DOI: 10.21683/1729-2646-2018-18-3-31-38

On an approach to the evaluation of the latent risk of expert assessment of roadbed seismic stability[1*]

Sergey K. Dulin[2], *JSC NIIAS, Moscow, Russia*
Igor N. Rozenberg[3], *JSC NIIAS, Moscow, Russia*
Vladimir I. Umansky[4], *JSC NIIAS, Moscow, Russia*

Sergey K. Dulin

Igor N. Rozenberg

Vladimir I. Umansky

Abstract. *The paper aims to examine the problem of integration of the opinions of a group of experts regarding a certain probabilistic distribution for the purpose of its evaluation by an analyst. It is implied that the decision-maker will use the result to evaluate the target risks and take according decisions. This problem may arise in many areas of risk analysis. For the purpose of this paper, the stability of various structures (buildings, railways, highways, etc.) against external mechanical effects, e.g. earthquakes, is chosen as the application object domain. As the primary research tool it is suggested to use the probabilistic method of decision-making risk calculation associated with involving experts into the analysis of risk of roadbed and other structures destruction in case of earthquakes. The evaluation of the seismic stability of rail structures using expert opinions is based on the Bayesian approach. The proposed method of estimation by analyst of the probabilistic distribution (fragility curve) on the basis of the opinions of a group of experts allows, using the obtained results, formalizing and explicitly expressing the latent risk of expert assessment. The procedure developed subject to a number of limitations allowed obtaining an explicit expression for the latent risk of expert assessment. The theoretical constructs presented in this paper can be easily implemented as software that will enable interactive input of parameters and data of the model under consideration and obtaining the desired distribution and the value of "risk in risk". Such system, on the one hand, will allow verifying some intuitive assumptions regarding the behavior of results depending on the variation of parameters, and on the other hand, will be able to be used as the tool of expert assessment automation and analysis of its quality that helps making grounded decisions under risk. Further development of the proposed method may involve the elimination of the dependence of the value of "risk in risk" from the expert assessment. Implicitly, this dependence is present in the final expression, while ideally this risk is to be determined only by the expert ratings. The proposed approach can serve as the foundation of some practical optimization problems, e.g. the selection of the best group of involved experts from the point of view of minimization of this share of risk in cases of restricted funding of expert assessment (obviously, the higher the expert's competence, the more accurate his/her estimates are and, subsequently, the lower is the risk, yet the higher is the cost of such expert's participation). An associated problem can be considered as well. It consists in the optimal selection of experts for the purpose of minimization of assessment costs under the specified maximum allowable level of "risk in risk". As a whole, the proposed method of evaluation of an unknown distribution and calculation of risk is sufficiently universal and can be used in the context of mechanical stability of structures, but also a wide class of problems that involve the assessment of a certain probabilistic distribution on the basis of subjective data about it.*

Keywords: *latent risk of expert assessment, Bayesian approach, fragility curve, quantiles.*

For citation: *Dulin SK, Rozenberg IN, Umansky VI. On an approach to the evaluation of the latent risk of expert assessment of roadbed seismic stability. Dependability 2018;3: 31-38. DOI: 10.21683/1729-2646-2018-18-3-31-38*

[1*] Work performed with the financial support of RFBR (grant no. 17-20-02153 ofi_m_rzd).
[2] Institute of Informatics Problems of the Russian Academy of Sciences, Federal Research Center Information Technology and Control of the Russian Academy of Sciences, JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), s.dulin@ccas.ru
[3] JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), i.rozenberg@vniias.ru
[4] JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), v.umanskiy@vniias.ru

Introduction

When analyzing various risks using probabilistic approaches, one often deals with events, of which the frequency is extremely low, e.g. various catastrophic phenomena. In addition, experimenting on real objects is normally either impossible in principle (usually, in cases of natural disasters), or extremely costly. Consequently, analysts have to deal with the situation of acute shortage, inconsistency or sometimes complete absence of direct experimental data. This forces the decision-makers (DMs or analysts) to construct risk analysis procedures solely on the basis of specifically invited subject matter experts, i.e. individuals who possess specific knowledge.

That causes the problem of optimal consideration and integration of all presented opinions obtained using different methods that probably contradict each other. Naturally, the DM must somehow classify the experts depending on the degree of trust. Additionally, he/she must be able to evaluate how close to reality the obtained result is, i.e. how satisfactory the conducted expert assessment is. In other words, while involving experts into the process of risk analysis, he/she must have an idea of the magnitude of the risk of wrong results and what possible negative consequences their use might have. This latent risk of expert assessment, i.e. the risk associated with the very fact of experts' involvement in the risk analysis is the main focus of this paper. The aim is to create the perfect tool for its assessment. Naturally, it is not supposed to appear out of nowhere, but be based on some procedure that transforms the information obtained from the experts into the final aggregated DM opinion. Further, a specific problem will be formulated, of which the solution will form the foundation of a practical method of calculation of the latent risk of expert assessment.

