Empirical Bias Corrections for Fitting Multilevel Models under Informative Sampling

Solange Correa[1,3] and Danny Pfeffermann[2]
[1]University of Southampton, Southampton, United Kingdom
[2]University of Southampton, Southampton, United Kingdom and
Hebrew University of Jerusalem, Jerusalem, Israel
[3]Corresponding author: Solange Correa, email: s.correa-onel@soton.ac.uk

Abstract

Survey data are generally obtained via a complex sampling design involving clustering, stratification and unequal sample inclusion probabilities. When the inclusion probabilities are correlated with the model outcomes after conditioning on the auxiliary variables, the sampling process is informative, and the model holding for the sample data is different from the model holding in the population from which the sample is taken. Standard estimation methods for multilevel models may provide severely biased estimates of the model parameters under informative sampling, especially when the cluster sample sizes are small, yielding to erroneous interpretation of the phenomenon studied. In this article, we apply a new approach for bias correction based on resampling procedures proposed by the same authors to correct for the bias of multilevel model parameter estimates under informative sampling of the first and second levels of the model hierarchy. The performance of the new method and alternative bias correction approaches proposed in the literature are assessed via an extensive simulation study and an application to a real data set.

Key Words: bootstrap, complex survey data, probability weighting.

1. Introduction

Unweighted multilevel analysis (Goldstein, 2003) of complex survey data may lead to severely biased estimates (Korn and Graubard, 1995) if the inclusion probabilities are related to the model response variable even after conditioning on the design variables, known in the sampling literature as informative sampling design (Pfeffermann, Krieger and Rinott, 1998). Under such schemes, the model holding for the population values is likely to be different from the model holding for the sample data, defined as sample model by Pfeffermann et al. (1998a). Therefore, the sample model needs to be estimated from the sample data in order to perform inferential statistical analyses based on the sample values.

Another important issue when fitting multilevel models to sample survey data is how to account for the sampling weights in multilevel analysis estimation. A large number of studies on how to do this have been proposed lately in the literature (Pfeffermann, Skinner, Holmes, Goldstein and Rabash, 1998); Korn and Graubard, 2003; Grilli and Pratesi, 2004; Rabe-Hesketh and Skrondal, 2006). Most of them are based on incorporating the sampling weights in the likelihood function and maximising it via numerical integration since closed expression for the estimators are not available. Pfeffermann et al. (1998b) propose a probability weighted iterative generalised least squares approach (PWIGLS), which is an adaptation of the iterative generalised least squares (IGLS) method (Goldstein, 1986) by analogy to the pseudo maximum likelihood principle (Binder, 1983; Skinner, 1989; Chambers, 2003). The PWIGLS approach basically consists of probability weighting of first and higher level units with weights equal to the reciprocal of the corresponding sampling inclusion probabilities. However, as shown in that article, the use of this approach, although reducing the bias of unweighted parameter estimators very substantially, does not eliminate it completely, unless in large samples.

Classical bootstrap bias corrections (Efron, 1979) involve estimating the bias of an estimator by apriorily choosing a function that depends only on the original and bootstrap estimates of the parameter of interest. In analyses that involve more than one parameter, it could well be that the bias of an estimator in estimating one parameter may depend on the value of that parameter and on the bias in estimating the other parameters. Pfeffermann and Correa (2012) proposed a general approach for bias correction, entitled the Empirical Bootstrap Bias Correction (EBC), based on the bootstrap resampling procedure and on a parametric model.

The aim of this article is to compare the performance of the EBC method to classical bootstrap bias corrections via an extensive Monte Carlo study. Unweighted (thereafter naïve) and PWIGLS estimators are assessed when fitting two-level models to survey data under informative sampling of first units with small sample sizes at both levels.

