Journal of Official Statistics, Vol. 24, No. 4, 2008, pp. 513–516

Discussion

David A. Binder[1]

This article provides several interesting and thought-provoking situations where Sedransk argues that a Bayesian approach is best for making statistical inference on finite population quantities. He considers establishment surveys, where with a vast amount of prior data, highly skewed distributions and plausible models concordant with data, the Bayesian approach seems natural. He argues that for inference for small subpopulations, there are good foundational reasons for preferring a Bayesian approach, and that large sample approximations are necessary but hard to obtain in a frequentist analysis. Another situation he considers where a Bayesian framework may be suitable is for analyses to improve survey designs prior to selecting the sample. Sedransk also considers pooling data from several sources, and using survey data for provider profiling. For population-based social surveys with small to moderate-sized subpopulations of interest, and for other cases, the large sample approximations may not hold, and Sedransk argues that the design-based approach would be less desirable. As a drawback, however, he admits that a limiting factor may be that time is needed to develop appropriate models.

Recently, there has been renewed interest in using Bayesian methods for survey data. For example, Little (2006) advocates what he refers to as a "calibrated Bayes" approach. Focusing on the goal of predicting the nonsampled values, he advocates using a prediction approach that captures design information, possibly by using the weights explicitly in the model.

On considering the appropriateness of a Bayesian approach in general, I suggest that such an approach is indeed appropriate in certain circumstances, and I have used such an approach myself in the past for some of these. Some situations where I have actually used a Bayesian approach and the reasons for adopting such an approach are as follows:

Too many parameters: In Binder (1978a, 1981) a Bayesian approach was used for cluster analyses where for each new observation there was a new parameter. This is similar to the problem of multiple comparisons. In Binder (1978b) the problem arose in the context of switching regressions.

Small sample size: Small samples can not only invalidate any asymptotic arguments, they can also lead to widely varying results according to the particular prior distribution selected in a Bayesian setting. Many non-Bayesian approaches may have an implicit prior distribution assumption. An extreme case of small samples was mentioned in Binder (1993) for the so-called "Exchange Paradox."

[1] Statistics Canada (Retired), 49 Bertona Street, Nepean, ON, Canada K2G 4G7. Email: dbinder49@hotmail.com

Clear decision problem: The decision-theoretic approach can be quite appropriate when the decision problem is clear. See, for example, Binder (1978a) for the case of a decision-theoretic approach in cluster analysis.
Sensitivity to the prior: When the inferences are sensitive to the prior, not using a Bayesian approach explicitly may imply that a particular prior is being used implicitly. For example, in Binder (1978a) it was found that certain clustering rules were equivalent to a particular Bayesian approach with strong priors in favour of almost equal-sized clusters. Examples were given in Binder (1981) where the conclusion would depend on the particular prior chosen. The sensitivity of the results to the prior arises more frequently when the sample sizes are small.
Insight into problem or addressing a new problem: In Binder (1982), the problem of percentile estimation for stratified samples was considered in a Bayesian framework and contrasted to the frequentist approach taken by Sedransk and Meyer (1978). In Binder and Dick (1989, 1990), an empirical Bayes approach – not strictly speaking a Bayes approach – was used to fit time series models to a series of possibly correlated survey estimates. The problem of multiple comparisons was considered in Binder (1978a).

In general, when considering a frequentist approach, as opposed to a likelihood-based approach, it is often the case that the frequentist approaches can be ad hoc, unless there are optimality principles, and even with the optimality principles, counter-examples can be concocted where results are absurd; see, for example, several examples in Cox and Hinkley (1979). When a frequentist approach has a Bayesian analogue, the underlying assumptions can be clearer. The notion of proper multiple imputation given by Rubin (1987) is a good example of using a Bayesian approach to provide insight into a problem and then providing conditions under which such an approach has appropriate frequentist properties.

