PENALIZED CHI SQUARE DISTANCE FUNCTION IN SURVEY SAMPLING

Patrick J. Farrell[1] and Sarjinder Singh[2]

[1]School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.
[2]Depratment of Statistics, St. Cloud State University, 720 Fourth Avenue South, St. Cloud, MN 56301-4498, USA.

SUMMARY

In the present investigation, we propose a penalized chi square distance function for estimating the total/mean of a finite population. The function produces a general class of estimators for the population total that includes, among others, the estimators of Searls (1964), Singh and Srivastava (1980), and the famous unbiased ratio estimator of Hartley and Ross (1954). Several of the estimators in the resulting general class are not members of the family based on the pioneer technique of Deville and Sarndal (1992). The existing gap in the GES developed at Statistics Canada to study Searls' (1964) estimator and unbiased estimation through calibration could be filled with the help of technology developed here.

Key words: Auxiliary information; Calibration; Estimation of total/mean; Model-Assisted approach; Ratio and regression type estimators; Searls' estimator.

1. INTRODUCTION

When auxiliary information is available, the most commonly used estimator of the population total/mean is the generalized linear regression (GREG) estimator. In what follows, we consider the simplest case of GREG, where information on only one auxiliary variable has been collected. Suppose that from a population $\Omega = \{1,2,..,i,..,N\}$, a probability sample $s(s \subset \Omega)$ is drawn with a given sampling design, $p(.)$. The inclusion probabilities $\pi_i = \Pr(i \in s)$ and $\pi_{ij} = \Pr(i \,\&\, j \in s)$ are assumed to be strictly positive and known. Let $(x_i,\ y_i)$ be a bivariate observation consisting of the values of the auxiliary variable and the variable of interest for the i-th population element. The Horvitz-Thompson (1952) estimator of the population total $Y = \sum_{i=1}^{N} y_i$ is:

$$\hat{Y}_{HT} = \sum_{i=1}^{n} d_i y_i \qquad (1.1)$$

where $d_i = 1/\pi_i$ are the Horvitz-Thompson weights. The variance of (1.1) is

$$V\left(\hat{Y}_{HT}\right) = \frac{1}{2}\sum_{i \neq j \in \Omega}\sum \Theta_{ij}\left(d_i y_i - d_j y_j\right)^2 \qquad (1.2)$$

where $\Theta_{ij} = \left(\pi_i \pi_j - \pi_{ij}\right)$.An unbiased estimator for (1.2) is

$$\hat{V}\left(\hat{Y}_{HT}\right) = \frac{1}{2}\sum_{i \neq j \in s}\sum D_{ij}\left(d_i y_i - d_j y_j\right)^2 \qquad (1.3)$$

where $D_{ij} = \Theta_{ij}/\pi_{ij}$. Deville and Sarndal (1992) proposed a new estimator for the total, naming it GREG:

$$\hat{Y}_{ds} = \sum_{i=1}^{n} w_i y_i \qquad (1.4)$$

Here the w_i are weights that, for a given metric, are as close as possible in an average sense to the d_i while respecting the calibration equation

$$\sum_{i=1}^{n} w_i x_i = X \qquad (1.5)$$

Let q_i be a set of suitably chosen weights. Minimizing the chi square type distance function

$$\sum_{i=1}^{n} \frac{\left(w_i - d_i\right)^2}{d_i q_i} \qquad (1.6)$$

subject to (1.5), yields new weights as

$$w_i = d_i + \left\{ d_i q_i x_i \Big/ \sum_{i=1}^{n} d_i q_i x_i^2 \right\}\left(X - \sum_{i=1}^{n} d_i x_i \right) \qquad (1.7)$$

Particular choices for q_i yield different forms of the estimator in (1.4). Substituting w_i in (1.7) into (1.4), yields the generalized regression estimator of the population total

$$\hat{Y}_{ds} = \sum_{i=1}^{n} d_i y_i + \hat{\beta}_{ds}\left(X - \sum_{i=1}^{n} d_i x_i \right) \qquad (1.8)$$

where $\hat{\beta}_{ds} = \sum_{i=1}^{n} d_i q_i x_i y_i \Big/ \sum_{i=1}^{n} d_i q_i x_i^2$, with approximate variance:

$$V\left(\hat{Y}_{ds}\right) = \frac{1}{2}\sum_{i \neq j \in \Omega}\sum \Theta_{ij}\left(d_i e_i - d_j e_j\right)^2 \qquad (1.9)$$

