Journal of Official Statistics, Vol. 12, No. 3, 1996, pp. 241–251

Cross-Classified Sampling

Esbjörn Ohlsson[1]

From a finite, two-dimensional population a sample of rows and a sample of columns are drawn independently. A sample from the two-dimensional population is formed by cross-classifying these samples. We call this procedure *cross-classified sampling* (CCS). This article is concerned with the problem of calculating the variance of estimators based on CCS samples. A general variance decomposition is presented. This decomposition leads to our main result, which simplifies the variance derivation when the one-dimensional samples are (separately) stratified. Finally, we apply the general results to the special case with the Horvitz-Thompson estimator in unequal probability sampling. The results in this article have been used to derive variance estimators for the Swedish Consumer Price Index.

Key words: Two-dimensional sampling; variance decomposition; variance estimation; Consumer Price Index.

1. Introduction

We consider a finite, two-dimensional population where we refer to the two dimensions as *rows* and *columns*, respectively. A sample of rows and a sample of columns are drawn independently. The cross-classification of these samples is our sample from the two-dimensional population, cf. Figure 1. We call this procedure *cross-classified sampling* (CCS). Note that the design of the row sample, on the one hand, and the column sample, on the other hand, may be arbitrarily complex here.

This article focuses on deriving a formula for the variance of an estimator based on a CCS sample. In Section 1.1 we give a decomposition of the variance that is applicable to any estimator with a CCS design. In Section 2 we consider the situation where each of the one-dimensional samples of rows and columns is stratified, cf. Figure 1. The cross-classification of these strata does *not* yield an independent stratification of the two-dimensional population. Our main result shows how we can utilize the variance decomposition of Section 1 to simplify variance calculations in the stratified case. It turns out that, in spite of the dependence of the cross-classified strata, we do not have to bother about any covariance terms. In Section 3 we specialize to the case with the Horvitz-Thompson estimator in unequal probability sampling.

The background to our interest in CCS is that large parts of the Swedish Consumer Price

[1] Mathematical Statistics, Stockholm University, S-106 91 Stockholm, Sweden[2]

[2] Large parts of this work were done while the author served as senior statistician at Statistics Sweden.

Acknowledgement: The author is sincerely grateful to Jörgen Dalén, Statistics Sweden, for many valuable discussions during the course of this work.

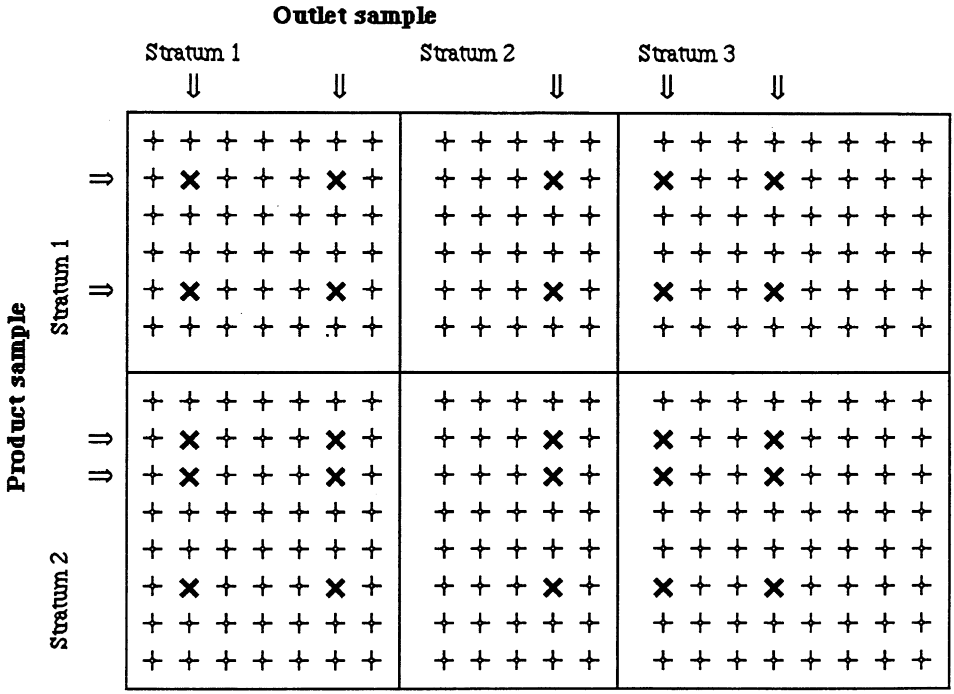

Fig. 1. Schematic view of Cross-Classified Sampling. Arrows show the product sample and outlet sample, respectively. Light plus signs indicate the two-dimensional population while bold crosses indicate the cross-classified sample

