Linearization Variance Estimators for Longitudinal Survey Data

A. Demnati and J. N. K. Rao

A. Demnati, Social Survey Methods Division,
Statistics Canada, Ottawa, Ontario, Canada, K1A 0T6
Abdellatif.Demnati@statcan.ca

J. N. K. Rao, School of Mathematics and Statistics,
Carleton University, Ottawa, Ontario, Canada, K1S 5B6
JRao@math.carleton.ca

Summary

In survey sampling, Taylor linearization is often used to obtain variance estimators for nonlinear finite population parameters, such as ratios, regression and correlation coefficients, which can be expressed as smooth functions of totals. Taylor linearization is generally applicable to any sampling design, but it can lead to multiple variance estimators that are asymptotically design unbiased under repeated sampling. The choice among the variance estimators requires other considerations such as (i) approximate unbiasedness for the model variance of the estimator under an assumed model, (ii) validity under a conditional repeated sampling framework. Demnati and Rao (2001) proposed a new approach to deriving Taylor linearization variance estimators that leads directly to a unique variance estimator that satisfies the above considerations. In this paper, we extended the work of Demnati and Rao (2001) to deal with longitudinal surveys which lead to dependent observations and to multiple weights on the same unit. We consider a variety of longitudinal sampling designs, covering panel surveys, household panel surveys as well as rotating surveys.

Key Words: Multiple weights; Repeated surveys; Taylor linearization.

1. Introduction

Taylor linearization is a popular method of variance estimation for complex statistics such as ratio and regression estimators and logistic regression coefficient estimators. It is generally applicable to any sampling design that permits unbiased variance estimation for linear estimators, and it is computationally simpler than a resampling method such as the jackknife. However, it can lead to multiple variance estimators that are asymptotically design unbiased under repeated sampling. The choice among the variance estimators, therefore, requires other considerations such as (i) approximate unbiasedness for the model variance of the estimator under an assumed model, (ii) validity under a conditional repeated sampling framework. For example, in the context of simple random sampling and the ratio estimator, $\hat{Y}_R = (\bar{y}/\bar{x})X$, of the population total Y, Royall and Cumberland (1981) showed that a commonly used linearization variance estimator $\vartheta_L = N^2(n^{-1} - N^{-1})s_z^2$ does not track the conditional variance of $\hat{Y}_R$ given $\bar{x}$, unlike the jackknife variance estimator ϑ_J. Here $\bar{y}$ and $\bar{x}$ are the sample means, X is the known population total of an auxiliary variable x, s_z^2 is the sample variance of the residuals $z_i = y_i - (\bar{y}/\bar{x})x_i$ and (n, N) denote the sample and population sizes. By linearizing the jackknife variance estimator, ϑ_J, we obtain a different linearization variance estimator, $\vartheta_{JL} = (\bar{X}/\bar{x})^2 \vartheta_L$, which also tracks the conditional variance as well as the unconditional variance, where $\bar{X} = X/N$ is the mean of x. As a result, ϑ_{JL} or ϑ_J may be preferred over ϑ_L. Valliant (1993) obtained ϑ_{JL} for the post-stratified estimator and conducted a simulation study to demonstrate that both ϑ_J and ϑ_{JL} possess good conditional properties given the estimated post-strata counts. Särndal, Swensson and Wretman (1989) showed that ϑ_{JL} is both asymptotically design unbiased and asymptotically model unbiased in the sense of $E_m(\vartheta_{JL}) = Var_m(\hat{Y}_R)$, where E_m denotes model expectation and $Var_m(\hat{Y}_R)$ is the model variance of $\hat{Y}_R$ under a Aratio model≅: $E_m(y_i) = \beta x_i$; $i = 1,...,N$ and the y_i's are independent with model variance $Var_m(y_i) = \sigma^2 x_i$, $\sigma^2 > 0$. Thus, ϑ_{JL} is a good choice from either the design-based or the model-based perspective. Demnati and Rao (2001) proposed a new approach to variance estimation that is theoretically justifiable and at the same time leads directly to a ϑ_{JL}-type variance estimator for general designs. Afterward, Demnati and Rao (2002) extended their

method to the case of missing responses when adjustment for complete nonresponse and imputation based on smooth functions of observed values, in particular ratio imputation, are used.