As an estimator for (1.9) considered by Sarndal *et al.* (1989), Deville and Sarndal (1992) and Sarndal (1996), is

$$\hat{V}\left(\hat{Y}_{ds}\right)=\frac{1}{2}\sum_{i\neq j\in s}\sum D_{ij}\left(w_i e_i - w_j e_j\right)^2 \tag{1.10}$$

where $e_i = y_i - \hat{\beta}_{ds}x_i$.The estimator in (1.8) is quite general and includes different estimators as particular cases. For example, if $q_i = 1/x_i$ then (1.8) reduces to the ratio estimator studied by Cochran (1977) as

$$\hat{Y}_{ds} = \hat{Y}_{HT}\left(X/\hat{X}_{HT}\right) \tag{1.11}$$

where $\hat{X}_{HT} = \sum_{i\in s} d_i x_i$. If $q_i = 1$, then (1.8) reduces to general regression estimator

$$\hat{Y}_{ds} = \hat{Y}_{HT} + \hat{\beta}_{ds}\left(X - \hat{X}_{HT}\right) \tag{1.12}$$

Unfortunately there is no choice of q_i that results in (1.8) matching the linear regression estimator of Hansen, Hurwitz and Madow (1953) given by

$$\hat{Y}_{hhm} = \hat{Y}_{HT} + \hat{\beta}_{ols}\left(X - \hat{X}_{HT}\right) \tag{1.13}$$

where $\hat{\beta}_{ols}=\sum_{i\in s} d_i(x_i - \bar{x})(y_i - \bar{y})\Big/\sum_{i\in s} d_i(x_i - \bar{x})^2$, $\bar{x}=\sum_{i\in s} d_i x_i \Big/ \sum_{i\in s} d_i$ and $\bar{y}=\sum_{i\in s} d_i y_i \Big/ \sum_{i\in s} d_i$. Wu and Sitter (2001) added the constraint $N^{-1}\sum_{i\in s} w_i = 1$ on the weights and developed a new estimator of the population total

$$\hat{Y}_{ws} = \hat{Y}_{HT} + \left(N - \sum_{i\in s} d_i\right)\hat{A} + \hat{B}_{ols}\left(X - \hat{X}_{HT}\right) \tag{1.14}$$

where $\hat{B}_{ols} = \sum_{i\in s} d_i q_i\left(x_i - \bar{x}^*\right)\left(y_i - \bar{y}^*\right)\Big/\sum_{i\in s} d_i q_i\left(x_i - \bar{x}^*\right)^2$ and $\hat{A} = \bar{y}^* - \hat{B}_{ols}\bar{x}^*$ with $\bar{x}^* = \sum_{i\in s} d_i q_i x_i \Big/ \sum_{i\in s} d_i q_i$ and $\bar{y}^* = \sum_{i\in s} d_i q_i y_i \Big/ \sum_{i\in s} d_i q_i$. Note that if $q_i = 1$ then $\hat{B}_{ols} = \hat{\beta}_{ols}$, but for any IPPS scheme (except simple random sampling without replacement) it is the case that $\hat{Y}_{ws} \neq \hat{Y}_{hhm}$. Wu and Sitter (2001) made $\hat{Y}_{ws} = \hat{Y}_{hhm}$ by neglecting a small term, $\left(N - \sum_{i\in s} d_i\right)\hat{A}$, from their estimator (equation no. 9, page 187, JASA). In short, the estimator (1.14) rediscovered by Wu and Sitter (2001) is a special case of the Deville and Sarndal (1992) estimator obtained by setting one auxiliary variable out of *p* at a fixed level. This demonstrates that, for any unequal IPPS sampling scheme, the calibration technique cannot achieve the lower bound for the variance of the traditional linear regression estimator. Using a priori information on $C_y = \sqrt{V\left(\hat{Y}_{HT}\right)/Y^2}$, Searls (1964, 1967) suggested the following estimator of the population total

$$\hat{Y}_{searl} = \hat{Y}_{HT}\Big/\left(1 + C_y^2\right) \tag{1.15}$$

with mean squared error given by

$$MSE\left(\hat{Y}_{searl}\right) = V\left(\hat{Y}_{HT}\right)\Big/\left(1 + C_y^2\right) \tag{1.16}$$

