A Method for Estimating Fragility Curves Using Expert Opinions

A. Mosleh, G. Apostolakis
School of Engineering and Applied Science, University of California, 5532 Boelter Hall, Los Angeles, California 90024, U.S.A.

Two models are presented for using expert opinions to estimate fragility curves for components and structures of nuclear power plants. The models are based on a Bayesian method of handling expert opinion and provide formal techniques for constructing fragility curves from a set of subjectively estimated percentiles. The two methods differ in the way the evidence is interpreted and used. The effects of the assumptions of each model are demonstrated through an example.

In method 1 it is assumed that the experts are estimating a lognormal fragility curve and that each expert presents his opinion in the form of a set of percentiles. These percentiles are assumed to represent a lognormal curve. Therefore only two percentiles are used in the model and the rest, if any, are ignored. The model explicitly accounts for the fact that each expert's percentiles are ordered, that is, the higher percentiles are larger than the lower percentiles. Also considered in the model are dependencies among similar percentiles estimated by the experts. These dependencies are modeled by allowing similar percentiles to be correlated by a positive or negative correlation coefficient. No correlations are assumed among percentiles estimated by the same expert and between dissimilar percentiles of different experts.

Method 2, on the other hand, allows for dependence among all percentiles and does not assume that each expert's percentiles represent a lognormal distribution. An important simplification used in method 2, as compared to method 1, is that the ranges of variation of higher percentiles for each expert are not limited to account for the fact that the percentiles are ordered.

In both methods a probability distribution is obtained for the parameters of the fragility curve, which is assumed to be lognormal. This probability distribution represents the likelihood of different fragility curves which can be deduced from the estimated percentiles. This distribution is then used to construct a family of fragility curves at various levels of confidence. In addition an expected fragility curve is obtained by finding the average of a lognormal distribution over the entire set of possible values for its parameters. The probability distribution of the parameters of the fragility curve is maximized to obtain the parameters of the most probable curve.

1. Introduction

In the assessment of the seismic risk of nuclear power plants the failure of various structures and components due to earthquakes is studied. The performance of a structure or component under seismic loads is usually characterized by a fragility curve which gives the frequency of failure at different peak horizontal ground accelerations [1]. It is assumed that failure occurs when the ground acceleration, a, exceeds the seismic capacity, c, of the component. Ideally one would need to test a large number of similar components for which a fragility curve is desired at various accelerations to obtain the fraction of components which fail at each acceleration. Such fractions in the limit of very large number of tested components form the fragility curve for that type of equipment and can be used to predict the behavior of similar equipment under seismic loads.

In reality, however, such tests for some components and structures are not practical. In those cases the opinions of experts about the seismic capacity of the component are used as a substitute for actual test data. The expert estimates may come in several different forms such as a set of parameters for the fragility curve or a set of percentiles and sometimes a "free hand" assessment of the curve. Most often, however, the experts provide a "best estimate" and a range at a certain level of confidence. These numbers then can be interpreted as percentiles of the fragility curve.

This paper presents a general method for estimating fragility curves using a set of expert-estimated percentiles. The method is then used to obtain the fragility curve of a component based on the opinions of two experts.

2. Fragility Curves

If the characteristics of the component under study were perfectly known, that is, if its seismic capacity had a known value, C_o, and also if the earth movement in earthquakes could be fully characterized by the ground peak acceleration, a, then the fragility curve would be a step function at $a = C_o$.

In reality, however, two kinds of factors cause the shape of a fragility curve to deviate from a step function. Firstly, many different types of ground motions have the same peak acceleration. That is, for the same "a" we can have, for example, different spectral content, duration, etc. Thus, the response of the component at a fixed peak ground acceleration is also a function of other parameters characterizing the earthquake. Secondly, due to variations in the material properties, design and manufacturing the seismic capacity of a generic category of equipment is not a fixed value. Therefore, even when a large number of similar equipment are tested to determine the fragility curve only a fraction of them fail at a given acceleration. The fragility curve therefore is a cumulative distribution function which gives, at any given acceleration, the fraction of components that fail at accelerations equal to or less than that acceleration.