While several methods have been proposed to correctly estimate the variance of an estimator under cross-sectional data, methods of variance estimation from longitudinal data are limited, despite an increase in recent years in the number of longitudinal surveys. The dynamic nature of longitudinal populations such as births and deaths, the limitation of longitudinal samples to take into account such features of longitudinal populations, and multiple weights are one of the difficulties inherent to longitudinal data in addition to the customary cross-sectional difficulties such as complex designs, missing data and so on. This paper is a first attempt to extend the Demnati and Rao (2001,2002) method to the case of longitudinal survey data. Section 2 gives a brief account of the method for the case of cross-sectional data, and section 3 presents the extension to longitudinal survey data.

2. Cross-Sectional Data

We give a brief account of the Demnati and Rao (2001) method under full response. Suppose an estimator $\hat{\theta}$ of a parameter θ can be expressed as a differentiable function $g(\hat{\boldsymbol{Y}})$ of estimated totals $\hat{\boldsymbol{Y}} = (\hat{Y}_1,...,\hat{Y}_m)^T$, where $\hat{Y}_j = \sum_{i\in U} d_i(s) y_{ij}$ is an estimator of the population Y_j, $j = 1,..,m$, where $d_i(s) = 0$ if the unit i is not in the sample s, U is the set of population units, and $\theta = g(\boldsymbol{Y})$ with $\boldsymbol{Y} = (Y_1,...,Y_m)^T$. We may write $\hat{\theta}$ as $\hat{\theta} = f(\boldsymbol{d}(s), \boldsymbol{A}_y)$ and $\theta = f(\boldsymbol{1}, \boldsymbol{A}_y)$, where $\boldsymbol{A}_y$ is an $m \times N$ matrix with j^{th} column $\boldsymbol{y}_j = (y_{1j},..., y_{mj})^T$, $j = 1,..., N$, $\boldsymbol{d}(s) = (d_1(s),..., d_N(s))^T$ and $\boldsymbol{1}$ is the N-vector of 1's. For example, if $\hat{\theta}$ denotes the ratio estimator $\hat{Y}_R = X(\sum_{i\in U} d_i(s) y_i)/(\sum_{i\in U} d_i(s) x_i)$, then $m = 2$, $y_{1i} = y_i$, $y_{2i} = x_i$ and $f(\boldsymbol{1}, \boldsymbol{A}_y)$ reduces to the total Y, noting that $X(Y/X) = Y$. Note that $\hat{Y}_R$ is a function of $\boldsymbol{d}(s)$, $\boldsymbol{y}$, $\boldsymbol{x}$ and the known total X, but we dropped X for simplicity and write $\hat{Y}_R = f(\boldsymbol{d}(s), \boldsymbol{y}, \boldsymbol{x})$. If the Horvitz-Thompson weights are used, then $d_i(s) = 1/\pi_i$ for $i \in s$, where π_i is the probability of selecting unit i in the sample s.

Let $f(\boldsymbol{b}, \boldsymbol{A}_y) = f(\boldsymbol{b})$ for arbitrary real numbers $\boldsymbol{b} = (b_1,..., b_N)^T$. Demnati and Rao (2001) showed that the Taylor linearization of $\hat{\theta} - \theta$, namely

$$\hat{\theta} - \theta = g(\hat{\boldsymbol{Y}}) - g(\boldsymbol{Y}) \approx \left(\partial g(\boldsymbol{a}) / \partial \boldsymbol{a}\right)^T \Big|_{\boldsymbol{a}=\boldsymbol{Y}} \left(\hat{\boldsymbol{Y}} - \boldsymbol{Y}\right),$$

is equivalent to

$$\begin{aligned} \hat{\theta} - \theta &\approx \textstyle\sum_{k=1}^{N} \left(\partial f(\boldsymbol{b}) / \partial b_k\right)\Big|_{\boldsymbol{b}=\boldsymbol{1}} \left(d_k(s) - 1\right) \\ &= \tilde{\boldsymbol{z}}^T \left(\boldsymbol{d}(s) - \boldsymbol{1}\right), \end{aligned} \tag{2.1}$$

