CONF-800820--1

Statistical Problems in the Assessment of Nuclear Risks

Robert G. Easterling
Sandia National Laboratories

1. Introduction and Background

Risk assessment goes to the heart of many fundamental problems of science and statistics -- inference, extrapolation, modeling, estimation. Suppose, for the sake of discussion, that we say there have been two serious accidents in 500 reactor-years of U.S. experience. What can we, or should we, infer about serious accidents, say in the next ten years? Who is comfortable with Poisson assumptions and the resulting statistical prediction or confidence intervals? Who can argue convincingly that the past is a random sample of the union of the past, present, and future? How much better off are we if we express accident occurrences as a function of various physical and human failures and bring the attendant data to bear? Difficult questions, yet ones we cannot ignore or dismiss. As John Tukey has said, "We should: Measure what is needed for policy guidance, even if it can only be measured poorly." But what is needed? Policy makers seem to be saying risk measurements are needed. Thus, if professional statisticians don't address these and related questions, others will (and already have).

The Reactor Safety Study (RSS), initiated by the Atomic Energy Commission in 1972 and completed under the auspices of the Nuclear Regulatory Commission (NRC), had as its objective a quantitative assessment of the risk to the public from reactor accidents. This assessment was to be "realistic" (as opposed to earlier studies which were "worst-case" or "conservative") and "uncertainty" was to be considered. Directed by Professor Norman Rasmussen of MIT, the study, which has become known as the Rasmussen Report, was a large effort, costing about four million dollars. A draft report was issued in late August, 1974, circulated for comment, revised, and then a final report (about 2000 pages in length) was issued in October, 1975. Both the draft and final versions were the subject of considerable praise and considerable criticism. The continuing criticism led the NRC to establish a Risk Assessment Review Group. This seven man group, chaired by Harold W. Lewis, Professor of Physics at the University of California, Santa Barbara, had the task of reviewing the criticisms of the RSS. After about a year of hearings and study, the Lewis Committee issued a report [2] in September, 1978.

The Review Group generally supported the RSS's probabilistic approach, but found that the RSS analysis included "the invention and use of wrong statistical methods." In January, 1979, the NRC issued a statement [3] accepting the findings of the Lewis Committee, withdrawing any endorsement of the report's Executive Summary (which had been the focus of much of the criticism), and declaring that "the Commission does not regard as reliable the Reactor Safety Study's numerical estimate of the overall risk of reactor accident."

This situation provides the impetus for statistical involvement in risk analysis. In fact, because criticisms of the RSS statistical methods reflect adversely on the statistical profession, we are already involved -- by implication if not by invitation. To make that involvement effective, there are several problems, both attitudinal and methodological, which need to be addressed. Some of these problems are discussed in the following sections.

2. Attitudinal Problems

Statisticians in all areas of application have encountered resistance and even resentment toward statistical "meddling." Perhaps because of proximity, the problem seems worse than usual in risk analysis. Much of the criticism of the RSS has been directed at the probabilistic and statistical methods used and the nuclear engineers and nonstatisticians who are the authors and supporters of the RSS results have naturally become defensive. The Lewis Committee, which incidentally included no statisticians, noted the "siege mentality" of the RSS staff. This problem needs to be overcome.

In addition to this natural resistance to criticism, there are other contributing factors to the conflicts between risk analysts and statisticians. One is the perception that engineering is pragmatic and practical while statistics is theoretical and esoteric. A corollary to this belief is that statisticians spend most of their time arguing about Bayesianism. We could all speculate about the extent to which these views are justified, but the important fact is that they are widely held.

My own view is that too often these views are adopted as a cop-out -- an excuse not to think carefully about data and the information they contain. These points won't be belabored, but it is important to point out that these problems exist and that it will take cooperative efforts to resolve them.