2. The Classical Bootstrap Bias Corrections

Efron (1979) proposed to estimate the bias by use of bootstrap samples as obtained by drawing units with replacement from the original sample. These resampling methods have become very popular in statistical inference and are applied in many diverse applications in order to obtain estimates of standard errors, confidence intervals, biases, etc. (Shao and Tu, 1995). In what follows, parametric and nonprametric bootstrap are reviewed along with the classical bootstrap bias corrections.

Let $z_1,...,z_n$ be the outcomes of independent and identically distributed (i.i.d.) random variables $Z_1,Z_2,...,Z_n$ having distribution F. Denoting the observed data by $\underline{z}=(z_1,...,z_n)$, the objective is to assess the accuracy of a statistic $\hat{\psi}=t(\underline{z})$ in estimating the unknown parameter of interest $\psi=t(F)$ .Let $\underline{z}_1^*,...,\underline{z}_B^*$ be B independent (parametric or nonparamentric) bootstrap samples and $\hat{\psi}_1^*,...,\hat{\psi}_B^*$ the corresponding bootstrap replications of the statistic $\hat{\psi}$, where $\hat{\psi}_i^*=t(\underline{z}_i^*)$. Thus, measures of accuracy of the statistic of interest are inferred from the observed values of the bootstrap replications $\hat{\psi}_1^*,...,\hat{\psi}_B^*$. In particular, the bootstrap estimation of bias is straightforward, as shown in Efron and Tibshirani (1986) and described as folows. The bias of the statistic $\hat{\psi}=t(\underline{z})$ in estimating the true value $\psi=t(F)$ is

$$bias_F = bias_F(\hat{\psi},\psi)=E_F[t(\underline{z})]-t(F) \qquad (1)$$

where $E_F[\cdot]$ is the expectation under the distribution F. Replacing F by the estimated distribution $\hat{F}$ in equation (1), we find the bootstrap estimate of bias:

$$bias_{\hat{F}} = E_{\hat{F}}\left[t(\underline{z}^*)\right]-t(\hat{F}). \qquad (2)$$

In practice, $E_{\hat{F}}\left[t(\underline{z}^*)\right]=E_{\hat{F}}\left[\psi^*\right]$ is approximated by averaging $\hat{\psi}_1^*,...,\hat{\psi}_B^*$ over a large number B of bootstrap replications yielding

$$\hat{bias}_B = \overline{\hat{\psi}}^* - t(\hat{F}) \qquad (3)$$

where $\overline{\hat{\psi}}^* = B^{-1}\sum_{b=1}^{B}\hat{\psi}_b^*$. As B tends to infinity, $bias_B$ tends to $bias_{\hat{F}}$ (Efron and Tibshirani, 1986).

Once an estimate of the bias is available, one can correct the original estimate by subtracting the estimated bias from it. Hence, the bootstrap bias-corrected estimate of the parameter of interest ψ, also known as additive correction, is given by

$$\hat{\psi}^{BC} = \hat{\psi}-\hat{bias}_B = \hat{\psi}-\left[\overline{\hat{\psi}}^* - t(\hat{F})\right]=2\hat{\psi}-\overline{\hat{\psi}}^*. \qquad (4)$$

Similarly, the multiplicative bias correction (Hall and Maiti, 2006) is given by $\hat{\psi}^{BC}=\hat{\psi}^2/\overline{\hat{\psi}}^*$.

3. The Empirical Bootstrap Bias Correction Approach

The main idea of the EBC approach (Pfeffermann and Correa, 2012) is to use data generated under an assumed model and a plausible parameter space to identify the relationship between the true parameter value and its estimates from the original and bootstrap samples. Hence, the functional relationship between the error of the estimator under study and its original and bootstrap estimates is extracted from the data themselves, rather than arbitrarily chosen. Besides, not only original and bootstrap estimates of the target parameter are included in that relationship but corresponding estimates of other model parameters can possibly be included in the function as well. To allow for the fact that the bias may depend on the true value of the parameter, the procedure explicitly takes into account a set of plausible parameter values in the process of identification of the function. The EBC approach has two main advantages. The first one is that it provides not only a bias-corrected estimator of the target parameter but also the bootstrap distribution of the bias-corrected estimator, allowing estimation of its measures of accuracy. The second advantage is that the EBC approach is not restricted to a particular bias correction formula, permitting to express the bias of the target estimator as a function of the biases of other estimators involved in the analysis.

