i \in S} \frac{y_i}{\pi_i}.$$

The variance is

$$V(\hat{Y}_{\text{SC}}) = \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij} - \pi_i \pi_j) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j},$$

and an unbiased variance estimator is

$$\hat{V}(\hat{Y}_{\text{SC}}) = \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij} - \pi_i \pi_j) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \frac{I_i I_j}{\pi_{ij}}.$$

For the combination of independent samples with positive first order inclusion probabilities we always have  $\pi_{ij} > 0$  for all pairs  $(i, j)$ , which is the requirement for the above variance estimator to be unbiased. In terms of MSE it may be beneficial not to use the single count estimator, but instead use an estimator that accounts for the random sample size. However, here we restrict ourselves to using only unbiased estimators.

### 3.2 Combining with multiple count

We first look at how to combine two independent samples  $S^{(1)}$  and  $S^{(2)}$  selected from the same population, where we allow for each unit to possibly be included multiple times. The number of inclusions of unit  $i$  in the combined sample is denoted by  $S_i^{(1, 2)}$ , and it is the sum of the number of inclusions of unit

$i$  in the two samples we combine, i.e.,  $S_i^{(1,2)} = S_i^{(1)} + S_i^{(2)}$ , where  $S_i^{(\ell)}$  is the number of inclusions of unit  $i$  in sample  $\ell$ . The expected number of inclusions of unit  $i$  in the combination is given by

$$E(S_i^{(1,2)}) = E_i^{(1,2)} = E_i^{(1)} + E_i^{(2)}, \tag{3.3}$$

where  $E_i^{(\ell)} = E(S_i^{(\ell)})$  is the expected number of inclusions for unit  $i$  in sample  $S^{(\ell)}$ ,  $\ell = 1, 2$ . The (possibly random) sample size is the sum  $\sum_{i=1}^N S_i^{(1,2)}$  of all individual inclusions and the expected sample size is the sum  $\sum_{i=1}^N E_i^{(1,2)}$  of all individual expected number of inclusions. It can be shown that

$$E(S_i^{(1,2)} S_j^{(1,2)}) = E_{ij}^{(1,2)} = E_{ij}^{(1)} + E_i^{(1)} E_j^{(2)} + E_i^{(2)} E_j^{(1)} + E_{ij}^{(2)}, \tag{3.4}$$

where  $E_{ij}^{(\ell)} = E(S_i^{(\ell)} S_j^{(\ell)})$ ,  $\ell = 1, 2$  are the second order of expected number of inclusions in sample  $\ell$ . Obviously  $E_{ij}^{(\ell)} = \pi_{ij}^{(\ell)}$  if the design for sample  $\ell$  is without replacement. Note that as  $S_i^{(\ell)}$  may take other values than 0 or 1 we have that  $E_{ii}^{(\ell)}$  is generally not equal to  $E_i^{(\ell)}$ , but  $\pi_{ii}^{(\ell)} = \pi_i^{(\ell)}$ . The equations (3.3) and (3.4) can be used recursively to obtain  $E_i^{(k)}$  and  $E_{ij}^{(k)}$  for the combination of an arbitrary number  $k$  of independent samples. We then get the recursive formulas

$$E_i^{(1, \dots, \ell+1)} = E_i^{(1, \dots, \ell)} + E_i^{(\ell+1)}$$

and

$$E_{ij}^{(1, \dots, \ell+1)} = E_{ij}^{(1, \dots, \ell)} + E_i^{(1, \dots, \ell)} E_j^{(\ell+1)} + E_j^{(1, \dots, \ell)} E_i^{(\ell+1)} + E_{ij}^{(\ell+1)}.$$