The objective of this work is to estimate this cumulative distribution using a set of M percentiles estimated by N experts. The unknown, therefore, is a distribution and, depending on our state of knowledge and strength of the evidence, there would be some degree of uncertainty associated with the estimated curve.

A fragility curve can take any form. The only physical requirement is that seismic capacities have a natural lower bound which is zero. In this paper it is assumed that fragility curves are members of the lognormal family of distributions, which meet the above requirements and with two parameters have a reasonable degree of flexibility in taking different shapes. Furthermore, as it will be seen, this choice brings a great deal of mathematical convenience to our modeling techniques. Similar assumptions about the shape of fragility curves have been made by others [1].

3. General Approach

The models of this paper are based on Bayesian methods for handling expert opinion [2-5]. In this approach, expert opinions are treated as evidence which is incorporated into the analyst's or decision maker's own state of knowledge via Bayes' theorem:

$$\pi(x|E) = k^{-1} L(E|x)\ \pi_0(x) \tag{3-1}$$

where $\pi_0(x)$ is the decision maker's "prior" state of knowledge about the unknown quantity, x, i.e., prior to receiving the opinions of the experts; E is the set of such opinions, and $L(E|x)$ is the likelihood of the evidence E given that the true value of the unknown quantity is x. Finally $\pi(x|E)$ is the decision maker's posterior state of knowledge about x after receiving the set of expert opinions. The quantity k is a normalizing factor which makes $\pi(x|E)$ a probability distribution.

The likelihood term is the most important element of this method. It is through the likelihood that the decision maker models the experts, shows his evaluation of their credibility and determines the level of impact of their opinion on forming his posterior state of knowledge. It essentially shows how much the decision maker believes in the accuracy of each expert's estimate. For instance, if he thinks that the i-th expert is perfect, then the likelihood of the expert's estimate, x_i, being anything other than the true value, x, is zero. This situation is modeled by a likelihood function for that expert which is a delta-function concentrated about x. The resulting posterior distribution would also be a delta-function concentrated about x_i, which means that the decision maker entirely accepts the opinion of that expert. More detailed discussion of the meaning of the likelihood function and various special cases can be found in references [2-4].

4. Method 1

Suppose that we have a set of M percentiles from N experts. The evidence is then

$$E = \{x_{ij} \quad ; \quad i=1,\dots,N \quad , \quad j=1,\dots,M\} \tag{4-1}$$

where x_{ij} is the estimate of the i-th expert for the jth percentile. The objective is to find the assumed true lognormal fragility curve

$$f(x|\theta,\omega) = \frac{1}{\sqrt{2\pi}\ x\ \omega} \exp\left\{-\frac{1}{2}\left(\frac{\ln x-\theta}{\omega}\right)^2\right\} \tag{4-2}$$

whose parameters θ and ω, need to be estimated based on the evidence E. Bayes' theorem in this case is written as

$$\pi(\theta,\omega|E) = k^{-1} L(E|\theta,\omega)\ \pi_o\ (\theta,\omega) \tag{4-3}$$

The joint distribution of θ and ω determines the likelihood of different pairs of θ and ω and thus different fragility curves as possible candidates for being the true curve.

One representative of the entire set of such distributions is an average curve defined as

$$\bar{f}(x) = \int_\omega\int_\theta f(x|\theta,\omega)\pi(\theta,\omega|E)\,d\theta\ d\omega \tag{4-4}$$

The most probable fragility curve is the one whose parameters θ_m and ω_m maximize the distribution $\pi(\theta,\omega|E)$ i.e. the distribution which has the highest chance of being the true distribution. The main problem to be dealth with is the construction of the likelihood term, L, in equation (4-3). The principal question to be answered is how good each expert is in estimating each of the M percentiles he is providing. In addition, the issue of dependence among various expert estimates as well as between different percentiles estimated by the same expert must be investigated and, if possible, modeled explicitly.

