Survey Methodology, December 2002
Vol. 28, No. 2, pp. 209-212
Statistics Canada, Catalogue No. 12-001-XPB

On the Use of Generalized Inverse Matrices in Sampling Theory

Robbert H. Renssen and Gerard H. Martinus [1]

Abstract

In theory, it is customary to define general regression estimators in terms of full-rank weighting models, *i.e.*, the design matrix that corresponds to the weighting model is of full rank. For such weighting models, it is well known that the general regression weights reproduce the (known) population totals of the auxiliary variables involved. In practice, however, the weighting model often is not of full rank, especially when the weighting model is for incomplete post-stratification. By means of the theory of generalized inverse matrices, it is shown under which circumstances this consistency property remains valid. As a non-trivial example we discuss the consistent weighting between persons and households as proposed by Lemaître and Dufour (1987). We then show how the theory is implemented in Bascula.

Key Words: Bascula; General regression estimator; Weighting.

1. Introduction

Weighting methods that are based on the general regression estimator are commonly used in sample surveys to adjust for both sampling error and non-sampling error, see *e.g.*, Bethlehem and Keller (1987) and Särndal, Swensson, and Wretman (1992). One complication in the use of general regression estimators, however, is that many weighting models are based on incomplete post-stratification, resulting in design matrices that are not of full rank. Usually, this problem is solved by using a reduced design matrix. Such a reduced design matrix can be constructed by deleting redundant columns and properly adjusting the population totals. Often, the redundancy can be recognized rather easily beforehand by the specification of the weighting model. However, for some weighting models such a redundancy check may be impractical.

For example, suppose that we have a post-stratification based on the complete crossing between two categorical variables *A* and *B*, with known counts for the population of each cell. We may obtain small sample counts or no sample in some cells. Then we may derive new classifications, A' from *A* and B' from *B*, by merging categories, and define the following more parsimonious scheme: $A + B + A' \times B'$. According to this incomplete post-stratification we simultaneously calibrate on three sets of counts, namely the marginal counts of *A*, the marginal counts of *B*, and the cell counts of $A' \times B'$. Since *A* and A' (and also *B* and B') appear in different weighting terms, it is difficult to recognize redundancy by the specification of the weighting model. This paper gives the theoretical background, which is based on generalized inverse matrices, of reducing such a design matrix.

In section 2 we briefly describe some properties of generalized inverse matrices. In section 3 we define the general regression estimator for weighting models that need not be of full rank. Given a regularity condition that can be nicely interpreted in a calibration estimation context (see Deville and Särndal 1992) it is shown that this estimator is invariant with respect to the choice of the generalized inverse. At the end of section 3 the fulfillment of this regularity condition is discussed for some well-known weighting models, such as incomplete post-stratification and consistent weighting between persons and households. In section 4 we describe the algorithm, which is implemented in Bascula (see Nieuwenbroek 1997; Renssen, Nieuwenbroek and Slootbeek 1997) for calculating the regression weights. Finally, in section 5 we briefly discuss the weighting model of the Dutch Labour Force Survey.

2. Generalized Inverse Matrices

We are mainly interested in the use of generalized inverses within the framework of the general regression estimator. Hence, we only give some properties of a generalized inverse of the form $\mathbf{X}^t \Lambda \mathbf{X}$, where Λ is a diagonal matrix of order $n \times n$ with strictly positive diagonal entries and $\mathbf{X}$ a design matrix of order $n \times p$ that results from the weighting model. For a more extensive discussion on generalized inverse matrices we refer to Searle (1971) and Rao (1973).