On a more technical level, another prevalent misconception of statistics is that a lot of data are needed before statistical methods are useful. Often one sees apologies for the lack of "statistically significant data" (?). With respect to nuclear risk estimation, I have often encountered the (mis)impression that the RSS shortcomings were in the limited amount of data involved rather than in the methods used to analyze those data. Again, these misconceptions must be overcome if we are to make progress.

The realization that meager data generally lead to imprecise estimates has led to dissatisfaction and has fed the impression that statistical methods are inadequate for risk analysis. For example, Apostolakis [4] calculates a statistical confidence limit on the rate of core melt accidents (Poisson assumptions) based on no core melts in 300 reactor-years, then dismisses the result for "misrepresenting our true beliefs." To obtain true beliefs, Apostolakis proposes the subjectivist Bayesian approach and illustrates it in a later paper [5], coauthored by Mosleh.

In that paper, the authors consider the following information:

A. The RSS "modal" estimate of the core melt rate is $\lambda^* = 1.5 \times 10^{-5}$ core melts per reactor-year.

B. The authors believe λ^* is low by a factor of 10.

C. Various critics of the RSS have indicated beliefs that λ^* is low by a factor of 12, 30, 33, 50, or 150.

D. There have been no core melts in 310 reactor-years of operation.

E. There was one "near-miss," the Browns Ferry fire (the paper was written pre-Three Mile Island). The probability that this accident might have led to a core melt has been estimated as 0.03.

In their analysis, Apostolakis and Mosleh use the beliefs A, B, and C to derive a "likelihood," and the data, D and E, to determine a "prior," which is certainly an unusual twist. The likelihood function obtained is the same function as the Poisson likelihood given one core melt in 6667 reactor-years. (This function is maximized at $\lambda = 10\lambda^*$.) The authors count the Browns Ferry fire as .03 core melts, and after some difficulties with gamma functions, arrive at a prior distribution equivalent to the pseudo-data of -.88 core melts in 120 reactor-years. Combining prior and likelihood leads to a function maximized at $\lambda = (1.0 - .88)/(6667 + 120) = 1.8 \times 10^{-5}$, remarkably close to the RSS value of $\lambda^* = 1.5 \times 10^{-5}$. Thus, Professor Rasmussen [6] was able to testify before Congress that even after the RSS results are modified by critics' opinions and the occurrence of the Browns Ferry fire, little change is noted.

It is fair to say, I think, that the analysis by Apostolakis and Mosleh is not an accurate summary of the information considered -

a view that should hold regardless of one's degree of Bayesianism. That it would be accepted as readily as it was by the nuclear community is disturbing. If Bayesian methods are to be useful in risk analyses (and I have my reservations about whether quantified true beliefs will help resolve safety issues) the example set by such analyses will have to be overcome. One difficulty encountered is that statistical criticism of infelicities committed in the name of Bayesian methods tend to be discounted as just being anti-Bayesian.

3. Methodological Problems

3.1 Expert Opinion

Expert, and inexpert, opinions have a role in risk assessment. Statistical problems arise in eliciting, characterizing, and summarizing opinions and in balancing these opinions against other data-based information. The nuclear industry needs to become aware of the statistical work which has been done in this area and I have the feeling, though this is not an area I'm familiar with, that there is an opportunity here for further development of methods.

3.2 Human Error Probabilities

One area in which an attempt has been made to use expert opinion is in the estimation of human error probabilities. At a conference sponsored by the IEEE in December, 1979, attendees were asked to give personal estimates of the probabilities of events such as failure to re-open one of several valves which had been closed for test or maintenance. Many "experts" felt a great reluctance, or even refused, to make such estimates and the estimates obtained varied widely. The effort can be regarded as a success only in that it demonstrated how difficult the problem of obtaining useful quantified opinion is. Such Delphi-like experiments would seem to provide more information on experts than on the events of interest. Surely statistical ideas in the design and analysis of experiments can help improve the situation. But can the improvement be enough to be useful? Even if a well-designed, well-run experiment in expert opinion is conducted and leads to the conclusion that 9 out of 10 experts believe that serious accidents will happen less than once per 1000 reactor-years, how is our state of knowledge advanced beyond that gleaned from the empirical evidence of 2 accidents in 500 reactor-years?

