## DISCUSSION

Donald A.S. FRASER, *University of Toronto, University of Waterloo,* and *York University*

Godambe (1979, 1982) presents an ancillarity principle and then gives examples in which statistical inference is obtained from ancillary variables. This paradoxical result is discussed by Genest and Schervish, and resolutions are offered: for a first example in terms of the disguising of a randomized procedure, and for the second in terms of the lack of equivalence of two inequalities. I suggest that there are deeper aspects to the resolution, and that Godambe has fortunately forced our attention to these.

For the first example, the model is isomorphic to pure randomization and has a redundant or noneffective parameter space. The procedure offered by Godambe uses a valid pivotal quantity, and the inversion gives a valid confidence region. Thus the paradox.

The example should force us to address the question: what is statistical inference (SI)? Until the 1950s SI was commonly viewed as just decision theory. The decision theorist A. Rapaport currently states that SI is just decision theory. Fisher, however, asserted otherwise, and the field now exists largely through his initiatives and yet without clear delineations. Some attempts towards delineations have been made (Birnbaum 1962, 1972; Evans, Fraser, and Monette 1986). For our purposes here, however, it suffices to view SI as embracing just estimation, tests, and confidence procedures.

For the two examples some additional notation is useful. Let $\boldsymbol{I}$ be the random $n \times N$ incidence matrix containing $n$ 1's, in different columns and in different rows randomly chosen, and with 0's otherwise; then $S = \boldsymbol{I}(1, \ldots, N)'$ is the sample vector of labels.

For the first example, the ancillary variable is $\boldsymbol{I}$ or $S$ and the parameter is $\boldsymbol{\theta}$. The basic pivotal quantity is $\boldsymbol{t} = \boldsymbol{I\theta}$, which has a *fixed* hypergeometric-type distribution; the confidence region is $C = \{\boldsymbol{\theta} : |\mathbf{1}'\boldsymbol{t}/n| \leq k\}$. This is a valid confidence procedure and is thus part of SI. Note that it is based only on the pure randomization variable $\boldsymbol{I}$. Also note more interestingly that it is a *noneffective* procedure as judged, say by the fact that, the coverage probability for the true value is equal to that for any false value. For a variant let $h$ be any 1-1 function on the set of permutations of $n$ from $\{1, \ldots, N\}$; then $C = \{\boldsymbol{\theta} : |\mathbf{1}'h(\boldsymbol{t})/n| \leq k\}$ is also a confidence region, one that is ineffective and has coverage probability for a true value equal to that for a false value.

For the second example, the yields are parameters and are more appropriately designated as $\eta_1, \ldots, \eta_N$. The formula (3.1) then defines $\boldsymbol{\theta} = \boldsymbol{\eta} - \mathbf{a}\phi$ where $\phi = \Sigma\eta_i/\Sigma a_i$; this relation is part of the *given* for the example. Note that $\boldsymbol{I\eta}$ is the sample sequence, which we designate here as $\boldsymbol{y}$. For the convenience of matrix notation, it suffices to examine the case $a_i = 1$; this avoids determining the correct weights for a sample, and the general case follows easily.

Genest and Schervish note that the "purported (paradox) results from rewriting the inequality" $|\mathbf{1}'\boldsymbol{I\eta}/n - \phi| \leq k$ "as" $|\mathbf{1}'\boldsymbol{I\theta}/n| \leq k$ (in matrix notation) "and then hiding the fact that the equivalence only holds if (3.1) is true . . . . That is, Godambe has disguised a calculation based on both $[\boldsymbol{I}]$ and $[\boldsymbol{y}]$ as equivalent to a calculation based only on $[\boldsymbol{I}]$." The equivalence (3.1) is part of the given and thus holds without question. Their conclusion then does not follow as argued.

