C. J. Skinner

Cross-classified sampling: some estimation theory

Article (Accepted version) (Refereed)

Original citation:
Skinner, C. J. (2015) *Cross-classified sampling: some estimation theory.* Statistics and Probability Letters. ISSN 0167-7152

DOI: 10.1016/j.spl.2015.06.001

Cross-classified Sampling: Some Estimation Theory

C. J. Skinner[a,]

[a]*Department of Statistics, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, U.K.*

Abstract

For a population represented as a two-way array, we consider sampling via the product of independent row and column samples. Theory is presented for the estimation of a population total under alternative methods of sampling the rows and columns.

Keywords: plane sampling, survey sampling, two dimensional sampling

1. Introduction

In some surveys the population can be represented by the elements of a two-way array, $(i, j), i = 1, \ldots, N, j = 1, \ldots, M$, and it is natural to take a sample as a Cartesian product $S = \{(i, j) : i \in S^R, j \in S^C\}$, where S^R and S^C are samples selected from the rows $\{1, \ldots, N\}$ and columns $\{1, ..., M\}$, respectively. A procedure in which S^R and S^C are selected independently by probability sampling schemes is called *cross-classified sampling*, following Ohlsson (1996).

A typical application of cross-classified sampling is to a survey of businesses which each handle a large number of products. Data is then collected from a sample of businesses on a sample of products.

We take the inferential objective to be to estimate the total of a variable y across units in the finite population, given only data on the values of y for units in the sample. In the business application, y might denote the value of the sale or purchase of a product in some time period and there may be interest in the total sales in the population. In some applications, there might be subunits, such as transactions, in which case y may be the sum across such subunits. This framework can also allow for cases where a

Email address: c.j.skinner@lse.ac.uk (C. J. Skinner)

product is not handled by a business in the time period by setting $y = 0$. Dalén and Ohlsson (1995) present an application of the use of cross-classified sampling in the construction of the Swedish Consumer Price Index.

Although cross-classified sampling can be treated within the general framework of finite population sampling, so that, for example, a Horvitz-Thompson estimator can be defined (Ohlsson, 1996), many specific aspects of this method are not covered by standard theory. For example, the latter typically treats sample quantities in different strata as independent, whereas we shall have to allow for dependence induced between strata in one dimension by sampling in the other dimension. Apart from the two key papers cited so far, the literature on the theory of cross-classified sampling is very limited. Vos (1964) provides some results for simple random sampling. There is a rather more extensive literature on the special case when the row and columns are ordered, typically in space, but possibly in time. This is usually called two-dimensional sampling or plane sampling. See e.g. Quenouille (1949), Bellhouse (1977), Iachan (1985) and Stevens and Olsen (2004). We shall, however, not assume an ordering of rows or columns and shall not refer further to this literature.

In this paper, we extend the theory in Ohlsson (1996) in a number of ways. First, we provide more explicit results on stratified sampling both from design-based and model-based perspectives. Second, we present results for with replacement unequal probability sampling. These results may be of interest in their own right, since it is common in business surveys for either businesses or the volume of product sales to vary considerably by size and for probability proportional to size sampling to be employed. However, in addition, we shall make use of the theory for with replacement sampling to construct bootstrap procedures for variance estimation. Such procedures may prove simpler to implement in practice than the more direct procedures we describe first.

2. Estimation for Simple Random and Stratified Sampling

We consider the estimation of the finite population total

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

where y_{ij} denotes the value of y for population unit (i, j) and is observed only for units $(i, j) \in S$. We consider two particular sampling schemes.

2.1. Simple Random Sampling

We first consider the prototypical case where S^R and S^C are selected by simple random sampling without replacement (SRSWOR). The natural unbiased estimator of Y here is the Horvitz-Thompson estimator given by

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

where n and m are the sizes of S^R and S^C respectively. The (design-based) variance of this estimator is now presented together with an unbiased estimator of this variance. The first of these results is given in Ohlsson (1996).