The previous results and (3.4) follow from the fact that  $S_i^{(1, \dots, \ell+1)} = S_i^{(1, \dots, \ell)} + S_i^{(\ell+1)}$  and that  $S_i^{(1, \dots, \ell)}$  and  $S_i^{(\ell+1)}$  are independent. For example, we have

$$\begin{aligned} E_{ij}^{(1, \dots, \ell+1)} &= E(S_i^{(1, \dots, \ell+1)} S_j^{(1, \dots, \ell+1)}) = E((S_i^{(1, \dots, \ell)} + S_i^{(\ell+1)})(S_j^{(1, \dots, \ell)} + S_j^{(\ell+1)})) = \\ &E(S_i^{(1, \dots, \ell)} S_j^{(1, \dots, \ell)} + S_i^{(1, \dots, \ell)} S_j^{(\ell+1)} + S_j^{(1, \dots, \ell)} S_i^{(\ell+1)} + S_i^{(\ell+1)} S_j^{(\ell+1)}) = \\ &E_{ij}^{(1, \dots, \ell)} + E_i^{(1, \dots, \ell)} E_j^{(\ell+1)} + E_j^{(1, \dots, \ell)} E_i^{(\ell+1)} + E_{ij}^{(\ell+1)}. \end{aligned}$$

For the combination of  $k$  independent samples we now use the simplified notation  $E_i = E_i^{(1, \dots, k)}$ ,  $E_{ij} = E_{ij}^{(1, \dots, k)}$ , and  $S_i = S_i^{(1, \dots, k)}$ . The total  $Y$  can be estimated without bias with the multiple count (MC) estimator, of which the Hansen-Hurwitz estimator (Hansen and Hurwitz, 1943) is a special case. It is given by

$$\hat{Y}_{MC} = \sum_{i=1}^N \frac{y_i}{E_i} S_i.$$

We get the Hansen-Hurwitz estimator if  $E_i = np_i$ , where  $n$  is the number of units drawn and  $p_i$ , with  $\sum_{i=1}^N p_i = 1$ , are probabilities for a single independent draw. The variance of  $\hat{Y}_{MC}$  can be shown to be

$$V(\hat{Y}_{MC}) = \sum_{i=1}^N \sum_{j=1}^N (E_{ij} - E_i E_j) \frac{y_i}{E_i} \frac{y_j}{E_j}.$$

A variance estimator is

$$\hat{V}(\hat{Y}_{MC}) = \sum_{i=1}^N \sum_{j=1}^N (E_{ij} - E_i E_j) \frac{y_i}{E_i} \frac{y_j}{E_j} \frac{S_i S_j}{E_{ij}}.$$

It follows directly that the above variance estimator is unbiased, because when combining independent samples with positive first order inclusion probabilities we always have  $E_{ij} > 0$  for all pairs  $(i, j)$ .

### 3.3 Comparing the combined and separate estimators

Two examples that illustrate that the combined estimator is not necessarily as good as the best separate estimator.

**Example 3:** Assume that the first sample,  $S^{(1)}$ , is of fixed size with  $\pi_i^{(1)} \propto y_i$ , and that the second is a simple random sample with  $\pi_i^{(2)} = n/N$ . Then the Horvitz-Thompson estimator  $\hat{Y}_1 = \sum_{i \in S^{(0)}} y_i / \pi_i^{(1)}$ , has zero variance, but the combined single count estimator with  $\pi_i = \pi_i^{(1)} + \pi_i^{(2)} - \pi_i^{(1)} \pi_i^{(2)}$  has positive variance. Thus the combined estimator is worse than the best separate estimator.

**Example 4:** Assume that the design for the first sample is stratified in such a way that there is no variation within strata. Then the separate estimator  $\hat{Y}_1 = \sum_{i \in S^{(0)}} y_i / \pi_i^{(1)}$  has zero variance. If the first sample is combined with a non-stratified second sample, then the resulting design does not have fixed sample sizes for the strata. Thus, the combined estimator has a positive variance.

These examples tell us that we need to be careful before combining very different designs, such as an unequal probability design with an equal probability design or a stratified with a non-stratified sampling design. Especially, we need to be careful if we plan to estimate the total directly based on the combined sample. When combining samples from relatively similar designs, it is however likely that the combined estimator becomes better than the best of the separate estimators.

Next, we investigate how to use the combined approach for estimation of the separate variances and then use the linear combination estimator. In fact, as we will see later, using the combined approach for variance estimation of separate variances can act stabilizing for the weights in the linear combination with weights based on estimated variances. There is a sort of pooling effect for the variance estimators when they are estimated with the same set of information.