One of the lessons learned from the Browns Ferry and Three Mile Island accidents is that more attention needs to be paid to human performance. Because of this, there is increased emphasis on using control room simulators to train operators. These simulators have the potential for providing useful estimates of human error probabilities and it is hoped that statisticians will be involved in collecting and analyzing data for this purpose. The opportunity is there.

3.3 Nonindependent Events

One problem in estimating human error probabilities is that of correlated performance. For example, multiple tasks may be performed by a single person and because skill levels vary across the population of operators, over this population the performance on various tasks will be correlated. Accounting for this correlation and obtaining the data (or expert opinion) necessary for estimating the relevant probabilities pose some considerable problems.

It was in this area of estimating the joint probability of dependent failure events that the RSS garnered some of its severest criticism. Consider the expression $P(AB) = P(A)\ P(B|A)$, where A and B are two failure events. Available data may yield a reasonably precise estimate of $P(A)$, but because A happens rarely one may not be able to obtain a useful estimate of $P(B|A)$. However, one may have an adequate estimate of the unconditional probability, $P(B)$, so what the RSS does, in effect, to make use of this estimate, is assume $P(B|A) = \sqrt{P(B)}$. This assumption is not stated this directly in the RSS (if it had been, it might have died on the drawing board), but instead two contradictory rationales are advanced for what the RSS calls its "lognormal bounding technique." This label and the fact that in the "uncertainty analysis" carried out in the RSS, $P(A)$ and $P(B)$ were assumed to be lognormally distributed, in which case $P(AB) = P(A)\sqrt{P(B)}$ is lognormal, has created the misimpression, shared by the Lewis Committee, that this square root assumption follows from lognormality. It doesn't.

This situation, as was the case with the Apostolakis and Mosleh Bayesian analysis and as is shown by my review [7] of an important book on risk analysis, show that the primary need in risk analysis is to introduce statistical fundamentals, rather than develop new statistical methods. One area in which there is a need, I believe, to develop statistical methods is that called "uncertainty analysis."

3.4 Uncertainty Analysis

In the Reactor Safety Study, fault trees and event trees were used to obtain models by which the probabilities of various accidents were expressed as functions of the probabilities of other events such as equipment failures and human errors. These latter probabilities were treated as random variables and Monte Carlo techniques were used to approximate the resulting distribution of the accident probability. The assumed distributions were based partially on data from one year of reactor experience (1972) and from experience with similar equipment in other applications. In Appendix II of the RSS, pp. II-40, various interpretations are offered. It is said that these distributions represent "data variability from component to component and plant to plant," or that due to "differing applicable environments," or a Bayesian representation of "the assessed knowledge of the true-value probabilities" and it is claimed that these interpretations are equivalent. They aren't. The interpretation of such an analysis is thus up for grabs.

Another problem with this approach is unwarranted precision. For example, to pass from vague information expressed as, "We're pretty sure p_1 is between .001 and .010," to the very precise, "p_1 is lognormally distributed with a 5th percentile of .001 and a 95th percentile of .010," is to introduce a considerable amount of precision into the analysis - precision which may not be justified. The Lewis Committee concluded that the RSS "uncertainty bounds" (the 5th and 95th percentiles from the Monte Carlo analysis) were too narrow, but not because of this probabilistic treatment. They had in mind the uncertainty due to imperfect models, such as the square root assumption. With respect to the probabilistic treatment, the Lewis Committee, as did most reviewers of the RSS, focused on the secondary question of whether all these distributions were lognormal. In contrast, one reviewer, Amory Lovins, aptly noted that this treatment is a "method of substituting arbitrary (and weighted) compromise for technical ignorance."