Reddy (1978) studied the properties of (1.15) and has found it to be useful if C_y is large and the sample size is small. A similar conclusion regarding sample size was also reached by Searls (1964) and Arnholt and Hebert (1995). Fay and Herriot (1979) have discussed the importance of the James-Stein (1961) procedure while estimating the income for small places from census data. Note that the estimators due to James-Stein (1961) and Searls (1964) can be shown to be from the same family. Prasad (1989) proposed a ratio type estimator of the population total under an SI design as

$$\bar{y}_{p(r)} = N\ \bar{y}_{searl}\left(\bar{X}/\bar{x}\right) \tag{1.17}$$

where $\bar{x}=n^{-1}\sum_{i=1}^{n} x_i$ is an unbiased estimator of $\bar{X}=N^{-1}\sum_{i=1}^{N} x_i$, while Jain (1987) considered the application of Searls' (1964) estimator to the usual linear regression estimator

$$\bar{y}_{jain} = \alpha\bar{y} + \beta\left(\bar{X} - \bar{x}\right) \tag{1.18}$$

Note that it can be shown that (1.18) has mean squared error:

$$MSE\left(\bar{y}_{jain}\right)=\begin{cases}\left(\dfrac{1-f}{n}\right)\dfrac{S_y^2 S_x^2\left(1-\rho_{xy}^2\right)}{S_x^2 + 2\rho_{xy}S_x S_y + S_y^2}, if \alpha + \beta = 1\\ \left(\dfrac{1-f}{n}\right)\dfrac{S_y^2\left(1-\rho_{xy}^2\right)}{1+\left(\dfrac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)}, \text{otherwise}\end{cases} \tag{1.19}$$

Singh and Srivastava (1980) investigated an unbiased regression estimation strategy for the population total, *Y*, as:

$$\hat{y}_{ss} = \frac{n(N-1)}{(n-1)}\left[\bar{y} - \frac{\sum_{i=1}^{n} y_i\left(x_i - \bar{X}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}\left(\bar{x} - \bar{X}\right)\right] \tag{1.20}$$

In the present investigation, we propose a penalized chi square distance function that produces a general class of estimators for the population total that includes, among others discussed here, the estimators of Searls (1964), Singh and Srivastava (1980), and the famous unbiased ratio estimator of Hartley and Ross (1954) given by

$$\hat{y}_{HR} = \frac{n(N-1)}{(n-1)}\bar{y} + \bar{r}\left[X - \frac{n(N-1)}{n-1}\bar{x}\right] \tag{1.21}$$

where $\bar{r} = \frac{1}{n}\sum_{i=1}^{n} y_i/x_i$.

2. PENALIZED CHI SQUARE DISTANCE FUNCTION

We suggest here the penalized chi square distance function

$$D = \frac{1}{2}\sum_{i \in s}\frac{\left(w_i^* - d_i\right)^2}{d_i q_i^*} + \frac{1}{2}\Phi^2 \sum_{i \in s}\frac{\left(w_i^*\right)^2}{d_i q_i^*} \qquad (2.1)$$

where q_i^* are weights and Φ is a positive quantity that reflects a penalty to be decided by the investigator based on prior knowledge, or the desire for certain levels of efficiency and bias. Different choices for q_i^* result in different estimators, while increasing Φ results in a decrease in the mean square error of the estimator; unfortunately has the side effect of increasing the bias. If $\Phi \to 0$, then the penalized chi square distance function (2.1) reduces to the Deville and Sarndal (1992) distance function. Considering the estimator of population total to be of the form

$$\hat{Y}_{new} = \sum_{i=1}^{n} w_i^* y_i \qquad (2.2)$$

We shall minimize the penalized chi square distance function for five different situations:

1. No auxiliary information is available.
2. Calibration of Deville and Sarndal (1992).
3. Penalized calibration constraint.
4. Unbiased strategy of Singh and Srivastava (1980).
5. Unbiased ratio estimator by Hartley and Ross (1954).