### 3.4 Using the combined sample for estimation of variances of separate estimators

An alternative to estimating directly the total  $Y$  based on the combined design is to use the combined design to estimate the variances of the separate estimators, and then proceed with a linear combination of the separate estimators. We assume access to  $k$  independent samples and that we want to estimate the variance of a separate estimator, whose variance is a double sum over the population units. There are two main options for the variance estimator; multiply by

$$\frac{I_i I_j}{\pi_{ij}} \text{ or } \frac{S_i S_j}{E_{ij}}$$

in the variance formula to obtain an unbiased estimator of the variance based on the combination of all the  $k$  samples  $S^{(\ell)}$ ,  $\ell = 1, \dots, k$ . For example, assuming that the variance of  $\hat{Y}_1$  is

$$V(\hat{Y}_1) = \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij}^{(1)} - \pi_i^{(1)} \pi_j^{(1)}) \frac{y_i}{\pi_i^{(1)}} \frac{y_j}{\pi_j^{(1)}}$$

we can use the combination of  $S^{(\ell)}$ ,  $\ell = 1, \dots, k$ , to estimate  $V(\hat{Y}_1)$  by the single count estimator

$$\hat{V}_{SC}(\hat{Y}_1) = \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij}^{(1)} - \pi_i^{(1)} \pi_j^{(1)}) \frac{y_i}{\pi_i^{(1)}} \frac{y_j}{\pi_j^{(1)}} \frac{I_i I_j}{\pi_{ij}}$$

or the multiple count estimator

$$\hat{V}_{MC}(\hat{Y}_1) = \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij}^{(1)} - \pi_i^{(1)} \pi_j^{(1)}) \frac{y_i}{\pi_i^{(1)}} \frac{y_j}{\pi_j^{(1)}} \frac{S_i S_j}{E_{ij}}$$

Note that  $\pi_{ij} = \pi_{ij}^{(1, \dots, k)}$ ,  $I_i = I_i^{(1, \dots, k)}$ ,  $E_{ij} = E_{ij}^{(1, \dots, k)}$  and  $S_i = S_i^{(1, \dots, k)}$ , so the above variance estimators use all available information on the target variable. Hence, these variance estimators can be thought of as general pooled variance estimators. It follows directly that both estimators are unbiased because all designs have positive first order inclusion probabilities, which imply that all  $\pi_{ij}$  and all  $E_{ij}$  are strictly positive. Interestingly, the above variance estimators are unbiased even if the separate design 1 has some second order inclusion probabilities that are zero, which prevent unbiased variance estimation based on the sample  $S^{(1)}$  alone.

Despite the appealing property of producing an unbiased variance estimator for any design, the above variance estimators cannot be recommended for designs with a high degree of zero second order inclusion probabilities (such as systematic sampling). The estimators can be very unstable for such designs and can produce a high proportion of negative variance estimates.

As we will see, if we intend to use a linear combination estimator, it is important that all variances are estimated in the same way. Then it is likely that the ratios, e.g.,

$$\frac{\hat{V}_{SC}(\hat{Y}_1)}{\hat{V}_{SC}(\hat{Y}_2)} \text{ and } \frac{\hat{V}_{MC}(\hat{Y}_1)}{\hat{V}_{MC}(\hat{Y}_2)}$$

become stable (have small variance). The ratios become more stable because the estimators in the numerator and denominator are based on the same information and are estimated with the same weights for all the pairs  $(i, j)$  in all estimators. With estimated variances we get

$$\hat{\alpha}_i = \left[ \sum_{j=1}^k \frac{\hat{V}(\hat{Y}_j)}{\hat{V}(\hat{Y}_j)} \right]^{-1}$$

so if the ratios of variance estimators have small variance then  $\hat{\alpha}_i$  has small variance. The weighting in the linear combination  $\hat{Y}_L^*$  then becomes stabilized. As the following example demonstrates, the ratio of the variance estimators can even have zero variance. Thus it can sometimes provide the optimal weighting even if the variances are unknown.