The RSS analysis is an example of a class of analyses called uncertainty analyses. One quantity of interest, call it the output, is expressed as a function of other quantities, call then inputs, concerning which there is uncertainty of some sort or sorts. How then does uncertainty about the inputs translate into uncertainty about the output? This terminology is used because in some applications, such as modeling the course of a reactor accident or of leakage from a waste storage site, the function is a computer program.

The conventional approach to this problem, as in the RSS, is to choose probability distributions to represent the input uncertainties and then by Monte Carlo, Taylor series, or some other technique, approximate the resulting distribution of the output. It is amazing to see how readily this probabilistic treatment of imprecise information is performed and accepted. Were it not for the misleading or uninterpretable results which have been obtained, one would be tempted to become a Bayesian. Physicists and engineers learn about tolerancing and propagation of measurement error. These problems are usually worked by assuming known distributions, or at least variances, for the random variables involved. If in fact these assumed distributions are actually data-based estimates, the estimation error is generally ignored. It therefore is a natural step to apply the same methods when a parameter of interest, say system failure probability, is a function of component failure probabilities, the distinction between random variables and parameters being lost. Statisticians need to look very carefully at these analyses. They may provide some useful information, of a conditional type: If X varies as assumed, then Y varies (approximately) as derived, but they are quite subject to mis- and over-interpretation. For example, the 5th and 95th Monte Carlo percentiles in the RSS were called confidence limits.

As an alternative to probabilistic uncertainty analysis, I would like to see development of statistical uncertainty analysis. To give the problem some form and to provide a framework for some very preliminary results, the problem will be expressed as follows. Let η be the parameter of interest and suppose η is related to other parameters $\underline{\theta}$ through $\eta = h(\underline{\theta})$, where $h(\cdot)$ is known, but perhaps untractable. Suppose data, $\underline{X}$, are available from which $\underline{\theta}$ can be estimated and denote this estimate by $\underline{\tilde{\theta}}$. Then a natural point estimate of η is $\tilde{\eta} = h(\underline{\tilde{\theta}})$. Statistical problems immediately posed are those of obtaining the standard error of $\tilde{\eta}$ and confidence intervals for η.

Considerable attention has been given to some special cases of this problem. For example, when $\underline{\theta}$ is a vector of component reliabilities, η is the system reliability, and binomial data on the components are available, numerous methods for obtaining approximate confidence limits on the system reliability have been developed. Parametric tolerance limits are another example. In the normal case, the parameter of interest is $\eta = h(\mu,\sigma) = \mu + Z_p\sigma$. Still another example is the classic Behrens-Fisher problem. In this case, $\eta = \theta_1 - \theta_2$ and the data, X_{ij} are $NID(\theta_i, \sigma_i^2)$. It is somewhat embarassing that a problem which appears so simple is not simple to solve statistically and in fact has been the source of controversy. It gives one pause before proceeding to tackle much more complicated h- functions, but proceed we must.

Consider just the problem of obtaining a standard error of $\tilde{\eta}$. Several approaches are possible. In what follows these will be discussed briefly and illustrated by application to the Behrens-Fisher problem.

1. Parametric. Under this approach, the first step is to estimate, according to some well-defined statistical method, the sampling distribution of $\tilde{\underline{\theta}}$. Then, assume $\tilde{\underline{\theta}}$ has this distribution and derive or approximate the resulting distribution of $\tilde{\eta}$. Use the square root of the variance of that distribution as the standard error of $\tilde{\eta}$. In the Behrens-Fisher example, the natural assumption is $\tilde{\theta}_i \sim NID(\bar{x}_i, \frac{s_i^2}{n_i})$, which leads to

$$V_p = \mathrm{var}_p(\tilde{\eta}) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} ,$$

where the subscript denotes "parametric."

Mechanically, this is the same analysis as probabilistic uncertainty analysis. It differs, however, in that the data have a well-defined role -- which means the operating characteristics of the method are approachable -- and in the interpretation of the result. No probabilistic or confidence interpretation is ascribed to the derived distribution of $\tilde{\eta}$. Only the variance is used as an estimated variance of $\tilde{\eta}$.