Theorem 2.1. *Under the SRSWOR design above, the estimator $\hat{Y}_{srs}$ is unbiased for Y, with variance*

$$var(\hat{Y}_{srs}) = N^2 M^2 \left\{ (1 - f_R) \frac{\sigma_R^2}{n} + (1 - f_C) \frac{\sigma_C^2}{m} + (1 - f_C)(1 - f_R) \frac{\sigma_{RC}^2}{nm} \right\}, \tag{1}$$

where

$$\sigma_R^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\bar{Y}_{i.} - \bar{Y}_{..})^2, \qquad \sigma_C^2 = \frac{1}{M-1} \sum_{j=1}^{M} (\bar{Y}_{.j} - \bar{Y}_{..})^2,$$

$$\sigma_{RC}^2 = \frac{1}{N-1} \frac{1}{M-1} \sum_{i=1}^{N} \sum_{j=1}^{M} (y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2,$$

$f_R = n/N$, $f_C = m/M$, $\bar{Y}_{i.} = \sum_{j=1}^{M} y_{ij}/M$, $\bar{Y}_{.j} = \sum_{i=1}^{N} y_{ij}/N$ *and* $\bar{Y}_{..} = \sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij}/NM$.

An unbiased estimator of $var(\hat{Y}_{srs})$ is obtained by replacing σ_R^2, σ_C^2 and σ_{RC}^2 in (1) by

$$\hat{\sigma}_R^2 = \frac{1}{n-1} \sum_{i \in S^R} (\bar{y}_{i.} - \bar{y}_{..})^2 - (1 - f_C) \frac{\hat{\sigma}_{RC}^2}{m},$$

$$\hat{\sigma}_C^2 = \frac{1}{m-1} \sum_{j \in S^C} (\bar{y}_{.j} - \bar{y}_{..})^2 - (1 - f_R) \frac{\hat{\sigma}_{RC}^2}{n},$$

$$\hat{\sigma}_{RC}^2 = \frac{1}{(n-1)(m-1)} \sum_{i \in S^R} \sum_{j \in S^C} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2,$$

where $\bar{y}_{i.} = \sum_{j\in S^C} y_{ij}/m$, $\bar{y}_{.j} = \sum_{i\in S^R} y_{ij}/n$ *and* $\bar{y}_{..} = \sum_{i\in S^R}\sum_{j\in S^C} y_{ij}/nm$.

Proof. The variance expression in (1) is given by Ohlsson (1996), with the basis of its proof indicated. The unbiasedness of $\hat{\sigma}^2_R$, $\hat{\sigma}^2_C$ and $\hat{\sigma}^2_{RC}$ may be shown by taking successive expectations with respect to the two sampling schemes and using standard results in sampling theory (Cochran, 1977, Theorem 2.4). □

2.2. Stratified Random Sampling

As noted in the introduction, variance expressions are less staightforward to obtain under stratification in cross-classified sampling than in standard theory. We now suppose that the rows and columns are stratified into G and H strata respectively and relabel the elements of the population as quadruples (g, h, i, j), where $g = 1, \dots, G, h = 1, \dots, H, i = 1, \dots, N_g$ and $j = 1, \dots, M_h$. The values N_g and M_h denote the stratum sizes, with $\sum N_g = N$ and $\sum M_h = M$. We suppose that samples S^R_g and S^C_h are selected independently by SRSWOR from the row and column strata respectively, with sizes n_g and m_h. The value of y for unit (g, h, i, j) is denoted y_{ghij} and the total of interest becomes

$$Y = \sum_{g=1}^{G}\sum_{h=1}^{H}\sum_{i=1}^{N_g}\sum_{j=1}^{M_h} y_{ghij}.$$

The natural unbiased estimator of Y is again the Horvitz Thompson estimator, which may be expressed as

$$\hat{Y}_{str} = \sum_{g=1}^{G}\sum_{h=1}^{H}\frac{N_g M_h}{n_g m_h}\sum_{i\in S^R_g}\sum_{j\in S^C_h} y_{ghij}. \qquad (2)$$