Index (CPI) are based on price quotations from CCS samples. These are cross-classifications of an outlet sample and a product sample. In Dalén and Ohlsson (1995) our general results on CCS are used to derive variance estimators for various parts of the CPI. The resulting numerical estimates have been used both to improve sample allocation and to obtain confidence intervals for the index.

Another application of our results is in the UK Retail Price Index, see Central Statistical Office (1995). For references to other papers on the variance of price indexes see Dalén and Ohlsson (1995).

In the particular case when the samples in both dimensions are drawn by simple random sampling without replacement (srswor) the variance of the sample mean can be obtained from the results in Vos (1964), though not explicitly given there. The outlet sample and the product sample for the CPI are stratified and drawn with unequal probabilities within strata. The starting point for the present work was a need to extend results derived from Vos (1964) to the more complicated CPI design.

Much work on two-dimensional sampling is concerned with the problem of area sampling (plane sampling). Quenouille (1949) deals with two-dimensional samples that are obtained by using stratified simple random sampling or equal probability systematic sampling in each dimension at a time; variances are derived under a superpopulation model. In the terminology of Quenouille, CCS can be described as the case with *aligned* samples in both dimensions. For references on extensions of this work, see Iachan

(1982) and Ripley (1981). To the best of our knowledge, none of these results can be used to solve the general CCS variance problem. Other articles on two-dimensional sampling discuss sampling among the cells in the crossing of two stratifications. This typically leads to "Latin square" type samples rather than CCS samples; for an overview and references see Cochran (1978).

1.1. *Definition of CCS and a variance decomposition*

We consider a two-dimensional population with $M \cdot N$ units, arranged in a matrix with N rows and M columns. For each population unit we have a value of our target variable y; these values form the matrix $\{y_{ij}; i = 1, 2. \ldots, N; j = 1, 2, \ldots, M\}$. A survey is carried out with the object of estimating the population total

$$Y = \sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij} \tag{1.1}$$

To this end, a random sample S^R of rows and a random sample S^C of columns are drawn. The sampling procedures for rows and columns are assumed to be independent; in all other respects S^R and S^C are arbitrary here.

Definition 1.1. A cross-classified sample S from a two-dimensional population, indexed by $\{(i,j)\}$ is the cross-classification of S^R and S^C, i.e., $S = \{(i,j): i \in S^R, j \in S^C\}$, where S^R and S^C are independent samples from the rows and columns.

On the basis of the y-values for the units in the sample S, we form some estimator $\hat{Y}$ of Y. The problem we focus on in this article is the derivation of an explicit formula for $V(\hat{Y})$.

We now present a decomposition of $V(\hat{Y})$, which will be useful in the further derivations. Let E^R denote *conditional* expectation, given the outcome of S^R. By the independence of S^R and S^C, E^R is simply the expectation over column samples. Conditional expectation given S^C is denoted E^C, while E is overall expectation. For variances we analogously define V^R, V^C and V. Let

$$\hat{Y}^R = E^R(\hat{Y}) \quad \hat{Y}^C = E^C(\hat{Y}) \tag{1.2}$$

We call $\hat{Y}^R$ the *row estimator* of Y. The reason for this is that $\hat{Y}^R$ depends solely on the outcome of the row sample: $\hat{Y}^R$ is an estimator which could be used if the columns were completely enumerated. Similarly, $\hat{Y}^C$ is the *column estimator* of Y. If $\hat{Y}$ is unbiased, then obviously both the row and column estimators are, too.