2. Bootstrap. This approach, investigated by Efron [8], is a nonparametric version of the preceding approach. Rather than assume a distribution for $\underline{\tilde{\theta}}$, $\underline{X}$ is assumed to follow its observed empirical distribution and the resulting sampling distribution of $\tilde{\eta}$ is derived or approximated. As before, the variance of that distribution is used as the estimated variance of $\tilde{\eta}$. In the Behrens-Fisher case, this analysis yields

$$V_B = \mathrm{var}_B(\tilde{\eta}) = \frac{s_1^2}{n_1}\left(\frac{n_1-1}{n_1}\right) + \frac{s_2^2}{n_2}\left(\frac{n_2-1}{n_2}\right),$$

which is less than the variance obtained from the parametric approach.

3. Jackknife. In the Behrens-Fisher example, the jackknife analysis, where first single observations pertaining to θ_1 are omitted, then single observations pertaining to θ_2, leads to

$$V_J = \frac{n-1}{n}\left[\frac{{s_1}^2}{n_1-1} + \frac{{s_2}^2}{n_2-1}\right],$$

where $n = n_1 + n_2$. Note that V_J exceeds V_P.

4. Weighted Jackknife. This modification, proposed by Hinkley [9] for the case of linear models and so perhaps not applicable for general h- functions, yields, in the Behrens-Fisher case,

$$V_w = \frac{n}{n-2}\left[\frac{(n_1-1)s_1^2}{n_1} + \frac{(n_2-1)s_2^2}{n_2}\right]$$

$$= \frac{n}{n-2} V_B.$$

5. Bayes/fiducial. Under this approach, which is again probabilistic, but with a defined role for data, the data would be used to obtain the posterior or fiducial distribution of $\underline{\theta}$, given $\underline{\tilde{\theta}}$. (A diffuse prior might be assumed so there would be no difference between these approaches.) Then the resulting distribution of η would be derived or approximated and the variance of that distribution taken as the estimated variance of $\tilde{\eta}$. As with the parametric and Bootstrap approaches, no probabilistic, confidence, or fiducial interpretation would be imputed to the distribution of η because of well-known problems with nonlinear functions of multiple parameters. For the Behrens-Fisher problem, the fiducial distribution of θ_i is that of a scaled and shifted Student's t variable with n_i-1 degrees of freedom and this leads to

$$V_F = \frac{s_1^2}{n_1}\left(\frac{n_1-1}{n_1-3}\right) + \frac{s_2^2}{n_2}\left(\frac{n_2-1}{n_2-3}\right) \qquad (n_i > 3)$$

6. Linear Approximation. In this approach, the function h would be replaced, if necessary, by a simpler function, say one amenable to differentiation. (Methods for doing so are the topic of Hunter and Mitchell's paper, presented at this session.) Let $h^{(i)}$ denote the (approximated) derivative of $h(\theta)$ with respect to θ_i, and let v_i^2 and c_{ij} denote the estimated variances and covariances of $\underline{\tilde{\theta}}$. Then, the variance of $\tilde{\eta}$ can be estimated by

$$V_L = \sum_i [h^{(i)}]^2 v_i^2 + 2 \sum_{i<j}\sum h^{(i)} h^{(j)} c_{ij}$$

For the Behrens-Fisher problem, one would surely discover the simple model, $\eta = \theta_1 - \theta_2$, in which case, $V_L = V_P$.

