LETTERS TO THE EDITOR

COMMENTS ON "THE UNCERTAINTY IN ACCIDENT CONSEQUENCES CALCULATED BY LARGE CODES DUE TO UNCERTAINTIES IN INPUT' "

We would like to make some comments on Ref. 1. It is our belief that certain errors in the paper have led the author to an erroneous result.

One of his principal conclusions is that the probability distribution of calculated accident consequences is quite sensitive to the precise shapes of the input distributions, even when input means and standard deviations are assumed known. This conclusion, if true, would greatly reduce the value of many probabilistic analyses. This is because, in practice, there are rarely enough data available to decide whether a given input should be assigned, for example, a uniform or a normal distribution. Often the best that can be done is to provide a reasonable estimate of the mean and standard deviation and some justification for the general shape assumed (symmetric, skewed, etc.).

The basis for the author's conclusion is Fig. 3 of Ref. 1. However, there are serious discrepancies between Fig. 3 and Eqs. (14) and (15), on which it is ostensibly based. Consider Eq. (14):

$$p(x) = K(C_0 + C_1 x)^m \exp(-x/C_1) \ ,$$

where

$K = 1.369$

$C_0 = 0.9833$

$C_1 = 0.1639$

$m = 35.61.$

The heavy line in Fig. 3 cannot be the graph of $p(x)$, as we see by plotting a few points. Next, the value of the normalization constant K, given by the author, is not correct. The true value is

$$K = [C_1^{2M+1} \exp(C_0/C_1^2)\Gamma(M + 1)]^{-1} = 0.73 \ .$$

However, even when the correct value of K is used, we do not obtain the author's graph. Indeed, the (correctly normalized) density function $p(x)$ has a standard deviation of $C_1(M + 1)^{1/2} = 0.992$, which is clearly incompatible with Fig. 3. Similar remarks apply to Eq. (15).

To obtain some insight into the problem, we performed two 5000-run Monte Carlo simulations using the response surface [Eq. (13)] of the paper. We employed the input variable distributions assigned by the author in Table IV and following his Eq. (5), assumed a mean of 0 and standard deviation of 1 for each. The calculated distributions are shown in Figs. 1 and 2. The difference between the two distributions is slight, and the 95% confidence intervals for both cases are about equal.

This computational result is supported by more theoretical considerations. The response surface [Eq. (13)] consists principally of a reasonably large number of linear terms whose distributions are assumed symmetric and unimodal about 0. There is, therefore, a strong expectation that the response distribution is normal or nearly so. In other words, one does not expect the precise *type* of symmetric unimodal distribution assigned, whether uniform or normal or whatever, to have an important influence on the response distribution. Moreover, the intuitive explanation given by

Fig. 1. The probability density function (pdf) of time-to-failure for mixed input, based on 5000 Monte Carlo runs.

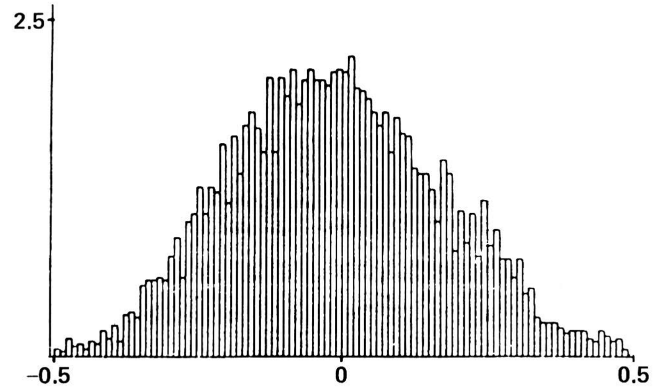

Fig. 2. The pdf of time-to-failure for uniform input, based on 5000 Monte Carlo runs.

the author for the difference between the graphs in Fig. 3 (namely, that "a uniform distribution is narrower than a normal distribution with the same standard deviation") is not convincing. To the contrary, see Fig. 3 of this Letter; the normal distribution is actually the more peaked of the two and, indeed, has the larger kurtosis (3 as compared to 1.8). Thus, if anything, one would expect the output distribution for "mixed" input to be somewhat more sharply peaked than that for "uniform" input. Comparison of Figs. 1 and 2 reveals such an effect. Its minor nature, however, only reinforces our conclusion that the output distribution is *not* particularly sensitive to the precise shapes of the input distributions.

