AN EMPIRICAL LIKELIHOOD APPROACH FOR COMPLEX SAMPLING

Yves G. Berger¹ and Melike Oğuz–Alper²

ABSTRACT

Survey data are often collected with unequal probabilities from a stratified population. We propose an empirical likelihood approach for sample data selected with unequal probabilities. We show that the empirical likelihood ratio statistic follows a chi-squared distribution asymptotically. The approach proposed does not rely on variance estimates, re-sampling or joint-inclusion probabilities, even when the parameter of interest is not linear. Standard confidence intervals based on variance estimates may give poor coverages, when normality does not hold. This can be the case with skewed data and outlying values. This paper contains the main results of Oğuz-Alper & Berger (2016a,c) published by Oxford University Press. Oğuz-Alper & Berger’s (2016a) empirical likelihood confidence interval has good coverages, even when the sampling distribution of the point estimator is not normal.

KEY WORDS: Design-based inference, Empirical likelihood, Estimating equation, Inclusion probability, Regression parameter, Unequal probability sampling.

RÉSUMÉ

Les données d’enquêtes sont souvent collectées à probabilités inégales à partir de populations stratifiées. Nous proposons une nouvelle approche basée sur la vraisemblance empirique pour des données sélectionnées avec des probabilités inégales. Nous montrons que le ratio de vraisemblance empirique suit une distribution chi-carré asymptotiquement. L’approche proposée ne repose pas sur des estimations de variance, le ré-échantillonnage ou les probabilités d’inclusion jointes, même lorsque le paramètre d’intérêt n’est pas linéaire. Les intervalles de confiance basés sur l’estimation de variance peuvent donner des couvertures pauvres en l’absence de normalité. Cela peut être le cas avec des données asymétriques et des valeurs aberrantes. Dans ce article, nous donnons les résultats de Oğuz-Alper & Berger (2016a,c). L’intervalle de confiance proposé par Oğuz-Alper & Berger (2016a) a des bonnes couvertures, même si l’estimateur ponctuel n’est pas normal.


1 INTRODUCTION

1.1 Description of the Problem

Statistical models are widely used in numerous fields of science. When statistical models are fit to sample data selected randomly with unequal probabilities (e.g. Brewer & Gregoire, 2009; Berger & Tillé, 2009), the assumption of independent and identically distributed (i.i.d.) observations will not be fulfilled. The estimators based on this assumption may be inconsistent and produce invalid inference (e.g. Binder & Roberts, 2009) if the random selection of the sample is ignored, especially when the sampling design is informative (e.g. Skinner, 1994; Pfeffermann, 1993; Pfeffermann & Sverchkov, 1999, 2003, 2009; Pfeffermann, 2011).

We propose a design-based empirical likelihood approach to make inference for a general class of multidimensional parameters defined as the solution of a set of estimating equations. This class includes complex parameters such as ratios, (non)linear regression parameters, generalised linear regression parameters and multilevel regression parameters with random intercept. We consider a semiparametric approach, where the model is specified by estimating equations. We consider that the underlining distribution is unknown.

¹Yves G. Berger, University of Southampton, United Kingdom of Great Britain and Northern Ireland, SO17 1BJ, Y.G.Berger@soton.ac.uk
²Melike Oğuz–Alper, Statistics Norway, Postboks 8131 Dep, NO-0033 Oslo, Kingdom of Norway, melike.oguz.alper@ssb.no
Let $U$ be a finite population of $N$ units labelled $i = 1, \ldots, N$. Consider that $n$ units are selected independently with replacement with unequal probabilities $p_i$ (e.g. Hansen & Hurwitz, 1943) from $U$, where $\sum_{i\in U} p_i = 1$. Let

$$\pi_i = np_i,$$

where $n$ is the fixed number of draws. The sampling design may also be stratified (see §4.2). Let $\bar{\pi}$ denote the sampling fraction or the mean of the $\pi_i$: $\bar{\pi} = \frac{N}{n} \sum_{i\in U} \pi_i = n/N$. The approach proposed is also valid under without-replacement sampling with first-order inclusion probabilities given by (1), when $n/N$ is negligible, as sampling with and without-replacement are asymptotically equivalent in this case. This is usually the case with large social surveys.