The EBC method generates a set of plausible parameter values $\boldsymbol{\psi}_1,\ldots,\boldsymbol{\psi}_g,\ldots,\boldsymbol{\psi}_G$ based on the original sample estimate and, for each of those, generates one pseudo original sample from $f_\xi(\mathbf{z};\boldsymbol{\psi}_g)$. Bootstrap samples are then generated from this pseudo original sample. As a result, each plausible parameter value generates one pseudo original estimate and corresponding bootstrap estimates. A mathematical relationship for the bias of the estimator under study can then be identified. A bias-corrected estimator for the target parameter is obtained by applying this function to the original and bootstrap estimates obtained from the original sample. Assume that the original sample yields original and bootstrap estimates, $\hat{\boldsymbol{\psi}}$ and $\overline{\hat{\boldsymbol{\psi}}}^*$, of the parameter $\boldsymbol{\psi}$. A single component of the vector of parameters $\boldsymbol{\psi}=(\varphi_1,\ldots,\varphi_K)$, say φ_1, is assumed to be the target for bias-corrected estimation. The approach is applied to the other model parameters in an identical manner. See Pfeffermann and Correa (2012) for a detailed description of the method.

4. Bias Corrections of Unweighted and PWIGLS Estimators of a Two-level model Under an Informative Sampling Design

In this section, the EBC approach is applied in order to reduce the bias of unweighted and PWIGLS (Pfeffermann et al., 1998b) estimators of two-level model parameters under an informative sampling scheme with small sample sizes of the upper level units. The study is based on simulated data adopting the same sampling scheme considered by Pfeffermann et al. (1998b) and under which the PWIGLS estimators showed large biases.

4.1. Population Model, Sampling Design and Estimation of Sample Models

Consider the following two-level random intercept model (denoted ξ):

$$y_{ij} \mid \beta_j = \beta_j + \varepsilon_{ij},\ i=1,\ldots,N_j \text{ (level 1 model)}$$
$$\beta_j = \beta + u_j,\ j=1,\ldots,M \text{ (level 2 model)}, \quad (5)$$

where u_j and ε_{ij} are independent random errors such that $u_j \overset{i.i.d.}{\sim} N(0,\sigma_u^2)$, $\varepsilon_{ij} \overset{i.i.d.}{\sim} N(0,\sigma_\varepsilon^2)$. Let $\boldsymbol{\psi}=(\beta,\sigma_\varepsilon^2,\sigma_u^2)^t$ be the vector of population parameters.

Consider a two-stage disproportional stratified clustered sampling design with informative sampling only at level 1 (elementary units). At the first stage, m level 2 units are selected by a probability proportional to size without replacement design. The measures of size are the level 2 sizes N_j, which are assumed to be uncorrelated with the random intercepts β_j, such that the sampling design is noninformative at level 2. At the second stage, level 1 units in selected level 2 unit j are partitioned into 2 strata according to whether $\varepsilon_{ij} > 0$ or $\varepsilon_{ij} \le 0$ and simple random sampling without replacement of sizes $n_{j.1}$ and $n_{j.2}$ (assuming $n_{j.1} \neq n_{j.2}$) are drawn from stratum 1 and 2 of level 2 unit j, respectively. In this case, the first level inclusion probability is related to the level 1 random error ε_{ij} and, consequently, to the outcome y_{ij}, featuring an informative sampling design at level 1. The sample distribution (Pfeffermann, Krieger and Rinott, 1998) for the level 1 measurements y_{ij} given the random intercept β_j and inclusion in the sample is