The purpose of the above procedure would be to solve the problem of integration of the opinions of a group of experts regarding a certain probabilistic distribution for the purpose of its evaluation by an analyst. It is assumed that the DM will subsequently use the result to evaluate the target risks and take according decisions.

This problem can arise in many areas of risk analysis. For the purpose of this paper, the stability of various structures (buildings, railways, highways, etc.) against external mechanical effects, e.g. earthquakes, is chosen as the application object domain. It is described with the so-called fragility curves (per [1]). According to the definition, the indicator of fragility of a structure or its component is the probability of its destruction (or failure) under the specified value of the parameter that characterizes an external effect (e.g. in case of an earthquake this parameter is the peak horizontal acceleration of ground). Thus, the fragility curve can be considered as the distribution function (integral) of a random value that reflects the ability of a structure to withstand mechanical stress with this parameter as the argument.

In order to formalize the concept of "expert opinion", the so-called quantile approach was used that is described, for example, in [2] and consists in the following. A certain finite set of distribution function values is defined. The experts are to express their opinion regarding under what values of the variable the distribution function is equal to each of the proposed values. These values of the variable are the distribution quantiles that correspond to the specified probabilities.

Several alternative methods were proposed for the solution of the problem. In this case, taking into account the initial goal, i.e. the definition of the concept of "latent risk of expert assessment", the Bayesian approach should be chosen, as it is based on the idea that experts are in principle imperfect sources of information, and attempts to take this imperfection into consideration. As part of this approach, each estimate received from an expert is interpreted as a result of an experiment and, therefore, is considered a random value that is specifically the main object of analysis.

The theoretical foundations of the Bayesian approach were laid in the mid-1970's to early 1980's ([2-5]), after which the practical applications started to develop in different areas, including the one at hand ([6-8]). In order to ensure the consistency of the presentation of the proposed method of evaluation of the fragility curve, it will start with the Bayes' theorem that, at the same, will be described briefly in the aspects that were earlier described in literature.

Bayesian formula

As part of the chosen approach, the opinions of the experts are considered input data, point estimates of the quantile that have an effect of the DM's "state of knowledge" of the Bayesian distribution:

$$\pi\left(x^t \mid E\right)= k^{-1} L\left(E \mid x^t\right) \pi_0\left(x^t\right), \qquad (1)$$

where the following designations are used:

$\pi\left(x^t \mid E\right)$ is the DM's a posteriori notion of the distribution (specifically, the value of the variable corresponding to one quantile or another) after studying the expert opinions E (here, distribution density is involved);

$\pi_0\left(x^t\right)$ is the DM's initial (a priori) notion of the unknown distribution before studying the expert opinions;

E is the experts' opinion on the distribution;

$L\left(E \mid x^t\right)$ can be called "plausibility function" of the input data E provided that the true value of the unknown (estimated) quantity is x^t; the meaning of this formula will be clarified below[1*];

k^{-1} is the normalization constant.

Thus, the problem comes down to the estimation of the a priori distribution $\pi_0\left(x^t\right)$ and plausibility function $L\left(E \mid x^t\right)$. The latter is the key element and its correct interpretation

[1*] In the object domain under consideration (mechanical resistance of structures) the fact that the unknown quantity equals x^t means that the structure will be destroyed with the probability 1 under the maximum value of vertical acceleration equal to *xt*. The superscript "*t*" here means "true".

is vital to the understanding of the whole method. For the simplest case of one expert and one assessment of quantile x_1 we have:

$$\pi\left(x^t \mid x_1\right)=k^{-1}L\left(x_1 \mid x^t\right)\pi_0\left(x^t\right).$$

In this expression the quantity $L\left(x_1 \mid x\right)dx_1$ is a subjectively estimated by the analyst probability that the value received from the expert will be between x_1 and $x_1 + dx_1$ provided that the true value of the variable corresponding to the quantile equals x^t. Obviously, this notion is true for the case of several experts. Thus, the plausibility function is in some way the measure of accuracy of the expert's opinion from the point of view of the DM who uses it to construct his/her own subjective model of the former's ability to give a quantitative evaluation of an unknown quantity.

