DISCUSSION

V.P. Godambe, University of Waterloo

We should all be thankful to the authors for a very lucid presentation of their views on the subject. Particularly valuable must be the details concerning the origin of 'unequal probability sampling', coming from one of its authors. This technique surely has become over the years one of the cornerstones of survey sampling. Today's sampling practice owes considerably to the authors and their collaborators; notable amongst whom was the late William Hurwitz.

This morning's paper clearly illustrates the great ingenuity with which the authors over the years have approached every practical survey problem they have come across. But the paper also emphasises the fact that the authors have considered each survey situation mostly in an adhoc or piece-meal manner. This surely must have had an effect on the progress of the subject: While the progress in other areas of statistics looked like a scientific development under the conceptual unifying structures provided primarily by K. Pearson, R.A. Fisher, Neyman and others, survey sampling looked like a trivial area of statistics having no challenging profound scientific problems of its own which could influence the entire body of statistical thinking. This is evidenced by the fact that only exceptionally does one come across a statistics department here which provides graduate courses in survey sampling. Fortunately now there are indications of a possible change.

Now I proceed to elaborate on some of the remarks I have just made.

In illustrations the authors compare the plausibility of some estimators against that of some other estimators. But why should the practitioner restrict to the estimators they compare? From Neyman's quotation the authors seem to imply that the use of an improper estimate would only produce unnecessarily broad confidence intervals. Actually the situation could be worse. In the absence of a generally well defined optimality criterion, it is easy to produce two plausible looking estimates which would imply seriously conflicting conclusions (confidence intervals). Where is then the scientific objectivity of statistical inference? As I have several times maintained, including at the discussions at Chapel Hill, Washington, New Delhi, the scientific objectivity of estimation in survey sampling can be preserved only by a well defined criterion of optimal estimation, which enables us to choose from all estimators and all sampling designs. The authors say "minimisation methods had been used only within specific classes of estimators or other aspects of designs." From the point of view of scientific objectivity just mentioned, I find this restriction to 'specific classes of estimators' without any foundation. On the other hand obviously the optimum design will have to be found subject to administrative and cost considerations.

The authors often refer to 'consistent' estimators. But the criterion of consistency for a finite population (size N) is very ill-defined. If one defines an estimate to be consistent if it gives the true value when the sample becomes as large as the population, i.e. for a sample of size N it is no restriction on how the estimate should behave for samples of sizes N-1, N-2, etc. Indeed there is some literature on applications of the law of large numbers and central limit theorems (Madow 1948, Rosen 1974) to sampling finite populations. But the results are hardly of any practical meaning or value because of their dependence on sequences of finite populations satisfying certain conditions. One does not know how to interpret or verify these conditions in practice. In addition a basic point here is that the sequence in which individuals are drawn is not a part of sample core or sufficent statistic. Furthermore in other areas of statistics, with the introduction of the concepts like sufficiency, ancillarity, power, etc, by Fisher, Neyman and others it is generally accepted that valid statistical inference can be made for small samples. Why then should the authors be required to rely exclusively on large sample properties for the validity of their estimates? Certainly they deny any (small sample) overall optimal properties for their estimates.

I consider Hansen and Hurwitz (1943 Annals) paper as a landmark for its introduction of unequal probability sampling designs. However I believe that their mention (page 336, 1943 Annals) of 'best linear unbiased estimate' was unfortunate and particularly puzzling in view of the authors' statement in the present paper, "Survey statisticians early recognised that there was no best estimator ..." If this recognition had received clear expression earlier indeed much of the ensuing confusion could have been avoided. Any way since this point is by now generally well taken I will go to the other aspects of Hansen & Hurwitz's papers and the book by Hansen, Hurwitz and Madow mentioned in the present paper.

The estimation and unequal probability sampling schemes proposed in the just mentioned references by Hansen, Hurwitz and Madow (1943, 1949, 1953) are quite intuitive. Yet it was necessary to subject these schemes to some generally applicable well defined optimality criterion, referred to above, before they could be scientifically meaningful. Whether they should be rejected or accepted and with what modifications and extensions is the purpose of this scientific scrutiny. In the present paper the authors propose a probability sampling scheme and the estimator $\hat{Y}_p$. In the following pages they defend the proposal against the estimator $\hat{Y}_M$. Though all that the authors say in this respect is quite intuitive, the authors introduce the estimator and sampling scheme in an adhoc manner. The estimator $\hat{Y}_M$ is at least justified by a well-defined criterion of least-squares. (Now, that the least squares criterion is not general enough because of its restriction to linearity of estimation and more seriously its restriction to implied normality is another matter). But the authors do not put forward any such justification for $\hat{Y}_P$ apart from arguing how under certain situations $\hat{Y}_P$ would do better than $\hat{Y}_M$. I accept the argument. But this argument can be made more generally applicable and scientifically compelling within the framework of the Unified Theory (Godambe (1955), Godambe & Joshi (1965)). I already have demonstrated the near optimality of $\hat{Y}_P$ with appropriate stratified sampling, under a sufficiently flexible model taking into account possible departures of the regression line from the origin which the authors discuss, in my talks last year at Chapel Hill, Washington, New Delhi ('Robust near optimal estimation in survey practice' by Godambe and Thompson, ISI, 1977). This is done with more elaboration now in my paper (1978) to which the authors refer.