Theorem 1.1. The variance of an estimator $\hat{Y}$, which is based on a cross-classified sample, can be decomposed as follows

$$V(\hat{Y}) = VR + VC + VRC \tag{1.3}$$

where

$$VR = V(\hat{Y}^R), \quad VC = V(\hat{Y}^C), \quad VRC = V(\hat{Y} - \hat{Y}^R - \hat{Y}^C) \tag{1.4}$$

We call VR the row variance, VC the column variance and VRC the row and column *interaction variance*. Before proving the theorem we illustrate it in a simple example.

Example 1.1. Suppose that S^R is a simple random sample drawn without replacement (srswor) and having size n, and that S^C is an srswor of size m. Let

$$y_{i\cdot} = \sum_{j=1}^{M} y_{ij} \quad y_{\cdot j} = \sum_{i=1}^{N} y_{ij} \quad y_{\cdot\cdot} = Y = \sum_{i=1}^{M} \sum_{j=1}^{M} y_{ij} \tag{1.5}$$

$$\bar{y}_{i\cdot} = y_{i\cdot}/M \quad \bar{y}_{\cdot j} = y_{\cdot j}/N \quad \bar{y}_{\cdot\cdot} = y_{\cdot\cdot}/NM \tag{1.6}$$

The conventional unbiased estimator of Y is given by

$$\hat{Y} = \frac{NM}{nm} \sum_{i \in S^R} \sum_{j \in S^C} y_{ij} \tag{1.7}$$

In this case the row and column estimators are readily found as

$$\hat{Y}^R = \frac{N}{n} \sum_{i \in S^R} y_i \quad \hat{Y}^C = \frac{M}{m} \sum_{j \in S^C} y_{\cdot j} \tag{1.8}$$

Furthermore, let

$$\sigma_R^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 \quad \sigma_C^2 = \frac{1}{M-1} \sum_{j=1}^{M} (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2$$

$$\sigma_{RC}^2 = \frac{1}{N-1} \frac{1}{M-1} \sum_{i=1}^{N} \sum_{j=1}^{M} (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2 \tag{1.9}$$

The variance of $\hat{Y}$ in (1.7) is given by (1.3) and

$$VR = \frac{N^2 M^2}{n} \left(1 - \frac{n}{N}\right) \sigma_R^2 \tag{1.10}$$

$$VC = \frac{N^2 M^2}{m} \left(1 - \frac{m}{M}\right) \sigma_C^2 \tag{1.11}$$

$$VRC = \frac{N^2 M^2}{nm} \left(1 - \frac{n}{N}\right) \left(1 - \frac{m}{M}\right) \sigma_{RC}^2 \tag{1.12}$$

Note the similarity between the variances in (1.9) and the mean sums of squares in a two-way analysis of variance table. This is one reason for using the terminology "row, column and interaction variance."

Equations (1.10)–(1.12) are readily derived from Theorem 3.1 below; we omit the details, though. The expression for $V(\hat{Y})$ in the srswor case can alternatively be derived by using the results in Vos (1964, Method C.1). Note that Vos (1964) cannot be used to cover more general sampling procedures S^R and S^C; the need for a generalization from srswor to pps and stratified samples was the starting point for the work reported here.

Proof of Theorem 1.1. Let $Z = \hat{Y} - \hat{Y}^R - \hat{Y}^C$ so that

$$\hat{Y} = \hat{Y}^R + \hat{Y}^C + Z. \tag{1.13}$$

It is sufficient to prove that $\hat{Y}^R$, $\hat{Y}^C$ and Z are mutually uncorrelated. From the independence between S^R and S^C it follows that $C(\hat{Y}^R, \hat{Y}^C) = 0$. Furthermore

$$C(\hat{Y}^R, Z) = C(\hat{Y}^R, \hat{Y} - \hat{Y}^R) - C(\hat{Y}^R, \hat{Y}^C) = C(\hat{Y}^R, \hat{Y} - \hat{Y}^R) \tag{1.14}$$

Let $\mu = E(\hat{Y})$. From (1.2) we get

$$E^R[(\hat{Y}^R - \mu)(\hat{Y} - \hat{Y}^R)] = (\hat{Y}^R - \mu) \cdot E^R[\hat{Y} - \hat{Y}^R] = 0 \tag{1.15}$$

By taking expectations in (1.15) we find that $C(\hat{Y}^R, \hat{Y} - \hat{Y}^R) = 0$; together with (1.14), this yields $C(\hat{Y}^R, Z) = 0$, as desired. By symmetry, we also have $C(\hat{Y}^C, Z) = 0$, which concludes the proof.

