if needed. The uncertainties were so great that the Navy's initial requirement for realiability would have been extremely costly.

Hodges' paper is a very welcome addition to the literature.

ADDITIONAL REFERENCES

BROWN, G. F., JR. and ROGERS, W. F. (1973). A Bayesian approach to demand estimation and inventory provisioning. *Naval Res. Logist. Quart.* **20** 607–624.

CHUANG, D. (1984). Further theory of stable decisions. In *Robustness of Bayesian Analysis* (J. B. Kadane, ed.) 165–228. North-Holland, Amsterdam.

EDWARDS, W., LINDMAN, H. and SAVAGE, L. J. (1963). Bayesian statistical inference for psychological research. *Psychol. Rev.* **70** 193–242. Reprinted in *Robustness of Bayesian Analysis* (J. B. Kadane, ed.). North-Holland, Amsterdam, 1984.

GOLDSTEIN, M. (1986). Separating beliefs. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti* (P. K. Goel and A. Zellner, eds.) 197–215. North-Holland, Amsterdam.

HILL, B. (1980a). On some statistical paradoxes and non-conglomerability (with discussion). In *Bayesian Statistics* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 39–49. University Press, Valencia.

HILL, B. (1980b). Robust analysis of the random model and weighted least square regression. In *Evaluation of Econometric Models* (J. Kmenta and J. Ramsey, eds.). Academic, New York.

KADANE, J. B. (1980). Predictive and structural methods for eliciting prior distributions. In *Bayesian Analysis in Econometrics and Statistics: Essays in Honour of Harold Jeffreys* (A. Zellner, ed.) 89–93. North-Holland, Amsterdam.

KADANE, J. B. (1986). Progress toward a more ethical method for clinical trials. *J. Med. Philos.* **11** 385–404.

KADANE, J. B. and CHUANG, D. (1978). Stable decision problems. *Ann. Statist.* **6** 1095–1110.

KADANE, J. B., DICKEY, J., WINKLER, R. L., SMITH, W. and PETERS, S. (1980). Interactive elicitation of opinion for a normal linear model. *J. Amer. Statist. Assoc.* **75** 845–854.

KADANE, J. B., SCHERVISH, M. and SEIDENFELD, T. (1986). Statistical implications of finite additivity. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti* (P. K. Goel and A. Zellner, eds.) 59–76. North-Holland, Amsterdam.

KADANE, J. B. and WINKLER, R. L. (1987a). de Finetti's methods of elicitation. To appear in *Probability and Bayesian Statistics* (R. Viertl, ed.). Plenum Press, New York.

KADANE, J. B. and WINKLER, R. L. (1987b). Separating probability elicitation from utilities. *J. Amer. Statist. Assoc.* To appear.

NOVICK, M. R. and RAMSEY, J. O. (1980). PLU robust Bayesian decision theory: Point estimation. *J. Amer. Statist. Assoc.* **75** 901–907.

SCHERVISH, M., SEIDENFELD, T. and KADANE, J. B. (1984). The extent of non-conglomerability in finitely additive probabilities. *Z. Wahrsch. verw. Gebiete* **66** 205–226.

WINKLER, R. L. (1980). Prior information, predictive distributions, and Bayesian model-building. In *Bayesian Analysis in Econometrics and Statistics: Essays in Honour of Harold Jeffreys* (A. Zellner, ed.) 95–109. North-Holland, Amsterdam.

Comment

Albert Madansky

The best way to referee a mathematics paper is to read only the statement of the theorem and then proceed along the lines of the flow chart given in Figure 1. The process of reading the entire paper through and checking its work in detail is clearly a second-best approach to refereeing. In the same vein, the best way to read a paper whose title is "Uncertainty, Policy Analysis, and Statistics" is to stop at the title and try to construct the list of questions that a paper with such a title should answer. And this I did.

My first question was: "Why is the area of policy analysis different from all other areas in which statistics is applied?"

I next speculated on how the word "statistics" in the title is to be used—to denote "things statisticians know" (i.e., the corpus of knowledge classified by *Mathematical Reviews* into category 62) or "things statisticians do?" And if the latter, when is it that the statistician crosses the invisible line between "doing statistics" and "doing something else?" Indeed, how is that invisible line defined? Finally, who today is classifiable as a statistician, now that our profession and the computer revolution have jointly made our wares as available as over the counter nonprescription drugs?