Let $s$ denote the sample containing the labels of the units selected after $n$ draws. The probability distribution of $s$ is called the sampling design. Let $v_i$ be a vector that contains the values of a set of variables for a unit $i \in U$. The sample data, given by $\{v_i : i \in s\}$, is a set of not identically distributed observations due to unequal probabilities. Under the design-based framework, the $v_i$ are fixed (non-random) vectors and the sampling distribution of the sample data is solely driven by the random selection of the sample (Neyman, 1938).

Consider that the parameter of interest is a sub-vector of the parameters. The remaining parameters, which are not part of the parameter of interest, are called ‘nuisance’ parameters. The nuisance parameters are unknown by definition. Nuisance parameters are found in numerous situations (e.g. Binder & Patak, 1994; Qin & Lawless, 1994; Godambe & Thompson, 2009; Zheng et al., 2012). For example, when testing the significance of parameters given two nested models, the parameters of the basic model are considered as nuisance while the parameters of interest are the additional parameters to be tested. In variance modelling, the variance is a function of nuisance parameters (e.g. Owen, 1991, p.1740). In regression estimation of means, the regression parameter is a nuisance parameter.

A valid inference in the presence of nuisance parameters require profiling, or equivalently, minimising the empirical log-likelihood ratio function over the nuisance parameters. This technique allows us to test and derive confidence intervals for the sub-vector of parameters by taking the randomness due to the nuisance parameters into account. Qin & Lawless (1994) showed that the profile empirical log-likelihood ratio function follows a $\chi^2$-distribution under the assumption of i.i.d. observations. Qin & Lawless’s (1994) approach does not directly apply to complex sampling designs. Oğuz-Alper & Berger (2016a) fill this gap. Berger & De La Riva Torres’s (2016) approach is limited to single parameters and cannot be straightforwardly extended for profiling. Oğuz-Alper & Berger (2016a) showed that the empirical log-likelihood ratio function that was proposed by Berger & De La Riva Torres (2016) can be profiled out over the nuisance parameter. Oğuz-Alper & Berger (2016a) showed that the profile empirical log-likelihood ratio function follows a $\chi^2$-distribution asymptotically. The approach proposed is different from Qin & Lawless’s (1994) approach, because we shall take the information about the sampling design into account, and we assume that the observations are selected with unequal probabilities.

### 1.2 Organization of the paper

In §2, we have a literature review. In §3, we defined the population parameter. In §4, we introduce the maximum empirical likelihood point estimator. In §4.1, we show how to incorporate a population level information. In §5, we define the profile empirical log-likelihood ratio function and show how it can be used for testing and constructing confidence intervals.

### 2 LITERATURE REVIEW

Standard design-based approaches involve linearisation (Binder, 1983; Deville, 1999; Demmati & Rao, 2004) or resampling techniques for variance estimation. Standard variance estimators may treat the nuisance parameters fixed by ignoring the fact that they are estimated. For example, this is the case for the regression estimation of means. Variance estimators may be biased when the randomness of the nuisance parameter is ignored.

Confidence intervals are constructed assuming that point estimators have normal distribution. However, point estimators may not follow normal distribution, when we have outlying observations, and linearised variance estimators may be biased for moderate sample sizes. Empirical likelihood inference does not rely on neither variance estimation nor the normality of the point estimator. It can also handle nuisance parameters; that is, empirical likelihood confidence intervals take into account the randomness of the nuisance parameter.
Empirical likelihood approach is a well established topic under the classical i.i.d. framework (e.g. Owen, 1991; Qin & Lawless, 1994; Wu & Sitter, 2001; Chaudhuri et al., 2008; Kim & Zhou, 2008; Zheng et al., 2012). The paper by Chen & Van Keilegom (2009) includes an elaborate review on the existing literature. Chen & Sitter (1999) proposed a pseudoempirical likelihood under unequal probability sampling. Wu & Rao (2006) showed that pseudoempirical likelihood confidence intervals can be constructed by using the design effect, which is estimated from variance estimates. However, this approach is limited to univariate parameters. Kim (2009) and Chen & Kim (2014) proposed an empirical likelihood approach under Poisson sampling. Berger & De La Riva Torres (2016) proposed an empirical likelihood approach for unequal probability sampling. This approach does not rely on design effects, variance estimates, linearisation or resampling. It deals with a wide range of non-linear finite population parameters.