$$f_{s_j}(y_{ij} \mid \beta_j, \boldsymbol{\psi}) = f(y_{ij} \mid i \in s_j, \beta_j, \boldsymbol{\psi}) = \begin{cases} \dfrac{(2n_{j.1}/N_j)\phi(y_{ij};\beta_j,\sigma_\varepsilon^2)}{n_j/N_j}, & if \ y_{ij} > \beta_j \\ \dfrac{(2n_{j.2}/N_j)\phi(y_{ij};\beta_j,\sigma_\varepsilon^2)}{n_j/N_j}, & if \ y_{ij} \le \beta_j \end{cases} \quad (6)$$

$$= \begin{cases} 2n_{j.1}n_j^{-1}\phi(y_{ij};\beta_j,\sigma_\varepsilon^2), & if \ y_{ij} > \beta_j \\ 2n_{j.2}n_j^{-1}\phi(y_{ij};\beta_j,\sigma_\varepsilon^2), & if \ y_{ij} \le \beta_j. \end{cases}$$

Since the sampling design is noninformative at level 2, the sample distribution for the level 2 measurements β_j given inclusion in the sample is the same as the distribution of β_j in the population, i.e., $f_s(\beta_j \mid \boldsymbol{\psi}) = f(\beta_j \mid j \in s, \boldsymbol{\psi}) = \phi(\beta_j;\beta,\sigma_u^2)$. (7)

4.2. Monte Carlo Study

In this section, the performance of the EBC approach is assessed via a Monte Carlo study for the population model, sampling scheme and sample model described in the previous section. The experiment mimics the simulation study performed by Pfeffermann et al. (1998b) with an additional step for adjusting the bias of unweighted and PWIGLS estimators by applying the EBC approach. The scaled 2 PWIGLS estimators proposed by Pfeffermann et al. (1998b) is adopted in this study.

Let $\hat{\boldsymbol{\psi}} = (\hat{\beta}, \hat{\sigma}_\varepsilon^2, \hat{\sigma}_u^2)^t$ and $\overline{\hat{\boldsymbol{\psi}}}^* = (\overline{\hat{\beta}}^*, \overline{\hat{\sigma}}_\varepsilon^{2,*}, \overline{\hat{\sigma}}_u^{2,*})^t$ be the respective vectors of original estimates and bootstrap means of the vector of population parameters $\boldsymbol{\psi} = (\beta, \sigma_\varepsilon^2, \sigma_u^2)^t$. The experiment involves generating populations from the model in (11) with parameters $\beta = 1$, $\sigma_u^2 = 0.2$, $\sigma_\varepsilon^2 = 0.5$ and $M = 300$ second level units. The second level sizes N_j were determined by $N_j = 75\exp(\tilde{u}_j)$, where $\tilde{u}_j \sim N(0, \sigma_u^2)$ truncated below by $-1.5\sigma_u$ and above by $1.5\sigma_u$. The values of N_j lie in the interval [38;147] with average around 80. The sample size at the first stage is $m = 35$ level 2 units. At the second stage, simple random samples of level 1 units of sizes $n_{j.1} = 2$ and $n_{j.2} = 7$ are drawn from strata 1 and 2, respectively. Therefore, the total sample sizes n_j are fixed ($n_j = 9$) for all level 2 units j.

Generation of bootstrap samples from the original and the pseudo original samples is carried out by both parametric and nonparametric bootstrap. The latter involves selecting the second level units by a simple random sampling design, with all the first level units (individual level) from the sampled level 2 units being included in the sample. The sampling weights used in the replicated values (bootstrap samples) are identical to the original sampling weights. For simplicity, generation of the plausible parameter values was performed from a uniform distribution between two boundaries determined by the original estimates and their corresponding standard errors.

