methodology but that it would also stimulate statisticians to make improvements so that it can bear comparison with the best statistical practice.

These, and many examples like them, might well be held to support the proposition that although social science has produced a considerable amount of statistical methodology it has not sufficiently influenced the thought or direction of mainstream statistics. If this discussion helps to bring that about it will have been well worthwhile.

Comment: It's the Interplay That's Important

Paul W. Holland

While he hides it well behind the mask of scholarly indifference, Clogg is hopping mad. He has had it up to here with effete mathematical snobs (couldn't data analyze their way out of a paper bag) telling him that, at bottom, all the really good ideas, even in statistics, come from mathematics and the physical sciences. The last straw was reading the same bilge in, of all places, R. A. Fisher! An understandable fury, but two things that it is good for applied statisticians to develop are a tolerance of foolishness and a very thick skin.

I have been "doing statistics" in the social sciences for most of my life, but on occasion the opportunity to examine raw data from the physical sciences arises, and it has always struck me on these occasions just how familiar they appear, even though they hail from an allegedly distant part of the scientific landscape. At the forefront of research, the high signal-to-noise ratio that some associate with physical science data simply isn't there. If it is, then we aren't seeing the new stuff, just the well-established old stuff. What does seem to distinguish the social from the physical sciences is that, in the latter, progress is always being made in instrumentation and the noise levels eventually get lower. Signals eventually stand out with great precision, but when they do, it is old news; it's just Kuhn's "normal science."

Most social scientists get used to the fact that, generally speaking, the noise level does not decrease (increasing n just introduces new ways to slice up the data, and we are then often back to our noisy little samples again). In this respect, progress in the social sciences is hard to make. Learning to live with this fact is part of the early socialization of social scientists and of the statisticians who work with them. Most learn the hard way that a correlation over 0.8 is probably an error, and one exceeding 0.95 is an error for sure. (These limits must be modified somewhat for those who routinely correlate a variable with a slightly modified version of itself, but the same basic fact must be learned even there.)

The ubiquity of noise is why statisticians and statistics are useful to science—if there is no noise, no uncertainty, then there is no need for us or it.

Oh Statist! seek Unruly's feast,
and shun Perfection's meagre fare.

Human variability is a root source of the wide application of statistics to social, behavioral and medical science, and the lack of such a reliable source of noise is why statistical science has a more limited role in routine, empirical, physical science.

Aside from these comments, I have little to say about the endless and, to me, the totally sterile "social versus physical science" debate. I do have opinions about some of the points and examples that Clogg mentions in his interesting paper, so I'll concentrate the rest of my comments on them.

There is no question in my mind that real problems, based on real data, influence the development of statistical procedures. How can it be otherwise? Example: structural zeros in multi-way contingency tables. Categorical variables can have impossible combinations—for example, male hysterectomies. Ignoring the fact that there will never be a non-zero entry in such a cell can lead to a false impression of the association in such tables, but I doubt if anyone ever thought much about the implications of structural zeros until they saw them in real data. However, structural zeros have long appeared in continuous distributions, for example, the uniform distribution on the unit disk, but not much is usually made of this, essentially mathematical, example except to counter the proposition that a zero correlation implies independence. This is a continuous example of the "quasi-independence" that Clogg discusses because the two variables are as independent as they

Paul W. Holland is Distinguished Research Scholar at the Educational Testing Service and Fellow at the Center for Advanced Study in the Behavioral Sciences. Address correspondence to ETS, 21-T, Rosedale Road, Princeton, New Jersey 08541.