In the survey context, the frequentist, design-based approach is often preferred, especially when the data are being exploited for several purposes, rather than being used for one easily identifiable purpose. Also, there is a concern as to whether the use of a model-based approach is robust to violation of the model-based assumptions. However, in many circumstances some use of model-based approaches is definitely appropriate. An example of this, as Sedransk has pointed out, is the case where aggregated data, arising from more than one survey, are being analyzed. For analyzing pooled data from several survey sources or from more than one cycle of a longitudinal survey, or for analyzing time series data from several points in time a Bayes, or empirical Bayes approach can be appropriate. A pure design-based approach could be used, but care is needed for defining the finite population about which inferences are being made. Also, the parameters of interest may not necessarily be finite population quantities, especially when there is no clearly defined finite population of interest.

The argument of asymptotic design consistency, used by advocates of the design-based approach (including myself), is not as relevant for nonresponse treatment, and small domain estimation. As well, when an analysis is being performed to design or plan a survey, the fact that the sample size is small argues in favour of using models to perform the analysis. In addition, there are some cases where models have been considered, even by practitioners who strongly prefer the design-based approach. For example, in social surveys, external information is not usually incorporated in the estimates, except perhaps for calibration such as age-sex poststratification. However, what happens when the survey results are counter-intuitive? For example, Hunter (1996) discusses a situation where a survey showed that smoking prevalence rates decreased after the price of cigarettes decreased, when the experts expected an increase in the prevalence rates. It may be suitable to use modelling methods if we have additional information on the nature of response errors or other nonsampling errors. For some analyses, such as those associated with item response theory models, the presence of latent variables may preclude the use of pure design-based methods. Another example in a social survey context could be where survey variables, such as income, may be collected using ranges rather than exact values and some model-based method may be required to impute the exact value; see, for example, Schenker et al. (2006).

Sedransk also makes a case for using Bayesian methods for economic surveys, where there may be valuable historical information available. For aggregate data, given the temporal correlation of the economic survey variables, there may be a case for using such information to attenuate the pure design-based estimates.

Bayesian methods, which are model-based, can be useful and should be learned and be part of the statistician’s toolbox. In general, however, in spite of some of the arguments favouring a Bayesian or a model-based approach, I believe that for inferences from survey data, design-based methods should be the dominant approach, especially when sample sizes are large and when the survey may be used for several purposes. Researchers should be aware, however, of the circumstances under which a design-based approach can be inappropriate, even though such cases are not as common.

References

Binder, D.A. (1978a). Bayesian Cluster Analysis. Biometrika, 65, 31–38.

Binder, D.A. (1978b). Comment to Estimating Mixtures of Normal Distributions and Switching Regressions by R.E. Quandt and J.B. Ramsey. Journal of the American Statistical Association, 73, 746–747.

Binder, D.A. (1981). Approximations to Bayesian Clustering Rules. Biometrika, 68, 275–285.

Binder, D.A. (1982). Non-parametric Bayesian Models for Samples From Finite Populations. Journal of the Royal Statistical Society, Series B, 44, 388–393.

Binder, D.A. (1993). Comment to Christensen, R., and Utts, J. (1992) Bayesian Resolution of the “Exchange Paradox.” The American Statistician, 46, 274–276, Letters to the Editor Section, The American Statistician, 47, 160.

Binder, D.A. and Dick, J.P. (1989). Modelling and Estimation for Repeated Surveys. Survey Methodology, 15, 29–45.

Binder, D.A. and Dick, J.P. (1990). A Method for the Analysis of Seasonal ARIMA Models. Survey Methodology, 16, 239–253.

Cox, D.R. and Hinkley, D.V. (1979). Theoretical Statistics. London: CRC Press.

Hunter, L. (1996). Survey on Smoking in Canada. Symposium 96, Nonsampling Errors: Proceedings, Catalogue No. 11-522-XPE. Ottawa: Statistics Canada, 181–189.

Little, R.J. (2006). Calibrated Bayes: A Bayes/Frequentist Roadmap. The American Statistician, 60, 213–223.

Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: Wiley.
Schenker, N., Raghunathan, T.E., Chiu, P.-L., Makuc, D.M., Zhang, G., and Cohen, A.J. (2006). Multiple Imputation of Missing Income Data in the National Health Interview Survey. Journal of the American Statistical Association, 101, 924–933.
Sedransk, J. and Meyer, J. (1978). Confidence Intervals for the Quantiles of a Finite Population: Simple Random and Stratified Simple Random Sampling. Journal of the Royal Statistical Society, Series B, 40, 239–252.

Received April 2008