SSC Annual Meeting, June 2000
Proceedings of the Survey Methods Section, pp. 209-214

OBTAINING RANGE RESTRICTED WEIGHTS IN REGRESSION ESTIMATION

Jiahua Chen, Randy R. Sitter and Changbao Wu[1]

ABSTRACT

In this work we present algorithms for computing empirical likelihood estimators and model-calibrated empirical likelihood estimators. Both estimators are asymptotically equivalent to the generalized regression estimator but with intrinsically positive weights. The first algorithm solves the computational problem of the empirical likelihood method in general, both in survey and non-survey settings, and theoretically guarantees its convergence. The second algorithm takes advantage of special properties of the model calibration method and is particularly simple and efficient. The algorithms are then adapted to handle more general pre-specified range restrictions on the weights.

KEY WORDS: Complex survey; Empirical likelihood methods; Generalized regression estimator; Model-calibration.

RÉSUMÉ

Nous présentons ici des algorithmes pour le calcul d'estimateurs empiriques de vraisemblance et de calage par modèle. Les deux estimateurs sont asymptotiquement équivalents à l'estimateur de régression généralisée mais avec intrinsèquement des poids positifs. Le premier algorithme résouds des problèmes de calcul de la méthode empirique de vraisemblance en général, pour les enquêtes probabilistes et non-probabilistes, et garantie théoriquement sa convergence. Le deuxième algorithme tire avantage des propriétés spéciales de la méthode de calage et est particulièrement simple et efficace. Les algorithmes sont alors modifiés pour tenir compte d'un intervalle de restrictions pré-spécifées sur les poids.

MOTS CLÉS : Enquête complexe; méthodes empiriques de vraisemblance; estimateur de régression generalisé; calibration par modèle.

1. INTRODUCTION

There are three major issues associated with estimation procedures in using auxiliary information: efficiency, simplicity and consistency. Efficiency is measured by the overall performance of the estimator in terms of bias and variance or mean square error, simplicity concerns the computational aspect of the procedure, while consistency refers to some internal conditions and requirements imposed by the surveyor. There are two commonly used consistency requirements, benchmark constraints and range restrictions, on the adjusted weights. The benchmark constraints require the adjusted weights w_i give perfect estimates when applied to auxiliary variables, i.e. $\sum_{i\in s} w_i x_i = X$, where x is a vector of auxiliary variables and X is the vector of the known corresponding population totals; the range restrictions require $\gamma_1 \le w_i / d_i \le \gamma_2$ for some pre-specified constants $0 \le \gamma_1 < 1$ and $\gamma_2 > 1$. That is, the adjusted weights are not allowed to deviate too far away from the basic design weights. The choices of $\gamma_1 = 0$ and $\gamma_2 = \infty$ represent the requirement of positive weights.

The generalized regression estimator (GREG) for finite population means or totals has been widely used in survey sampling when known totals of auxiliary

[1] J. Chen and C. Wu, Department of Statistics & Actuarial Science, University of Waterloo, Waterloo, ON, Canada, N2L 3G1, cbwu@uwaterloo.ca; R.R. Sitter, Department of Statistics & Actuarial Science, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

variables are available from complex survey data. The GREG adjusted weights $w_i^{(r)}$ satisfy the benchmark constraints, but some of them can be negative, or may not meet the pre-specified range restrictions. Negative weights could result in a negative estimate for a known positive quantity, and "negative sample weights are apt to breed mistrust of one's estimates in one's colleagues and clients!" (Bardsley & Chambers, 1984). This drawback of GREG's has long been recognized in the literature and the problem has been attacked by several authors, notably the work by Huang & Fuller (1978), Singh & Mohl (1996) and Rao & Singh (1997).

It should be noted that the problem of positive weights or range-restricted weights is of practical importance only. One could simply use a calibration estimator with a distance measure that forces the weights to be positive or within certain ranges. The problem is the level of difficulty of the corresponding constrained minimization problem, i.e. the simplicity of the method. The range restrictions are not a problem in large samples either, since the GREG weights have the property that $w_i^{(r)} = d_i\{1+o(1)\}$ under proper moment conditions. Hence, asymptotically any range restrictions will be satisfied with large enough samples. Once this is realized, and since the methods are all asymptotically equivalent, the choice becomes one of operational ease, irrespective of the philosophical motivations behind the various methods.