We use notation $\bar{Y}_{g.i.}, \bar{Y}_{g...}$, etc. to denote population means over the 'dot' subscript so that, for example, $\bar{Y}_{g.i.} = \sum_{h=1}^{H}\sum_{j=1}^{M_h} y_{ghij}/M$. We similarly use notation $\bar{y}_{g.i.}, \bar{y}_{g...}$ etc. to denote sample means, subject to weighting means across strata g and h by N_g/n_g and M_h/m_h, as in (2). Thus, $\bar{y}_{ghi.} = m_h^{-1}\sum_{j\in S^C_h} y_{ghij}$, whereas $\bar{y}_{g.i.} = M^{-1}\sum_{h=1}^{H}\sum_{j\in S^C_h}(M_h/m_h)y_{ghij}$. The stratum sampling fractions are denoted $f_{Rg} = n_g/N_g$ and $f_{Ch} = m_h/M_h$. We now present the (design-based) variance of this estimator and an unbiased estimator of this variance.

Theorem 2.2. *Under the stratified random sampling design above, $\hat{Y}_{str}$ is unbiased for Y, with variance*

$$var(\hat{Y}_{str}) = \sum_{g=1}^{G} N_g^2 M^2 (1 - f_{Rg}) \frac{\sigma_{Rg}^2}{n_g} + \sum_{h=1}^{H} N^2 M_h^2 (1 - f_{Ch}) \frac{\sigma_{Ch}^2}{m_h} + \sum_{g=1}^{G} \sum_{h=1}^{H} N_g^2 M_h^2 (1 - f_{Rg}) (1 - f_{Ch}) \frac{\sigma_{RCgh}^2}{n_g m_h}, \quad (3)$$

where

$$\sigma_{Rg}^2 = \frac{1}{N_g - 1} \sum_{i=1}^{N_g} (\bar{Y}_{g.i.} - \bar{Y}_{g...})^2, \qquad \sigma_{Ch}^2 = \frac{1}{M_h - 1} \sum_{j=1}^{M_h} (\bar{Y}_{.h.j} - \bar{Y}_{.h..})^2,$$

$$\sigma_{RCgh}^2 = \frac{1}{N_g - 1} \frac{1}{M_h - 1} \sum_{i=1}^{N_g} \sum_{j=1}^{M_h} (y_{ghij} - \bar{Y}_{ghi.} - \bar{Y}_{gh.j} + \bar{Y}_{gh..})^2.$$

An unbiased estimator of $var(\hat{Y}_{str})$ is given by

$$\hat{V}(\hat{Y}_{str}) = \sum_{g=1}^{G} N_g^2 M^2 (1 - f_{Rg}) \frac{\hat{\sigma}_{Rg}^2}{n_g} + \sum_{h=1}^{H} N^2 M_h^2 (1 - f_{Ch}) \frac{\hat{\sigma}_{Ch}^2}{m_h} + \sum_{g=1}^{G} \sum_{h=1}^{H} N_g^2 M_h^2 (1 - f_{Rg}) (1 - f_{Ch}) \frac{\hat{\sigma}_{RCgh}^2}{n_g m_h}, \quad (4)$$

where

$$\hat{\sigma}_{Rg}^2 = \frac{1}{n_g - 1} \sum_{i \in S_g^R} (\bar{y}_{g.i.} - \bar{y}_{g...})^2 - \sum_{h=1}^{H} \frac{M_h^2}{M^2 m_h} (1 - f_{Ch}) \hat{\sigma}_{RCgh}^2, \quad (5)$$

$$\hat{\sigma}_{Ch}^2 = \frac{1}{m_h - 1} \sum_{j \in S_h^C} (\bar{y}_{.h.j} - \bar{y}_{.h..})^2 - \sum_{g=1}^{G} \frac{N_g^2}{N^2 n_g} (1 - f_{Rg}) \hat{\sigma}_{RCgh}^2, \quad (6)$$

$$\hat{\sigma}_{RCgh}^2 = \frac{1}{(n_g - 1)(m_h - 1)} \sum_{i \in S_g^R} \sum_{j \in S_h^C} (y_{ghij} - \bar{y}_{ghi.} - \bar{y}_{gh.j} + \bar{y}_{gh..})^2. \quad (7)$$