We shall also discuss the estimators belonging to each one of the above situations.

3. NO AUXILIARY INFORMATION AVAILABLE

In the absence of auxiliary information, minimizing (2.1) with respect to w_i^* gives

$$\frac{\partial D}{\partial w_i^*} = \frac{\left(w_i^* - d_i\right)}{d_i q_i^*} + \frac{w_i^* \Phi^2}{d_i q_i^*} = 0 \qquad (3.1)$$

which implies that

$$w_i^* = d_i \Big/ \left(1 + \Phi^2\right) \qquad (3.2)$$

On substituting (3.2) into (2.2), we obtain a new estimator of population total given by

$$\hat{Y}_{new} = \sum_{i \in s} d_i y_i \Big/ \left(1 + \Phi^2\right) = \hat{Y}_{HT} \Big/ \left(1 + \Phi^2\right) \qquad (3.3)$$

The penalized estimator $\hat{Y}_{new}$ is independent of the choice of q_i^*, which indicates that Searls(1964) estimator is a unique estimator in its class. Following Searls (1964), the mean squared error of the penalized estimator is

$$MSE\left(\hat{Y}_{new}\right) = V\left(\hat{Y}_{HT}\right) \Big/ \left(1 + \Phi^2\right) \qquad (3.4)$$

so that the efficiency of (3.3) relative to the Horvitz-Thompson estimator is given by

$$RE = 1 + \Phi^2 \qquad (3.5)$$

The bias in (3.3) is given by

$$B\left(\hat{Y}_{new}\right) = -\left\{\Phi^2 \Big/ \left(1 + \Phi^2\right)\right\} Y \qquad (3.6)$$

Interestingly if $\Phi \to \infty$, then $RE \to \infty$ but $RB\left(\hat{Y}_{new}\right) = B\left(\hat{Y}_{new}\right) \big/ Y \to -1$; thus an increase in the penalty may be advantageous in that the gain in relative efficiency may outweigh the increase in bias. If $\Phi \to 0$ then $RE \to 1$ and $RB\left(\hat{Y}_{new}\right) \to 0$. If $\Phi = C_y$, the penalized estimator reduces to Searls (1964) estimator. In practice the best choice of value for the penalty is in the range from 0 to 1. If $\Phi = 1$, the penalized estimator is 200% more efficient than the Horvitz-Thompson estimator with $RB\left(\hat{Y}_{new}\right) = -0.5$. If $\Phi = 0.5$, the relative efficiency is 125% with $RB\left(\hat{Y}_{new}\right) = -0.2$.

James-Stein (1961) Estimator

If the Y_i are assumed to be independent and identically distributed according to a normal distribution with mean θ_i and variance D, then each Y_i is the obvious estimate of its respective θ_i. For k > 3, James and Stein (1961) defined

$$\delta_i^{'} = \left[1 - \frac{(k-2)D}{S}\right] Y_i \qquad (3.7)$$

as an estimator of θ_i, with $S = \sum_i Y_i^2$. Note that if $\hat{Y}_{HT} = Y_i$ and $\Phi^2 = \{(k-2)D/S\} / \{1 - (k-2)D/S\}$, then the penalized estimator $\hat{Y}_{new}$ in (3.3) reduces to James and Stein (1961) estimator.

Fay and Herriot (1979) Estimator

Following James and Stein (1961), if $Y_i \sim_{ind} N(\theta_i, D)$ and $\theta_i \sim_{ind} N(X_i \beta, A)$, Fay and Herriot (1979) combined the regression estimator $Y_i^* = X_i (X'X)^{-1} X'Y$ with the direct estimator Y_i to form an empirical Bayes estimator of θ_i as

$$\theta_i^* = Y_i^* + \left(1 - \frac{D}{D + A}\right)\left(Y_i - Y_i^*\right) \qquad (3.8)$$

Similarly, the convex combination of the proposed penalized estimator with the empirical Bayes estimator will lead to the estimator of Fay and Herriot (1979)

$$\hat{Y}_{hf} = \hat{Y}_{new} + \left(1 - \frac{1}{1+\Phi^2}\right)\hat{Y}_{bayes} \qquad (3.9)$$

The estimator (3.9) is also called a composite estimator and plays an eminent role in small area estimation.

