Estimation under restrictions built upon biased initial estimators

Natalja Lepik

University of Tartu, Estonia, email: natalja.lepik@ut.ee

Keywords: restriction estimator, survey sampling.

The users of official statistics often require that sample-based estimates satisfy certain restrictions. In the domain's case it is required that the estimates of domain totals sum up to the population total or to its estimate. For example, in time domains, quarterly estimates have to sum up to the yearly total. The relationships holding for the true population parameters do not necessarily hold for the respective estimates. This inconsistency of estimates is annoying for statistics users. On the other hand, known relationships between population parameters is a kind of auxiliary information. Involving this information into estimation process presumably improves estimates. Our goal is to define consistent domain estimators that are more accurate than the initial inconsistent domain estimators.

One solution to the problem of finding estimates under restrictions is the general restriction estimator (GR) proposed by Knottnerus (2003). His estimator is based on the unbiased initial estimators and is unbiased itself. The advantage of the GR estimator is the variance minimizing property in a class of linear estimators. Sõstra (2007) has developed the GR estimator for estimating domain totals under summation restriction. Optimality property of the domain GR estimator is studied in Sõstra and Traat (2009). In all these works, the unbiasedness or asymptotic unbiasedness of initial estimators is assumed.

It is well known that there are many useful estimators that are biased. For example, the model-based small area estimators are design-biased. The synthetic estimator can be biased on the domain level. Even the widely used GREG estimator is only asymptotically unbiased. In this thesis we will allow the vector of initial estimators $\hat{\boldsymbol{\theta}}$ to be biased, and will construct three new restriction estimators, based on the biased initial estimators:

$$\hat{\boldsymbol{\theta}}_{GR1} = (\mathbb{I} - \mathbf{KR})(\hat{\boldsymbol{\theta}} - \boldsymbol{b}),$$

$$\hat{\boldsymbol{\theta}}_{GR2} = (\mathbb{I} - \mathbf{K}^*\mathbf{R})\hat{\boldsymbol{\theta}},$$

$$\hat{\boldsymbol{\theta}}_{GR3} = (\mathbb{I} - \mathbf{K}^*\mathbf{R})(\hat{\boldsymbol{\theta}} - \boldsymbol{b}),$$

where $\mathbf{K} = \mathbf{VR}'(\mathbf{RVR}')^{-1}$, $\mathbf{K}^* = \mathbf{MR}'(\mathbf{RMR}')^{-1}$, $\mathbf{V}$ and $\mathbf{M}$ are accordingly the covariance and the MSE matrices of the initial estimator-vector; $\boldsymbol{b}$ is its bias.

References

[1] Knottnerus, P. (2003). *Sample Survey Theory. Some Pythagorean Perspectives.* Wiley, New York.

[2] Sõstra, K. (2007). Restriction estimation for domains. *Doctoral Dissertation.* Tartu University Press, Tartu.

[3] Sõstra, K., Traat, I. (2009). Optimal domain estimation under summation restriction. *Journal of Statistical Planning and Inference* **139**, 3928-3941.