Before giving these properties, we briefly review the definition of a generalized inverse. Consider a $p \times q$ matrix $\boldsymbol{A}$ of any rank and let $\boldsymbol{Ax} = \boldsymbol{y}$ be a system of consistent equations, *i.e.*, any linear relationship existing among the rows of $\boldsymbol{A}$ also exists among the corresponding elements of $\boldsymbol{y}$. A generalized inverse of $\boldsymbol{A}$ is a $q \times p$ matrix $\boldsymbol{A}^-$ such that $\boldsymbol{x} = \boldsymbol{A}^- \boldsymbol{y}$ is a solution of this system of equations. It is easy to verify that the existence of $\boldsymbol{A}^-$ implies $\boldsymbol{AA}^-\boldsymbol{A} = \boldsymbol{A}$ (choose $\boldsymbol{y}$ as the i^{th} column of $\boldsymbol{A}$). Conversely, if $\boldsymbol{A}^-$ satisfies $\boldsymbol{AA}^-\boldsymbol{A} = \boldsymbol{A}$ and $\boldsymbol{Ax} = \boldsymbol{y}$ is consistent, then $\boldsymbol{A}(\boldsymbol{A}^-\boldsymbol{y}) = \boldsymbol{A}(\boldsymbol{A}^-\boldsymbol{Ax}) = \boldsymbol{Ax} = \boldsymbol{y}$ and hence $\boldsymbol{A}^-\boldsymbol{y}$ is a solution. Thus, as an alternative definition, a generalized

1. Robbert H. Renssen and Gerard H. Martinus, Department of Statistical Methods, Statistics Netherlands, P.O. Box 4481, 6401 CZ Heerlen, The Netherlands.

inverse matrix of $\boldsymbol{A}$ is any matrix $\boldsymbol{A}^-$ such that $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}$.

Now, if $\mathbf{G}$ denotes a generalized inverse of $\boldsymbol{X}^t\Lambda\boldsymbol{X}$, then the following properties of $\mathbf{G}$ are proven in Searle (1971) for $\Lambda = \mathbf{I}_n$:

(P1) $\mathbf{G}^t$ is also a generalized inverse of $\mathbf{X}^t\Lambda\mathbf{X}$,

(P2) $\mathbf{XGX}^t\Lambda\mathbf{X} = \mathbf{X}$ *i.e.*, $\mathbf{GX}^t\Lambda$ is a generalized inverse of $\mathbf{X}$,

(P3) $\mathbf{XGX}^t$ is invariant to the choice of $\mathbf{G}$,

(P4) $\mathbf{XGX}^t = \mathbf{XG}^t\mathbf{X}^t$ whether $\mathbf{G}$ is symmetric or not.

The proofs of (P1) to (P4) for diagonal are almost identical to those of Searle (1971, chapter 1.5, theorem 7) and therefore not repeated here.

3. The General Regression Estimator

Consider a finite population U of N units from which a sample S of n units is drawn without replacement. Let π_k denote the first order inclusion probability of the k^{th} unit, $k = 1, ..., N$. We associate with each unit a vector of study variables $\mathbf{y}_k$. Then, the data matrix for the sampled units is given by $\boldsymbol{Y}_S = (\boldsymbol{y}_1, ..., \boldsymbol{y}_n)^t$. We distinguish between study variables with known population totals (auxiliary variables) and study variables with unknown population totals. The start in the definition of a general regression estimator (Särndal *et al.* 1992) is the specification of the weighting model, *i.e.*, the choice of the set of auxiliary variables to be used in the estimation. Denoting this specific set of p variables by $\mathbf{x}$, we call the $n \times p$ matrix $\mathbf{X}_S = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)^t$ the design matrix, which is, by definition, a column subset of $\mathbf{Y}_S$. The vector of known population totals of $\mathbf{x}$ is denoted by $\mathbf{t}_x$. Let $\mathbf{x}_{\text{HT}} = \sum_{k\in S}\pi_k^{-1}\mathbf{x}_k$ denote the Horvitz- Thompson estimator for $\mathbf{t}_x$, then, given $\mathbf{x}$, the general regression estimator of the vector of population totals of the i^{th} study variable $\mathbf{y}_k^{(i)}$ is defined as

$$\hat{\mathbf{t}}_{\text{greg}}^{(i)} = \mathbf{y}_{\text{HT}}^{(i)} + \hat{\mathbf{B}}^t(\mathbf{t}_x - \mathbf{x}_{\text{HT}}) \qquad (1)$$

with

$\hat{\mathbf{B}} = \mathbf{G}_S\mathbf{X}_S^t\Lambda_S\mathbf{Y}_S^{(i)}$.