4. CALIBRATION OF DEVILLE AND SARNDAL

In order to minimize the penalized chi-square distance function subject to the calibration constraint of Deville and Sarndal (1992), we consider the Lagrange function

$$L = \frac{1}{2}\sum_{i\in s}\frac{\left(w_i^* - d_i\right)^2}{d_i q_i^*} + \frac{1}{2}\Phi^2 \sum_{i\in s}\frac{w_i^{*2}}{d_i q_i^*} - \lambda\left\{\sum_{i\in s} w_i^* x_i - X\right\} \qquad (4.1)$$

Setting $\partial L / \partial w_i^* = 0$, yields

$$w_i^* = \frac{1}{\left(1+\Phi^2\right)}\left[d_i + \frac{d_i q_i^* x_i}{\sum_{i\in s} d_i q_i^* x_i^2}\left\{\left(1+\Phi^2\right)X - \sum_{i\in s} d_i x_i\right\}\right] \qquad (4.2)$$

so that a penalized estimator of population total *Y* is given by

$$\hat{Y}_{new} = \frac{\sum_{i\in s} d_i y_i}{\left(1+\Phi^2\right)} + \frac{\sum_{i\in s} d_i q_i^* x_i y_i}{\sum_{i\in s} d_i q_i^* x_i^2}\left\{X - \frac{\sum_{i\in s} d_i x_i}{\left(1+\Phi^2\right)}\right\} \qquad (4.3)$$

Note that if $q_i^* = 1/x_i$, then (4.3) reduces to the usual ratio estimator of the population total, namely

$$\hat{Y}_{new} = \hat{Y}_{HT}\left(X / \hat{X}_{HT}\right) \qquad (4.4)$$

The mean squared error of the estimator in (4.3) is given by

$$MSE\left(\hat{Y}_{new}\right) = V\left(\hat{Y}_{ds}\right) / \left(1+\Phi^2\right) \qquad (4.5)$$

where $V\left(\hat{Y}_{ds}\right) = \frac{1}{2}\sum_{i\neq j}\sum_{\in\Omega}\Theta_{ij}\left(d_i e_i - d_j e_j\right)^2$ and $\Phi^2 = V\left(\hat{Y}_{ds}\right) / \left(Y - \beta X\right)^2$. It is interesting to note that if $\beta \to Y/X$, then $\Phi \to \infty$ and the relative efficiency also approaches infinity; however this may have a serious adverse effect on bias. If Φ is known, then an estimator for the mean squared error of (4.3) is

$$MS\hat{E}\left(\hat{Y}_{new}\right) = \hat{V}\left(\hat{Y}_{ds}\right) / \left(1+\Phi^2\right) \qquad (4.6)$$

where $\hat{V}\left(\hat{Y}_{ds}\right) = \frac{1}{2}\sum_{i\neq j}\sum_{\in s} D_{ij}\left(w_i^* e_i - w_j^* e_j\right)^2$ is a new penalized estimator of variance of the Deville and Sarndal (1992) estimator.

PRODUCT METHOD OF ESTIMATION

The problem of estimating the product of two variables is well known when the two variables are negatively correlated. For example, an estimate of the force, *F*, of certain objects is given by $\hat{F} = \hat{m} \times \hat{a}$, where $\hat{m}$ and $\hat{a}$ are the average mass and acceleration. For further details on product estimation, see Robson (1957) and Murthy (1964). Minimization of (2.1) subject to a new calibration constraint, defined as

$$X \sum_{i\in s} w_i^* = \hat{X}_{HT} \qquad (4.7)$$

leads to calibrated weights given by

$$w_i^* = \frac{1}{1+\Phi^2}\left[d_i + \frac{d_i q_i^*\left\{\frac{\hat{X}_{HT}}{X}\left(1+\Phi^2\right) - \sum_{i\in s} d_i\right\}}{\sum_{i\in s} d_i q_i^*}\right] \qquad (4.8)$$

and the following penalized estimator of the population total

$$\hat{Y}_{new} = \frac{1}{\left(1+\Phi^2\right)}\left[\sum_{i\in s} d_i y_i + \frac{\sum_{i\in s} d_i q_i^* y_i}{\sum_{i\in s} d_i q_i^*}\left(\frac{\hat{X}_{HT}}{X}\left(1+\Phi^2\right) - \sum_{i\in s} d_i\right)\right] \qquad (4.9)$$