where $\partial g(\boldsymbol{a}) / \partial \boldsymbol{a} = (\partial g(\boldsymbol{a}) / \partial a_1,..., \partial g(\boldsymbol{a}) / \partial a_m)^T$ and $\tilde{\boldsymbol{z}} = (\tilde{z}_1,..., \tilde{z}_N)^T$ with $\tilde{z}_k = \partial f(\boldsymbol{b}) / \partial b_k \big|_{\boldsymbol{b}=\boldsymbol{1}}$. It follows from (2.1) that a variance estimator of $\hat{\theta}$ is approximately given by the variance estimator of the estimated total $\sum d_i(s) \tilde{z}_i = \hat{Y}(\tilde{z}, \boldsymbol{d}(s))$; that is, $\mathrm{var}(\hat{\theta}) \approx \vartheta(\tilde{z}, \boldsymbol{d}(s))$, where $\vartheta(y, \boldsymbol{d}(s))$ denotes the variance estimator of $\hat{Y} = \hat{Y}(y, \boldsymbol{d}(s))$ in operator notation using the vector of design weights $\boldsymbol{d}(s)$. Now we replace $\tilde{z}_k$ by $z_k = \partial f(\boldsymbol{b}) / \partial b_k \big|_{\boldsymbol{b}=\boldsymbol{d}(s)}$, since $\tilde{z}_k$ =s are unknown, to get a linearization variance estimator

$$\vartheta_L(\hat{\theta}) = \vartheta(z, \boldsymbol{d}(s)). \tag{2.2}$$

Note that $\vartheta_L(\hat{\theta})$ given by (2.2) is simply obtained from the formula $\vartheta(y, \boldsymbol{d}(s))$ for $\hat{Y} = \hat{Y}(y, \boldsymbol{d}(s))$ by replacing y_i by z_i for $i \in s$. Note that we do not first evaluate the partial derivatives $\partial f(\boldsymbol{b}) / \partial b_k$ at $\boldsymbol{b} = \boldsymbol{1}$ to get $\tilde{z}_k$ and then substitute estimates for the unknown components of $\tilde{\boldsymbol{z}}$. Our method, therefore, is similar in spirit to Binder(1996)'s approach. The variance estimator $\vartheta_L(\hat{\theta})$ is valid because z_k is a consistent estimator of $\tilde{z}_k$.

Suppose $\hat{\theta}$ is the ratio estimator $\hat{Y}_R = X(\sum d_i(s)y_i)/(\sum d_i(s)x_i)$, where $\sum$ denotes summation over $i \in U$. Then $f(\boldsymbol{b}) = X(\sum b_i y_i)/(\sum b_i x_i) = X\hat{Y}(y,\boldsymbol{b})/\hat{Y}(x,\boldsymbol{b})$ and

$$z_k = \partial f(\boldsymbol{b})/\partial b_k\big|_{\boldsymbol{b}=\boldsymbol{d}(s)} = \frac{X}{\hat{X}}\left(y_k - \hat{R}x_k\right).$$

For simple random sampling, $\vartheta_L(\hat{Y}_R) = \vartheta(z,\boldsymbol{d}(s))$ agrees with $\vartheta_{JL} = (\overline{X}/\overline{x})^2\vartheta_L$.

The above derivation is simple and natural. On the other hand, in the standard linearization method, $\hat{\theta}$ is first expressed in terms of elementary components $\hat{Y}_1,...,\hat{Y}_m$ as $g(\hat{\boldsymbol{Y}})$ and the partial derivatives $\partial g(\boldsymbol{a})/\partial a_j$ are then evaluated at $\boldsymbol{a} = \boldsymbol{Y}$. It is interesting to note that all the components $\hat{Y}_j$ use the same weights $d_i(s)$ and our approach always takes partial derivatives of $f(\boldsymbol{b})$ with respect to b_k at $\boldsymbol{b} = \boldsymbol{d}(s)$. It is not necessary to first express $\hat{\theta}$ in terms of elementary components. Demnati and Rao (2001) applied the method to a variety of problems, covering regression calibration estimators of a total Y and other estimators defined either explicitly or implicitly as solutions of estimating equations. They obtained a new variance estimator for a general class of calibration estimators that includes generalized raking ratio and generalized regression estimators. They also extended the method to two-phase sampling and obtained a variance estimator that makes fuller use of the first phase sample data compared to traditional linearization variance estimators.