In terms of regression weights, this general regression estimator can also be written as

$$\hat{\mathbf{t}}_{\text{greg}}^{(i)} = \sum_{k\in S} w_k\mathbf{y}_k^{(i)} \qquad (2)$$

with

$$w_k = \pi_k^{-1} + \lambda_k\mathbf{x}_k^t\mathbf{G}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}}).$$

Here, $\mathbf{G}_S$ denotes a generalized inverse of $\mathbf{X}_S^t\Lambda_S\mathbf{X}_S$ and $\Lambda_S = \text{diag}(\lambda_1, ..., \lambda_n)$ is some diagonal matrix with strictly positive entries.

Like the weighting model, the diagonal matrix Λ_S has to be specified by the user. Often, one takes $\Lambda_S = \Pi_S^{-1}\Sigma_S^{-1}$, where $\Pi_S = \text{diag}(\pi_1, ..., \pi_n)$ and $\Sigma_S = \text{diag}(\sigma_1^2, ..., \sigma_n^2)$ with σ_k^2 interpreted as the variance of independent random variables of which some of the study variables are supposed to be the outcome according to some super-population model, see Särndal *et al.* (1992). It is required that all σ_k^2 be known up to a common scale factor. An important special case is $\sigma_k^2 = \sigma^2$, *i.e.*, all the modeled variances are the same. This results in the regression estimator proposed by Bethlehem and Keller (1987). If the population units represent households (of size m_k) and if we take $\sigma_k^2 = m_k\sigma^2$ we arrive at the estimator proposed by Lemaître and Dufour (1987) to obtain consistent weights between person and households. From a different point of view, Alexander (1987) derived the GLS-P estimate, which results in essentially the same estimator.

Below we show that the regression weights are invariant to the choice of $\boldsymbol{G}_S$. To that purpose we make the following assumption:

(A1) there exists a n-vector $\mathbf{w}$ such that $\mathbf{X}_S^t\mathbf{w} = \mathbf{t}_x$.

Clearly, this assumption states that $\mathbf{X}_S^t\mathbf{w} = \mathbf{t}_x$ is a system of consistent equations. It is interesting to note that this system precisely corresponds to the set of calibrations equations when considering the general regression estimator as a special case of the calibration estimator (see *e.g.* Deville and Särndal 1992). If $\mathbf{X}_S^t\mathbf{w} = \mathbf{t}_x$ is a system of consistent equations, then so is $\mathbf{X}_S^t\mathbf{v} = (\mathbf{t}_x - \mathbf{x}_{\text{HT}})$. This is easily seen by taking $\mathbf{v} = \mathbf{w} - \mathbf{d}_S$ with $\mathbf{d}_S = (\pi_1^{-1}, ..., \pi_n^{-1})^t$. The invariance of the regression weights to the choice of $\mathbf{G}_S$, and hence the invariance of the general regression estimator can be shown as follows. Let $\mathbf{F}_S$ be some other generalized inverse of $\mathbf{X}_S^t\Lambda_S\mathbf{X}_S$, different from $\mathbf{G}_S$. Then, we have

$$\begin{aligned}
\mathbf{X}_S\mathbf{G}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}}) &= \mathbf{X}_S\mathbf{G}_S\mathbf{X}_S^t\mathbf{v} && \text{by (A1)}\\
&= \mathbf{X}_S\mathbf{F}_S\mathbf{X}_S^t\mathbf{v} && \text{by (P3)}\\
&= \mathbf{X}_S\mathbf{F}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}}) && \text{by (A1).}
\end{aligned}$$

So, it holds that $\mathbf{x}_k^t\mathbf{G}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}})$ is invariant to $\mathbf{G}_S$ for all $k \in S$, implying that the regression weights are invariant to the choice $\mathbf{G}_S$.