If $q_i^* = 1$, then

$$\hat{Y}_{new} = \hat{Y}_{HT}\left(\hat{X}_{HT} / X\right) \qquad (4.10)$$

which is same product estimator studied by Murthy (1964).

5. PENALIZED CALIBRATION CONSTRAINT

We suggest here a new penalized calibration constraint as:

$$\sum_{i\in s} w_i^* x_i = X\left(1+\Phi^2\right)^{-1} \qquad (5.1)$$

Minimization of the penalized chi square distance function (2.1) subject to (5.1) leads to the calibrated weights

$$w_i^* = \frac{1}{\left(1+\Phi^2\right)}\left[d_i + \frac{d_i q_i^* x_i}{\sum_{i\in s} d_i q_i^* x_i^2}\left(X - \sum_{i\in s} d_i x_i\right)\right] \qquad (5.2)$$

The resultant penalized estimator of the population total is:

$$\hat{Y}_{new} = \frac{1}{\left(1+\Phi^2\right)}\left[\sum_{i\in s} d_i y_i + \frac{\sum_{i\in s} d_i q_i^* x_i y_i}{\sum_{i\in s} d_i q_i^* x_i^2}\left(X - \sum_{i\in s} d_i x_i\right)\right] \qquad (5.3)$$

while the minimum mean square error of (5.3) is

$$MSE\left(\hat{Y}_{new}\right) = V\left(\hat{Y}_{ds}\right) / \left(1+\Phi^2\right) \qquad (5.4)$$

where $\Phi^2 = V\left(\hat{Y}_{ds}\right) / Y^2$. An estimator of the variance of (5.3) is

$$MS\hat{E}\left(\hat{Y}_{new}\right)=\hat{V}\left(\hat{Y}_{ds}\right)\Big/\left(1+\Phi^{2}\right) \tag{5.5}$$

where $\hat{V}\left(\hat{Y}_{ds}\right)=\frac{1}{2}\sum_{i\neq j}\sum_{\in s}D_{ij}\left(w_i^{*}e_i-w_j^{*}e_j\right)^{2}$. Note that (5.3) is similar to the estimator studied by Jain (1987). If $q_i^{*}=1/x_i$, then it reduces to a penalized ratio estimator, as

$$\hat{Y}_{new}=\frac{1}{\left(1+\Phi^{2}\right)}\hat{Y}_{HT}\left(X/\hat{X}_{HT}\right) \tag{5.6}$$

which is similar to the estimator studied by Prasad (1989). Following Prasad (1989), the penalty in (5.6) can be taken as $\Phi^{2}=V\left(\hat{Y}_{HT}\right)/Y^{2}$. Similarly, minimization of (2.1) with respect to the calibration constraint $X\sum_{i\in s}w_i^{*}=\hat{X}_{HT}\left(1+\Phi^{2}\right)$ leads to

$$\hat{Y}_{new}=\frac{1}{\left(1+\Phi^{2}\right)}\hat{Y}_{HT}\left(\frac{\hat{X}_{HT}}{X}\right) \tag{5.7}$$

which is a new estimator of the population total.

6. SINGH AND SRIVASTAVA's UNBIASED ESTIMATION STRATEGY

Here we consider a slightly different penalized distance function given by

$$D^{*}=\frac{1}{2}\sum_{i\in s}\frac{\left(w_i^{*}-d_i\right)^{2}}{d_iq_i^{*}}-\frac{1}{2}\Phi^{2}\sum_{i\in s}\frac{(w_i^{*})^{2}}{d_iq_i^{*}} \tag{6.1}$$

and suggest a new calibration constraint

$$\sum_{i\in s}w_i^{*}\left(x_i-\overline{X}\right)=0 \tag{6.2}$$

where $\overline{X}=\sum_{i\in\Omega}d_ix_i\Big/\sum_{i\in\Omega}d_i$, so that the Lagrange function is

$$L=\frac{1}{2}\sum_{i\in s}\frac{\left(w_i^{*}-d_i\right)^{2}}{d_iq_i^{*}}-\frac{\Phi^{2}}{2}\sum_{i\in s}\frac{w_i^{*2}}{d_iq_i^{*}}-\lambda\sum_{i\in s}w_i^{*}\left(x_i-\overline{X}\right) \tag{6.3}$$