Binder & Patak (1994) proposed a nonparametric version of the likelihood based score statistics that can be used with nuisance parameters (see also Godambe & Thompson, 1999, p.162). They proposed a method of inverse testing to construct confidence intervals. They used a pivotal statistics defined by the square of the estimating function divided by its variance. The bounds of a confidence interval are the solutions of a system of equations that can be solved numerically. Godambe & Thompson (2009, p.92) pointed out that solutions may not exist when the variance is a function of the population parameters instead of their estimates. This approach also relies on variance estimation.

Chen & Sitter (1999) and Zhong & Rao (2000) proposed an algorithm based on profiling the pseudoempirical likelihood ratio function when strata totals of auxiliary variables are unknown. For confidence intervals, the pseudoempirical likelihood needs to be adjusted by variance estimates. This approach is limited to estimation of totals and univariate parameters. There is no general theory on profiling for the pseudoempirical likelihood approach.

Pfeffermann & Sverchkov (1999, 2003) considered a parametric approach that takes into account the sampling design by adjusting the likelihood. They also considered a semiparametric approach that may require modelling the survey weights. The variances of the model parameters are estimated through linearisation or resampling techniques (e.g. Pfeffermann, 2011). Oğuz-Alper & Berger (2016a) compared numerically their approach with the semiparametric approach that was proposed by Pfeffermann & Sverchkov (1999, 2003).

Inferences about the population parameters can be improved by incorporating population level information, which may be available from administrative data, census data and/or population projections (e.g. Deville & Särndal, 1992; Chaudhuri et al., 2008). The empirical likelihood approach proposed allows incorporating population level information in the presence of nuisance parameters.

3 PARAMETERS AND ESTIMATING EQUATIONS

The parameter \( \psi_N \in \Psi \subset \mathbb{R}^b \) is the \( b = O(1) \) finite population vector that is the unique solution of the population estimating equation (2) (Godambe, 1960), where \( \Psi \) is compact.

\[
G(\psi) = \sum_{i \in U} g_i(\psi) = 0_b;
\]

where \( g_i(\psi) = g(v_i, \psi) \) is a \( b \times 1 \) vector of estimating functions and \( v_i \) is the vector of variables for unit \( i \). Here, \( 0_b \) is a \( b \times 1 \) vector of zeros. As we consider a design-based approach, the parameter \( \psi_N \) is a fixed (non-random) unknown quantity. The parameter \( \psi_N \) shall be estimated from the sample data. Note that \( g_i(\psi) \) may also depend on some known population parameters (see § 4.1).

Most finite population parameters can be defined by estimating equations (e.g. Binder, 1983; Binder & Patak, 1994; Qin & Lawless, 1994; Godambe & Thompson, 2009). For example, \( \psi_N \) can be a vector of population means, totals, ratios, quantiles, low income measures, regression coefficients (e.g. Qin & Lawless, 1994; Binder & Kovacević, 1995) or measures of poverty (e.g. Berger & De La Riva Torres, 2014). The equation (2) can be also viewed as a population moment condition (Hansen, 1982). In this case, (2) is the derivative of some objective function with respect to the parameter (e.g Godambe & Thompson, 2009).
3.1 Examples: Regression parameters

Consider \( v_i = (y_i, x_i^\top) \), where \( y_i \) is a scalar response variable and \( x_i \) are some explanatory variables. Consider a non-linear model with a smooth model (scalar) function \( \mu(\cdot) \). The parameter \( \psi_N \) can be the non-linear least squares parameter defined by

\[
\psi_N = \arg \min_{\psi \in \Psi} \sum_{i \in U} \sigma_i^{-1} \left\{ y_i - \mu(h(x_i)^\top \psi) \right\}^2,
\]

where \( h(\cdot) \) is a known vector function and \( \sigma_i \) is a known variance function. In this case, \( \psi_N \) is the solution of the estimating equation (e.g. Chen & Van Keilegom, 2009).

\[
\sum_{i \in U} \frac{\partial h(x_i)^\top \psi}{\partial \psi} \left\{ y_i - \mu(h(x_i)^\top \psi) \right\} \sigma_i^{-2} i = 0
\] (3)

For ordinary least squares parameters, we have \( h(x_i) = x_i \), \( \mu(h(x_i)^\top \psi) = x_i^\top \psi \) and \( \sigma_i^2 = \sigma \), for some \( \sigma \). Hence, equation (3) reduces to the normal equation.

\[
\sum_{i \in U} x_i (y_i - x_i^\top \psi) = 0_b.
\] (4)

Equation (4) can be extended to include instrumental variables.