The experiment consists in replicating the steps of the EBC approach a large number of times in order to assess the bias and mean squared error of the EBC estimators. The method is described below for the PWIGLS estimator $\hat{\boldsymbol{\psi}}^{pw}$, with an identical process applied to the naïve estimator $\hat{\boldsymbol{\psi}}^{naive}$. The process was repeated $R = 100$ times. The number of plausible parameter values $G = 400$ was chosen in order to allow a reasonable number of parameter values for estimation (350 values) and validation (50 values) of the candidate bias correction functions. The classical additive and multiplicative bootstrap bias corrections are special case of the functions in equations (8) and (9) below respectively, taking $a_1 = 1$ and $a_0 = a_2 = a_3 = 0$.

$$\frac{\hat{\sigma}^2_{u,g}}{\sigma^2_{u,g}} = a_0 + a_1 \cdot (\frac{\bar{\hat{\sigma}}^{2,*}_{u,g}}{\hat{\sigma}^2_{u,g}}) + a_2 \cdot (\frac{\bar{\hat{\sigma}}^{2,*}_{\varepsilon,g}}{\hat{\sigma}^2_{\varepsilon,g}}) + a_3 \cdot (\frac{\bar{\hat{\beta}}^*_g}{\hat{\beta}_g}) \quad (8)$$

$$\hat{\sigma}^2_{u,g} - \sigma^2_{u,g} = a_0 + a_1 \cdot (\bar{\hat{\sigma}}^{2,*}_{u,g} - \hat{\sigma}^2_{u,g}) + a_2 \cdot (\bar{\hat{\sigma}}^{2,*}_{\varepsilon,g} - \hat{\sigma}^2_{\varepsilon,g}) + a_3 \cdot (\bar{\hat{\beta}}^*_g - \hat{\beta}_g) \quad (9)$$

4.3. Results

Table 1 shows summary statistics for the naïve, PWIGLS and respective bias-corrected estimators for parameter σ^2_u under the classical corrections and the EBC approach. Results are reported for nonparametric bootstrap only. Similar conclusions are valid for the parametric case and for the other model parameters: β and σ^2_ε. The number of original samples considered for naïve and PWIGLS estimators is $R = 72$ and $R = 95$, respectively.

Let $\hat{\psi}_r$ denote an estimator (naïve, PWIGLS, EBC or classical bias-corrected estimator) of the parameter ψ (known in the simulation study) for the original sample $r = 1,...,R.$ The following summary statistics were computed for $\hat{\psi}_r$: simulation mean (Mean): $\bar{\hat{\psi}} = R^{-1}\sum_{r=1}^{R}\hat{\psi}_r$; simulation standard deviation (SD): $\sqrt{(R-1)^{-1}\sum_{r=1}^{R}(\hat{\psi}_r - \bar{\hat{\psi}})^2}$; empirical bias (Bias): $R^{-1}\sum_{r=1}^{R}(\hat{\psi}_r - \psi)$; empirical relative bias (RB): $R^{-1}\sum_{r=1}^{R}(\frac{\hat{\psi}_r - \psi}{\psi})$ and empirical root mean squared error (RMSEemp): $\sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\psi}_r - \psi)^2}$.

Table 1: Naïve and PWIGLS estimators. Nonparametric bootstrap. True value $\sigma^2_u = 0.2$.

Estimator	Mean	SD	Bias	RB	RMSE^{emp}
$\hat{\sigma}^{2,naive}_{\varepsilon}$	0.169	0.046	-0.031	15%	0.055
Additive Correction	0.165	0.049	-0.035	17%	0.060
Multiplicative Correction	0.165	0.049	-0.035	17%	0.060
EBC	0.187	0.071	-0.013	6%	0.071
$\hat{\sigma}^{2,pw}_{u}$	0.158	0.052	-0.042	-20%	0.065
Additive Correction	0.165	0.054	-0.035	-15%	0.064
Multiplicative Correction	0.165	0.054	-0.035	-15%	0.064
EBC	0.196	0.051	-0.004	-2%	0.050

As anticipated, the naïve and PWIGLS estimators based on the original sample are highly biased in the present scenario. The EBC bias-corrected estimators show very good performance for all model parameters, including the variance component estimators, which are expected to be most problematic to estimate due to their high sensitivity to small sample sizes. Classical bias corrections, however, perform poorly with small or no reduction in the biases.