As to the a priori knowledge of the analyst, in this paper it will be described with a uniform distribution. That was done for the sake of simplicity and corresponds to the situation when before receiving the expert assessments the DM does not have any information on the nature of the desired distribution. In this case from his/her point of view the probability of the unknown quantity being between x^t and x^t+Δx^t does not depend on x^t, which exactly corresponds to the absence of any knowledge.

Limitations of the model

The problem of construction of the desired probabilistic distribution based on experts' opinions is in general extremely complicated. Therefore, in order to obtain a practically applicable result, some simplifying assumptions must be made regarding both the nature of the distribution itself and the properties of the plausibility unction.

First, it will be assumed that the desired distribution belongs to the lognormal family, i.e. its density is defined by two parameters, и and щ:

$$f\left(x\right)=\frac{1}{\sqrt{2\pi}\ \omega\ x}\exp\left\{-\frac{1}{2}\left[\frac{\ln x-\theta}{\omega}\right]^2\right\}. \qquad (2)$$

Accordingly, the fragility curve is determined by integral of (2). Taking into account the selected subject field, this assumption is completely valid. A number of research programs dedicated to the study of real fragility curves (e.g. see [1]) indicate that the integral of the function (2) approximates them with good accuracy.

In the context of this assumption the problem of finding the distribution is significantly simplified and is reduced to the estimation of its parameters. Bayes's theorem is rewritten as follows:

$$\pi\left(\theta,\omega \mid E\right)=k^{-1}L\left(E \mid \theta,\omega\right)\pi_0\left(\theta,\omega\right). \qquad (3)$$

Under known distribution of parameters, the final estimate of the fragility curve, i.e. a specific pair of parameters, must be chosen. It appears that the most logical choice is the most probable distribution. Its parameters are found from the maximum condition based on the parameters of a posteriori distribution density $\pi\left(\theta,\omega \mid E\right)$ and are, therefore, the roots of the system:

$$\begin{cases}\dfrac{\partial\pi\left(\theta,\omega \mid E\right)}{\partial\theta}=0,\\ \dfrac{\partial\pi\left(\theta,\omega \mid E\right)}{\partial\omega}=0.\end{cases} \qquad (4)$$

The second hypothesis concerns the input data that are the set of estimates:

$$E=\left\{x_{ij},\ i=\overline{1,\ N};\ j=\overline{1,\ M}\right\},$$

where x_{ij} is the estimate by the i-th expert for the j-th quantile. It will be assumed that the estimates for all quantiles given by all experts are mutually independent. Certainly, this is a very strong assumption that can only be approximately true, and even then under a small number of quantiles. However, in the simple model under consideration it provides satisfactory results. Accounting for the dependences between the estimates, while radically increasing the inconvenience of calculations and reducing the illustrative qualities of the model, does not always have a significant effect on the result.

Subject to the second assumption, the general plausibility function is simply the product of the individual ones:

$$L\left(E \mid \theta,\omega\right)=\prod_{i=1}^{N}\prod_{j=1}^{M}L_{ij}\left(x_{ij} \mid \theta,\omega\right) \qquad (5)$$

where $L_{ij}\left(x_{ij} \mid \theta,\omega\right)\Delta x_{ij}$ is the probability that the estimate of the value of the variable corresponding to the j-th quantile by the i-th expert will fall into the small interval $\left[x_{ij},\ x_{ij}+\Delta x_{ij}\right]$ provided that the parameters of true distribution are equal to и and щ.

And finally the last assumption concerns the description of the analyst's expectations regarding the result obtained in the process of the expert opinion formation. There are two models of experts (additive and multiplicative), in which the probability of the deviation of the expert's opinion (from the DM's point of view) about the unknown value from the true value is explicitly expressed, i.e. the basic idea of the Bayesian approach to accounting for the inaccuracy of the information obtained from the expert is implemented. In this paper, the multiplicative model will be used. It is briefly presented below.

According to this model, the analyst examines the i-th expert's estimate of the value of the variable corresponding to the j-th quantile as random variable X_{ij} that is the product of two terms:

$$X_{ij}=x_j^t B_{ij},$$

where x_j^t is the true value (defined by the unknown parameters of the lognormal distribution), while B_{ij} is the random variable that corresponds to the error. Taking the logarithm, we will obtain:

$$\ln X_{ij}=\ln x_j^t+\ln B_{ij}.$$

Assuming that random variable $\ln B_{ij}$ is distributed over the normal law with the mathematical expectation $\ln b_{ij}$ and dispersion σ_{ij}^2, we will obtain, as it is easy to show, the lognormal distribution of the expert estimate:

$$L\left(x_{ij}|\theta,\omega\right)=\frac{1}{\sqrt{2\pi}\sigma_{ij}x_{ij}}\exp\left\{-\frac{1}{2}\left[\frac{\ln x_{ij}-\left(\ln x_j^t+\ln b_{ij}\right)}{\sigma_{ij}}\right]^2\right\}. \quad (6)$$

This function well describes the behaviour of experts and is widely used. This is that "building block" (since it is the plausibility function for the case of one estimate given by one expert) that will be the foundation for the construction of the general aggregated plausibility function that is in equation (1).