For generalised linear models, we have \( \mu(h(x_i)^\top \psi) = F^{-1}(x_i^\top \psi) \), where \( F(\cdot) \) is a link function. For example, \( F(\mu) = \log(\mu(1 - \mu)^{-1}) \) with a logistic regression model. In this case, equation (3) reduces to (e.g. Binder, 1983, p.285)

\[
\sum_{i \in U} x_i \left\{ y_i - \frac{\exp(x_i^\top \psi)}{1 + \exp(x_i^\top \psi)} \right\} = 0_b.
\]

4 EMPIRICAL LIKELIHOOD UNDER UNEQUAL PROBABILITY SAMPLING

Consider the following empirical log-likelihood function.

\[
\ell(m) = \sum_{i \in s} \log (m_i),
\]

where the \( m_i \) are unknown scale-loads allocated to data points \( i \in s \) (Hartley & Rao, 1968) and \( m \) is the \( n \times 1 \) vector of the \( m_i \) \( (i \in s) \).

Let \( \hat{m}_i^*(\psi) \) maximises \( \ell(m) \) subject to the constraints \( m_i > 0 \) and

\[
\sum_{i \in s} m_i \ c_i^*(\psi) = C^*
\] (5)

with

\[
c_i^*(\psi) = (c_i^\top, g_i(\psi)^\top)^\top \quad \text{and} \quad C^* = (C^\top, 0^\top)^\top,
\]

for a given vector \( \psi = (\theta^\top, \nu^\top)^\top \); where \( g_i(\psi) \) is defined in § 3. Here, the \( c_i \) are \( r \)-vectors, which satisfy the constraint

\[
\sum_{i \in s} m_i c_i = C, \quad \text{with} \quad C = \sum_{i \in U} c_i.
\] (6)

The \( c_i \) incorporate the information about the sampling design, such as stratification and unequal inclusion probabilities, and the population level information (see §§ 4.1 and 4.2). When we have a single stratum and do not use any population level information, we use \( c_i = \pi_i \) and \( C = n \). In the case of a stratified design and the use of a population level information, the \( c_i \) is defined in a different way (see §§ 4.1 and 4.2).
We assume that the $C$ is an inner point of the conical hull formed by $\sum_{i \in s} m_i c_i$, so that the set of $\hat{m}_i$ is unique. We assume that $c_i$ and $C$ satisfy the regularity conditions given in Oğuz-Alper & Berger (2016a). We also assume that there exist an $r$-vector $t$ such that $t^\top c_i = \pi_i$. By using (6), we have that $\sum_{i \in s} m_i \pi_i = n$, which specifies the fact that a sample of size $n$ is selected.

We assume that $c_i(\psi)$ is differentiable with respect to $\nu$ for all $i \in s$ in a neighbourhood around the true population value $\nu_N$. The maximum value of $\ell(m)$ under $m_i > 0$ and (5) is given by

$$\ell(\psi) = \sum_{i \in s} \log (\hat{m}_i(\psi)).$$

### 4.1 Incorporating known population level information

Since Hartley & Rao (1968) first introduced population level information within the empirical likelihood framework, it became a key feature of empirical likelihood inference (e.g. Owen, 2001; Chaudhuri et al., 2008; Rao & Wu, 2009). Let $\varphi_N$ be a vector of deterministic known population level parameters. We consider that $\varphi_N$ can be defined as the unique solution of the population estimating equation

$$\sum_{i \in U} f_i(\varphi) = 0,$$

where $f_i(\varphi) = f(v_i, \varphi)$. Here $f(v_i, \varphi)$ is a function of $v_i$ and $\varphi$. This function does not depend on $\psi_N$. For example, $\varphi_N$ may be known population means, totals, ratios, proportions, variances, quantiles or distribution functions of some of the variables within $v_i$. If the $\varphi_N$ are the population means of an auxiliary variable $x_i$ (a sub-vector of $v_i$), we use $f_i(\varphi) = x_i - \varphi$. If $\varphi_N$ is a vector of population totals, we use $f_i(\varphi) = x_i - \varphi p_i n^{-1} N$. The $f_i(\varphi)$ do not have to be differentiable. For example, $f_i(\varphi)$ is not differentiable when $\varphi_N$ contains population quantiles. Note that $g_i(\psi)$ can be a function that depends on $\varphi_N$. With known population-level information, we shall use