3. Longitudinal Data

As in the usual finite population situation, we identify a population of size N during a fixed period of time by a set of indices $P = \{1,...,N\}$. Let $P^{(t)}$ denote the cross-sectional population at time t and let $J_i^{(t)}$ denote the cross-sectional population membership indicator for unit i, $i = 1,...,N$, i.e. $J_i^{(t)} = 1$ if $i \in P^{(t)}$ and $J_i^{(t)} = 0$ if $i \notin P^{(t)}$ with $J_+^{(t)} = \sum_i J_i^{(t)} = N^{(t)}$ the size of the cross-sectional population $P^{(t)}$, where $\sum_i$ denotes summation over the population units i. In this section, we consider variance estimation of a differentiable function of estimated cross-sectional totals under a series of longitudinal sampling designs assuming full response and no post-stratification. It is shown that a variance estimator can be obtained through variance estimation of a design-weighted estimator of total of a synthetic population. The case of estimators obtained as solution of estimating equations can be obtained along the lines of Demnati and Rao (2001), and the case of estimators obtained from survey data with missing responses can be obtained along the lines of Demnati and Rao (2002). In this version, details of these extensions are omitted for space reason.

To discuss multiple weights inherent in repeated surveys, suppose an estimator $\hat{\theta}$ of a parameter θ can be expressed as a differentiable function $\hat{\theta} = g(\hat{\boldsymbol{Y}})$ of estimated cross-sectional totals $\hat{\boldsymbol{Y}} = (\hat{Y}^{(1)},...,\hat{Y}^{(T)})^T$, where $\hat{Y}^{(t)}$ is an estimator of the population total $Y^{(t)} = \sum_i J_i^{(t)} y_i^{(t)}$ at time t, $t = 1,...,T$, and $\theta = g(\boldsymbol{Y})$ with $\boldsymbol{Y} = (Y^{(1)},...,Y^{(T)})^T$. We may write $\hat{Y}^{(t)}$ as $\hat{Y}^{(t)} = \sum_g^{G^{(t)}} \sum_i \tilde{d}_{gi}^{(t)}(s_g^{(t)}) y_i^{(t)} = \boldsymbol{1}_{G^{(t)}}^T \boldsymbol{A}_{\tilde{d}}^{(t)} \boldsymbol{y}^{(t)}$ with specified constants $G^{(t)}$, where $\boldsymbol{A}_{\tilde{d}}^{(t)} = (\tilde{\boldsymbol{d}}_1^{(t)}(s_1^{(t)}),...,\tilde{\boldsymbol{d}}_{G^{(t)}}^{(t)}(s_{G^{(t)}}^{(t)}))^T$, $\tilde{\boldsymbol{d}}_g^{(t)}(s_g^{(t)})$ are vectors of cross-sectional weights at time t, $\boldsymbol{1}_{G^{(t)}}$ is the $G^{(t)}$-vector of 1's and $\boldsymbol{y}^{(t)} = (y_1^{(t)},...,y_N^{(t)})^T$. In this section we show that the estimator $\hat{\theta}$ can be written as $\hat{\theta} = f(\boldsymbol{d}(s))$ under a wide class of longitudinal survey designs. Therefore, it follows from Demnati and Rao (2001) that a linearization variance estimator of $\hat{\theta}$ is given by

$$\vartheta_L(\hat{\theta}) = \vartheta(z,\boldsymbol{d}(s)), \tag{3.1}$$

where

$$\boldsymbol{d}(s) = vech(\boldsymbol{A}_d^T), \tag{3.2}$$

$\boldsymbol{A}_d = (\boldsymbol{A}_d^{(1)T},...,\boldsymbol{A}_d^{(T)T})^T$, $\boldsymbol{A}_d^{(t)} = (\boldsymbol{d}_1^{(t)}(s_1^{(t)}),...,\boldsymbol{d}_{G^{(t)}}^{(t)}(s_{G^{(t)}}^{(t)}))^T$, $\boldsymbol{d}_g^{(t)}(s)$ are vectors of design weights, $vech(\boldsymbol{A})$ denotes the vector obtained by stacking distinct column of $\boldsymbol{A}$ underneath each other, in order from left to right, and

$$z_k = \partial f(\boldsymbol{b})/\partial b_k \big|_{\boldsymbol{b}=\boldsymbol{d}(s)} . \tag{3.3}$$