2. Cross-Classified Sampling and Stratification

If we imposed an ordinary two-way stratification on our two-dimensional population, sampling independently in each cell, then the results of the preceding section could, of course, be used inside each cell, and $V(\hat{Y})$ could be found by simply adding the cell variances. In the CPI case (and possibly others), however, the rows and columns are separately stratified. The cross-classification of these stratifications yield cells in which the samples are actually *dependent*. In this section we discuss the derivation of variances with the latter type of stratification.

Let us assume, then, that the rows are divided into G strata of sizes $N_1, N_2, \ldots, N_G$. As usual, S^R is the union of independent samples from each of the strata. Similarly, the columns are divided into H strata of sizes $M_1, M_2, \ldots, M_H$. We still assume S^R and S^C to be independent, and Theorem 1.1 is still valid.

By crossing the row-stratification and the column-stratification we get a division of the population into $G \cdot H$ cells. Let Y_{gh} denote the population total of the y's in cell (g, h), for $g = 1, 2, \ldots, G$; $h = 1, 2, \ldots, H$. Then Y in (1.1) can be rewritten

$$Y = \sum_{g=1}^{G} \sum_{h=1}^{H} Y_{gh} \tag{2.1}$$

We assume that the estimator of Y is composed of some estimators $\hat{Y}_{gh}$ of the cell totals Y_{gh}, i.e.,

$$\hat{Y} = \sum_{g=1}^{G} \sum_{h=1}^{H} \hat{Y}_{gh} \tag{2.2}$$

We assume that $\hat{Y}_{gh}$ is computed from the sampled units in cell (g, h) only. In other respects, the $\hat{Y}_{gh}$'s are arbitrary. Note that a pair of $\hat{Y}_{gh}$'s are not independent if they are from the same row or the same column. Hence, $V(\hat{Y})$ cannot simply be calculated as the sum of the $V(\hat{Y}_{gh})$'s, as with ordinary two-way stratification. On the contrary, we have to add a number of covariance terms to the sum, making variance estimation very cumbersome in practice. However, invoking the decomposition of Theorem 1.1 makes the situation much simpler, as we now see.

Introduce the within-cell row and column estimators

$$\hat{Y}_{gh}^R = E^R(\hat{Y}_{gh}) \quad \hat{Y}_{gh}^C = E^C(\hat{Y}_{gh}) \tag{2.3}$$

The within row stratum g (column stratum h) estimators are defined as

$$\hat{Y}_{g\cdot}^R = \sum_{h=1}^{H} \hat{Y}_{gh}^R \quad \hat{Y}_{\cdot h}^C = \sum_{g=1}^{G} \hat{Y}_{gh}^C \tag{2.4}$$

Trivially, we have the following relations

$$\hat{Y}^R = \sum_{g=1}^{G} \hat{Y}_{g\cdot}^R = \sum_{g=1}^{G}\sum_{h=1}^{H} \hat{Y}_{gh}^R \quad \hat{Y}^C = \sum_{h=1}^{H} \hat{Y}_{\cdot h}^R = \sum_{g=1}^{G}\sum_{h=1}^{H} \hat{Y}_{gh}^C \tag{2.5}$$