The equivalence (3.1) gives that the pivotal quantity $\mathbf{1}'\boldsymbol{I\theta}/n = \mathbf{1}'\boldsymbol{I\eta}/n - \phi$, the latter for cases $\bar{y} - \phi$ where sample values of $\boldsymbol{y} = \boldsymbol{I\eta}$ are observable. The pivotal quantity thus can be presented in terms of a parameter $\boldsymbol{\theta}$ and a variable $\boldsymbol{I}$ or in terms of a parameter $\phi$ and variable $\boldsymbol{y} = \boldsymbol{I\eta}$, This then gives the conclusion mentioned above. The statement following (3.2) is thus incorrect; the kind of inferences available depend directly on the definition of the observable variable, here $\boldsymbol{I}$ or $\boldsymbol{I}\boldsymbol{\eta}$.

## ADDITIONAL REFERENCES

Birnbaum, A. (1962). On the foundations of statistical inference. *J. Amer. Statist. Assoc.*, 57, 269–332.
Birnbaum, A. (1972). More on concepts of statistical evidence. *J. Amer. Statist. Assoc.*, 67, 858–886.
Evans, M.; Fraser, D.A.S., and Monette, G. (1986). On the sufficiency-conditionality to likelihood argument. *Canad. J. Statist.*, 14, to appear.

Vinayak M. JOSHI, *University of Western Ontario*

Godambe's two examples—according to him—contradict the ancillarity principle and are therefore paradoxical. The short point of this note is that neither of them contradicts the principle. Hence no paradox exists.

In this example a sample of fixed size $n$ is drawn by SRS, the $\theta_i$, $i \in s$, being unobserved, and inference is made as follows: on observing a sample $s'$, consider as implausible all $\boldsymbol{\theta}' \in \Omega$ such that

$$\left| \frac{1}{n} \sum_{i \in s'} \theta_i' \right| > k \qquad \text{for suitably chosen } k. \tag{1}$$

According to Godambe the inference in (1) is intuitively appealing because for all $\boldsymbol{\theta} \in \Omega$, $E(\bar{\theta}_s) = 0$, and $Var(\bar{\theta}_s)$ is small, where $\bar{\theta}_s = (1/n)\Sigma_{i \in s}\theta_i$, but it contradicts the ancillarity principle, as the observed data $s'$ are distributed independently of $\boldsymbol{\theta}$.

The fallacy in this argument lies in the assumption that (1) is the unique mode of inference which is intuitively appealing. Actually, for given $k$ there are $\binom{N}{n}!$ such modes, all having exactly the same degree of appeal as (1) on the basis of the twin criteria of $E(\bar{\theta}_s)$ and $Var(\bar{\theta}_s)$. For let $S$ be the set of all samples of size $n$, and $s^*(\cdot)$ one particular 1-1 mapping from $S$ to $S$. Consider the following alternative mode: On observing $s'$, consider as implausible all $\boldsymbol{\theta}' \in \Omega$ for which

$$\left| \frac{1}{n} \sum_{i \in s^*(s')} \theta_i' \right| > k.$$

The $\binom{N}{n}!$ such modes are all equally valid. Hence there exists no inferential reason for preferring any one of them over another. It follows that for any given $\boldsymbol{\theta}' \in \Omega$, on observing $s'$, an inference that $\boldsymbol{\theta}'$ is implausible has just as much validity (support) as the contrary inference that $\boldsymbol{\theta}'$ is *not* implausible. But this is just another way of stating that no inference regarding the true value of $\boldsymbol{\theta}$ is possible on the basis of an observed $s'$. Since this is what the ancillarity principle provides, there is no contradiction with it.

Godambe's second example does not require detailed consideration, because (as pointed out by Genest and Schervish) the observed data in this case are not $s'$ but $\{s'; y_i, i \in s'\}$, the probability distribution of which is obviously not independent of $\boldsymbol{\theta}$.