It is worth emphasising that the bias-corrected PWIGLS estimators obtained by the EBC approach perform well even in the case where the non-corrected estimator is practically unbiased, which is a desirable characteristic of a bias correction procedure. This is the case of the PWIGLS estimator of the intercept β (Pfeffermann et al., 1998b). In addition, the trade-off bias-variance does not seem to be an issue for the EBC bias-corrected estimators, since the mean squared errors (RMSEemp) show minimum increase or even reduction when compared to the non-corrected estimators.

5. Conclusions and Remarks

In this article the extended bootstrap bias correction (EBC) approach was applied to bias adjustment of unweighted and PWIGLS estimators of linear two level model parameters under informative sampling of level 1 units with small sample sizes at both levels. The EBC procedure was assessed by Monte Carlo study, evaluating the behaviour of the EBC estimators through mean squared errors and biases estimates.

The main finding of this article is that the EBC approach performs very well for all the scenarios considered (naïve and PWIGLS estimators under nonparametric bootstrap). A range of factors can be changed to improve the EBC approach proposed in this article. First, the best function can be chosen such that the estimated MSE (not the bias) of the EBC estimator is minimum. See Pfeffermann and Correa (2012). Another issue to be explored is the generation of the parameter values. Adopting wider intervals and distributions other than the uniform is an important issue to be considered. The results presented in this paper are preliminary. An application of the EBC approach to a real survey data is in progress.

References

Chambers, R.L. (2003). Introduction to Part A. In Analysis of Survey Data (eds. R.L. Chambers and C. J. Skinner). 13-27, Chichester: Wiley.

Binder, D.A. (1983). On the variances of asymptotically normal estimators from complex surveys. International Satistical Review, 51, 279-292.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. The annals of Statistics, 7, 1-26.

Efron, B. and Tibshirani, R.J. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science, 1, 54-75.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. Biometrika, 73, 43-56.

Goldstein, H. (2003). Multilevel statistical models. 3rd edn. London: Hodder Arnold.

Grilli, L. and Pratesi, M. (2004). Weighted estimation in multilevel ordinal and binary models in the presence of informative sampling designs. Survey Methodology, 30, 93-103.

Hall, P. and Maiti, T. (2006). On bootstrap methods for small area prediction. Journal of Royal Statistical Society B, 68, 221-238.

Korn, E.L. and Graubard, B.I. (1995). Examples of differing weighted and unweighted estimates from a sample survey. American Statistician, 49, 291-295.

_____, (2003). Estimating variance components by using survey data. Journal of the Royal Statistical Society B, 65, 175-190.

Pfeffermann, D. and Correa, S. (2012). Empirical bootstrap bias correction and estimation of prediction mean square error in small area estimation. *Biometrika*, 99 457–472.

Pfeffermann, D., Skinner, C.J., Holmes, D.J., Goldstein, H. and Rabash, J. (1998b). Weighting for unequal selection probabilities in multilevel models. Journal of the Royal Statistical Society B, 60, 23-40.

Pfeffermann, D., Krieger, A. and Rinott, Y. (1998a). Parametric distributions of complex survey data under informative probability sampling. Statistica Sinica, 8, 1087-1114.

Rabe-Hesketh, S. and Skrondal, A. (2006). Multilevel modelling of complex survey data. Journal of the Royal Statistical Society A, 169, part 4, 805-827.

Shao, J. and Tu, D. (1995). The Jackknife and the Bootstrap. New York: Springer-Verlag.

Skinner, C.J. (1989). Domain means, regression and multivariate analysis. In Analysis of Complex Surveys (eds. C.J. Skinner, D. Holt and T. M. F. Smith), 59-87. Chichester: Wiley.