The fact that these weights reproduce the population totals of the auxiliary variables follows from the following series of equations:

$$\begin{aligned}
\sum_{k\in S} w_k\mathbf{x}_k &= \mathbf{x}_{\text{HT}} + \sum_{k\in S}\mathbf{x}_k\lambda_k\mathbf{x}_k^t\mathbf{G}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}})\\
&= \mathbf{x}_{\text{HT}} + \left(\mathbf{X}_S^t\Lambda_S\mathbf{X}_S\right)\mathbf{G}_S(\mathbf{t}_x - \mathbf{x}_{\text{HT}})\\
&= \mathbf{x}_{\text{HT}} + \left(\mathbf{X}_S^t\Lambda_S\mathbf{X}_S\right)\mathbf{G}_S\mathbf{X}_S^t\mathbf{v} \quad \text{by (A1)}\\
&= \mathbf{x}_{\text{HT}} + \mathbf{X}_S^t\mathbf{v} && \text{by (P2) and (P4)}\\
&= \mathbf{x}_{\text{HT}} + (\mathbf{t}_x - \mathbf{x}_{\text{HT}}) = \mathbf{t}_x && \text{by (A1).}
\end{aligned}$$

We close this section by having a closer look at the stated assumption for some well-known weighting models. In case of post-stratification in which the weighting model is described by a complete crossing of categorical variables, (A1) has a simple interpretation. Namely (A1) is satisfied if and only if empty post-strata in the sample correspond to empty post-strata in the population. Next, we consider incomplete post-stratification in which the weighting model consists of several terms, each term describing a complete crossing of categorical variables and so each term corresponding to a post-stratification. Then, a necessary condition for (A1) to be satisfied is that empty post-strata in the sample correspond to empty post-strata in the population for each of these terms. Unfortunately, this condition is not sufficient. For example, inconsistencies may still occur when we attempt to calibrate on a number of complete crossings larger than the sample size.

The assumption is less straightforward in case of consistent weighting between persons and households (see *e.g.*, Lemaître and Dufour 1987). This is due to the redefinition of the auxiliary variable. For example, if $\mathbf{x}_k$ is a variable defined at the person level, and from this variable a new variable is defined on the household level, say $\mathbf{z}_k$, then (A1) should be defined in terms of $\mathbf{Z}_S = (\mathbf{z}_1, ..., \mathbf{z}_n)^t$ instead of $\mathbf{X}_S$, *i.e.*, (A1) is satisfied if there exists an n-vector $\mathbf{w}$ such that $\mathbf{Z}_S^t \mathbf{w} = \mathbf{t}_x$. In many (regular) situations, the linear manifold spanned by $\mathbf{Z}_S$ will coincide with the linear manifold spanned by $\mathbf{X}_S$. In such situations the method of Lemaître and Dufour does not affect the validity of (A1). However, in specific cases this may not be true. The following simplified example illustrates this.

Let $\mathbf{x}_k$ denote sex of the k^{th} person, say $\mathbf{x}_k = (0, 1)^t$ if the k^{th} person is a female and $\mathbf{x}_k = (1, 0)^t$ if the k^{th} is a male. According to the method of Lemaître and Dufour (1987), let $\mathbf{z}_k$ denote the j^{th} household mean for $\mathbf{x}_k$ whenever k belongs to the j^{th} household. Furthermore, let the population consists of N_1 males and N_2 females, from which a sample of 10 households is drawn. Suppose that each sampled household consists of two persons, namely one male and one female. This gives $\mathbf{z}_k = (1/2, 1/2)^t$ for all $k \in S$. For this example the linear manifold spanned by $\mathbf{Z}_S$ is a linear subspace of the linear manifold spanned by $\mathbf{X}_S$. If $N_1 = N_2$ then (A1) is satisfied. Otherwise, if $N_1 \neq N_2$ then (A1) is not satisfied. Especially, when the method of Lemaître and Dufour is applied on a relatively large weighting model, the linear manifold spanned by $\mathbf{Z}_S$ may be a proper subspace of the linear manifold spanned by $\mathbf{X}_S$. Then, (A1) only is satisfied if $\mathbf{t}_x$ accidentally belongs to this subspace.