In this paper we propose two simple algorithms using the recent development of the pseudo-empirical likelihood method (Chen & Sitter, 1999; see also Zhong & Rao, 1996) and the model-calibration method (Wu & Sitter, 2000). The pseudo-empirical maximum likelihood estimators (PEML) and the model-calibrated pseudo-empirical maximum likelihood estimators (MCPE) are asymptotically equivalent to the GREG but with intrinsically positive weights. Our algorithms for obtaining PEML and MCPE are simple and convergences are guaranteed. We also develop a simple adaptation of the algorithms to handle the more general range restrictions on the weights. The simplicity of the algorithms and the theoretical support for the convergences make these two methods serious competitors to the existing methods.

2. PEML AND MCPE ESTIMATORS

Consider a finite population consisting of N identifiable units. Associated with the ith unit are, the study variable, y_i, and a vector of auxiliary variables, $\boldsymbol{x}_i$. The values of $\{y_i, \boldsymbol{x}_i, i\in s\}$ and the population mean of the $\boldsymbol{x}_i$, $\overline{\boldsymbol{X}} = N^{-1}\sum_{i=1}^{N}\boldsymbol{x}_i$, are known. Assume the inclusion probabilities $\pi_i = \Pr(i\in s)$ are strictly positive. We will restrict attention to estimating the population mean $\overline{Y} = N^{-1}\sum_{i=1}^{N} y_i$.

A GREG estimator of $\overline{Y}$ is given by

$$\hat{\overline{Y}}_{GR} = \frac{1}{N}\sum_{i\in s} w_i y_i = \hat{\overline{Y}}_{HT} + (\overline{\boldsymbol{X}} - \hat{\overline{\boldsymbol{X}}}_{HT})'\hat{\boldsymbol{\theta}}, \tag{1}$$

where $\hat{\overline{\boldsymbol{X}}}_{HT} = N^{-1}\sum_{i\in s} d_i\boldsymbol{x}_i$ and $\hat{\overline{Y}}_{HT} = N^{-1}\sum_{i\in s} d_i y_i$ are the usual Horvitz-Thompson estimators,

$$\hat{\boldsymbol{\theta}} = \{\sum_{i\in s} d_i(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})'\}^{-1}\sum_{i\in s} d_i(\boldsymbol{x}_i - \overline{\boldsymbol{x}})y_i, \tag{2}$$

and

$$w_i = d_i[1 + (\overline{\boldsymbol{X}} - \hat{\overline{\boldsymbol{X}}}_{HT})'\{N^{-1}\sum_{i\in s} d_i(\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})'\}^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})]. \tag{3}$$

Note that these weights only depend on the $\boldsymbol{x}_i$'s and not the y_i's and satisfy the benchmark constraints $N^{-1}\sum_{i\in s} w_i\boldsymbol{x}_i = \overline{\boldsymbol{X}}$.

The pseudo-empirical maximum likelihood estimator proposed by Chen & Sitter (1999) is $\hat{\overline{Y}}_{PE} = \sum_{i\in s}\hat{p}_i y_i$. The weights, $\hat{p}_i$'s, are obtained by maximizing

$$\hat{l}(\boldsymbol{p}) = \sum_{i\in s} d_i \log p_i, \tag{4}$$

subject to

$$\sum_{i\in s} p_i = 1,\ \sum_{i\in s} p_i\boldsymbol{x}_i = \overline{\boldsymbol{X}}\ \ (0\le p_i\le 1). \tag{5}$$

The pseudo empirical likelihood $\hat{l}(\boldsymbol{p})$ is a design-unbiased estimator of the log-empirical likelihood one would use if one had the entire population: $E_p(\sum_{i\in s} d_i\log p_i) = \sum_{i=1}^{N}\log p_i$. Here, E_p refers to the expectation under the sampling design. The adjusted weights $\hat{p}_i$ are obtained using a vector-valued Lagrange multiplier approach, which yields

$$\hat{p}_i = \frac{d_i^*}{1+\lambda'(\boldsymbol{x}_i - \overline{\boldsymbol{X}})}\quad \text{for } i\in s, \tag{6}$$

where $d_i^* = d_i/\sum_{i\in s} d_i$ and the vector Lagrange multiplier, λ, is the solution to

$$g_1(\lambda)=\sum_{i\in s}\frac{d_i^*(x_i-\overline{X})}{1+\lambda'(x_i-\overline{X})}=0.$$

Chen & Sitter (1999) show that the resulting estimator, $\hat{\overline{Y}}_{PE}$, is asymptotically equivalent to $\hat{\overline{Y}}_{GR}$, but by definition the weights $\hat{p}_i>0$.

