DISCUSSION

Malay Ghosh* University of Florida

The following three papers are discussed:

(1) A Survey Practitioner's Viewpoint by Barbara A. Bailar, Cary T. Isaki, and Kirk M. Wolter

(2) Survey Sampling - Modelling, Randomization and Robustness: A Subjective Bayesian Approach by W. A. Ericson and

(3) Survey Sampling: Modelling, Randomization and Robustness: A Unified Theory View by V. P. Godambe

Discussion of (1):

I agree with Dr. Bailar as to the role of randomization. It "removes personal biases in the selection of a sample, and thus increases public acceptance". The role of randomization in designing a survey is valued even in a Bayesian framework (see e.g. the recent article of Royall and Pfefferman (Biometrika (1982) V69, #2, pp.401-409). However, I am rather suspicious about the utility of randomization at the inference stage.

I sympathize with her hesitance in accepting superpopulation models in drawing inferences from finite populations, or in the selection of sampling units. There is a usual "lack of robustness against model misspecification" argument which goes against estimators derived from superpopulation models. However, I will not be averse to the use of such estimators if they prove to be robust against model misspecification. Some recent attempts in the derivation of such robust estimators are due to Royall and Pfefferman, but, definitely, more remains to be done.

I find in this article too much emphasis on "consistency" of estimators. Consistency is a large sample property, and may not be adequate to depend on if convergence is slow. It would be appropriate in this context to remind ourselves once again about the famous example of Sir R. A. Fisher.

Example. Suppose $\{T_n\}$ is a sequence of consistent estimators of θ. Let

$$T_n' = 0 \text{ if } n < 10^{10}$$
$$= T_n, \text{ otherwise.}$$

Then $\{T_n'\}$ is also consistent for θ. But, it would be very undesirable to use T_n' for estimating θ.

Discussion of (2):

Basically, I agree with Professor Ericson's subjective Bayesian stand, and like his linear Bayes estimators which can be derived by knowing only the mean vector and the variance-covariance matrix. From my own point of view, I would possibly go one step further when these parameters are unknown, and advocate use of empirical Bayes estimators by substituting data based estimators of the prior parameters in the Bayes estimators. The resulting estimators will then be non-linear, but from my own personal experience in some other related areas of estimation, the resulting empirical Bayes estimators might prove to be more robust against misspecified priors either from a Bayesian or a frequentist risk criterion.

I found the derivation of the linear Bayes estimators of Professor Ericson quite interesting. An alternate interesting (though parametric) approach of deriving such estimators could be extension of the ideas of Carl Morris (Ann. Statist. (1983), V11, #2, 515-529) to a multiparameter exponential family with certain structure on the variance-covariance matrix. Morris considered a restricted one parameter exponential family with variance a quadratic function of the mean.

The model which Professor Ericson assumes to incorporate response bias is a good start, but more general models should be looked into.

Discussion of (3):

The estimator $e_{\underset{\sim}{a}}$ introduced in (2.1) of this paper can be appreciated also in a Bayesian framework. Recall that

$$e_{\underset{\sim}{a}} = \sum_{i\in s} y_i + \hat{\beta}_{\underset{\sim}{a}} \sum_{i\notin s} x_i + \sum_{i\notin s} a_i,$$

where

$$\hat{\beta}_{\underset{\sim}{a}} = \frac{\sum_s x_i(y_i - a_i)/\sigma_i^2}{\sum_s x_i^2/\sigma_i^2}.$$

A more general ratio type estimators for estimating the population total is as follows.

$$e^*_{\underset{\sim}{a}} = \sum_{i\in s} y_i + \sum_{i\notin s} a_i + \Big\{ \sum_{i\in s} \Big(\frac{y_i - a_i}{m_i} \frac{c_i}{\sum_{i\in s} c_i}\Big)\Big\} \sum_{i\notin s} m_i. \quad (1)$$

(See e.g. Meeden and Ghosh: "On the admissibility and uniform admissibility of ratio type estimators" (Proceedings of the Golden Jubilee Conference of the Indian Statistical Institute)).

We furnish below several examples of the ratio type estimators introduced in (1).

Example 1. Put $a_i=0$, $m_i=c_i=x_i(>0)$. This leads to the classical "ratio estimator".

Example 2. Put $a_i=0$, $m_i=\pi_i$, $c_i=1-\pi_i$, $\sum_1^N \pi_i=n$. This leads to the Horvitz-Thompson estimator.

Example 3. Put $a_i=0$, $c_i=1/n$. This leads to the estimator

$$e^*_o = \sum_{i\in s} y_i + \{n^{-1} \sum_{i\in s} (y_i/m_i)\} \sum_{i\notin s} m_i$$

of the population total as introduced by Basu ((1971) Foundations of Statistical Inference

Edited by Godambe and Sprott pp 202-233). The further special cases $m_1=\ldots=m_N=1$ gives rise to the classical estimator $\frac{N}{n(s)} \sum_{i\varepsilon s} y_i$ of the population total.

Example 4. Put $m_i=x_i$, $c_i=x_i^2/\sigma_i^2$. This leads to the estimator $e_{\hat{a}}$ of this paper.

Very often the estimator $e^*_{\hat{a}}$ is not a Bayes estimator, as it cannot be obtained by using a single prior distribution. However, the following Bayes like interpretation of $e^*_{\hat{a}}$ can be given as follows.

Let a_i denote the prior guess about y_i, and m_i the degree of uncertainity about the prior guess. We observe the y_i ($i\varepsilon s$), while a_i's and m_i's are all known. It is assumed that for any $j\not\varepsilon s$, the ratios $(y_j-a_j)/m_j$ assume the values $(y_i-a_i)/m_i$ ($i\varepsilon s$) with "posterior probabilities" $c_i/(\sum_{i\varepsilon s} c_i)$. Then the "posterior expectation" of the parameter $\sum_1^N y_i$ (the population total) given s and y_i, $i\varepsilon s$ is given by (1). However, there may not be a single prior generating such "posterior probabilities".

Back to the article of Professor Godambe, he gives a nice presentation of the unified theory view, an area which evolved out of his pioneering 1955 JRSS B paper. However, I have some difficulty in accepting certain basic assumptions in his paper.

First he assumes the $y_1,\ldots,y_N$ when distributed as ξ to be probabilistically mutually independent. For me, inference in finite population sampling is primarily a study of the conditional distribution of the unseen y_i's given the observed ones. Assumption of independence of the y_i's in this approach means that one learns nothing about the unseen y_i's by knowing the seen ones. This does not seem very satisfactory, and an assumption of exchangability of the y_i's seems to be more suitable than the assumption of independence.

Second, there could be situations (for example with certain prior information) where estimation of the nuisance parameters could be more meaningful than an attempt to eliminate such parameters.

Finally, I do not find any compelling reason why attention should be restricted to unbiased estimators. Biased estimators have made their way eminently into the statistical literature, and should not be ignored in any unified theory development.

* Research supported by the NSF Grant Number MCS-8218091.