In conclusion, we can only speculate as to the source of the errors described. The method of second-order error propagation, followed by moment matching to a member of a prescribed family of distributions, does not seem especially suited for use in this study. Its application involves the assumption that two distributions with the same first four moments do not differ substantially. However, it is well known that a distribution may not be determined even by its entire infinite sequence of moments. This is by no means merely a theoretical possibility, since the popular log-normal distribution is a case in point (see Fig. 4). Moreover, it is not easy to measure the size of the potential errors that can be introduced by the method. The Monte Carlo method for calculating output distributions seems preferable in this type of study for two reasons. First, it is simpler and more flexible. Second, one can easily

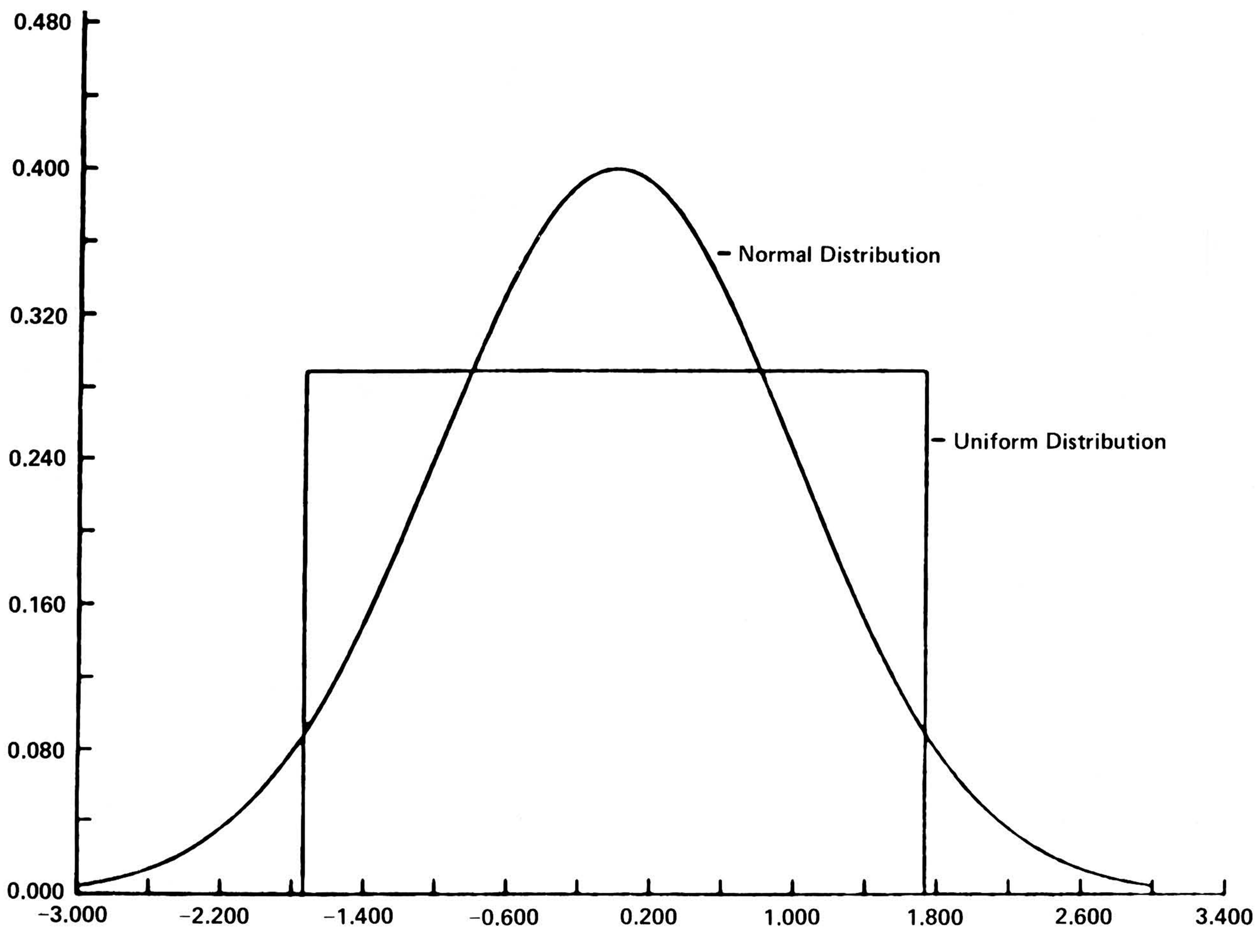

Fig. 3. Normal and uniform densities each with standard deviation of 1.

