DISCUSSION

Oscar Kempthorne, Iowa State University

1. It is highly appropriate to applaud the authors for their efforts. The statistical profession has seen papers by various writers, particularly by Basu, to the effect that the sampling practices of the U. S. Government are strongly defective in basic logic. The late L. J. Savage made the same comment in an oral presentation some 14 years ago.

2. The issue is, of course, philosophical. One group of protagonists insists on the need for its processes to possess a repeated sampling property — rather like the view that a measurement process is useless unless one knows how that measurement process behaves in repetitions of "equal" status. The other group says that the issue is one of belief on the part of the individual stating the belief.

3. With the polarity stated thus (and I hope, more or less correctly), the issue boils down to the question of what significance, or weight, I or Society should place on the belief of statistician X, say Basu, for example. I have to state that I have no basis for attaching weight to a Bayesian statistician. Note that I am not saying that no weight should be attached to his belief. But in the absence of a basis which is compelling to me, why should I attach weight to the belief of statistician X?

4. An interesting feature of the situation is that there is little evidence on how the Bayesian statistician actually tackles a real problem in sample design and inference. When attacked on this basis, the reply of statistician X is that a particular real problem that is posed to him is too complex for him to react to.

5. The upshot for me is, then, that the Bayesian sampler should get out of his office and show us how he does real problems.

6. Whether Royall and his collaborators fall in this class of statisticians I am attacking, I do not know. My mind is open to data. I would like references to real surveys that have actually been done.

7. The problem in the sampling profession that underlies the paper is highly important for society and not just the profession.

8. I find myself rather uncomfortable with the use of the idea of consistency on the first page. Indeed, I am strongly of the opinion, as of now, that it is irrelevant. What would happen if I draw samples of increasing size up to that of the population is, I think, irrelevant to the interpretation for a given size.

9. The critical matter in the context of confidence interval procedures is that $(\hat{\theta} - \theta)/s(\hat{\theta})$ be known to be a pivotal function to an adequate extent. This is the problem of pivotality I have mentioned in the literature and which, as far as I can see, is essentially unmentioned in all the books. One may be able to use asymptotic theory to give support to the idea that a proposed pivotal function is in fact a pivotal. But the relevance of such proof with assumptions on the underlying sequence of populations used in the proof of the desired asymptoticity is, I think, doubtful with respect to the existent real population that one has with the actual sample size.

10. The authors have accepted confidence interval procedures. In doing so, they have, of course, placed themselves in total opposition to the Bayesians. It is at this point, I think, that the real controversy occurs.

11. Whether one should, with Neyman, grant "that the problem of confidence intervals is solved correctly" is moot. I find certain difficulties in granting this, though I do not reject the process entirely. I suspect, however, that some of my own concerns do not bear on sampling a finite population.

12. On the matter of model-dependent processes, it is obvious, of course, that if one can supply a conventional parametric probability model and uses it to draw inferences, those inferences will be sharper than without such data, i. e., a confidence interval will be shorter. In contrast by assuming only that there is a population set $\{y_i,\ i = 1, 2, \ldots, N\}$ which is arbitrary except that for some $\hat{\theta}$ and some $s(\hat{\theta})$, $(\hat{\theta} - \theta)/s(\hat{\theta})$ is a t-pivotal, one is surely not assuming as much, so one gets less. But I think I represent Hansen et al correctly by saying that these workers claim to know better what they have in respect to the existent real population. I agree with that claim.

13. It is interesting and, I suggest, important, that the role of use of model in analysis of a probability sampling is not at all as important as the role in the mode of choice of the sample. Even there it is not negligible as Hansen et al show. But the role of assumption in the design is huge. The extreme is that some community, say Nevada, Iowa, reflects perfectly the voting preferences of the whole U. S. nation. Everyone knows such a model can be disastrous. The ideas put forward by model-dependent surveys are not as ludicrous as this, of course. If Nature has put a simple i. i. d. random process, say, on deviations from some simple linear relation, such as $y_i = \alpha + \beta x_i$, why should the sampler not use it? He will then pick half of his "sample" at the smallest x-value and the other half at the largest, with of course extensions when the x's have a known distribution. The whole point of the disagreement is obvious, of course. We do not know that Nature has been so beneficient.

14. So one then goes into balancing, an idea that has been around for, I guess, centuries. The only problem is that no one knows how to apply random sampling theory; what is the population of balanced samples, etc? What is the variability among the population of balanced samples?

15. The irony is that, as the authors state, we do not have theory and practice for balanced sampling. We stratify and then use random sampling within strata. We then have the result that we do have approximately balanced sampling. So, ideas of balancing support a widely accepted version of probability sampling, as the authors say.

15. On the matter of superpopulation ideas, I have to express my opinion, held for a long time, that assumption of superpopulation is totally invalid. It may be that the population of the U. S. is an outcome of a stochastic process which could have led to a population of populations. But this is irrelevant, even if the case, because I want to form judgments about the existent outcome, i. e.,

the population as it is, here and now. And I really could not care less about what might have been with the supposed stochastic process. It is, I suppose, a curosity that assumption of a super-population gives the right answer for a fixed population sometimes.

16. On the matter of "causal analysis", which is, I suppose, the main ingredient of "analytical surveys", I find the literature totally unconvincing. That one should pay attention to clustering or covariance in attempted causal modelling even if one has a simple random sample, seems obvious. It is not, I think, entirely clear that one should take account of the clustering given by a sampling process, except for the fact that one will have used clustering in the sampling because the individuals in a cluster are more nearly alike than a random group of the same size.

17. It is obvious that the probability sampler (and the experiment designer) uses a model in formulating his design. What distinguishes his ideation is, however, that he does not believe his model. It is his best guess of an approximative model. When he turns to analysis he works within his best approximate guessing with a process of inference that uses the repeated sampling principle. I have made before the comment: "The trouble with Bayesians is that they really believe their beliefs and models". There are, of course, deep problems in putting into effect a logic of beliefs.

18. The problem of "outliers" that the authors discuss is related, of course, to the question of pivotality. And as has been said, by myself and others, it is easy to envisage a population in which a 95 percent confidence interval is wrong 100 percent of the time. A real problem exists, and it is good to see it addressed. On lack of response, one has to use a model and that is why in a tight epidemiological study, or a surgical follow-up study, one makes a massive effort to obtain all responses.

19. I close with stating my opinion that, with some points of exception that I raise above, I find the guiding principles at the end of the paper to be compelling for me. There are problems, but I am glad that our Government uses probability sampling. I do have some confidence in our federal statistics. So I judge do the Bayesians.