Wu & Sitter (2000) introduce a model calibration idea to the PEML approach to reduce the high dimensional computation problem to a scalar one. When the working model is linear, they obtain predicted values $\hat{\mu}_i=x_i'\hat{\theta}$ for y_i, where $\hat{\theta}$ is given in (2), and then calibrate on the predicted values using the pseudo-empirical likelihood approach. That is, choose the $\hat{p}_i$ to maximize (4) but replace the constraints in (5) by

$$\sum_{i\in s}p_i=1 \quad\text{and}\quad \sum_{i\in s}p_i\,\hat{\mu}_i=N^{-1}\sum_{i=1}^{N}\hat{\mu}_i \qquad (7)$$

It is important to note that the second constraint in (7) reduces to $\sum_{i\in s}p_ix_i'\hat{\theta}=\overline{X}'\hat{\theta}$ and thus only $\overline{X}$ need be known for the entire population, despite being motivated by predicting y_i at each x_i in the population. This constrained maximization can now be solved using a scalar Lagrange multiplier. It can be shown that the resulting $\hat{p}_i$'s satisfy

$$\hat{p}_i=\frac{d_i^*}{1+\lambda(x_i-\overline{X})'\hat{\theta}}\quad\text{for } i\in s,$$

and the scalar Lagrange multiplier, λ, is the solution to

$$g_2(\lambda)=\sum_{i\in s}\frac{d_i^*(x_i-\overline{X})'\hat{\theta}}{1+\lambda(x_i-\overline{X})'\hat{\theta}}=0. \qquad (8)$$

The general results of Wu & Sitter (2000) then imply that the resulting estimator of $\overline{Y}$, given by $\hat{\overline{Y}}_{MC}=\sum_{i\in s}\hat{p}_iy_i$ is asymptotically equivalent to $\hat{\overline{Y}}_{GR}$, but by definition the weights $\hat{p}_i>0$.

Thus, the PEML approach of Chen & Sitter (1999) combined with the model-calibration strategy of Wu & Sitter (2000) will yield an approximate GREG with strictly positive weights. As we will show in Section 4, the use of model-calibration simplifies the maximization problem to one that can be solved using a particularly simple and elegant algorithm. Purely from an efficiency and simplicity point of view this makes the MCPE combined with this algorithm a superior method. However, since the constraint is imposed on the predicted values rather than on each of the x variables, the benchmark constraints may be slightly violated for individual auxiliary variables. Also, the resulting weights depend upon the y variable implicitly through the estimated model parameters and therefore different sets of weights have to be attached to different response variables. If exact benchmark constraints and a single set of weights for many response variables are of great interest, the extra computational complexity and effort in PEML of Section 3 as well as those methods like Huang & Fuller (1978) or Singh (1993) may be necessary.

3. AN ALGORITHM TO OBTAIN WEIGHTS IN PEML

Numerically, the key to the pseudo empirical likelihood problem is to find the vector solution of $g_1(\lambda)=\mathbf{0}$ within the range of λ such that the resulting $\hat{p}_i>0$ for all i. A necessary and sufficient condition for the existence of the solution is that the convex hull of $\{x_i : i\in s\}$ contains $\overline{X}$ as an inner point. This condition is satisfied with probability approaching one for most sampling designs and conceivable populations as the sample and population sizes increase (see Chen & Sitter, 1999, for details).

For notational simplicity and without loss of generality, assume $\overline{X}=\mathbf{0}$. If not, replace x_i by $x_i-\overline{X}$ throughout. Let

$$\tilde{l}(\lambda)=\sum_{i\in s}d_i^*\log(1+\lambda'x_i).$$

Then $\tilde{l}(\lambda)$ is a concave function and its maximum point λ satisfies the same equation $g_1(\lambda)=\mathbf{0}$, and $\hat{p}_i$ given by (6) will satisfy $\hat{p}_i>0$ and $\sum_{i\in s}\hat{p}_i=1$. That is, maximizing $\tilde{l}(\lambda)$ is a dual problem of maximizing $\tilde{l}(p)$ subject to constraints (5).

We now present an algorithm for finding the solution to $g_1(\lambda)=\mathbf{0}$. The algorithm can be viewed as a modified Newton-Raphson algorithm. We use $\|\cdot\|$ for the Euclidean norm.

Step 0: Let $\lambda_0=\mathbf{0}$. Set $k=0, \gamma_0=1$ and $\varepsilon=10^{-8}$.