As to that word "uncertainty," I mused about whether Hodges is referring to the Knightian use of the term, as contrasted with "risk," to distinguish between subjective and objective probability? Or does he have a different use for that well-worn term?

Consistent with my paradigm for mathematical refereeing, I did not pass the title page until I had constructed answers to these questions, after which I dived into the paper. To my surprise I found none of my questions answered. Instead I found yet another list of the steps in the process that a statistician goes through when dealing with an applied problem, along

Albert Madansky is Professor of Business Administration and Associate Dean for Ph.D. Studies, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637.

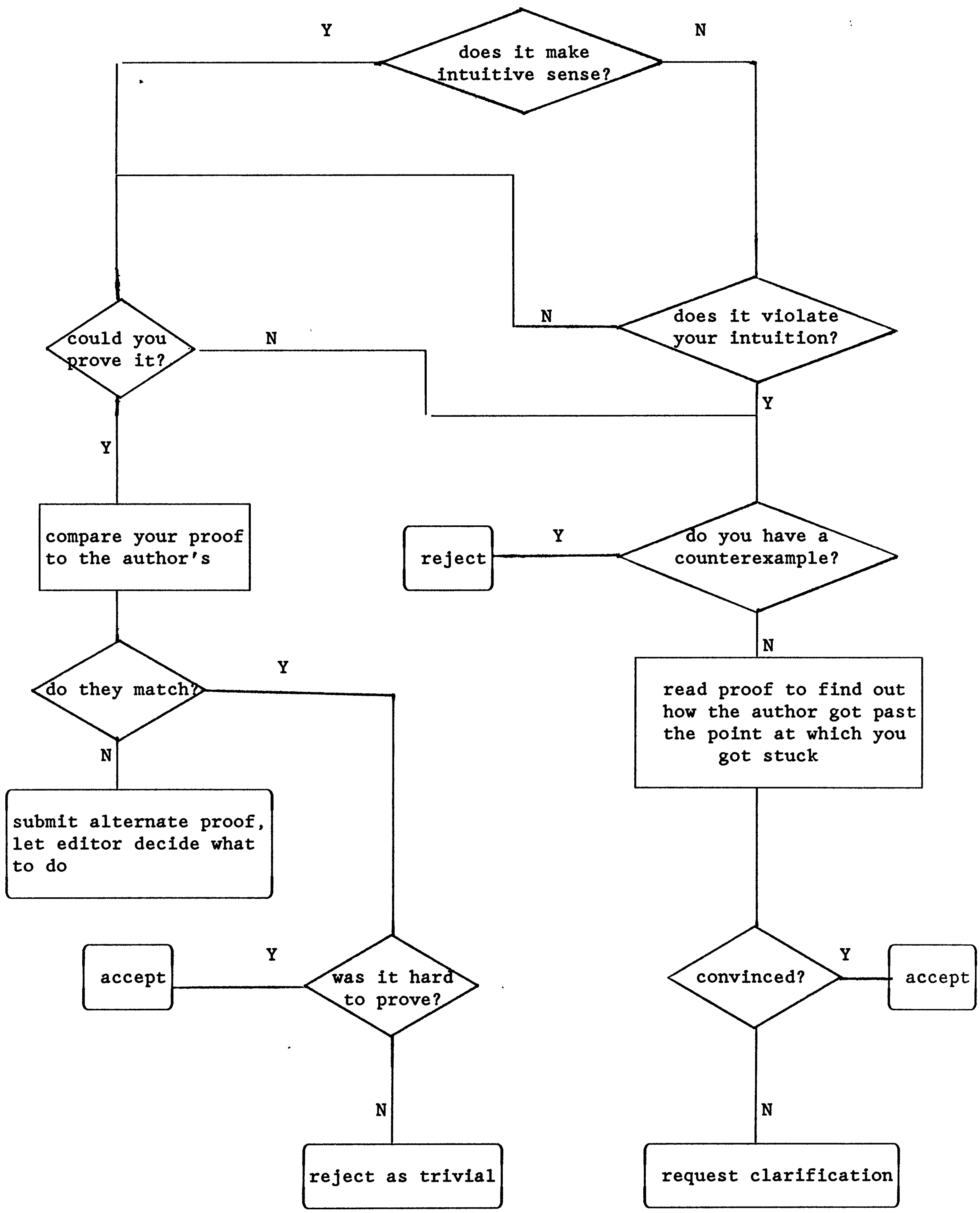

FIG. 1. *How to referee a mathematics paper.*

with the associated potential sources of error:

Step	Source of Error
Discovery/imposition of structure	Structural uncertainty
Assessment of variation conditional on structure	Risk (estimation/prediction)
Execution of techniques selected	Technical uncertainty

Them's mighty big words for the various errors that statisticians can make! I think in more earthy terms. The model created by the statistician (or his client) may be "off-base," the procedure recommended by the statistician may be "dead wrong" and for a variety of reasons the statistician may "drop the ball" in implementing his recommended procedure even if it (and the model) are quite good. Statisticians have concerned themselves with all these problems, as Hodges has noted. And what his paper does is give us a "scorecard" so that we can tell which statistical players are working on aspects of each of these errors. Others who have put into print their thoughts about aspects of these errors are Kimball (1957), who defined the "error of the third kind" as that of finding the right answer to the wrong problem, and Good (1980), and many other writings both before and since, who tries to construct a philosophy of data analysis. I welcome Hodges' scorecard, especially as it has led me to a number of interesting papers published in "off the beaten track" places.

But the paper left me with many more questions upon which to muse, some of which are the following:

1. Are these steps the only junctures at which a statistician can make an error? A full taxonomy of where statisticians can err even within the structure discovery step would undoubtedly entail such substeps as "developing a model based on subject-matter theory," "developing a model based on exploratory data analysis" and "checking a theory-based model against the data."

2. How much do each of these errors matter? Perhaps a theory of "the professional statistician's lifetime loss function" is needed. The statistician can no longer get away with such statements as "In my lifetime I will bat .950 with respect to my confidence intervals bracketing the parameters I will estimate and will field .950 with respect to the error of rejecting a null hypothesis when it is true."

3. How does a statistician remain objective in the face of his everdeeper involvement in exploratory data analysis, pretesting and the setting of "prior" probability distributions on model parameters? And how objective were they in the "good old days"? ("What used to be called prejudice is now called a null hypothesis," (Edwards, 1971).)

4. How "structural" are the structures found by statisticians? Often the theory with which to model real data is quite well-developed, but our statistical procedures aren't. For example, multiple linear regression is used to fit linear-in-the-parameters approximations to complicated functions dictated by theory only because our statistical technology is not so fully developed that it can work directly with the appropriate functions. The statistician's error exists because it is inherent in the process the statistician invokes in searching for structure. And the statistician's structure is at best an approximation to what would be considered by the client as a true structure.

5. Is the statistician's search for "structure" a search for "reality substitutes" or a search for "perspectives" (cf. Strauch (1983) for a discussion of this issue, especially in light of policy analysis)? Should statistical practice differ in these two contexts? And if so, how?

6. Where is mention made of the time-honored (but of anonymous authorship, hence not bibliographable) approach to the search for structure involving use of a "hold-out sample?" Has this become a casualty of the bootstrap?

Hodges devotes a great deal of attention to "prediction risk." This is quite reasonable, as the statistician (and also the soothsayer) are called upon to answer the question "Based on the past, what can I say about the future?" But look carefully at the "prediction" example given by Hodges. Why should one use $\bar{x}$ as the basis for predicting the next observation?

Let us step back from the fact that μ is unknown and ask what the prediction would be if μ were in fact known. Moreover, let us ask the question more generally, for the case where x is drawn from an arbitrary distribution with known mean μ and unit standard deviation. What is lacking in the problem formulation is a loss function. If the loss function is $E(x - p)^2$, where p is the prediction, then, because this loss function is the moment of inertia of $f(x)$ about p, it is minimized by taking the population mean as the prediction. But suppose the loss function were 0 if the prediction is "close" and $c > 0$ if it is not "close" to the new observation. Then, the population mode would be the best prediction, based not merely on the mathematics that leads to this conclusion but on the more intuitive dictum of Damon Runyon, "The race is not always to the swift, but that's the way to bet." Worse yet, for certain loss functions a "bold play" prediction, i.e., predicting a low probability observation (a "long shot") is the optimal prediction (cf. Dubins and Savage, 1965).