Proof. The unbiasedness of $\hat{Y}_{str}$ follows straightforwardly, as in Theorem 2.1. The variance follows by evaluating moments of $\sum_{i\in S_g^R}\sum_{j\in S_h^C} y_{ghij}$ with respect to each of the row and column sampling schemes separately, combining these moments within general expressions provided by Ohlsson (1996, Theorem 2.1) and then evaluating further moments as necessary with respect to the other dimension of sampling. The unbiasedness of $\hat{V}(\hat{Y}_{str})$ is shown by evaluating expectations successively with respect to the two dimensions of sampling.

□

3. Model-based Variances

We now consider a model-based analogue of the results in the previous section. This may serve various purposes, in particular it may enable established software to be employed for estimation. A standard model for values y_{gi} in a singly stratified population with strata $g = 1, \ldots, G$ and units $i = 1, \ldots, N_g$ within stratum g is $y_{gi} = \mu_g + e_{gi}$, where the e_{gi} are uncorrelated with $e_{gi} \sim (0, \tilde{\sigma}_g^2)$ (i.e. with mean 0 and variance $\tilde{\sigma}_g^2$) and the μ_g and $\tilde{\sigma}_g^2$ are fixed parameters (Valliant et al., 2000, Example 2.3.3).We propose to extend this model to a stratified cross-classified population by writing:

$$y_{ghij} = \mu_{gh} + u_{gi} + v_{hj} + e_{ghij}, \tag{8}$$

where the u_{gi}, v_{hj} and e_{ghij} are uncorrelated random effects, $u_{gi} \sim (0, \tilde{\sigma}_{Rg}^2)$, $v_{hj} \sim (0, \tilde{\sigma}_{Ch}^2)$ and $e_{ghij} \sim (0, \tilde{\sigma}_{RCgh}^2)$ and the μ_{gh}, $\tilde{\sigma}_{Rg}^2$, $\tilde{\sigma}_{Ch}^2$ and $\tilde{\sigma}_{RCgh}^2$ are fixed parameters. We denote expectation under this model by E_ξ.

Theorem 3.1. *Under model (8), the estimator $\hat{Y}_{str}$ is unbiased for Y, in the sense that $E_\xi(\hat{Y}_{str} - Y) = 0$, and $E_\xi(\hat{Y}_{str} - Y)^2$ takes the same form as $var(\hat{Y}_{str})$ in (3) with $\sigma_{Rg}^2, \sigma_{Ch}^2$ and σ_{RCgh}^2 replaced by $\tilde{\sigma}_{Rg}^2$, $\tilde{\sigma}_{Ch}^2$ and $\tilde{\sigma}_{RCgh}^2$, respectively.*

Proof. This follows by replacing y_{ghij} in $\hat{Y}_{str} - Y$ by the right hand side of (8), noting that μ_{gh} drops out of this expression, and evaluating the model expectations of this function of the random effects and its square.

□

We can also show that the estimators $\hat{\sigma}_{Rg}^2$, $\hat{\sigma}_{Ch}^2$ and $\hat{\sigma}_{RCgh}^2$ in (5),(6) and(7) are unbiased for $\tilde{\sigma}_{Rg}^2$, $\tilde{\sigma}_{Ch}^2$ and $\tilde{\sigma}_{RCgh}^2$ under the model. These estimators are referred to as ANOVA estimators in the mixed model literature Searle et al. (1992, sect. 4.4). This literature also offers alternative estimation methods, including maximum likelihood under the assumption that the random effects are normally distributed (Searle et al., 1992, Goldstein, 2011) and a number of established software packages may be used to implement such methods.