**Example 5:** Assume we want to combine estimates resulting from two simple random samples of different sizes. This can of course be done optimally without estimating the variances, but as an example we will use the above approach to estimate the separate variances by use of the combined sample. In this case the use of the estimators  $\hat{V}_{SC}(\hat{Y}_1)$  and  $\hat{V}_{SC}(\hat{Y}_2)$  provides the optimal weighting, and so does  $\hat{V}_{MC}(\hat{Y}_1)$  and  $\hat{V}_{MC}(\hat{Y}_2)$ . This result follows from the fact that if both designs are simple random sampling we have

$$\frac{\hat{V}_{SC}(\hat{Y}_1)}{\hat{V}_{SC}(\hat{Y}_2)} = \frac{\hat{V}_{MC}(\hat{Y}_1)}{\hat{V}_{MC}(\hat{Y}_2)} = \frac{V(\hat{Y}_1)}{V(\hat{Y}_2)},$$

which is straightforward to verify. For two simple random samples the situation corresponds to using a pooled estimate for  $S^2$  (the population variance of  $y$ ) in the expressions for the variance estimates, and this pooled estimate is then cancelled out in the calculation of the weights.

The conclusion is that this procedure is likely to provide a more stable weighting also for designs that deviate from simple random sampling as long as the involved designs have large entropy (a high degree of randomness). The problem of bias for the linear combination estimator with estimated variances will be reduced compared to using separate and thus independent variance estimators.

We believe that this can be a very interesting alternative, because the estimator of the total based on the combined design does not necessarily provide a smaller variance than the best of the separate estimators. With this strategy we can improve the separate variance estimators, especially for a smaller sample (if data is available for a larger sample). Hence the resulting linear combination with jointly estimated variances can be a very competitive strategy.

With single count we might use a ratio type variance estimator such as the following

$$\hat{V}_R(\hat{Y}_1) = \frac{N^2}{\gamma_{1, \dots, k}} \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij}^{(1)} - \pi_i^{(1)} \pi_j^{(1)}) \frac{y_i}{\pi_i^{(1)}} \frac{y_j}{\pi_j^{(1)}} \frac{I_i I_j}{\pi_{ij}},$$

where  $\gamma_{1, \dots, k} = \sum_{i=1}^N \sum_{j=1}^N \frac{I_i I_j}{\pi_{ij}}$ . For multiple count we can replace  $I_i I_j / \pi_{ij}$  with  $S_i S_j / E_{ij}$ . This ratio estimator uses the known size of the population of pairs  $(i, j) \in \{1, 2, \dots, N\}^2$ , which is  $N^2$ , and divides by the sum of the sample weights for the pairs. Note that  $E(\gamma_{1, \dots, k}) = N^2$ . This correction is useful because the number of pairs in the estimator may be random (since the union of the samples may have random size). This rescales the sample (of pairs) weights to sum to  $N^2$ . This will introduce some bias (as usual for ratio estimators), but the idea is that this will reduce the variance of the variance estimator. However, this approach is only useful if we are interested in the separate variance as the correction term will be the same for all separate variance estimators. Hence it does not change the weighting of a linear combination estimator with estimated variances.

## 4 Simulation examples

Two Monte-Carlo simulation examples are presented here. In the first example we combine two Poisson samples using inclusion probabilities approximately proportional to the target variable. In the second example we combine an unstratified simple random sample with a stratified simple random sample.

### 4.1 Combining two Poisson samples

We generate a population of size  $N = 200$ , with an auxiliary variable  $X_i \sim N(\mu = 20, \sigma^2 = 16)$ . The target variable is generated as  $(Y_i | X_i = x_i) = x_i + \epsilon_i$ , where  $\epsilon_i \sim N(0, (x_i/20)^2)$ . Two sets of inclusion probabilities are generated  $\pi_i^{(1)} \propto \pi_i^{(2)} \propto x_i$ , where  $\sum_{i=1}^N \pi_i^{(1)} = n_1$  and  $\sum_{i=1}^N \pi_i^{(2)} = n_2$ . We let the expected sample sizes be  $n_1 = 15$  and  $n_2 = 25$ . For simplicity we let both designs be Poisson designs (where units are selected independently). This allows us to calculate exactly the variances for both separate estimators (and thus the optimal linear combination) and for the combined samples with single and multiple count. For the strategies with linear combination using estimated variances, we performed a Monte-Carlo simulation with 1,000,000 repeated sample selections. True variances for the two separate HT estimators, the SC/MC estimators for the combined samples and the optimal linear combination of the separate estimators are presented (Table 4.1). Simulation results for the different linear combinations with estimates variances are also presented (Table 4.1).