Now for the normal distribution the population mode is equal to μ, so when μ is unknown an estimate of the population mean (aka mode) is needed. Hence,

$\bar{x}$ is the prediction in Hodges' example. But in general, when the population mean is not equal to the population mode, it's moot whether one should use $\bar{x}$ or a sample-based estimate of the population mode as the prediction. Perhaps this "prediction" is not taught at Berkeley because it might gull statisticians into the bad habit of thinking that they should always predict the "average."

Enough of these questions. Since I so sorely missed answers to my original set of questions prompted by the title of Hodges' paper, I will use my remaining space to provide my answers to these questions.

ON POLICY ANALYSIS

There are a number of ways in which policy analysis is unique in its use of statistics. First of all, as Hodges say, in policy analysis "(w)e must act." Contrast this with other areas in which statistics is useful: as an extreme instance, in the discovery of theoretical constructs. (See the exchange between Freedman and Fienberg in Mason and Fienberg (1985) for some caricatures of the role of statistics in that context. I believe these are caricatures in the sense in which Hodges uses the term.) Secondly, because so much of policy analysis rests on cost/benefit analysis, there is a good deal of explicit concern given to loss functions here. Indeed, it is here more than in any other area of statistical application that the full panoply of concepts originating in economics and incorporated into statistics (e.g., expected utility maximization) come into explicit use. Unfortunately, the utility function is somewhat murky, ill-defined, not easily measurable and often measured either by surrogate measures or in indirectly observable quantities such as "opportunity costs."

ON STATISTICS

If the statistician were to circumscribe his domain merely to making statements such as "if the data $x_1, \ldots, x_n$ are independently drawn from a normal distribution with unknown mean μ and unit variance, then in repeated samples from this distribution the statistics $\bar{x} \pm 1.96/\sqrt{n}$ will bracket μ 95% of the time," there would be scant need for the kind of introspection given in Hodges' paper (and others, many of them cited by Hodges). It is precisely because the statistician's wares are used beyond the confines of the limitations expressed in such statements as quoted above, and by people we have trained in a limited fashion, that we as a profession must begin to "break mental set" in the way in which we teach applied statistics. It can no longer be taught as a "watered-down" parallel to a first course in mathematical statistics (as exemplified by both classical elementary textbooks as Dixon and Massey (1957) and modern elementary textbooks as Freedman, Pisani and Purves (1978)). Rather, it should take the student along Hodges' steps, searching for structure (perhaps including exercises in adducing utility functions and/or subjective probability distributions in addition to teaching exploratory data analysis), followed by training in assessment of variation, selection of techniques and their execution.

As to the related issue of the statistician's hubris to tread beyond the confines of the data, contrast the following quote from Chernoff and Moses (1959), "Years ago a statistician might have claimed that statistics deals with the processing of data . . . Today's statistician will be more likely to say that statistics is concerned with decision making in the face of uncertainty" with that of Kerridge (1968), "It is the statistician's job to inform, not to decide." To the extent that the statistician crosses the line and is more than a reporter and interpreter of data is it relevant to consider the issue about the role of statistics raised by Freedman in the papers cited in Hodges' references.

ON UNCERTAINTY

I find myself wishing that Tukey, with his penchant for inventing new words, had edited this paper. Given that Hodges felt the need to "give names to all the animals," the least he could have done is invented totally new names, not used old names, such as "risk" and "uncertainty," with pre-existent meanings, to mean new things.

ON THE PAPER ITSELF

I've spent about half my alloted space discoursing on the title of Hodges' paper, but what of the paper itself? Had I not been convinced from prior experience that indeed policy analysis is different from all other areas in which statistics is applied, I would not have found Hodges' arguments persuasive. Indeed, I would welcome a convincing essay confirming my priors about statistics in policy analysis. Had I not seen other lists of the steps in the applied statistician's process and the associated errors, I would have welcomed that of Hodges; given that I have seen such lists, I can find nothing in Hodges' paper that commends it above the others.