3.1 Independent Samples

Under independent samples we have $G^{(t)} = 1$, $\hat{Y}^{(t)} = \hat{Y}(y^{(t)}, \boldsymbol{d}^{(t)}(s^{(t)})) = \sum_i y_i^{(t)} d_i^{(t)}(s^{(t)})$, and

$$\boldsymbol{d}(s) = vech(\boldsymbol{d}^{(1)}(s^{(1)}),...,\boldsymbol{d}^{(T)}(s^{(T)})) , \tag{3.4}$$

where $\boldsymbol{d}^{(t)}(s^{(t)}) = (d_1^{(t)}(s^{(t)}),...,d_N^{(t)}(s^{(t)}))^T$ is the vector design weights at time t. If the Horvitz-Thompson weights are used then $d_i^{(t)}(s^{(t)}) = J_i^{(t)} a_i^{(t)}(s)/\pi_i^{(t)}$, $a_i^{(t)} = 1$ if unit i belongs to the cross-sectional sample $s^{(t)}$ and $a_i^{(t)} = 0$ if not, and $\pi_i^{(t)} = \Pr(i \in s^{(t)} / J_i^{(t)} = 1)$.

3.2 Panel Sample

Under panel sample situation, where observations are collected over time on the same sampled units, we have $G^{(t)} = 1$, $\hat{Y}^{(t)} = \hat{Y}(y^{(t)}, \boldsymbol{d}^{(0)}(s^{(0)}))$ and

$$\boldsymbol{d}(s) = \boldsymbol{d}^{(0)}(s) , \tag{3.5}$$

where $d_i^{(0)}(s^{(0)})$, $i = 1,...,N$, are the basic survey weights of the unique sample selected at the initial wave (say at time $t = 0$).

3.3 Household Panel Sample

Under household panel sample situation, where observations are collected over time not only on the sampled units selected at the initial wave but also on non-sampled units who join households containing at least one sampled unit, we have $G^{(t)} = 1$, $\hat{Y}^{(t)} = \sum y_i^{(t)} \tilde{d}_i^{(t)}(s^{(t)})$, and

$$\boldsymbol{d}(s) = \boldsymbol{d}^{(0)}(s) , \tag{3.6}$$

where $\tilde{d}_i^{(t)}(s^{(t)}) = \sum_j J_j^{(t)} I_{ji}^{(t)} d_j^{(0)}(s^{(0)}) \alpha_{ji}^{(t)}$, $\alpha_{ji}^{(t)} = \alpha_i^{(t)} = 1/\sum_k J_k^{(t)} I_{ki}^{(t)} a_k^{(0)}(s^{(0)})$, $I_{ji}^{(t)}$ is an indicator variable indicating if the unit j lives in the same household as unit i at time t, and $a_k^{(0)}(s^{(0)}) = d_k^{(0)}(s^{(0)}) E(a_k^{(0)}(s^{(0)}))$. The value for $\alpha_{ji}^{(t)}$ is obtained through the multiplicity approach, Sirken (1970).

3.4 Multiple Samples

New entrants who have zero probability to join the initial population units are not covered, unless a supplementary sample is taken. That's why multiple samples are used at the estimation stage at time t under household panel surveys. Under multiple samples situation each sample $s_g^{(t)}$, $g = 1,...,G^{(t)}$ can be viewed as a sample from the same cross-sectional population using a different sampling frame $F_g^{(t)}$. The frames are probably incomplete, but the union of the $G^{(t)}$ frames covers the entire cross-sectional population. The $G^{(t)}$ frames overlap. Let $\tilde{J}_{gi}^{(t)}$ be the conditional frame membership indicator variable, i.e., $\tilde{J}_{gi}^{(t)} = 1$ if $i \in F_g^{(t)} \mid J_i^{(t)} = 1$ and $\tilde{J}_{gi}^{(t)} = 0$ otherwise. When combining the $G^{(t)}$ samples, the weights have to be adjusted to account for the multiplicity of the sampling frames:

$$\tilde{d}_{gi}^{(t)}(s) = d_{gi}^{(t)}(s) \tilde{\alpha}_{gi}^{(t)} , \tag{3.7}$$

where $\tilde{\alpha}_{gi}^{(t)}$ are constants satisfying the constraints $\sum_g \tilde{J}_{gi}^{(t)} \tilde{\alpha}_{gi}^{(t)} = 1$ and $d_{gi}^{(t)}(s)$ are the cross-sectional weights resulting from frame $F_g^{(t)}$. A first choice for $\tilde{\alpha}_{gi}^{(t)}$ is $\tilde{\alpha}_{gi}^{(t)} = 1/\tilde{J}_{+i}^{(t)}$, where $\tilde{J}_{+i}^{(t)}$ is the multiplicity of unit i at time t, i.e., the total number of sampling frames reporting unit i at time t, (Sirken, 1970). A second choice for $\tilde{\alpha}_{gi}^{(t)}$ is