Concerning my 1978 paper the authors remark, "Godambe... achieves results consistent with those on the choice of optimum probabilities as discussed by Hansen & Hurwitz (1949) and Hansen, Hurwitz & Madow (1953). However the approach discussed by Hansen et al seems to provide more insight...". Indeed by some restricted minimisation procedure the authors show that Neyman allocation for the situation under study is given by $n_h \propto \sqrt{\bar{X}_h}$. This approximates my result if the x-values in the stratum are assumed to be nearly constant. But what the above quoted remark of the authors and the subsequent discussion of the authors ignore is the following: I establish near optimality, under a very general model also taking into account possible departures, of the estimate $e^* = (\sum_1^N \sqrt{x_i}/n)\sum_{i\in s} y_i/\sqrt{x_i}$, together with appropriate stratified sampling with inclusion probabilities proportional to $\sqrt{x_i}$. Now the estimate e^* is not a Neyman-type estimator. Neyman-type estimators depend on individual labels only through the strata to which the individuals belong. I therefore cannot see how Hansen et al could, using Neyman formalism, obtain the estimate e^*. So when the authors say my results are consistent with theirs, possibly they mean my results are practically acceptable to them. This then is a matter of great satisfaction for me. I surely cannot claim to possess their practical insight. When the authors say I do not take into account a more elaborate cost function I would say that in these situations I should investigate approximations along the line of already established optimal or near optimal results instead of relying on some adhoc restricted minimisation procedure as suggested by the authors. Incidentally our 'optimality' investigations (1977, 1978) also characterise possible departures from the model under which the combined ratio estimate $\hat{Y}_P$ mentioned earlier (or separate ratio estimate) with appropriate stratified sampling, instead of the estimate e^* and the corresponding design, provide robust near optimal estimation. Briefly under a given model, specified inclusion probabilities and the corresponding estimate provide optimum estimation. Additional stratification corresponding to the possible departures from the model render the optimality of estimation robust to these departures. In general our 'optimality' is defined such that if the underlying model happens to be valid the optimal estimation would agree with model-based estimation. In the extreme case if the assumed model is almost totally invalid or breaks down, design frequencies provide a meaningful interpretation of our estimation (Godambe & Thompson (1976)). Let me conclude by saying that in our approach to estimation the model and sampling design are treated on an equal logical footing in contrast to the approach of the authors where the sampling design receives the priority, or the approach of Royall and others where the model is given the priority. The authors' approach would yield misleading estimation if the model is valid and design adopted seriously disagrees with it. Similarly heavily model-based estimation would be without interpretation if the model turns out to be invalid in major respects. Both these possibilities of misleading estimation are avoided by our optimality criterion which, as said before, provides logically equal status to both, the model and the design. Our optimal estimation is obtained, from the class of all estimation strategies which are unbiased both under the model and under the design, by minimising the joint expectation under the design and model together, of the squared error. Such a unity of model based and design based estimation can alone provide an objective or scientific estimation.

REFERENCES

Godambe, V.P. (1955), "A unified theory of sampling from finite populations", *Journal of the Royal Statistical Society*, B, 17, 269-278.

Godambe, V.P. and Joshi, V.M. (1965), "Admissibility and Bayes estimation in sampling finite populations-I", *Annals of Mathematical Statistics*, 36, 1707-1722.

Godambe, V.P. and Thompson, M.E. (1976), *"Philosophy of survey sampling practice"*, *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science*. Eds. Harper & Hooker. D. Reidel Publishing Co., Dordrecht-Holland.

Godambe, V.P. and Thompson, M.E. (1977), "Robust near optimal estimation in survey-practice", to appear in *ISI Bulletin*, 47.

Godambe V.P. (1978), "Estimation in survey sampling: robustness and optimality".