4. Calculating the Regression Weights in Bascula

In the previous section we have shown that the general regression weights $w_k = \pi_k^{-1} + \lambda_k \mathbf{x}_k^t \mathbf{G}_S (\mathbf{t}_x - \mathbf{x}_{\text{HT}})$ are invariant to the choice of $\mathbf{G}_S$. In this section we show how to compute these weights. To do so, we start with the Cholesky decomposition of the positive (semi) definite matrix $\mathbf{X}_S^t \Lambda_S \mathbf{X}_S$, see Seber (1977, page 322). If $\mathbf{X}_S$ is of full rank, then $\mathbf{X}_S^t \Lambda_S \mathbf{X}_S$ is positive definite and it can be expressed uniquely in the form $\mathbf{X}_S^t \Lambda_S \mathbf{X}_S = \mathbf{U}^t \mathbf{U}$, where $\mathbf{U}$ is an upper triangular matrix with positive diagonal elements. Let a_{ij} denote the ij^{th} element of $\mathbf{X}_S^t \Lambda_S \mathbf{X}_S$, then $\mathbf{U}$ can be computed, row by row, according to

$$u_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ki}^2} \quad \text{for } i = 1, ..., p$$

and (3)

$$u_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj}}{u_{ii}} \quad \text{for } j = i+1, ..., p .$$

If $\mathbf{X}_S$ has rank $r < p$, then an application of (3) will give r non-zero and $p - r$ zero diagonal elements of $\mathbf{U}$. If we find a zero diagonal element then we put its corresponding row and column elements at zero. Subsequently, by elementary row and column interchanges, we obtain the following upper triangular matrix:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Accordingly to the elementary row and column interchanges we also interchange the elements of $\mathbf{X}_S$ and $(\mathbf{t}_x - \mathbf{x}_{\text{HT}}) : \mathbf{X}_S \mathbf{E}^t = (\mathbf{X}_{1S} \mathbf{X}_{2S})$ and

$$\mathbf{E}(\mathbf{t}_x - \mathbf{x}_{\text{HT}}) = \begin{pmatrix} (\mathbf{t}_{1x} - \mathbf{x}_{1\text{HT}}) \\ (\mathbf{t}_{2x} - \mathbf{x}_{2\text{HT}}) \end{pmatrix},$$

where, by construction, $\mathbf{X}_{1S}$ is of full rank and $\mathbf{E}$ is a non-singular matrix of order $p \times p$. But, since

$$\mathbf{G}_S' = \begin{pmatrix} (\mathbf{X}_{1S}^t \Lambda_S \mathbf{X}_{1S})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1^{-1} (\mathbf{U}_1^t)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

is a generalized inverse of $(\mathbf{X}_{1S} \mathbf{X}_{2S})^t \Lambda_S (\mathbf{X}_{1S} \mathbf{X}_{2S})$, we have that $\mathbf{G}_S = \mathbf{E}^t \mathbf{G}_S' \mathbf{E}$ is a generalized inverse of $\mathbf{X}_S^t \Lambda_S \mathbf{X}_S$. Inserting this generalized inverse into $w_k = \pi_k^{-1} + \lambda_k \mathbf{x}_k^t \mathbf{G}_S (\mathbf{t}_x - \mathbf{x}_{\text{HT}})$ gives

$$w_k = \pi_k^{-1} + \lambda_k (\mathbf{x}_{1k}^t \mathbf{x}_{2k}^t) \mathbf{G}_S' \begin{pmatrix} (\mathbf{t}_{1x} - \mathbf{x}_{1\text{HT}}) \\ (\mathbf{t}_{2x} - \mathbf{x}_{2\text{HT}}) \end{pmatrix}$$

$$= \pi_k^{-1} + \lambda_k \mathbf{x}_{1k}^t \mathbf{U}_1^{-1} (\mathbf{U}_1^t)^{-1} (\mathbf{t}_{1x} - \mathbf{x}_{1\text{HT}}),$$

which is computed as follows. First $\mathbf{z} = (\mathbf{U}_1^t)^{-1} (\mathbf{t}_{1x} - \mathbf{x}_{1\text{HT}})$ is computed by solving the lower triangular system $\mathbf{U}_1^t \mathbf{z} = (\mathbf{t}_{1x} - \mathbf{x}_{1\text{HT}})$. Thereafter $\mathbf{u} = \mathbf{U}_1^{-1} \mathbf{z}$ is computed by solving the upper triangular system $\mathbf{U}_1 \mathbf{u} = \mathbf{z}$. Once

$\mathbf{u} = \mathbf{U}_1^{-1}(\mathbf{U}_1^t)^{-1}(\mathbf{t}_{1x} - \mathbf{x}_{1\mathrm{HT}})$ is computed it is a simple matter to compute w_k.