$\tilde{\alpha}_{gi}^{(t)} = \tilde{J}_{gi}^{(t)} f_g^{(t)} / \sum_g \tilde{J}_{gi}^{(t)} f_g^{(t)}$ in which case one wants to take into account the sampling fraction $f_g^{(t)}$ of each sampling frame (Kalton and Anderson, 1986). Under multiple frames, the estimated total $\hat{Y}^{(t)}$ can be expressed as before as $\boldsymbol{1}_{G^{(t)}}^T \boldsymbol{A}_{\tilde{d}}^{(t)} \boldsymbol{y}^{(t)}$. Hence, estimation of the variance can be done along the line of sub-section 3.1-3.3. Details are omitted for simplicity.

3.5 Rotating Samples

Rotating samples are used as a compromise between independent samples and panel surveys. Rotating samples control overlap between successive time periods and provide unbiased cross-sectional estimates while taking advantage of the correlation between two time periods. For example, in labor-force surveys which are conducted monthly in many countries, overlap is controlled by re-interviewing a high fraction of previous selected households (say $G-1$ rotation groups) while new households are selected for a first interview(say one rotation group) for a total of G rotation groups or panel samples. Each household remains in the same rotation group for a predetermined number of months. Unlike multiple samples, rotating groups do not overlap. A simple example of a rotating group is given by the following two steps:

(a) Split the population at random into G groups U_g $(g = 1,..., G)$ each of size N / G.

(b) Split each group at random into N / n samples each of size n / G.

The following table illustrates a rotation pattern with $G = 2$, $N = 20$, $n = 4$ and 50% overlap.

Table 1 : Example of Rotating Scheme

Group	Sample	Time			
		1	2	3	…
U_1	s_{11}	√			
	s_{12}		√	√	
	…				
	s_{15}				
U_2	s_{21}	√	√		
	s_{22}			√	
	…				
	s_{25}				
Cross-sectional sample		$s^{(1)} = s_{11} \cup s_{21}$	$s^{(2)} = s_{12} \cup s_{21}$	$s^{(3)} = s_{12} \cup s_{22}$	

Each cross-sectional sample is composed of 2 samples. The sample s_{11} appears in the first cross-sectional sample (time $t = 1$), the samples s_{12} and s_{21} appear in two time periods(time 2 and 3), while the sample s_{22} appears only in the third cross-sectional sample (time $t = 3$).

The first step consists on selecting G non-overlap samples from the population with a complete coverage of the union of the G samples. Let $a_{g1i}(U_g)$ be the group U_g, $g = 1,..., G$, membership indicator variables with $\sum_i a_{g1i}(U_g) = N_g$, $\sum_g \sum_i a_{g1i}(U_g) = N$, and $a_{g1i}(U_g) a_{h1i}(U_h) = 1(g = h)$. The second step is a repetition of the first step within each group. Let $a_{gs2i}(s_{gs})$ be the conditional sample s_{gs} membership indicator variables with $\sum_i a_{gs2i}(s_{gs}) = n_{gs}$, $\sum_s \sum_i a_{gs2i}(s_{gs}) = N_g$, and $a_{gs2i}(s_{gs}) a_{gq2i}(s_{gq}) = 1(s = q)$. We have $G^{(t)} = G$, $\hat{Y}^{(t)} = \sum_g \hat{Y}(y^{(t)}, \boldsymbol{d}_g^{(t)}(s_g^{(t)})) / G$, and

$$\boldsymbol{d}(s) = vech(\boldsymbol{d}_1^{(1)}(s_1^{(1)}),..., \boldsymbol{d}_G^{(1)}(s_G^{(1)}),..., \boldsymbol{d}_1^{(T)}(s_1^{(T)}),..., \boldsymbol{d}_G^{(T)}(s_G^{(T)})), \tag{3.8}$$

with $\boldsymbol{d}_g^{(t)}(s_g^{(t)}) = (d_{g1}^{(t)}(s_g^{(t)}),..., d_{gN}^{(t)}(s_g^{(t)}))^T$, $d_{gi}^{(t)}(s_g^{(t)}) = d_{g1i}(U_g) d_{g2i}^{(t)}(s_g^{(t)})$, $d_{g1i}(U_g) = J_i^{(1)} a_{g1i}(U_g) / E(a_{g1i}(U_g))$,

$d_{g2i}^{(t)}(s_g^{(t)}) = a_{gi}^{(t)}(s_g^{(t)}) / E(a_{gi}^{(t)}(s_g^{(t)}))$, and $a_{gi}^{(t)}(s_g^{(t)}) = a_{gs2i}(s_{gs})$ if s_{gs} is the sample used at time t in group g.