One additional hypothesis is accepted in this paper: experts are considered to be sufficiently competent in their subject area to not make systematic errors. Therefore, in equation (6), the calculations will assume that

$$\ln b_{ij} = 0,$$

which corresponds to the absence of systematic shift. Thus, as part of the general idea of taking account of the inevitable inaccuracies in the obtained information, only one type of errors will be considered, the random ones. As the result, formula (6) is rewritten as follows:

$$L_{ij}\left(x_{ij}|\theta,\omega\right)=\frac{1}{\sqrt{2\pi}\,\sigma_{ij}x_{ij}}\exp\left\{-\frac{1}{2}\left[\frac{\ln x_{ij}-\ln x_j^t}{\sigma_{ij}}\right]^2\right\}. \quad (7)$$

Now, subject to the above assumptions, an a posteriori distribution of parameters $\pi\left(\theta,\omega\mid E\right)$, and, therefore, the desired estimation of the fragility curve can be constructed on the basis of experts' opinions.

Construction of the distribution

Let us first express the individual plausibility function that is defined by (7). The method of evaluation of the dispersions of estimationσ_{ij}^2 will be presented below. The true value of the variable that corresponds to the j-th quantile can be found using the assumption (2) of the true curve being part of the lognormal family. This value is associated with the lognormal distribution parameters as follows:

$$\ln x_j^t = \omega\, Z_j + \theta, \quad (8)$$

where Z_j is the value of the standard normal distribution variable (with zero mathematical expectation and the unit dispersion) corresponding to this quantile. Thus, formula (7) transforms into:

$$L_{ij}\left(x_{ij}|\theta,\omega\right)=\frac{1}{\sqrt{2\pi}\,\sigma_{ij}x_{ij}}\exp\left\{-\frac{1}{2}\left[\frac{\ln x_{ij}-\left(\omega Z_j+\theta\right)}{\sigma_{ij}}\right]^2\right\}. \quad (9)$$

Now, by assumption of mutual independence of all expert assessment, by substituting (9) into (5), and further (5) into (3) by virtue of the hypothesis of the uniformity of a priori distribution we will obtain:

$$\pi\,\left(\theta,\omega\mid E\right)=k_1^{-1}\exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\frac{\ln x_{ij}-\left(\omega Z_j+\theta\right)}{\sigma_{ij}}\right]^2\right\}.$$

This is the desired distribution of the lognormal distribution parameters. However, in this form it is inconvenient to examine it to the maximum on и and щ. By squaring the expression under the summation sign and extracting perfect squares by parameters, after sufficiently simple, yet tedious calculations we will obtain:

$$\pi\left(\theta,\omega|E\right)=K^{-1}\exp\left\{-\frac{1}{2}\left[\left(\frac{\omega-\Omega}{\sigma_\omega}\right)^2+\left(\frac{\theta-\Theta(\omega)}{\sigma_\theta}\right)^2\right]\right\}, \quad (10)$$

where

$$\Omega=\frac{\left(\sum_{ij}\sigma_{ij}^{-2}\right)\left(\sum_{ij}\sigma_{ij}^{-2}Z_j\ln x_{ij}\right)-\left(\sum_{ij}\sigma_{ij}^{-2}\ln x_{ij}\right)\left(\sum_{ij}\sigma_{ij}^{-2}Z_j\right)}{\left(\sum_{ij}\sigma_{ij}^{-2}\right)\left(\sum_{ij}\sigma_{ij}^{-2}Z_j^2\right)-\left(\sum_{ij}\sigma_{ij}^{-2}Z_j\right)^2}, \quad (11)$$

$$\sigma_\omega^2=\frac{\sum_{ij}\sigma_{ij}^{-2}}{\left(\sum_{ij}\sigma_{ij}^{-2}\right)\left(\sum_{ij}\sigma_{ij}^{-2}Z_j^2\right)-\left(\sum_{ij}\sigma_{ij}^{-2}Z_j\right)^2}, \quad (12)$$

$$\Theta\left(\omega\right)=\frac{\sum_{ij}\sigma_{ij}^{-2}\ln x_{ij}-\omega\sum_{ij}\sigma_{ij}^{-2}Z_j}{\sum_{ij}\sigma_{ij}^{-2}}, \quad (13)$$

$$\sigma_\theta^2=\left(\sum_{ij}\sigma_{ij}^{-2}\right)^{-1}, \quad (14)$$

while K^{-1} is the proportionality coefficient that does not depend on и and щ.