Step 1: Calculate $\Delta_1(\lambda_k)$ and $\Delta_2(\lambda_k)$ where

$$\Delta_1(\lambda) = \frac{\partial \tilde{l}}{\partial \lambda} = \sum_{i \in s} d_i^* \frac{x_i}{1 + \lambda' x_i} \text{ and}$$

$$\Delta_2(\lambda) = \left(\frac{\partial^2 \tilde{l}}{\partial \lambda \partial \lambda'} \right)^{-1} \Delta_1 = \left\{ -\sum_{i \in s} d_i \frac{x_i x_i'}{(1 + \lambda' x_i)^2} \right\}^{-1} \Delta_1(\lambda)$$

If , $\|\Delta_2(\lambda_k)\| < \varepsilon$ stop the algorithm and report λ_k; otherwise go to Step 2.

Step 2: Calculate $\delta_k = \gamma_k \Delta_2(\lambda_k)$. If $1 + (\lambda_k - \delta_k)' x_i \le 0$ for some i or $\tilde{l}(\lambda_k - \delta_k) < \tilde{l}(\lambda_k)$, let $\gamma_k = \gamma_k / 2$ and repeat Step 2.

Step 3: Set $\lambda_{k+1} = \lambda_k - \delta_k, k = k+1$ and $\gamma_{k+1} = (k+1)^{-1/2}$. Go to Step 1.

The above algorithm is most similar to the modified Newton's method described in Polyak (1987, pg. 63). Such algorithms for minimizing a convex function or maximizing a concave function almost always converge. Theoretically, some mild conditions are however needed to guarantee the convergence. These amount to boundedness of the first derivative and the inverse of the second derivative of the objective function and that the second derivative satisfies the Lipschitz condition. When these conditions are satisfied, a sufficiently short step size can be determined so that the norm of the first derivative is always reduced after each iteration. For details, we refer to Chen, Sitter and Wu (2000).

4. A SIMPLE ALGORITHM FOR OBTAINING WEIGHTS IN MCPE

One of the advantages of using the MCPE is the computational simplicity. There exists an extremely simple and stable algorithm which will identify when a solution exists and will always yield a solution if a solution exists.

To solve $g_2(\lambda) = 0$ in (8), let $b_i = (x_i - \overline{X})' \hat{\theta}$ for $i \in s$, and let $s^* = \{i : b_i \ne 0\}$,

$$g_2(\lambda) = \sum_{i \in s^*} \frac{d_i^*}{\lambda - a_i},$$

where $a_i = -1/b_i$ *for* $i \in s^*$. We observe that: (i) $g_2(\lambda)$ is a piece-wise monotone decreasing function of λ; and (ii) the a_i's are the singular points of $g_2(\lambda)$. There will be a unique solution to $g_2(\lambda) = 0$ in each of the intervals (a_j, a_{j+1}), $j = 1, \ldots, k$. Note that $p_i = d_i^* / \{1 + \lambda(x_i - \overline{X})' \hat{\theta}$, and $p_i > 0$ implies $\lambda > a_i$ for $a_i < 0$ and $\lambda < a_i$ for $a_i > 0$. So the unique solution to our problem lies in the unique interval (a_m, a_{m+1}) which contains 0. The solution can then be found using a bi-section method. If the $\max_{i \in s} a_i > 0$ or $\max_{i \in s} a_i < 0$, there will be no solution to the constrained maximization problem. This happens with probability $o(1)$.

The algorithm can be summarized as follows:

Step 1. Compute the estimated regression coefficients $\hat{\theta}$ as in (2); compute $b_i = (x_i - \overline{X})' \hat{\theta}$ for $i \in s$; and let $b_L = \min\{b_i : i \in s\}$ and $b_U = \max\{b_i : i \in s\}$;

Step 2. If $b_L \ge 0$ or $b_U \le 0$, stop. No solution to the constrained maximization problem exists. Otherwise proceed to Step 3;

Step 3. Find the solution to $g_2(\lambda) = 0$ in the interval $(-1/b_U, -1/b_L)$ using the bi-section method: (i) Let $L = -1/b_U, R = -1/b_L$ and $\varepsilon = 10^{-8}$; (ii) Let $M = (L+R)/2$. If $|g_2(M)| \le \varepsilon$, stop, report $\lambda = M$; otherwise let $L = M$ if $g_2 = (M) > 0$, or let $R = M$ if $g_2 = (M) < 0$; (iii) Repeat (ii) until $|g_2(M)| \le \varepsilon$ and report $\lambda = M$.