Under rotating samples, the variance of $\hat{Y}^{(t)}$ is given by $Var(\hat{Y}^{(t)}) = EVar(\hat{Y}^{(t)}) + VarE(\hat{Y}^{(t)})$. Now $VarE(\hat{Y}^{(t)}) = Var(Y^{(t)}) = 0$, so an estimator of $Var(\hat{\theta})$ is given by

$$\vartheta_L(\hat{\theta}) = \sum_g \vartheta_2(z_g, \boldsymbol{d}_{g2}(s_g)), \tag{3.9}$$

where

$$\boldsymbol{d}_{g2}(s_g) = vech(\boldsymbol{d}_{g2}^{(1)}(s_g^{(1)}),...,\boldsymbol{d}_{g2}^{(T)}(s_g^{(T)})), \tag{3.10}$$

and $\vartheta_2(.)$ denotes a conditional variance estimator of a total.

Composite estimation

Under rotating samples, it is advantageous to use a composite estimator. Suppose the following simple composite estimator

$$\widetilde{Y}^{(t)} = \hat{Y}^{(t)} - \hat{\alpha}^{(t)}(\hat{Y}_u^{(t)} - \hat{Y}_m^{(t)}) - \hat{\beta}^{(t)}(\hat{Y}_m^{(t-1)} - \hat{Y}^{(t-1)}). \tag{3.11}$$

where $\hat{Y}_u^{(t)}$ is an estimator of the population total using the unmatched sample, $\hat{Y}_m^{(t)}$ and $\hat{Y}_m^{(t-1)}$ are estimators of the population totals using matched samples, and $\hat{\alpha}^{(t)}$ and $\hat{\beta}^{(t)}$ are assumed to be for simplicity regression parameters estimators as in the GREG approach. The case of optimal regression is covered in Demnati and Rao (2001). A variance estimator of the composite estimator, $\widetilde{\theta} = g(\widetilde{\boldsymbol{Y}})$, can be obtained using (3.9) with

$$\boldsymbol{d}_{g2}(s_g) = vech(\boldsymbol{d}_{g2}^{(0)}(s_g^{(0)}),...,\boldsymbol{d}_{g2}^{(T)}(s_g^{(T)})). \tag{3.12}$$

The composite estimator (3.11) has three components as in the AK composite estimator, but it is not recursive and estimates for different domains add up unlike in the case of AK composite estimator.

Concluding Remarks

We have presented a new approach to variance estimation under longitudinal survey data. A valid variance estimator is given under a variety of longitudinal sampling designs. Extension to estimators defined as solutions to survey weighted estimating equations is currently under investigation.

References

Binder, D. (1996), ALinearization Methods for Single Phase and Two-Phase Samples: A Cookbook Approach≅, *Survey Methodology, 22*, 17-22.

Demnati, A. and Rao, J. N. K. (2001),A Linearization Variance Estimators for Survey Data≅, Methodology Branch Working Paper, SSMD-2001-010E. Statistics Canada.

Demnati, A. and Rao, J. N. K. (2002), ALinearization Variance Estimators for Survey Data With Missing Responses≅, *in Proceeding of the Section Survey Research Methods*, American Statistical Association.

Kalton, G. and Anderson, D. (1986), "Sampling Rare Populations", *Journal of the Royal Statistical Society*, Series A, 149, Part 1, pp.65-82.

Royall, R. M., and Cumberland, W. G. (1981), AAn Empirical Study of the Ratio Estimator and Estimators of its Variance≅, *Journal of the American Statistical Association, 76*, 66-77.

Särndal, C.-E., Swensson, B., and Wretman, J.H. (1989),AThe Weighted Residual Technique for Estimating the Variance of the General Regression Estimator of the Finite Population Total≅, *Biometrika, 76*, 527-537.

Sirken, M. G. (1970),"Household Surveys with Multiplicity", *Journal of the American Statistical Association*, 65, 257-266.

Valliant, R. (1993),APostsratification and Conditional Variance Estimation≅, *Journal of the American Statistical Association, 88*, 89-96.