Of the methods considered, only the jackknifes provide a measure of the reliability of the estimated variance of $\tilde{\eta}$, and that measure may be fairly crude. That is, it is sometimes recommended that the variance estimate be treated as being based on n-1 degrees of freedom, but it is not clear in what situations this is a good recommendation. An alternative approach would be to jackknife the variance estimate and use the usual moment matching method to obtain an effective degrees of freedom. If the computational problems could be overcome, this approach could be applied to any of the above methods. For the linear approximation approach, in the case of zero covariances, a simpler alternative is to use Satterthwaite's approximation to obtain an effective degrees of freedom. The linear approximation method has the additional advantage that the estimated variance of $\tilde{\eta}$ is partitioned according to the $\tilde{\theta}_i$. (In the case of separate data pertaining to each θ_i and jackknifing by dropping out one of the $n = \Sigma n_i$ observations at a time, the jackknifes also yield such a partitioning). This is useful, say in deciding what additional data are needed to improve the precision of $\tilde{\eta}$. It is possible the other approaches can be augmented or modified so as to effect a similar partitioning.

To carry the comparison of these methods further, in the case of the Behrens-Fisher problem, an example given by Hinkley [9] was considered. Let V_I denote the estimated variance of $\tilde{\eta}$ under method I. The variance of $\tilde{\eta} = \bar{X}_1 - \bar{X}_2$ is $V = \sigma_1{}^2/n_1 + \sigma_2{}^2/n_2$. As a measure of the performance of method I, we'll use $\rho_I = E(V_I)/V$. Also, because each V_I is a known linear combination of chi squared variables, by equating moments an effective degrees of freedom associated with V_I can be obtained. Following Hinkley, we consider the case of $n = 10$ and $\sigma_2{}^2 = 2\sigma_1{}^2$. Table 1 gives the resulting values of ρ and the effective degrees of freedom. The methods are listed in order of increasing ρ_I.

The Bootstrap method appears to be more nonconservative than I would like. For larger n, this problem might be alleviated, but still it is cause for concern. If the weighted jackknife variance estimate is treated as though it were based on 9 df, say for the purpose of calculating a confidence interval on η, it too is nonconservative. In contrast, the jackknife overestimates

TABLE 1

Values of $\rho_I = E(V_I)/V$ and the Effective Degrees of Freedom Associated with V_I (in parentheses): Behrens-Fisher Problem, $n = n_1 + n_2 = 10$, $\sigma_2^2 = 2\sigma_1^2$.

	n_1				
Method	3	4	5	6	7
Bootstrap	.75 (6.3)	.80 (8.0)	.80 (7.2)	.77 (5.2)	.70 (3.2)
Wtd. Jackknife*	.94 (6.3)	1.00 (8.0)	1.00 (7.2)	.96 (5.2)	.88 (3.2)
Parametric	1.00 (5.5)	1.00 (7.9)	1.00 (7.2)	1.00 (5.0)	1.00 (2.9)
Linear Approx.	1.00 (5.5)	1.00 (7.9)	1.00 (7.2)	1.00 (5.0)	1.00 (2.9)
Jackknife*	1.21 (4.8)	1.13 (7.8)	1.13 (7.2)	1.17 (4.8)	1.30 (2.7)
Bayes/Fiducial	----- -----	2.24 (6.8)	2.00 (7.2)	2.67 (4.1)	----- -----

*Values of ρ obtained from Hinkley [9].

the variance of $\tilde{\eta}$ but this would be offset somewhat by overstating the degrees of freedom. In this example, the Bayes/fiducial variance estimate is quite conservative. The effective degrees of freedom don't differ appreciably across methods. Whether these patterns hold over more complicated and realistic $h(\underline{\theta})$ remains to be investigated.

In some applications, evaluations of $h(\underline{\theta})$ may be expensive and so this factor would affect the choice of method. For the three methods involving sampling, normal theory results pertaining to estimating a variance with a predetermined precision should provide useful insight into how many evaluations of $h(\underline{\theta})$ are needed. This number would not depend on the dimensionality of $\underline{\theta}$ or $\underline{X}$. For the jackknifes, the number of evaluations depends on the nature of $\underline{X}$, but by deleting groups of more than one observation, rather than single observations, this problem can be alleviated. For the linear approximation method, the higher the dimension of $\underline{\theta}$, in general, the more evaluations of $h(\underline{\theta})$ will be required to obtain an approximate h. Even so, this might require fewer evaluations than the sampling methods or jackknifes. In all cases, it might be more efficient to obtain a simplified, inexpensive approximation for $h(\underline{\theta})$ before proceeding to the estimation of η. To do so, though, introduces an error of approximation which needs to be considered. Clearly, many problems remain to be investigated.