5. GENERAL RANGE RESTRICTIONS ON WEIGHTS

In some cases it is desirable to place restrictions on the adjusted weights so as not to allow them to be too different from the basic design weights. Suppose we wish to restrict the weights to fall in the range $\gamma_1 d_i^* \le \hat{p}_i \le \gamma_2 d_i^*$ where $0 \le \gamma_1 < 1 < \gamma_2$ and $d_i^* = d_i / \sum_{i \in s} d_i$ as before, it is possible for pre-specified γ_1 and γ_2 that some of the resulting $\hat{p}_i$'s from one or both of the algorithms presented in Sections 3 and 4 might not meet the requirement. In this section we present a simple adaptation of the algorithms of Sections 3 and 4 to handle this problem. We follow the idea of Rao & Singh (1997) by relaxing the benchmark constraints. Our method uses a minimum relaxation of benchmark constraints to meet the pre-specified range restriction requirement while maintaining the empirical likelihood interpretation. The simplicity of the algorithm remains as before and a solution is guaranteed.

5.1 PEML and MCPE

Without loss of generality, we assume $\overline{X} = \mathbf{0}$ and a solution to PEML exists. The benchmark constraints become $\sum_{i\in s} p_i \boldsymbol{x}_i = \boldsymbol{0}$ (or equivalently, replace $\boldsymbol{x}_i$ with $\boldsymbol{x}_i - \overline{\boldsymbol{X}}$ throughout, including in the definition of $\hat{\overline{\boldsymbol{X}}}_d$). If the weights under these benchmark constraints do not meet the range restrictions or the convex hull of $\{x_i : i \in s\}$ does not contain $\overline{\boldsymbol{X}}$ so that the benchmark constraints are unattainable, we would like to relax the benchmark constraints by using $\sum_{i\in s} p_i \boldsymbol{x}_i = \boldsymbol{t}$ for some $\boldsymbol{t}$ that differs from $\boldsymbol{0}$. Note that a solution always exists if we choose $\boldsymbol{t} = \hat{\overline{\boldsymbol{X}}}_d = \sum_{i\in s} d_i^* \boldsymbol{x}_i$, which amounts to removing the benchmark constraints, and the solution, which is given by $p_i = d_i^*$, will automatically meet the range restrictions. Thus we can choose $\boldsymbol{t} = \delta \hat{\overline{\boldsymbol{X}}}_d$ with the smallest possible δ such that the resulting weights meet the range restrictions. That is, we can move $\boldsymbol{t}$ away from $\boldsymbol{0}$ in the direction of $\hat{\overline{\boldsymbol{X}}}_d$. Since $0 \le \delta \le 1$, the simple and stable bi-section method can be used to search for this δ.

Step 0. Run the algorithm given in Section 3 to get the $\hat{p}_i$'s. If a solution exists and $\gamma_1 d_i^* \le \hat{p}_i \le \gamma_2 d_i^*$ for $i \in s$, stop. The PEML weights meet both the benchmark constraints and the range restrictions.

Step 1. Let $L = 0, R = 1$ and $\varepsilon = 10^{-5}$;

Step 2. Let $\delta = (L+R)/2$ and $\boldsymbol{t} = \delta \hat{\overline{\boldsymbol{X}}}_d$. Run the algorithm given in Section 3 with constraint $\sum_{i\in s} p_i \boldsymbol{x}_i = \boldsymbol{t}$ to get the p_i's.

Step 3. If $p_i \notin [\gamma_1 d_i^*, \gamma_2 d_i^*]$ for some $i \in s$, let $L = \delta$ and go to Step 2; otherwise, let $R = \delta$. If $R - L < \varepsilon$, stop and use these weights; otherwise go to Step 2.

One should note the above algorithm will always yield a solution. Also, it will find a point which represents the smallest possible departure from $\boldsymbol{t} = \boldsymbol{0}$, in the direction of $\hat{\overline{\boldsymbol{X}}}_d$. A similar adaptation can be made for the MCPE. See chen, Sitter and Wu (2000) for further discussions.