To address these issues we adopt the models presented in [4] and, in particular, the multiplicative error model. In this model an expert's estimate is assumed to be the product of the true value and a random error term whose logarithm is assumed to be normally distributed, i.e.,

$$L(x_{ij}|x_{tj}) = \frac{1}{\sqrt{2\pi}\,\sigma_{ij}\,x_{ij}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x_{ij} - \ln x_{tj}}{\sigma_{ij}}\right)^2\right\} \quad (4\text{-}5)$$

where x_{ij} is the i-th expert's estimate of the j-th percentile whose true value is x_{tj} and σ_{ij} is the multiplicative standard deviation for the expert's error as subjectively assessed by the decision maker to reflect his degree of confidence in the expert's competence.

This model is generalized using correlations between the estimates as measures of dependence [4].

In the most general case one would want to model the dependence among various percentiles of the same expert and among non-similar percentiles of different experts (e.g., between x_{ij} and x_{km} for $i \neq k$ and $j \neq m$). However a principal guiding rule in this, as in any other model, should be a balance between accuracy and simplicity. In other words, while it is important to consider all the significant aspects of the problem, the model should not be made unnecessarily complicated especially when more complexity would not add much to accuracy. In this context the following assumptions are made.

a. Correlation is only assumed between estimates of various experts for the same percentile.

b. It is assumed that the percentiles provided by each expert are coherently estimated consistent with the lognormality of the actual curve. In other words, it is assumed that each expert is aware of the lognormality of the distribution and his estimates represent a lognormal distribution. In that case, since the lognormal has only two parameters and only two conditions are needed to completely determine them, two of the M percentiles provided by each expert contain all the necessary information and the other percentiles are redundent. It is further assumed that the experts give estimates for at least two similar percentiles.

c. Finally it is assumed that the second percentile of each expert is estimated independent of the first one. However, is it is acknowledged that, for a coherent expert, the higher percentile is not smaller than the lower percentile.

Assumptions a and b can be relaxed rather easily but only at the expense of the coherence assumption in c. This model is also discussed in this paper.

Reference 5 presents the details of the model for a general case involving N dependent experts. The results for two experts are as follows: The posterior distribution is

$$\pi(\theta,\omega|x_{11}, x_{12}, x_{13}) = k^{-1} L_i \cdot L_2 \tag{4-6}$$

where L_i's have the form of bivariate lognormal distributions [6]. The variables of L_i are x_{11} and x_{21} (the estimates of the two experts for the first percentile) which are assumed to be correlated with ρ_1 as the correlation coefficient. Each estimate is assumed to be distributed about the true value x_{t1} which is related to θ and ω the parameters of the fragility curve, by

$$ln\ x_{t1} = \theta + \omega z_1 \tag{4-7}$$

where z_1 is the value of the standard normal vairable corresponding to the first percentile.

The term L_2 is also a bivariate lognormal distribution, but the variables are u_1 and u_2 defined as the difference between the first and the second percentiles of experts one and two respectively:

$$u_i = x_{i2} - x_{i1} \qquad i = 1,2 \tag{4-8}$$

These two variables are also assumed to be dependent with ρ_2 as the correlation coefficient. Both u_1 and u_2 are assumed to be distributed about u_t, the difference between the true values of the two percentiles, which is related to θ and ω by

$$\begin{aligned} ln u_t &= ln(x_{t2} - x_{t1}) \\ &= \theta + ln(e^{\omega z_2} - e^{\omega z_1}) \end{aligned} \tag{4-9}$$