The within row stratum g (column stratum h) variance is defined as

$$VR_g = V(\hat{Y}_{g\cdot}^R) \quad VC_h = V(\hat{Y}_{\cdot h}^C) \tag{2.6}$$

We also define the within cell (g, h) interaction as

$$VRC_{gh} = V(\hat{Y}_{gh} - \hat{Y}_{gh}^R - \hat{Y}_{gh}^C) \tag{2.7}$$

Theorem 2.1. Suppose we have a CCS procedure where the row and column samples are separately stratified. Then the variance of an estimator $\hat{Y}$, based on the CCS sample and structured as in (2.2), is given by

$$V(\hat{Y}) = VR + VC + VRC \tag{2.8}$$

where VR, VC and VRC are defined in (1.4) and can be expanded as

$$VR = \sum_{g=1}^{G} VR_g \quad VC = \sum_{h=1}^{H} VC_h \tag{2.9}$$

$$VRC = \sum_{g=1}^{G}\sum_{h=1}^{H} VRC_{gh} \tag{2.10}$$

Since *VR* and *VC* in (1.4) depend only on ordinary ''one-dimensional'' samples, it should be no surprise that they can be expanded as simply as in (2.9). The significance of the theorem lies in the fact that VRC can be equally simply expanded (without any covariance terms). We find that the decomposition of Theorem 1.1 provides us with the means of making the variance calculations with these dependent cells almost as simple as in the independent case.

Proof of Theorem 2.1. Equation (2.8) follows directly from Theorem 1.1 and (2.9) is an immediate consequence of the independence between the samples in different row (column) strata, and (1.4), (2.5) and (2.6).

It remains to show that *VRC*, as defined in (1.4), can be expressed as in (2.10). To this end, let $\mu_{gh} = E(\hat{Y}_{gh})$ and

$$Z_{gh} = \hat{Y}_{gh} - \hat{Y}_{gh}^R - \hat{Y}_{gh}^C + \mu_{gh} \tag{2.11}$$

Note that $E(Z_{gh}) = 0$. From (1.4), (2.2) and (2.5) we see that

$$VRC = V\left(\sum_{g=1}^{G}\sum_{h=1}^{H} Z_{gh}\right) \tag{2.12}$$

We must show that the right-hand sides in (2.12) and (2.10) are equal. By definition (2.7), $VRC_{gh} = V(Z_{gh})$ and it is sufficient to show that the Z_{gh}'s are mutually uncorrelated. By the independence of the sampling in different rows and column strata, it is immediate that

$$C(Z_{gh}, Z_{g'h'}) = 0 \qquad g \neq g' \quad h \neq h' \tag{2.13}$$

Next we turn to the case when $g = g'$ and $h \neq h'$. By the independence between S^R and S^C, we have $E^C(\hat{Y}^R_{gh}) = E(\hat{Y}^R_{gh}) = \mu_{gh}$, which together with the second equality in (2.3) yields that $E^C(Z_{gh}) = 0$. By the independence of the samples in the column strata h and h', we have

$$E^C(Z_{gh}Z_{gh'}) = E^C(Z_{gh}) \cdot E^C(Z_{gh'}) = 0 \tag{2.14}$$

By taking expectations in (2.14), we conclude that

$$C(Z_{gh}, Z_{gh'}) = E(Z_{gh}Z_{gh'}) = 0 \tag{2.15}$$

The case with $g \neq g'$ and $h = h'$ follows by symmetry. The proof of Theorem 2.1 is complete.

3. Cross-Classified Sampling and the Horvitz-Thompson Estimator

Here we discuss the important special case when $\hat{Y}$ is the unbiased Horvitz-Thompson estimator (also called the π-estimator). First we go back to the unstratified situation. Following Särndal et al. (1992) we introduce the following notation. For any i, i', j and j', set

$$\begin{aligned} &\pi_i^R = P(i \in S_R) && \pi_{ii'}^R = P(i, i' \in S_R) \\ &\pi_j^C = P(j \in S_C) && \pi_{jj'}^C = P(j, j' \in S_C) \\ &\Delta_{ii'}^R = \pi_{ii'}^R - \pi_i^R \pi_{i'}^R && \Delta_{jj'}^C = \pi_{jj'}^C - \pi_j^C \pi_{j'}^C \end{aligned} \tag{3.1}$$

where P denotes probability. Furthermore, let

$$\check{y}_{ij} = \frac{y_{ij}}{\pi_i^R \pi_j^C} \quad \check{y}_{i\cdot} = \frac{y_{i\cdot}}{\pi_i^R} \quad \check{y}_{\cdot j} = \frac{y_{\cdot j}}{\pi_j^C}$$