**Table 4.1**  
**Results for the combination of Poisson samples. True variances for the two separate HT estimators, the SC/MC estimators for the combined samples and the optimal linear combination of the separate estimators. Simulation results, in terms of estimated bias and MSE, for three linear combination estimators with estimated variances**

Estimator	Bias (Rel. bias)	MSE
$\hat{Y}_1$	0	1,053,083
$\hat{Y}_2$	0	596,069
$\hat{Y}_{sc}$	0	361,088
$\hat{Y}_{mc}$	0	380,929
$\hat{Y}_L$ Optimal	0	380,626
$\hat{Y}_L^*$ Separate	-92.8 (-2.24%)	412,248
$\hat{Y}_L^*$ Pooled SC	1.6 (+0.04%)	381,106
$\hat{Y}_L^*$ Pooled MC	1.6 (+0.04%)	381,106

Using combined (pooled) variance estimators reduced both the bias and the variance for a linear combination in comparison to using separate variance estimators. For this example, the linear combination with pooled variance estimation came very close to the optimal linear combination in performance. The negative bias with separate variance estimators is mainly due to a positive correlation between the total

estimator and its variance estimator under the Poisson design. For this setting, the best result was obtained by combining the samples using a single count.

## 4.2 Combining an unstratified SRS with a stratified SRS

Here we generated a population of size  $N = 1,000$ , with two strata of sizes  $N_1 = 600$  and  $N_2 = 400$ . The target variable  $y_i, i = 1, \dots, N$ , was generated as follows. In stratum 1 there were 500  $y_i$ 's equal to zero and the other 100  $y_i$ 's were drawn from  $N(\mu_1 = 10, \sigma^2 = 4)$ . In stratum 2 there were 300  $y_i$ 's equal to zero and the other 100  $y_i$ 's were drawn from  $N(\mu_2 = 15, \sigma^2 = 4)$ . The first sample is an unstratified simple random sample of size  $n = 50$  and the second sample is a stratified simple random sample with stratum sample sizes  $n_1 = 30$  and  $n_2 = 20$ . The variances for both separate HT estimators and for the combined samples with single and multiple count were calculated exactly. A Monte-Carlo simulation with 10,000 repetitions was performed to evaluate the performance of a linear combination estimator with estimated variances. The results are presented in Table 4.2. Bias is reduced by using a linear combination with pooled variance estimators compared with using separate variance estimators. Also for this setting, the best result was obtained by combining the samples using a single count.

**Table 4.2**

**Results for the combination of a SRS and a stratified SRS. True variances for the two separate HT estimators, the SC/MC estimators for the combined sample and the optimal linear combination of the separate estimators. Simulation results, in terms of estimated bias and MSE, for three linear combination estimators with estimated variances**

Estimator	Bias (Rel. bias)	MSE
$\hat{Y}_1$	0	516,835
$\hat{Y}_2$	0	498,321
$\hat{Y}_{SC}$	0	248,888
$\hat{Y}_{MC}$	0	253,789
$\hat{Y}_L$ Optimal	0	253,704
$\hat{Y}_L^*$ Separate	-77 (-3%)	287,680
$\hat{Y}_L^*$ Pooled SC	9 (+0.4%)	257,229
$\hat{Y}_L^*$ Pooled MC	9 (+0.4%)	257,217

## 5 Discussion

The simulation examples in the previous section are only intended to demonstrate the different approaches and we make no claim of the generality of the result. However, we find it very likely that using pooled variance estimators is better than using separate variance estimators in a linear combination estimator, especially in cases where the separate total estimators are highly correlated to their variance estimators.



We have presented in detail how to combine independent probability samples and derived corresponding design features needed to do unbiased estimation and variance estimation. The danger of using the combined sample approach for very different designs has been illustrated. Moreover, we have shown that there is often a risk for a strong positive correlation between the HT estimator and its variance estimator. Such dependence can be a source for bias if estimated variances are used in a linear combination. Thus, as an alternative approach, we have shown how to use the combined sample to estimate separate variances. This alternative approach can lead to more stable weights in a linear combination of separate estimators, and has potential to reduce both bias and variance.