$$c_i = (\pi_i, f_i(\varphi_N) \top)^\top.$$

### 4.2 Stratification

Suppose that a population $U$ is stratified into $H$, strata denoted by $U_1, \ldots, U_h, \ldots, U_H$, where $\cup_{h=1}^H U_h = U$. Suppose that a with-replacement sample $s_h$ of fixed size $n_h$ is selected from $U_h$, with unequal probabilities. We assume that the number of strata $H$ is bounded ($H = O(1)$). In this case, $\pi_i$ is substituted by

$$\pi_i = n_h p_i, \quad \text{for } i \in U_h.$$

Let $z_i$ be the values of the design (or stratification) variables defined by

$$z_i = (z_{i1}, \ldots, z_{iH}) \top,$$

where $z_{ih} = \pi_i$ for $i \in U_h$ and $z_{ih} = 0$ otherwise. We have $C = \nu$ where $\nu = (n_1, \ldots, n_H) \top$ denotes the vector of the strata sample sizes. With stratification and known population-level information, we shall use

$$c_i = (z_i \top, f_i(\varphi_N) \top)^\top.$$

### 4.3 Maximum empirical likelihood point estimator

The maximum empirical likelihood estimator $\hat{\psi}$ of $\psi_N$ is the vector that maximises $\ell(\psi)$ over $\psi$. Berger & De La Riva Torres (2016) showed that $\hat{\psi}$ is the solution of the sample estimating equation

$$\hat{G}(\psi) = 0_b, \quad \text{where} \quad \hat{G}(\psi) = \sum_{i \in s} \hat{m}_i g_i(\psi),$$

where $\hat{m}_i$ is given by

$$\hat{m}_i = (\pi_i + \eta \top c_i)^{-1}.$$
Here, the vector $\eta$ is such that (6) and $m_i > 0$ hold. A modified Newton-Raphson algorithm as in Chen et al. (2002) can be used to compute $\eta$. We assume that the $g_i(\psi)$ is such that equation (8) has a unique solution. Note that there exists situations when the solution is not unique (e.g. Domínguez & Lobato, 2004).

When we do not use any population level information, we have $c_i = \pi_i$ and $C = n$. It can be shown that $\eta = 0$, and $\hat{m}_i = \pi_i^{-1}$, which is the standard Horvitz & Thompson (1952) weight for unit $i$. When $\hat{m}_i = \pi_i$, $\hat{G}(\psi)$ is the Horvitz-Thompson estimator of $G(\psi)$, for a given $\psi$, and the estimator $\hat{\psi}$ is the pseudolikelihood estimator that was proposed by Binder (1983).

5 PROFILE EMPIRICAL LOG-LIKELIHOOD RATIO FUNCTION

Suppose that we would like to make inference about a $p \times 1$ sub-parameter $\theta_N \in \Theta \subset \mathbb{R}^p$, where $p < b$. Consider $\psi_N = (\theta_N^T, \nu_N^T)^T$, where $\nu_N$ is a $q \times 1$ sub-parameter ($\nu_N \in \Lambda \subset \mathbb{R}^q$) that are not of primary interest ($q = b - p$), where $\Theta$ and $\Lambda$ are compact set. The parameter $\nu_N$ is the nuisance parameter. In this §, we assume that we do not use any population level information. We propose to test and construct a confidence region for the parameter $\theta_N$ by using the profile empirical log-likelihood ratio function proposed by Oğuz-Alper & Berger (2016a) and defined by

$$\hat{r}(\theta) = 2 \left\{ \ell(\hat{\psi}) - \max_{\nu \in \Lambda} \ell(\theta, \nu) \right\},$$

(10)

where $\ell(\theta, \nu) = \ell(\psi)$ with $\psi = (\theta^T, \nu^T)^T$. Note that $\hat{r}(\theta)$ is a random function of $\theta$. An algorithm to compute (10) can be found in Oğuz-Alper & Berger (2016a). It can be shown that

$$\ell(\hat{\psi}) = \sum_{i \in s} \log(\hat{m}_i)$$

is the maximum value of $\ell(m)$ under the constraints $m_i > 0$ and (6), because (8) holds for $\hat{\psi}$. Note that the maximum empirical likelihood estimator of $\theta_N$ minimises the function (10).