Finally, Hodges leaves me with a great deal of "uncertainty" as to where he comes out on the issue raised by the Freedman papers. Sure, no one (not even Dempster or I) would disagree with the thesis that studies should be criticized "on the grounds that they take insufficient account for structural uncertainty." But does that mean we should drop our tools, as the quote from Freedman suggests we do, and pass the buck completely? Doesn't a statistician have something to contribute even to the Freedman-recommended "ad hoc analysis by experts?" I am sorry if my desired for clever turn of phrase, wherein I refer to Freedman as the Neturei Karta of statistics, conveyed to Hodges the caricature characterization of Freedman's position as merely that of a defender of our discipline's virtue. I hoped to engage the reader to think about the more pressing issue, whether or not a statistician qua statistician has a role in (if you will) policy analysis when "the basic theory is incomplete or the data sparse." And I was hard put to pin down Hodges' position on this issue in this paper.

ADDITIONAL REFERENCES

CHERNOFF, H. and MOSES, L. E. (1959). *Elementary Decision Theory.* Wiley, New York.

DIXON, W. J. and MASSEY, F. J. (1957). *Introduction to Statistical Analysis*, 2nd ed. McGraw-Hill, New York.

DUBINS, L. and SAVAGE, L. J. (1965). *How to Gamble If You Must.* McGraw-Hill, New York.

EDWARDS, A. W. F. (1971). Science, statistics and society. *Nature* **233** 17–19.

FREEDMAN, D., PISANI, R. and PURVES, R. (1978). *Statistics.* Norton, New York.

GOOD, I. J. (1980). The philosophy of exploratory datum analysis. *Proc. Amer. Statist. Assoc. Bus. Econ. Statist. Sec.* 1–7.

KERRIDGE, D. F. (1968). Cited in Bibby, J. (1983). *Quotes, Damned Quotes, and....* Demast, Halifax, England.

KIMBALL, A. W. (1957). Errors of the third kind in statistical consulting. *J. Amer. Statist. Assoc.* **52** 133–42.

MASON, W. M. and FIENBERG, S. E. (1985). *Cohort Analysis in Social Research.* Springer, New York.

STRAUCH, R. (1983). *The Reality Illusion.* Theosophical Publishing House, Wheaton, Ill.

Comment

Adrian F. M. Smith

On the one hand, this paper claims that, both in theory and in practice, statisticians currently fail to acknowledge and incorporate important aspects of uncertainty in their modeling and analysis methodology, thus potentially distorting the inference and decision making processes in many areas of application. On the other hand, it claims that the subjectivist approach of de Finetti provides the most promising general framework for developing a language and methodology that might overcome the defects of current approaches. I am entirely in agreement with these views and therefore naturally welcome Hodges' paper, both in its own right and as a focus for a general discussion of the issues raised.

However, the structuring of the paper left me a little unclear as to what particular emphasis was intended in various of its sections. Sometimes, the emphasis seemed to be on drawing a pragmatic boundary between those problems and activities that can and cannot be approached by using some kind of more-or-less formal statistical modeling and analysis. At other times, the emphasis seemed to be on drawing attention to the unique merits of the Bayesian approach in providing a natural and unified framework for the development of precisely those tools that Hodges seems to consider so desirable, including predictive forms of uncertainty statements and between- as well as within-model uncertainty evaluations, both as outputs in themselves and as the basis for sensitivity analysis. Policy analysis applications seemed to fall somewhat between these two tools. Were we supposed to see policy analysis as an archetypal area where the boundary problem is particularly acute? Or as an archetypal area where Bayesian methods particularly come into their own? I fully realize that Hodges is attempting a grand overview of a large number of conceptual and practical problems that are all too rarely discussed together, but I would welcome some clarification from him of the main messages he was hoping we would extract from all this.

What I certainly do recognize from Hodges' running example and his general discussion is the total inadequacy of any view of modeling and analysis that does not appreciate the sociologic and institutional dimensions of dealing with large, messy systems in large, messy organizations. In an unpublished joint study undertaken for a major government agency in the United Kingdom, Dr. Ray Paul, of the London School of Economics, and I considered similar broad issues of model building and validation in representing and summarizing uncertainties in the context of very large scale problems. I shall briefly describe some of our general perceptions and conclusions and would very much welcome Hodges' views as to whether and to what extent we are thinking along the same lines. A

Adrian F. M. Smith is Professor of Mathematical Statistics, Department of Mathematics, University Park, Nottingham, NG7 2RD, England.