By the independence of S^R and S^C, the two-dimensional inclusion probabilities are just products of the one-dimensional ones. Hence, the Horvitz-Thompson estimator takes the following form in the case of cross-classified sampling

$$\hat{Y} = \sum_{i \in S^R}\sum_{j \in S^C} \check{y}_{ij} \tag{3.2}$$

The estimators in (1.2) take the form, in the notation of (1.5)

$$\hat{Y}^R = \sum_{i \in S^R} \check{y}_{i\cdot} \quad \hat{Y}^C = \sum_{j \in S^C} \check{y}_{\cdot j} \tag{3.3}$$

Note that the row estimator $\hat{Y}^R$ is simply the Horvitz-Thompson estimator of Y in case the columns are completely enumerated, and vice versa for $\hat{Y}^C$.

3.1. *Variance decomposition*

Next we specialize Theorem 1.1 to the Horvitz-Thompson case.

Theorem 3.1. The variance of the Horvitz-Thompson estimator in (3.2) is given by (1.3) and

$$VR = \sum_{i=1}^{N}\sum_{i'=1}^{N} \Delta_{ii'}^{R} \check{y}_{i\cdot}\check{y}_{i'\cdot} \tag{3.4}$$

$$VC = \sum_{j=1}^{M}\sum_{j'=1}^{M} \Delta_{jj'}^{C} \check{y}_{\cdot j}\check{y}_{\cdot j'} \tag{3.5}$$

$$VRC = \sum_{i=1}^{N}\sum_{i'=1}^{N}\sum_{j=1}^{M}\sum_{j'=1}^{M} \Delta_{ii'}^{R}\Delta_{jj'}^{C} \check{y}_{ij}\check{y}_{i'j'} \tag{3.6}$$

Note: In a typical application, exact or approximate formulas for one-dimensional variances are available and can be applied directly to the one-dimensional estimators in (3.3), yielding *VR* and *VC*. Hence, the practical use of Theorem 3.1 is mainly for computation of *VRC*.

Proof. The formulas for *VR* and *VC* follow from well-known results for the one-dimensional Horvitz-Thompson estimator, see e.g., Särndal et al. (1992, p. 43), and we turn to *VRC*. Let I_i^R and I_j^C be indicators of the events that $i \in S_R$ and $j \in S_C$, respectively. Then by (1.1), (1.5), (3.2) and (3.3)

$$\hat{Y} - \hat{Y}^R - \hat{Y}^C + Y = \sum_{i=1}^{N}\sum_{j=1}^{M} (I_i^R - \pi_i^R)(I_j^C - \pi_j^C)\check{y}_{ij} \tag{3.7}$$

Furthermore, since $\hat{Y}$, $\hat{Y}^R$ and $\hat{Y}^C$ all have expectation Y in this case, we have

$$VRC = V(\hat{Y} - \hat{Y}^R - \hat{Y}^C) = E[(\hat{Y} - \hat{Y}^R - \hat{Y}^C + Y)^2] \tag{3.8}$$

By inserting (3.7) into the right-hand side of (3.8) we find

$$VRC = E\left[\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{i'=1}^{N}\sum_{j'=1}^{M} (I_i^R - \pi_i^R)(I_j^C - \pi_j^C)(I_{i'}^R - \pi_{i'}^R)(I_{j'}^C - \pi_{j'}^C)\check{y}_{ij}\check{y}_{i'j'}\right] \tag{3.9}$$

Now take term-wise expectations in (3.9), use the independence of S^R and S^C and note that

$$E[(I_i^R - \pi_i^R)(I_{i'}^R - \pi_{i'}^R)] = \Delta_{ii'}^R \quad E[(I_j^C - \pi_j^C)(I_{j'}^C - \pi_{j'}^C)] = \Delta_{jj'}^C \tag{3.10}$$

This yields (3.6) and the proof is complete.