5. The Dutch Labour Force Survey

To illustrate some of the issues stated in this paper, we briefly discuss the weighting model of the Dutch Labour Force Survey (LFS) of 1987 up to 2000. The target population of this survey consisted of the non-institutional population residing in the Netherlands and its sampling design was based on a stratified three-stage sampling with households as ultimate sampling units. For details we refer to Nieuwenbroek and Van der Valk (1996). Five categorical variables were involved into the weighting model, namely Sex (2 categories), Age (12 categories), Marital Status (2 categories), Region (15 categories), and Nationality (2 categories). Mainly based on consistency requirements, the desired weighting model was

$$\mathrm{Sex} \times \mathrm{Age} \times \mathrm{MaritalStatus} \times \mathrm{Region} \times \mathrm{Nationality}.$$

However, this weighting model resulted in too many small cell counts, which gave unstable estimators. Therefore, the reduced model

$$(\mathrm{Sex} \times \mathrm{Age} \times \mathrm{MaritalStatus} \times \mathrm{Region}) + (\mathrm{Sex} \times \mathrm{Age}^{+} \times \mathrm{Region} \times \mathrm{Nationality})$$

was used instead, where Age^{+} (2 categories) was obtained by grouping the categories in Age. This reduced weighting model resulted in a design matrix not of full rank for two reasons, namely 1) some columns of the design matrix completely consisted of zeros due to impossible combinations of the categorical variables and 2) there were linear combinations between the columns of the design matrix.

Now, the first kind of redundancy can be easily traced. If such columns are found, then their corresponding population totals should be zero. Bascula carries out a check on this condition. The second kind of redundancy is more difficult to trace. Linear combinations between columns may arise because one variable is incorporated into several weighting terms. For example, sex and region appear in both weighting terms of the LFS weighting model. The resulting linear combinations can be recognized beforehand by the name of the variable. For the age-variable, which also appears in both weighting terms, such a redundancy check beforehand is less obvious. These latter kinds of redundancy are traced by means of the Cholesky decomposition. Naturally, if any linear combinations are found, either by name beforehand or by the Cholesky decomposition, then the same linear combinations should also exists between the vector of population totals. Bascula also checks this condition.

Acknowledgements

The authors wish to thank Jan van den Brakel, Nico Nieuwenbroek, and Jeroen Pannekoek for their careful reading and helpful comments. The authors also wish to thank the referee for his careful reading and valuable suggestions. Especially his remarks on assumption (A1) of a previous version and his suggestions to simplify several proofs have led to a considerable improvement of the paper.

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands.

References

Alexander, C.H. (1987). A class of methods for using person controls in household weighting. *Survey Methodology*, 13, 183-198.

Bethlehem, J.G., and Keller, W.J. (1987). Linear weighting of sample survey data. *Journal of Official Statistics*, 3, 141-153.

Deville, J.-C., and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.

Lemaître, G., and Dufour, J. (1987). An integrated method for weighting persons and families. *Survey Methodology*, 13, 199-207.

Nieuwenbroek, N.J. (1997). General regression estimator in Bascula: Theoretical background. Research paper no. 9737, Statistics Netherlands.

Nieuwenbroek, N.J., and Van Der Valk, J. (1996). Research paper no. 9629, Statistics Netherlands.

Rao, C.R. (1973). *Linear Statistical Inference And Its Applications* (2nd edition). New York: John Wiley & Sons, Inc.

Renssen, R.H., Nieuwenbroek, N.J. and Slootbeek, G.T. (1997). Variance module in Bascula: Theoretical background. Research paper no. 9712, Statistics Netherlands.

Särndal, C.-E., Swensson, B. and Wretman, J.H. (1992). *Model-assisted Survey Sampling*. New York, Spinger-Verlag.

Searle, S.R. (1971). *Linear Models*. New York: John Wiley & Sons, Inc.

Seber, G.A.F. (1977). *Linear Regression Analysis*. New York: John Wiley & Sons, Inc.