Oğuz-Alper & Berger (2016c) showed that under a series of regularity conditions, the random variable $\hat{r}(\theta_N)$ asymptotically follows a $\chi^2$-distribution with $p$ degrees of freedom under unequal probability sampling, where $p$ denotes the dimension of $\theta_N$; that is,

$$\hat{r}(\theta_N) \xrightarrow{d} \chi^2_{df=p}.$$

Thus, $\hat{r}(\theta_N)$ is asymptotically pivotal statistics and can be used to make inference about the sub-parameter $\theta_N$. Suppose we wish to test $H_0 : \theta_N = \theta_N^0$ versus $H_1 : \theta_N \neq \theta_N^0$, by using $\hat{r}(\theta_N^0)$. The $p$-value is $\int_{\hat{r}(\theta_N^0)}^{\infty} \chi^2_{df=p}(x) dx$, where $\chi^2_{df=p}(x)$ is the density of a $\chi^2$-distribution with $p$ degrees of freedom.

5.1 Empirical likelihood confidence intervals

The pivotal statistics (10) can also be used to construct confidence intervals for a scalar ($p = 1$) sub-parameter $\theta_N$ of $\psi_N$. In this case, $\nu_N$ denotes the remaining parameters of $\psi_N$. Hence, $\hat{r}(\theta_N)$ follows asymptotically a $\chi^2$-distribution with one degree of freedom. Thus, the $\alpha$% empirical likelihood Wilks' (1938) type confidence interval for $\theta_N$ is given by

$$\left\{ \theta : \hat{r}(\theta) \leq \chi^2_{df=1}(\alpha) \right\},$$

where $\chi^2_{df=1}(\alpha)$ is the upper $\alpha$-quantile of the $\chi^2$-distribution with one degree of freedom. Note that $\hat{r}(\theta)$ is a convex function of $\theta$ with a minimum value when $\theta$ is equal to the empirical maximum likelihood estimator $\hat{\theta}$. Based on this property, the bisection method can be used to find the lower and upper bounds. This involves calculating $\hat{r}(\theta)$ for several values of $\theta$. 

6
CONCLUSION

There are numerous situations where the parameter of interest depends on nuisance parameters. The function (10) can be used to test the parameter of interest, and to construct confidence intervals that takes into account the estimation of the nuisance parameters. The empirical likelihood confidence intervals do not rely on variance estimates, linearisation (e.g Binder, 1983; Deville, 1999; Demnati & Rao, 2004) or resampling. They do not rely directly on the normality of the point estimator. Oğuz-Alper & Berger (2016a) provided a profile empirical log-likelihood ratio function under population level information and showed how the approach proposed can be extended for multistage sampling.

Oğuz-Alper & Berger’s (2016a) approach proposed can be used with non–linear models such as generalised linear models (see §§ 3.1). It can also be used for multilevel regression parameters (e.g. Oğuz-Alper & Berger, 2016b). The approach proposed is not limited to regression parameters. It can be applied to any finite population parameter that is uniquely defined as the solution of a set of estimating equations.

The approach proposed is less computer intensive than bootstrap and simpler to implement than linearisation, because it does not involve the derivation of linearised variables. Standard confidence intervals based on variance estimates may give poor coverages, when normality does not hold. This could be the case with skewed data and outlying values. Even when the normality holds, heteroscedasticity or model misspecification may affect the coverage of standard confidence intervals (e.g Owen, 1991; Rao & Wu, 2009). Furthermore, the coverage may be also affected by the bias or instability of linearised or resampling variance estimators.

Simulation studies provided by Oğuz-Alper & Berger (2016a) show that the observed coverage of empirical likelihood confidence interval is close to the nominal value. The empirical likelihood confidence interval based on the function (10) achieves better coverages and more balanced tail error rates than standard approaches, which involves linearisation or resampling, even when the point estimator is not normal or the data is skewed or includes outlying observations.

In § 4.1, we show that the population level information can be taken into account. The empirical likelihood survey weights (9) appear naturally because of the maximisation of empirical log-likelihood function (7), and the fact that a know population parameter is fixed within the function (7). The survey weights (9) are always positive and calibrated. There are some analogies between empirical likelihood and calibration (Deville & Särndal, 1992), although they are different. First, the empirical likelihood approach does not always requires population level information. Secondly, the calibration distance function is only used to derive calibration weights for point estimation, and plays no role in testing or constructing confidence intervals. The empirical log-likelihood function (7) is used for point estimation, testing and confidence intervals. The empirical likelihood weights are also asymptotically optimal for the estimation of totals and means. Calibration weights (Deville & Särndal, 1992) can be negative and not asymptotically optimal.

REFERENCES