3.5 Acceptable Risk Criteria

The ability, real or imagined, to estimate the probabilities and consequences of hazardous events has led to a desire to set standards of acceptable risk. That is, can we, as a society draw a line on risks, and if so, where? It is in this area that the U.S. Nuclear Regulatory Commission has requested ASA participation. My view is that before these questions can be addressed, we need to set some standards or guidelines for the estimation of risks. As this paper has indicated, there are many problems. How should expert opinion be counted? What is required in the way of an uncertainty analysis? What models can be used? To what extent should they be "validated'? What constitutes an adequate peer review (which the RSS didn't get)? If we don't set some standards for risk estimation, then no matter what acceptable risk criteria are set, proponents of a given endeavor will come up with estimates showing the criteria are met, opponents will get estimates showing they're not. This will lead to unending arguments about assumptions and methods, as is the case of the RSS.

Even if acceptable estimates of risk-related parameters can be obtained, I'm not yet convinced we need quantitative risk acceptance criteria. It is perhaps a utopian view, but I would prefer that careful analyses be required in order to estimate the risks, costs, and benefits of a proposed endeavor, as competently as possible, and then lay these out in an easily understood way before the public and responsible decision-makers. This information, along with other nonquantitative information, would be assimilated and discussed and a decision reached. As I see it, quantitative risk acceptance criteria have three major drawbacks:

1. They encourage charlatanism. Even "noncharlatans" will be pushed toward ways to "sharpen their pencil," which is not the same thing as making their endeavor safer.
2. They encourage laziness and self-delusion. The "decision-maker" can go by the numbers and ignore the subtler nonquantitative aspects of the decision.
3. They don't allow for a grey area. Suppose the acceptance limit is 10^{-6} (in some dimension, the choice of which may be a considerable problem). What if the analysis yields 1.1×10^{-6} or 2×10^{-6}. Should major decisions rest on one's rules for rounding numbers?

4. Summary

Risk analyses have had and will continue to have, I believe, an important role in policy-making with respect to nuclear energy. For example, the U.S. NRC is moving toward requiring a risk analysis of nuclear power plant license applicants. Statisticians have an opportunity, even a responsibility, to help put those analyses on a sound statistical basis. The purpose of this paper has been to identify some of the problems which need to be addressed.

REFERENCES

1. U.S. Nuclear Regulatory Commission, Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, NRC Report WASH-1400 (NUREG-75/014) October 1975.

2. Risk Assessment Review Group Report to the U.S. Nuclear Regulatory Commission, NUREG/CR-0400, September 1978.

3. U.S. Nuclear Regulatory Commission, Office of Public Affairs, Release No. 79-19, January 19, 1979.

4. Apostolakis, G., Probability and Risk Assessment: The Subjectivistic Viewpoint and Some Suggestions, Nucl. Safety, 19(3): 305-315 (May-June 1978).

5. Apostolakis, G. and Mosleh, A., Expert Opinion and Statistical Evidence: An Application to Reactor Core Melt Frequency, Nucl. Sci. Eng., 70(2): 135-149 (May 1979).

6. Statement of Norman C. Rasmussen before the Subcommittee on Energy and the Environment, House Committee on Interior and Insular Affairs, Feb. 26, 1979.

7. Easterling, R. G., Review of An Anatomy of Risk, by W. D. Rowe, Technometrics, May 1980.

8. Efron, B. A., Bootstrap Methods: Another Look at the Jackknife, Annals of Statistics, V. 7, No. 1 (1979) 1-26.

9. Hinkley, D. V., Jackknifing in Unbalanced Situations, Technometrics, V. 19, No. 3 (1977) 285-292.