4. With Replacement Sampling

We now suppose that each of S^R and S^C is selected by with replacement (WR) sampling with possibly unequal probabilities. Thus S^R is obtained by n independent draws from $\{1, \ldots, N\}$, at each of which unit i is drawn with probability p_i. Similarly, S^C is obtained by m independent draws from $\{1, \ldots, M\}$, at each of which unit j is drawn with probability q_j. Thus, $\sum_1^N p_i = \sum_1^M q_j = 1$.

The natural unbiased estimator of Y, extending the Hansen-Hurwitz estimator (Berger and Tillé, 2009), is

$$\hat{Y}_{wr} = \frac{1}{nm} \sum_{k=1}^{n} \sum_{l=1}^{m} \frac{y_{i(k)j(l)}}{p_{i(k)} q_{j(l)}}, \tag{9}$$

where $i(k)$ and $j(l)$ denote the units in $\{1, \ldots, N\}$ and $\{1, \ldots, M\}$ selected on the k^{th} and l^{th} draws, respectively. The (design-based) variance of this estimator and an unbiased estimator of this variance are given in the following theorem. We see that results are analogous to those in Theorem 2.1, where the finite population corrections disappear and variances and their estimators are weighted by the factors $1/(Np_i)$ and $1/(Mq_j)$, which each reduce to 1 in the equal probability case.

Theorem 4.1. *Under the WR design above, the estimator $\hat{Y}_{wr}$ is unbiased for Y, with variance*

$$var(\hat{Y}_{wr}) = N^2 M^2 \left\{ \frac{\sigma^2_{Rwr}}{n} + \frac{\sigma^2_{Cwr}}{m} + \frac{\sigma^2_{RCwr}}{nm} \right\}, \tag{10}$$

where

$$\sigma^2_{Rwr} = \sum_{i=1}^{N} \left(\frac{\bar{Y}_{i.}}{Np_i} - \bar{Y}_{..} \right)^2 p_i \ , \qquad \sigma^2_{Cwr} = \sum_{j=1}^{M} \left(\frac{\bar{Y}_{.j}}{Mq_j} - \bar{Y}_{..} \right)^2 q_j,$$

$$\sigma^2_{RCwr} = \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\frac{y_{ij}}{NMp_i q_j} - \frac{\bar{Y}_{i.}}{Np_i} - \frac{\bar{Y}_{.j}}{Mq_j} + \bar{Y}_{..} \right)^2 p_i q_j,$$

and $\bar{Y}_{i.}$, $\bar{Y}_{.j}$ and $\bar{Y}_{..}$ are defined in Theorem 2.1. Unbiased estimators of σ^2_{Rwr}, σ^2_{Cwr} and σ^2_{RCwr} are given by

$$\begin{aligned}
\hat{\sigma}^2_{Rwr} &= \frac{1}{n-1}\sum_{k=1}^{n}\left(\frac{\bar{y}_{i(k).wr}}{Np_{i(k)}} - \bar{y}_{..wr}\right)^2 - \frac{\hat{\sigma}^2_{RCwr}}{m}, \\
\hat{\sigma}^2_{Cwr} &= \frac{1}{m-1}\sum_{l=1}^{m}\left(\frac{\bar{y}_{.j(l)wr}}{Mq_{j(l)}} - \bar{y}_{..wr}\right)^2 - \frac{\hat{\sigma}^2_{RCwr}}{n}, \\
\hat{\sigma}^2_{RCwr} &= \frac{1}{(n-1)(m-1)}\sum_{k=1}^{n}\sum_{l=1}^{m}\left(\frac{y_{i(k)j(l)}}{NMp_{i(k)}q_{j(l)}} - \frac{\bar{y}_{i(k).wr}}{Np_{i(k)}} - \frac{\bar{y}_{.j(l)wr}}{Mq_{j(l)}} + \bar{y}_{..wr}\right)^2,
\end{aligned}$$

respectively, where $\bar{y}_{i(k).wr} = m^{-1}\sum_{l=1}^{m} y_{i(k)j(l)}/(Mq_{j(l)})$ for $k = 1,\ldots,n$, $\bar{y}_{.j(l)wr} = n^{-1}\sum_{k=1}^{n} y_{i(k)j(l)}/(Np_{i(k)})$ for $l = 1,\ldots,m$ and $\bar{y}_{..wr} = (MN)^{-1}\hat{Y}_{wr}$.