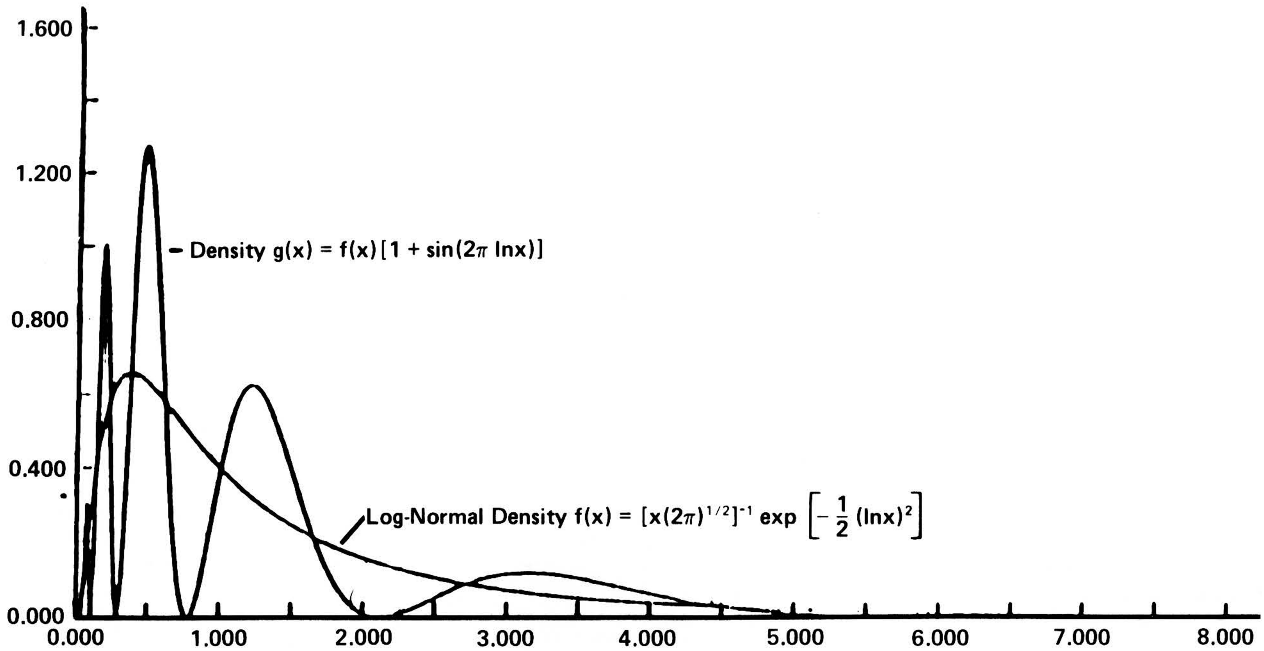

Fig. 4. Two different pdf's with the same moments.

obtain confidence bounds for any quantity of interest, and uncontrollable sources of error are thereby avoided.

David C. Cox
Paul Baybutt
Robert E. Kurth

Battelle Columbus Laboratories
Nuclear and Flow Systems Section
505 King Ave.
Columbus, Ohio 43201

July 11, 1980

REFERENCE

1. DONG H. NGUYEN, "The Uncertainty in Accident Consequences Calculated by Large Codes due to Uncertainties in Input," *Nucl. Technol.*, **49**, 80 (1980).

REPLY TO "COMMENTS ON 'THE UNCERTAINTY IN ACCIDENT CONSEQUENCES CALCULATED BY LARGE CODES DUE TO UNCERTAINTIES IN INPUT' "

The critique of Cox et al.[1] of my paper[2] can be summarized as follows:

1. There are discrepancies between Fig. 3 and Eqs. (14) and (15).

2. The probability density functions (pdf's) of the calculated output for the two cases considered in my paper (inputs with all uniform distributions and inputs with mixed uniform-normal distributions) should not differ significantly. The source of error was speculated to originate from the methods of error propagation, followed by moments matching to a member of a prescribed family of distributions.

3. In general, "the output distribution should not be particularly sensitive to the precise shapes of the input distributions."

Concerning the first point, it is recalled that in my work, the second-order error propagation was calculated by the program SOERP (Ref. 3), yielding the mean, the variance, and the coefficients of skewness and kurtosis. These properties were then used in the program PDFPLOT (Ref. 4) to match a theoretical distribution or a member of an empirical family of distributions. The PDFPLOT performed intermediate calculations using standardized random variable (mean = 0, variance = 1), then rescaled back to the actual random variable for plotting. The discrepancies between Fig. 3 and Eqs. (14) and (15) in Ref. 2 are caused by the fact that the constants given following these equations are for the standardized, not for the actual, variable, while Fig. 3 shows unnormalized distributions of the actual variable. Also, the normalization constant should be $K = 1/1.369$. For the given sets of inputs, PDFPLOT calculated the correct distributions, the discrepancies being caused by an inconsistency on my part. But, this is not a very important problem.

The major critique is in the shapes of the distributions shown in Fig. 3 of Ref. 2. The basic argument that the output distributions resulting from two sets of inputs, each having symmetric unimodal, although different, distributions should not differ significantly appears plausible. Therefore, the details of the calculations were carefully reviewed and an error was found in the inputs to SOERP for the moments of the standardized rectangular distribution. The central moments μ_r of the distribution