5.2 GREG

Interestingly, we can apply the idea of relaxing the benchmark constraints in the same way as in the previous sub-section to directly obtain an approximate GREG with range restricted weights. We recall that the calibration estimator is identical to the GREG under a chi-square distance measure. To remain consistent with the previous presentation we will consider estimation of $\overline{Y}$ with $\overline{X}$ known. Thus, the calibration weights, w_i, will be chosen to minimize

$$\sum_{i\in s} (w_i - d_i)^2 / d_i, \tag{1}$$

subject to

$$N^{-1} \sum_{i\in s} w_i \boldsymbol{x}_i = \overline{\boldsymbol{X}}$$

The solution yields the GREG, $\overline{Y}_{GR} = N^{-1} \sum_{i\in s} w_i y_i$ as given in (1) with w_i's given in (3). Suppose we take an approach similar to that in Section 5.1. As before, assume $\overline{\boldsymbol{X}} = \boldsymbol{0}$. Similarly to Section 5.1, define $\boldsymbol{t} = (\delta \hat{\overline{X}}_{HT1}, \ldots, \delta \hat{\overline{X}}_{HTk})'$, we can now present an algorithm for obtaining an approximate GREG under range restrictions.

Step 0. Calculate the weights, w_i, in (3). If $\gamma_1 d_i \le w_i \le \gamma_2 d_i$ for $i \in s$, stop. The GREG weights meet both the benchmark constraints and the range restrictions.

Step 1. Let $L = 0, R = 1$ and $\varepsilon = 10^{-5}$;

Step 2. Let $\delta = (L+R)/2$ and and calculate

$$w_i = d_i [1 + (\boldsymbol{t} - \hat{\overline{\boldsymbol{X}}}_{HT})' \{N^{-1} \sum_{i\in s} d_i (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})'\}^{-1} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})].$$

Step 3. If $w_i \notin [\gamma_1 d_i, \gamma_2 d_i]$ for some $i \in s$, let $L = \delta$ and go to Step 2; otherwise, let $R = \delta$. If $L - R < \varepsilon$, stop and use these weights; otherwise, go to Step 2.

This is a very simple method and for that reason may be very attractive to statistical agencies.

6. DISCUSSION

For most surveys, efficiency of estimation procedures is of primary concern. For some statistical agencies where surveys are conducted on a monthly or yearly basis over the same target population, benchmark constraints may be important. The minimum restriction of positive weights, however, should be imposed whenever is possible. The model-calibrated empirical likelihood method is ideal in terms of the efficiency of the estimate, the stability of the weights, and the simplicity of the algorithm. The benchmark constraints may not hold in general, although in practice the departure from exact benchmark constraints is small. The pseudo-empirical maximum likelihood estimator requires some extra computation, but the benchmark constraints are always satisfied. The problem of range-restricted weights can be handled quite easily using the proposed strategy, and the model-calibrated empirical likelihood estimator is impacted much less by range restrictions.

REFERENCES

Bardsley, P. & Chambers, R.L. (1984).« Multipurpose Estimation from Unbalanced Samples ». *Applied Statistics*, 33, No. 3, 290-299.

Chen, J.& Sitter, R.R. (1999). « A Pseudo Empirical Likelihood Approach to the Effective Use of Auxiliary Information in Complex Surveys ». *Statistica Sinica*, 9, 385-406.

Chen, J., Sitter, R.R.& Wu, C. (2000). « Using Empirical Likelihood Methods to Obtain Positive Weights in Regression Estimators for Complex Surveys ». Working paper 2000-03, Department of Statistics and Actuarial Science, University of Waterloo.

Huang, E.T.& Fuller, W.A. (1978). « Nonnegative Regression Estimation for Sample Survey Data ». *Proceedings of the Social Statistics Section*, ASA, Washington, D.C., 300-305.

Polyak, B. T. (1987). *Introduction to Optimization.* Optimization Software, Inc. Publications Division, New York.

Rao, J.N.K.& Singh, A.C. (1997). « A Ridge-Shrinkage Method for Range-Restricted Weight Calibration in Survey Sampling ». *Proceedings of the Section on Survey Research Methods*, American Statistics Association, Washington, D.C., 57-65.

Singh, A.C.& Mohl, C.A (1996).. « Understanding Calibration Estimators in Survey Sampling ». *Survey Methodology*, Vol. 22, No. 2, 107-115.

Singh, A.C. (1993). « On Weight Adjustment in Survey Sampling ». Discussion paper for the 18th meeting of the Advisory Committee on Statistical Methods, Statistics Canada, Ottawa, October 25-26.

Wu, C.& Sitter, R.R. (2000). « A Model-calibration Approach to Using Complete Auxiliary Information From Survey Data » . *J. Amer. Statist. Asso.*, accepted.

Zhong, C.X.B.& Rao, J.N.K. (1996). « Empirical Likelihood of Inference Under Stratified Random Sampling Using Auxiliary Information ». *Proceedings of the Section on Survey Research Methods*, ASA, Chicago, Illinois, August, 1996. 798-803.