Proof. The unbiasedness of $\hat{Y}_{wr}$ follows by first evaluating its expectation conditional on S^R, giving

$$E^R(\hat{Y}_{wr}) = \frac{M}{n}\sum_{k=1}^{n}\frac{\bar{Y}_{i(k).}}{p_{i(k)}}.$$

That the expectation of this quantity is Y follows as in the proof of the unbiasedness of the usual Hansen-Hurwitz estimator. The variance of $\hat{Y}_{wr}$ is similarly obtained in two steps, where, for example, the result of the first step, conditional on S^R, is:

$$E^R(\hat{Y}_{wr} - Y)^2 = \left[\left(\frac{M}{n}\sum_{k=1}^{n}\frac{\bar{Y}_{i(k).}}{p_{i(k)}}\right) - Y\right]^2 + \frac{1}{m}\sum_{j=1}^{M}\left[\frac{1}{n}\sum_{k=1}^{n}\left(\frac{y_{i(k)j}}{p_{i(k)}q_j} - \frac{M\bar{Y}_{i(k).}}{p_{i(k)}}\right)\right]^2 q_j.$$

The unbiasedness of $\hat{\sigma}^2_{Rwr}$, $\hat{\sigma}^2_{Cwr}$ and $\hat{\sigma}^2_{RCwr}$ follow similarly. □

5. Bootstrap Variance Estimation

Possible benefits of bootstrap variance estimation are (i) that it reduces the complexity of the programming requirements of the previous variance estimators in section 2 (and hence the potential for coding error in practice) and (ii) that it is applicable to a broader class of point estimators. We just consider the case of stratified random sampling as in section 2.2.

We propose the following algorithm for generating a single bootstrap replicate from a stratified random sample.

1. draw independent WR samples with equal probabilities ($p_i = 1/n_g$) of size n_g^* from S_g^R, $g = 1, \ldots, G$.
2. draw independent WR samples with equal probabilities ($q_j = 1/m_h$) of size m_h^* from S_h^C, $h = 1, \ldots, H$.

Let $\hat{Y}_{str}^*$ take the same form as $\hat{Y}_{str}$ in (2) but where the sums are across units selected in the bootstrap replicate, allowing for duplication, as in (9). We define the bootstrap estimator of the variance of $\hat{Y}_{str}$ as

$$\hat{V}_{boot}(\hat{Y}_{str}) = var_*(\hat{Y}_{str}^*),$$

where var_* denotes variance across repeated independent bootstrap replicates, as in Wolter (2007, Chapter 5). In practice, only a finite number of bootstrap replicates will be drawn. Some initial empirical work suggests that 200 replicates is sufficient to make the resulting bootstrap variance estimator an adequate approximation in practice to the 'infinite replicate' bootstrap variance estimator (c.f. Wolter, 2007, p. 195) but I have not undertaken any systematic assessment of the number of replicates needed and I shall ignore this approximation henceforth for simplicity.

The following result shows that the bootstrap variance estimator becomes equal to the estimator in Theorem 2.2, if finite population correction (fpc) terms $1 - f_{Rg}$ and $1 - f_{Ch}$ and the bias correction terms in $\hat{\sigma}_{Rg}^2$ and $\hat{\sigma}_{Ch}^2$ are removed. The fpc terms become negligible when the sampling fractions n_g/N_g and m_h/M_h are small. The bias correction terms may be expected to make smaller relative contributions as the stratum sample sizes n_g and m_h increase.