It is shown in [5] that the parameters of the fragility curve with the highest chance of being the true fragility are

$$\theta = \frac{z_1}{z_1-z_2}\ ln[\exp\ (\frac{Q_2'}{Q_2'} - \frac{Q_2}{Q_1}) + 1] + \frac{Q_2}{Q_1} \tag{4-10}$$

$$\omega = \frac{1}{Z_2-Z_1}\ ln\ [\exp\ (\frac{Q_2'}{Q_1'} - \frac{Q_2}{Q_1})] \tag{4-11}$$

where

$$Q_1 \equiv \frac{1}{(1-\rho_1^2)} \quad [\frac{1}{\sigma_{11}^2} + \frac{1}{\sigma_{21}^2} - \frac{2\rho_1}{\sigma_{11}\ \sigma_{21}}] \tag{4-12}$$

$$Q_2 = \frac{1}{(1-\rho_1^2)}\ [\frac{lnx_{21}}{\sigma_{21}^2} - 2\rho_1\ \frac{lnx_{11}+lnx_{21}}{\sigma_{11}\sigma_{21}} + \frac{lnx_{11}}{\sigma_{11}^2}] \tag{4-13}$$

and Q_1' and Q_2' have similar forms in terms of u_1 and u_2 with the corresponding σ's and ρ_2.

The expected curve is obtained by using eq. (4-6) in eq. (4-4). The case of two independent experts can be obtained by setting $\rho_1 = \rho_2 = 0$ in all of above equations.

In the next section a model will be presented in which the number of percentiles from each expert is not limited to two. However the condition $x_{i2} > x_{i1}$ is removed.

5. Method 2

In this model all of the evidence given by eq. (4-1) is used. Therefore the model is good for any number of percentiles. This means that we do not assume that the experts are aware of the lognormality of the fragility curve and the set of percentiles of each expert does not necessarily represent a lognormal distribution. To keep the model simple, however, we will not require that $x_{ij} \geq x_{ik}$ for $j > k$. The effect of this assumption will be observed in the example of the following section, where this method is compared to method 1.

The general case for N dependent experts and M percentiles is discussed in [5]. It is shown that the posterior distribution of θ and ω has the following form,

$$\pi(\theta,\omega|E) = k^{-1}\exp\{-\frac{1}{2}[ln\ \underline{Y} - (\theta\underline{I} + \omega\underline{z})]^T \underset{\sim}{\Sigma} [ln\underline{Y} - (\theta\underline{I} + \omega\underline{z})]\} \qquad (5\text{-}1)$$

where $\underline{Y}$ is the vector of all the expert percentiles, $\underline{z}$ is the vector of MN z_j's (the values of standard normal variable for each percentile), and $\underline{I}$ is a vector of NM unities. The quantity $\underset{\sim}{\Sigma}$ is the covariance matrix of elements of $\underline{Y}$ and represents the dependence among various estimated percentiles.

The expected curve is obtained by using eq. (5-1) in eq. (4-4). The parameters of the most probable fragility distribution is obtained by maximizing eq. (5-1) with respect to θ and ω. These results are derived and discussed in [5].

6. Example

The methods of sections 4 and 5 were applied to the estimates of two experts for the percentiles of the fragility curve of the "control" failure mode of a particular component used in nuclear power plants. The first expert's estimates are: 10th percentile: 1.50g, 50th percentile: 2.50 g and 90th percentile: 3.00g. The second expert's estimates are: 10th percentile: 1.25g, 50th percentile: 1.50g and 90th percentile: 2.00g.

As can be seen from the estimates, the experts were not trying to estimate a lognormal fragility curve. However, in method 1 only the 10-th and the 90th percentiles of each expert were used while in method 2 all three percentiles were used.

Figure 1 shows the resulting expected fragility curves based on both methods where no correlations are assumed between the experts. Also shown in the figure are bounds on the fragility curves for both methods. The methods of obtaining these bounds are discussed in detail in [5].

Figure 2 gives the results of similar calculations for the case where a correlation of 0.75 is assumed between similar percentiles of the two experts.

7. Discussion

Figures 1 and 2 show that method 1 results in tighter ranges compared to method 2. In other words, the upper and the lower bound fragility curves are closer to each other in the first method than in the second one. This is perhaps due to the fact that method 1 limits the variation of the second percentile for both experts and treats the information more realistically. In addition, the fact that the 50th percentiles of the two experts are not used in method 1 is not expected to have worked to the disadvantage of the accuracy of the method because the set of percentiles given by each expert seems to be fairly consistent with the lognormal distribution which in turn means that dropping the 50th percentiles is not a bad approximation.