Setting $\partial L\big/\partial w_i^{*}=0$ yields

$$w_i^{*}=\left\{d_i+\lambda\, d_iq_i^{*}\left(x_i-\overline{X}\right)\right\}\big/\left(1-\Phi^{2}\right) \tag{6.4}$$

Substituting (6.4) in (6.2) gives

$$\lambda=-\sum_{i\in s}d_i\left(x_i-\overline{X}\right)\Big/\sum_{i\in s}d_iq_i^{*}\left(x_i-\overline{X}\right)^{2} \tag{6.5}$$

so that the minimal weights are

$$w_i^{*}=\frac{1}{1-\Phi^{2}}\left\{d_i-\frac{d_iq_i^{*}\left(x_i-\overline{X}\right)}{\sum_{i\in s}d_iq_i^{*}\left(x_i-\overline{X}\right)^{2}}\sum_{i\in s}d_i\left(x_i-\overline{X}\right)\right\} \tag{6.6}$$

which satisfy the condition of minimal distance if

$$\partial^{2}L\big/\partial w_i^{*2}=\left(1-\Phi^{2}\right)>0 \tag{6.7}$$

Thus a new penalized estimator of the population total is

$$\hat{Y}_{new}=\frac{1}{\left(1-\Phi^{2}\right)}\left[\sum_{i\in s}d_iy_i-\frac{\sum_{i\in s}d_iq_i^{*}y_i\left(x_i-\overline{X}\right)}{\sum_{i\in s}d_iq_i^{*}\left(x_i-\overline{X}\right)^{2}}\sum_{i\in s}d_i\left(x_i-\overline{X}\right)\right] \tag{6.8}$$

Under SRSWOR sampling, $d_i=N/n$ and if $\Phi^{2}=\{n(N-2)+1\}/\{n(N-1)\}$ and $q_i^{*}=1$, then (6.8) becomes

$$\hat{Y}_{new}=\frac{n(N-1)}{(n-1)}\left[\overline{y}-\frac{\sum_{i=1}^{n}y_i\left(x_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(x_i-\overline{X}\right)^{2}}\left(\overline{x}-\overline{X}\right)\right] \tag{6.9}$$

which is identical to the unbiased regression type estimator proposed by Singh and Srivastava (1980). Note that $\Phi^{2}=\{n(N-2)+1\}/\{n(N-1)\}$ lies between 0 and 1 and hence satisfies the condition of minimal distance.

7. HARTLEY AND ROSS ESTIMATOR

In order to minimize (6.1) subject to the Deville and Sarndal (1992) calibration constraint, we consider the Lagrange function

$$L=\frac{1}{2}\sum_{i\in s}\frac{\left(w_i^{*}-d_i\right)^{2}}{d_iq_i^{*}}-\frac{1}{2}\Phi^{2}\sum_{i\in s}\frac{w_i^{*2}}{d_iq_i^{*}}-\lambda\left\{\sum_{i\in s}w_i^{*}x_i-X\right\} \tag{7.1}$$

Setting $\partial L\big/\partial w_i^{*}=0$, yields

$$w_i^{*}=\frac{1}{\left(1-\Phi^{2}\right)}\left[d_i+\frac{d_iq_i^{*}x_i}{\sum_{i\in s}d_iq_i^{*}x_i^{2}}\left\{\left(1-\Phi^{2}\right)X-\sum_{i\in s}d_ix_i\right\}\right] \tag{7.2}$$

so that a penalized estimator of the population total Y is:

$$\hat{Y}_{new}=\frac{\sum_{i\in s}d_iy_i}{\left(1-\Phi^{2}\right)}+\frac{\sum_{i\in s}d_iq_i^{*}x_iy_i}{\sum_{i\in s}d_iq_i^{*}x_i^{2}}\left\{X-\frac{\sum_{i\in s}d_ix_i}{\left(1-\Phi^{2}\right)}\right\} \tag{7.3}$$