Theorem 5.1. *The bootstrap variance estimator with $n_g^* = n_g - 1$ and $m_h^* = m_h - 1$ may be expressed as*

$$\hat{V}_{boot} = \sum_{g=1}^{G} N_g^2 M^2 \frac{s_{Rg}^2}{n_g} + \sum_{h=1}^{H} N^2 M_h^2 \frac{s_{Ch}^2}{m_h} + \sum_{g=1}^{G} \sum_{h=1}^{H} N_g^2 M_h^2 \frac{\hat{\sigma}_{RCgh}^2}{n_g m_h}, \qquad (11)$$

where

$$s_{Rg}^2 = \frac{1}{n_g - 1} \sum_{i=1}^{n_g} (\bar{y}_{g.i.} - \bar{y}_{g...})^2, \quad s_{Ch}^2 = \frac{1}{m_h - 1} \sum_{j=1}^{m_h} (\bar{y}_{.h.j} - \bar{y}_{.h..})^2. \qquad (12)$$

Proof. The proof follows as in the with replacement results by taking expectations in turn by the row and column bootstrap sampling. Letting E_*^R denote expectation over the column bootstrap sampling, we find that $E_*^R(\hat{Y}_{str}^*) = \sum_h (M_h/m_h) \sum_{j=1}^{m_h} A_{hj}$, where $A_{hj} = \sum_g (N_g/n_g^*) \sum_{k=1}^{n_g^*} y_{ghi(k)j}$. It follows that $E_*\left(\hat{Y}_{str}^*\right) = \hat{Y}_{str}$. We also obtain $var_*^R(\hat{Y}_{str}^*) = \sum_h (M_h^2 \sum_{j=1}^{m_h} (A_{hj} - \bar{A}_h)^2/(m_h m_h^*)$, where $\bar{A}_h = \sum_{j=1}^{m_h} A_{hj}/m_h$. The result follows by taking further expectations of the relevant terms over the row bootstrap sampling. □

6. Conclusion

The results in this paper may be useful for both design and estimation in surveys employing cross-classified sampling. Design issues may arise, for example, when sample sizes need to be determined for both rows and columns, under a joint resource constraint, and the relative contributions of row and column sampling to the variance need to be assessed.

References

Bellhouse, D., 1977. Optimal designs for sampling in two dimensions. Biometrika 64, 605–611.

Berger, Y., Tillé, Y., 2009. Sampling with unequal probabilities. In Sample Surveys: Design, Methods and Applications, Vol. 29A of Handbook of Statistics, eds. Pfeffermann, D. and Rao, C.R., Elsevier. chapter 2. pp. 39–54.

Cochran, W., 1977. Sampling Techniques. Wiley, New York. 3rd edition.

Dalén, J., Ohlsson, E., 1995. Variance estimation in the swedish consumer price index. J. Bus. Econ. Statist. 13, 347–356.

Goldstein, H., 2011. Multilevel Statistical Models, 4th edition. Wiley, Chichester.

Iachan, R., 1985. Plane sampling. Statist. Probab. Lett. 3, 151–159.

Ohlsson, E., 1996. Cross-classified sampling. J. Official Statist. 88, 241–251.

Quenouille, M., 1949. Problems in plane sampling. Ann. Math. Statist. 20, 335–375.

Searle, S., Casella, G., McCullagh, C., 1992. Variance Components. Wiley, New York.

Stevens, D., Olsen, A., 2004. Spatially balanced sampling of natural resources. J. Amer. Statist. Assoc. 99, 262–278.

Valliant, R., Dorfman, A., Royall, R., 2000. Finite Population Sampling and Inference. Wiley, New York.

Vos, J., 1964. Sampling in space and time. Int. Statist. Rev. 32, 226–241.

Wolter, K., 2007. Introduction to Variance Estimation. Springer, New York. 2nd edition.