The same reasons which have resulted in wider bounds in the second method have also contributed to the increase in variance of the expected curve based on the second method as opposed to that of the first method. This can be seen in both figures where the expected curves based on method 2 cover a wider range of accelerations compared with the expected curves obtained by method 1.

The effect of dependence when modeled by a positive correlation coefficient is shown in Figure 2. Although the difference between the results of the two methods as discussed above is preserved the bounds for the case where the experts are assumed to be dependent are wider as compared to the case where no dependence is assumed between the two experts

(Figure 1). Furthermore, both methods result in higher variances for the average curves for the dependent case. These results are intuitively expected, because one is more uncertain when dealing with two dependent experts as opposed to the case where the sources of information are independent.

Acknowledgment

This work was partially supported by the Seismic Safety Margins Research Program of Lawrence Livermore National Laboratory.

References

[1] Kennedy, R.P., C.A. Cornell, R.D. Campbell, S. Kaplan and H.F. Perla, "Probabilistic Seismic Safety Study of an Existing Nuclear Power Plant", Nuclear Engr. and Design 59, 315-338 (1980).

[2] Morris, P.A., "Decision Analysis Expert Use", Management Science, 20, 1233-1241 (1974).

[3] Apostolakis, G. and A. Mosleh, "Expert Opinion and Statistical Evidence: An Application to Reactor Core Melt Frequency", Nuclear Sci. and Engr. 70, 135 (1979).

[4] Mosleh, A., and G. Apostolakis, "Models for the Use of Expert Opinion", paper presented at the Workshop on Low-Probability/High-Consequence Risk Analysis, Society for Risk Analysis, Arlington, Virginia, June 1982.

[5] Mosleh, A. and G. Apostolakis, "Estimating Distributions Using Subjective Percentiles", to be published.

[6] Press, S.J., Applied Multivariate Analysis, Holt, Rinehart and Winston, (1973).

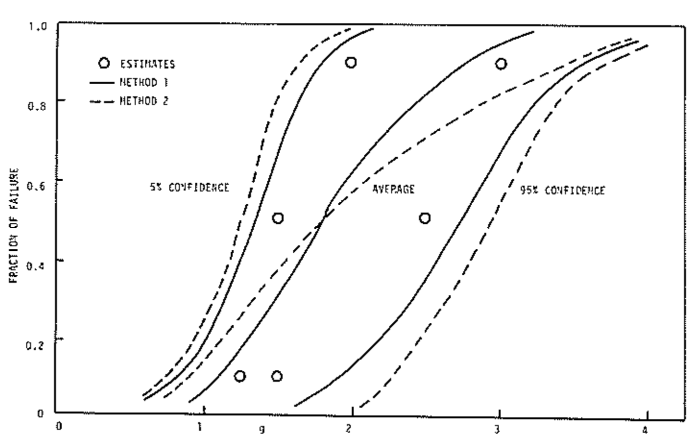

Fig. 1 Fragility Curve Based on Two Independent Experts

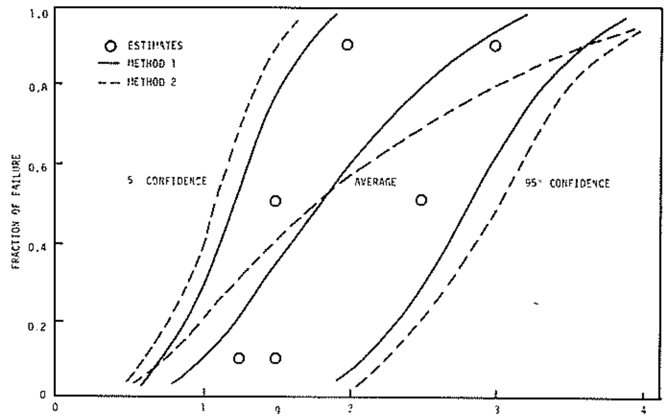

Fig. 2 Fragility Curve Based on Two Dependent Experts