Under SRSWOR sampling, if $q_i^{*}=1/x_i^{2}$ and $\Phi^{2}=\{n(N-2)+1\}/\{n(N-1)\}$ then (7.3) reduces to

$$\hat{Y}_{new}=\frac{n(N-1)}{(n-1)}\overline{y}+\overline{r}\left[X-\frac{n(N-1)}{n-1}\overline{x}\right] \tag{7.4}$$

which is equivalent to the Hartley and Ross (1954) estimator.

CONCLUSION

A penalized chi square distance function has been proposed that covers a wider variety of estimators than the original chi square function introduced by Deville and Sarndal (1992). The proposed function produces a general class of estimators for the population total that includes the estimators of Searls (1964), Singh and Srivastava (1980), and the famous unbiased ratio estimator of Hartley and Ross (1954), among others.

FURTHER STUDY

The extension of one-dimensional penalized chi-square distance function to two-dimensional penalized chi-square distance function and study of resultant estimators on the lines of Singh, Horn and Yu (1998) is in progress. Note that we have considered only simplest penalized chi-square distance function, but any one among the distance functions discussed by Deville and Sarndal (1992) can be penalized, and will lead to more interesting estimators.

ACKNOWLEDGEMENTS

This research was supported by the Natural Sciences and Engineering Council of Canada. The second author was post-doc at Carleton University during this investigation.

REFERENCES

Arnholt, A.T. and Hebert, J.L. (1995). Estimating the mean with known coefficient of variation. *American Statistician*, 49(4), 367-369.

Cochran, W.G. (1977). *Sampling Techniques, Third Edition*. New York: Wiley.

Deville, J.C. and Sarndal, C.E.(1992).Calibration estimator in survey sampling. *J. Amer. Statist. Assoc*., 87, 376-382.

Hartley, H.O. and Ross, A (1954). Unbiased ratio estimators. *Nature*, 174, 270-271.

Horvitz, D.G. and Thompson, D.J. (1952). A generalisation of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc*., 47, 663-685.

Jain, R.K. (1987). Properties of estimators in simple random sampling using auxiliary variable. *Metron,* 265-271.

James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability,* Vol. 1, University of California Press, 361-379.

Murthy, M.N. (1964). Product method of estimation. *Sankhya*, 26, 69-74.

Prasad, B. (1989). Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Commun. Statist. Theory Meth.*, 18(1), 379-392.

Reddy, V.N. (1979). A study on the use of prior knowledge on certain population parameters in estimation. *Sankhya*, C, 40, 29-37.

Robson, D.S. (1957). Application of multivariate polykays to the theory of unbiased ratio-type estimation. *J. Amer. Statist. Assoc.*, 52, 511-522.

Sarndal, C.E. (1996). Efficient estimators with simple variance in unequal probability sampling. *J. Amer. Statist. Assoc*., 91, 1289-1300.

Sarndal, C.E., Swensson, B. and Wretman, J.H. (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika*, 76(3), 527-537.

Searls, D.T. (1964). The utilization of a known co-efficient of variation in the estimation procedure. *J. Amer. Stat. Assoc*., 59, 1225-1226.

Searls, D.T. (1967). A note on the use of an approximately known co-efficient of variation. *American Statistician,*21 (2), 20-21.

Singh, P. and Srivastava, A. K. (1980). Sampling schemes providing unbiased regression estimators. *Biometrika*, 67, 205-209.

Singh, S., Horn, S. and Yu, F. (1998). Estimation of variance of general regression estimator : Higher level calibration approach. Survey Methodology, 24, 41-50.

Wu, C. and Sitter, R.R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *J. Amer. Statist. Assoc*., 96, 185-193.

-☺-Patrick (613)-520-2600 Ext. 1804 **E-mail** : pfarrell@math.carleton.ca	-☺-Sarjinder (320)-654-5324 **E-mail:** sarjinder@yahoo.com