There are of course limitations to when this methodology can be applied due to our assumption of fully known designs and use of the same frame with identifiable units. Sensitivity for deviations from some of these assumptions, such as having unidentifiable units or using approximate second order inclusion probabilities, needs further investigation.

In particular, knowledge of this methodology is important if an initial sampling effort was proven insufficient. Such situations are common in e.g., environmental monitoring (Christensen and Ringvall, 2013). Then a complementary sample may be designed in such a way that it allows for a combination with improved efficiency.

## Acknowledgements

We thank an anonymous reviewer and an associate editor for helpful comments that improved the paper. This work was funded by the Swedish Science Council grant 340-2013-5076.

## References

- Bankier, M.D. (1986). Estimators based on several stratified samples with applications to multiple frame surveys. *Journal of the American Statistical Association*, 81(396), 1074-1079.
- Christensen, P., and Ringvall, A.H. (2013). Using statistical power analysis as a tool when designing a monitoring program: Experience from a large-scale Swedish landscape monitoring program. *Environmental Monitoring and Assessment*, 185(9), 7279-7293.
- Cochran, W.G. (1954). The combination of estimates from different experiments. *Biometrics*, 10(1), 101-129.
- Cochran, W.G., and Carroll, S.P. (1953). A sampling investigation of the efficiency of weighting inversely as the estimated variance. *Biometrics*, 9(4), 447-459.
- Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A.H. and Ståhl, G. (2014). Adapting National Forest Inventories to changing requirements—the case of the Swedish National Forest Inventory at the turn of the 20<sup>th</sup> century. *Silva Fennica*, 48(3), article id 1095.

- Grafström, A. (2010). Entropy of unequal probability sampling designs. *Statistical Methodology*, 7(2), 84-97.
- Gregoire, T.G., and Schabenberger, O. (1999). Sampling-skewed biological populations: Behavior of confidence intervals for the population total. *Ecology*, 80, 1056-1065.
- Hansen, M.H., and Hurwitz, W.N. (1943). On the theory of sampling from finite populations. *The Annals of Mathematical Statistics*, 14(4), 333-362.
- Koricheva, J., Gurevitch, J. and Mengersen, K. (Eds.) (2013). Handbook of meta-analysis in ecology and evolution. Princeton University Press.
- Lohr, S.L. (2009). Multiple-frame surveys. *Handbook of Statistics*, 29, 71-88.
- Lohr, S.L. (2011). Alternative survey sample designs: Sampling with multiple overlapping frames. *Survey Methodology*, 37, 2, 197-213. Paper available at <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2011002/article/11608-eng.pdf>.
- Næsset, E., and Gjevestad, J.G. (2008). Performance of GPS precise point positioning under conifer forest canopies. *Photogrammetric Engineering & Remote Sensing*, 74(5), 661-668.
- O'Muircheartaigh, C., and Pedlow, S. (2002). Combining samples vs. cumulating cases: a comparison of two weighting strategies in NLSY97. In *Proceedings of the 2002 Joint Statistical Meetings*, American Statistical Association, 2557-2562.
- Roberts, G., and Binder, D. (2009). Analyses based on combining similar information from multiple surveys. In *Joint Statistical Meetings 2009 Survey Research Methods Section*, 2138-2147.
- Rubin, D.B., and Weisberg, S. (1974). The variance of a linear combination of independent estimators using estimated weights. *ETS Research Bulletin Series*, 1974(2), i-5.
- Schmidt, F.L., and Hunter, J.E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage publications.
- Ståhl, G., Allard, A., Esseen, P.-A., Glimskår, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Gallegos Torell, Å., Högström, M., Lagerqvist, K., Marklund, L., Nilsson, B. and Inghe, O. (2011). National Inventory of Landscapes in Sweden (NILS)-scope, design, and experiences from establishing a multiscale biodiversity monitoring system. *Environmental Monitoring and Assessment*, 173(1-4), 579-595.
- Tomppo, E., Gschwantner, T., Lawrence, M. and McRoberts, R.E. (Eds.) (2009). *National Forest Inventories: Pathways for Common Reporting*. Springer Science & Business Media.