مجلهی پژوهشهای آماری ایران
J. Statist. Res. Iran 4 (2007): 129-131
دورهی ۴، شمارهی ۱، بهار و تابستان ۱۳۸۶، صص ۱۲۹-۱۳۱

Calibration Weighting to Compensate for Extreme Values, Non-response and Non-coverage in Labor Force Survey

Arman Bidarbakht-nia[†,*] and Hamidreza Navvabpour[‡]

[†] Statistical Center of Iran
[‡] Allameh Tabatabaie University

1 Introduction

Frame imperfection, non-response and unequal selection probabilities always affect survey results. In order to compensate for the effects of these problems, Devill and Särndal (1992) introduced a family of estimators called calibration estimators. In these estimators we look for weights that have minimum distance with design weights based on a distance function and satisfy calibration equations.

In this paper after introducing generalized regression estimator, we explain general form of calibration estimators. Then special cases of calibration estimators due to using different distance functions, practical aspects and results of comparing the methods are considered.

2 Calibration Estimators

Devill and Särndal (1992) rewrited the ordinal regression estimator as a weighted estimator of observed values, y_k, to derive the general form of calibration estiomators.

The regression equation $\hat{t}_{y\mathrm{GREG}} = \sum_S a_k g_k y_k$ where $g_k = 1 + c_k\big(\sum_U X_k - \sum_S a_k X_k\big)'\big(\sum_S a_k c_k X_k X_k'\big)^{-1} X_k$, and a_k is the design weight, is equivalent

* Corresponding author

to finding the weights $w_k = a_k d_k$ such that a distance function, $\sum_S \frac{(w_k - a_k)^2}{2a_k}$, is minimized while control equations $\sum_S w_k X_k = \sum_U X_k$ are satisfied, where X_k is a vector of auxiliary variables and $\sum_U X_k$ a vector of known values of auxiliary variables for the population. In fact, in calibration estimators we are looking for weights which minimize a distance function and satisfy calibration equations.

The distance function can have different forms. Regression estimator, ratio estimator and raking estimator are all special cases of calibration estimators in which we use different distance functions.

Devill and Särndal (1992) introduced a distance function which avoids extreme calibration weights due to using constant bounds, $L < 1 < U$, as

$$D_k(w_k, a_k) = A^{-1} \sum_{k=1}^{n} a_k [(F_k - L) \log\{(1-L)^{-1}(F_k - L)\} + (U - F_k) \log\{(U-1)^{-1}(U - F_k)\}],$$

where $F_k(X_k'\lambda) = \frac{L(U-1)+U(1-L)\exp(AX_k'\lambda)}{(U-1)+(1-L)\exp(AX_k'\lambda)}$, λ is a vector of Lagrange multipliers, and $A = \frac{U-L}{(1-L)(U-1)}$.

If L and U are large negative and large positive values, respectively we are close to linear regression and for $L = 0$ and large U we are close to raking.

3 Discussion and Conclusion

Using four calibration estimators introduced in this paper, calibration weighting has been performed on labor force survey data in 2005 autumn and obtained estimates for unemployment rate and number of unemployed people in rural and urban areas in each province have been compared.

The results show that unemployed population and unemployment rate estimators from Devill and Särndal method have smaller standard errors in compare to other methods. The distribution of the adjustment factor for four estimators confirms the above issue.

As pointed, in the mentioned methods, calibration estimator via the Devill and Särndal distance function is the most precise estimator. It is more appropriate to use Devill and Särndal distance in surveys in which the mentioned problems occur to a considerable extent, because in other cases there is no tangible difference between estimators.

Keywords. design weights; over/under coverage; non-response; weighting methods, labor force survey.

Arman Bidarbakht-nia
Statistical Center of Iran,
Dr. Fatemi Ave.,
Tehran 14146 63111,
Iran.
e-mail: *bidar_a@yahoo.com*

Hamidreza Navvabpour
Department of Statistics,
Faculty of Economic,
Allameh Tabatabaie University,
Tehran, Iran.
e-mail: *h.navvabpour@srtc.ac.ir*

The full version of the paper, in Persian, appears on pages 1-14.