Vidyadhar P. GODAMBE, *University of Waterloo*

Since my formulation of a statistical paradox in 1979 and in more detail in 1982, several authors have commented, criticized, and offered resolutions: Dawid (1979), Good (1980), Basu (1981), and now Genest and Schervish. It is interesting, even amusing, or perhaps deeply meaningful that of the "resolutions" offered no two agree in any important respect. In fact, Genest and Schervish clearly say Good's resolution is "flawed" and "invalid". To see the meaning of these differences, let us have a quick look at the paradox. In the following the references are to Godambe (1982).

Let $s^0$ be the subset actually drawn from the distribution (2.2). Then for a given $\epsilon$, if $N$ and $n$ are sufficiently large, one would with considerable confidence assert "$|\bar{\theta}_{s^0}| \leq \epsilon$", where $\bar{\theta}_s$ is the same as in (2.7). From my lectures in several universities on this topic, I found that initially practically all statisticians accepted the validity of the above assertion—that is, until I pointed out to them that logically the assertion

$$[|\bar{\theta}_{s^0}| \leq \epsilon] \equiv [\boldsymbol{\theta} \notin \Omega',\ \Omega' \subset \Omega]$$

is equivalent to asserting that "some specific values of $\boldsymbol{\theta}$ in $\Omega$ are unlikely". This assertion about $\boldsymbol{\theta}$ was at once rejected by the statisticians [the distribution (2.2) of $s$ being independent of $\boldsymbol{\theta}$], and then they had second thoughts about their having accepted the initial assertion "$|\bar{\theta}_{s^0}| \leq \epsilon$".

Is this a "psychological trap" as Genest and Schervish say? Absurd. They, like the abovementioned statisticians, want to throw over board a "deep and widely used instance of statistical inference" because it does not fit within the "existing theory". On the other hand, sciences have progressed by modifying and extending theories to accommodate facts. Various mutually conflicting resolutions of the paradox mentioned earlier raise the question: "What are the facts of statistics?" Can we come to an agreement as to what, at the moment, is the subject matter of statistics?

Curiously, Genest and Schervish accept my second example, and Joshi seems to support them. But putting $a_i = 1$, $i = 1, \ldots, N$, we get as a special case in (2.7) $\bar{\theta}_s = \bar{y}_s - \phi$, where $\bar{y}_s = \Sigma_{i \in s} y_i/n$. I do not see how the fact that some $y_i : i \in s$ renders some $\boldsymbol{\theta}$ impossible, at all affects my argument. Similarly Joshi's resolution of my first example ignores the distinction between the "actually drawn" $s^0$ mentioned above and its permutations, which are equally probable (Godambe 1979a).

Perhaps the paradox could be better understood in the historical context. A few years ago, some statisticians energetically debated (for references, see Godambe 1975) whether individual "labels" are a relevant part of the "data" in survey sampling. By now the controversy seems to have been settled in favor of accepting "labels" as a relevant part of the "data"; almost all the works on survey-sampling are based on the 'label dependent model' I put forward long ago. Interesting references in this connection are Rao (1977, 1984). The paradox discussed above shows that "labels" all by themselves can provide some information.

## ADDITIONAL REFERENCES

Basu, D. (1981). Ancillary statistics, pivotal quantities and confidence statements. Presented at the Canadian Conference of Applied Statistics, Concordia University.

Godambe, V.P. (1975). A reply to my critics. *Sankhyā Ser. C*, 37, 53–76.

Godambe, V.P. (1979a). A comment on a point raised by C.R. Rao. *Sankhyā Ser. C*, 41, 156–158.

Lindley, D.V. (1979). Contribution to the discussion of Dawid (1979). *J. Roy. Statist. Soc.*, 41, 15–16.

Rao, C.R. (1977). Some problems of survey sampling. *Sankhyā Ser. C*, 39, 128–139.

Rao, C.R. (1984). Review of "Theory of point estimation" by E.L. Lehmann. *J. Amer. Statist. Assoc.*, 79, 942.