3.2. *Comparison with the direct approach*

Here we compare our approach to direct use of standard theory. The usual one-dimensional expression for the variance of a Horvitz-Thompson estimator (Särndal et al. 1992, p. 43) extends immediately to the two-dimensional case as follows

$$V(\hat{Y}) = \sum_{i=1}^{N}\sum_{i'=1}^{N}\sum_{j=1}^{M}\sum_{j'=1}^{M} \Delta_{(i,j),(i',j')}\check{y}_{ij}\check{y}_{i'j'} \tag{3.11}$$

where

$$\begin{aligned}\Delta_{(i,j),(i',j')} &= P((i,j) \in S; (i',j') \in S) - P((i,j) \in S) \cdot P((i',j') \in S) \\ &= \pi_{ii'}^{R}\pi_{jj'}^{C} - \pi_{i}^{R}\pi_{i'}^{R}\pi_{j}^{C}\pi_{j'}^{C}\end{aligned} \tag{3.12}$$

Since it may not be obvious that the approach of Theorem 3.1 is simpler than that of (3.11), we provide a simple example.

Example 3.1. Suppose both the row and column samples are unstratified Poisson samples. For a description of Poisson sampling see, e.g., Särndal et al. (1992, p. 85–87), from which we immediately get (recalling equations 3.3 and 1.4)

$$VR = \sum_{i=1}^{N}\left(\frac{1}{\pi_i^R} - 1\right)y_{i\cdot}^2 \quad VC = \sum_{j=1}^{M}\left(\frac{1}{\pi_j^C} - 1\right)y_{\cdot j}^2 \tag{3.13}$$

In this simple case we have

$$\pi_{ii'}^R = \pi_i^R\pi_{i'}^R \quad i \neq i' \qquad \pi_{jj'}^C = \pi_j^C\pi_{j'}^C \quad j \neq j' \tag{3.14}$$

so that

$$\Delta_{ii'}^R = \begin{cases} \pi_i^R(1-\pi_i^R) & i = i' \\ 0 & i \neq i' \end{cases} \qquad \Delta_{jj'}^C = \begin{cases} \pi_j^C(1-\pi_j^C) & j = j' \\ 0 & j \neq j' \end{cases} \tag{3.15}$$

From (3.6) and (3.15) we directly get

$$VRC = \sum_{i=1}^{N}\sum_{j=1}^{M}\left(\frac{1}{\pi_i^R} - 1\right)\left(\frac{1}{\pi_j^C} - 1\right)y_{ij}^2 \tag{3.16}$$

Putting (1.3), (3.13) and (3.16) together, we get

$$V(\hat{Y}) = \sum_{i=1}^{N}\left(\frac{1}{\pi_i^R} - 1\right)y_{i\cdot}^2 + \sum_{j=1}^{M}\left(\frac{1}{\pi_j^C} - 1\right)y_{\cdot j}^2 + \sum_{i=1}^{N}\sum_{j=1}^{M}\left(\frac{1}{\pi_i^R} - 1\right)\left(\frac{1}{\pi_j^C} - 1\right)y_{ij}^2 \tag{3.17}$$

We now turn to the direct approach using (3.11) and observe that, by (3.12) and (3.14)

$$\Delta_{(i,j).(i',j')} = \begin{cases} \pi_i^R\pi_j^C(1-\pi_i^R\pi_j^C) & i = i', j = j' \\ \pi_i^R(1-\pi_i^R)\pi_j^C\pi_{j'}^C & i = i', j \neq j' \\ \pi_j^C(1-\pi_j^C)\pi_i^R\pi_{i'}^R & i \neq i', j = j' \\ 0 & i \neq i', j \neq j' \end{cases}. \tag{3.18}$$

By inserting (3.18) into (3.11) we get

$$V(\hat{Y}) = \sum_i \sum_j \left(\frac{1}{\pi_i^R \pi_j^C} - 1 \right) y_{ij}^2 + \sum_i \sum_{j \neq j'} \left(\frac{1}{\pi_i^R} - 1 \right) y_{ij} y_{ij'}$$

$$+ \sum_{i \neq i'} \sum_j \left(\frac{1}{\pi_j^C} - 1 \right) y_{ij} y_{i'j} \tag{3.19}$$

After a good deal of algebra, (3.19) reduces to the expression in (3.17).

This example shows that already in the very simple case of Poisson sampling, our approach reduces the amount of algebra involved. We directly get the expression (3.17) which is obviously preferable to (3.19). With a more complex sampling procedure we can expect substantial gains by using Theorem 1.1 or its special case Theorem 3.1 instead of the direct approach. See Dalén and Ohlsson (1995) for such a case.

However, the most important advantage of our approach is in the case of Section 2, where rows and columns are separately stratified, yielding a cross-classification of cells which are *not* independent strata. With the direct approach we must use the usual general formula, recalling (2.2)

$$V(\hat{Y}) = V\left(\sum_g \sum_h \hat{Y}_{gh} \right) = \sum_g \sum_h \sum_{g'} \sum_{h'} C(\hat{Y}_{gh}, \hat{Y}_{g'h'}) \tag{3.20}$$

When $g \neq g'$, $h \neq h'$, the covariance is zero due to independence. We get the following structure

$$V(\hat{Y}) = \sum_g \sum_h V(\hat{Y}_{gh}) + \sum_g \sum_{h \neq h'} C(\hat{Y}_{gh}, \hat{Y}_{gh'}) + \sum_{g \neq g'} \sum_h C(\hat{Y}_{gh}, \hat{Y}_{g'h}) \tag{3.21}$$

In the next step the covariances must be calculated "from scratch" with sample membership indicators and inclusion probabilities. We do not pursue this approach further since our point is that this is unnecessarily complicated.

In contrast, by Theorem 2.1, we find the overall *VR*, *VC* and *VRC* simply by adding the corresponding within cell expressions so that no new calculations are involved compared to the unstratified case.

Furthermore, the variance structure of Theorem 2.1

$$V(\hat{Y}) = \sum_g VR_g + \sum_h VC_h + \sum_g \sum_h VRC_{gh} \tag{3.22}$$

is well-suited both for reporting and for sample allocation.

In the CPI case of Dalén and Ohlsson (1995), Theorems 1.1, 2.1 and 3.1 are applied to the CPI environment where we have stratified unequal probability sampling in both dimensions, using systematic πps selection (Särndal et al. 1992, p. 96) of products and sequential Poisson sampling (Ohlsson 1995a and 1995b) of outlets. Numerical variance estimates are reported as sums of "product variance," "outlet variance" and "outlet and product interaction variance." In this case, the use of (3.22) for allocation is simplified by the fact that the first two terms tend to dominate over the interaction term.

4. References

Central Statistical Office (CSO) (1995). Sampling Errors in the UK Retail Price Index. Paper No. CES/AC.49/54 presented by the United Kingdom Central Statistical Office at the Joint ECE/ILO meeting on consumer price indices, Geneva, November 20–24.

Cochran, R.S. (1978). Sampling in Two or More Dimensions. In David, H.A. (ed.): Contributions to Survey Sampling and Applied Statistics, New York: Academic Press, 113–119.

Dalén, J. and Ohlsson, E. (1995). Variance Estimation in the Swedish Consumer Price Index. Journal of Business & Economic Statistics, 13, 347–356.

Iachan, R. (1982). Systematic Sampling: A Critical Review. International Statistical Review 50, 293–303.

Quenouille, M.H. (1949). Problems in Plane Sampling. Annals of Mathematical Statistics 20, 335–375.

Ohlsson, E. (1995a). Coordination of Samples Using Permanent Random Numbers. In Business Survey Methods, eds Brenda Cox, David Binder, Nanjamma Chinnapa, Anders Christianson, Michael Colledge, and Phillip Kott, New York: John Wiley and Sons, 153–169.

Ohlsson, E. (1995b). Sequential Poisson Sampling. Research report No. 182, Institute of Actuarial Mathematics and Mathematical Statistics, Stockholm University.

Ripley, B.D. (1981). Spatial Statistics. New York: John Wiley and Sons.

Särndal, C.E., Swensson, B., and Wretman, J. (1992). Model Assisted Survey Sampling. New York: Springer Verlag.

Vos, J.W.E. (1964). Sampling in Space and Time. International Statistical Review, 32, 226–241.

Received September 1994
Revised October 1996