By substituting (10) in the system of equations (4) and solving it we will obtain the parameters of the most probable distribution:

$$\omega_m=\Omega,\quad \theta_m=\Theta\ (\Omega). \quad (15)$$

Thus, the desired distribution density function (the integral of which is the fragility curve) subject to all the above assumptions is as follows:

$$\pi\ (x)=\frac{1}{\sqrt{2\pi}\ \Omega\, x}\exp\left\{-\frac{1}{2}\left(\frac{\ln x-\Theta(\Omega)}{\Omega}\right)^2\right\}.$$

We should mention another type of expressions for ω_m and θ_m that will demonstrate the contribution of each expert assessment into the final result. By regrouping the terms in the sums we obtain:

$$\theta_m = \sum_{ij} c_{ij}^{\theta} \ln x_{ij}, \quad (16)$$

$$\omega_m = \sum_{ij} c_{ij}^{\omega} \ln x_{ij}, \quad (17)$$

where

$$c_{ij}^{\theta} = \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_{ij}^2} \left[\left(\sum_{ij} \sigma_{ij}^{-2} Z_j^2 \right) - \left(\sum_{ij} \sigma_{ij}^{-2} Z_j \right) Z_j \right], \quad (18)$$

$$c_{ij}^{\omega} = \frac{\sigma_\theta^2 \sigma_\omega^2}{\sigma_{ij}^2} \left[\left(\sum_{ij} \sigma_{ij}^{-2} \right) Z_j - \left(\sum_{ij} \sigma_{ij}^{-2} Z_j \right) \right]. \quad (19)$$

It is apparent that the contribution of the estimation of the j-th quantile given by the i-th expert is proportional to the value $\frac{\ln x_{ij}}{\sigma_{ij}^2}$. This fact will be used subsequently.

In the formulas that describe the resultant distribution there are values σ_{ji}^2, dispersions in the multiplicative model of errors that are part of the expression of plausibility function. The approach used for their estimation is also based on certain assumptions.

First, for one expert the dispersions of assessment for different quantiles are taken equal to each other. That means that the amount of random deviation of the estimate from the true value expected by the analyst does not depend on the quantile, but is defined only by the general degree of trust the DM has for this experts' opinion:

$$\sigma_{ij} = \sigma_i, \quad j = \overline{1, M}, \ i = \overline{1, N}. \quad (20)$$

These standard deviations of the opinions of each expert are to be estimated. To that effect, the concept of weight (or rating) w_i assigned to experts is introduced in this model. The value of this parameter determines the general degree of the analyst's trust in the i-th expert's opinion, the expected margin of error in the quantitative assessments he/she provides. Naturally, the higher is an expert's rating compared with the rest, the better the obtained distribution must correspond to his/her assessment.

As it is obvious from formulas (16)-(19), the terms are proportional to σ_{ij}^{-2}. Therefore, it appears to be natural to associate the dispersions with the weights in this way, i.e. taking into account (20),

$$w_i = \frac{\gamma}{\sigma_i^2},$$

where г is proportionality coefficient.

It should be noted, that, as it follows from (11) and (13), when identifying the parameters of the most probable distribution, only the relative values σ_i^2 matter, as in both formulas both the numerator and the denominator are dispersion-homogeneous and the degree of uniformity is identical. The absolute values affect the values σ_θ^2 and σ_ω^2, as it follows from (12) and (14). Thus, the scale of the weights of experts (under identical relationships among them) reflects only the quality level of the produced expert assessment, i.e. the expected probability that the obtained curve will be sufficiently close to the true fragility curve.

In order to obtain specific numerical estimations of the risk of involving experts, this scale, i.e. some "single", reference level of risk with which all values will be associated, must be identified at the beginning. There are problems that involve the minimization of the risk with some limitations, and in their context its magnitude is of no significance as regards the choice of the optimal solution. However, in some decision-making problems the value of latent risk of expert assessment is in itself an important indicator.

In this paper the scale will be identified as follows. Let us assume that w_i are known and let us examine the coefficient г. The "reference" value of dispersion (that corresponds to the weight of an expert's opinion equal to 1) can be evaluated by purely empirical methods.

Let us examine the graph of the lognormal distribution function (that describes the experts' errors) with the parameters y and b (Figure 1).

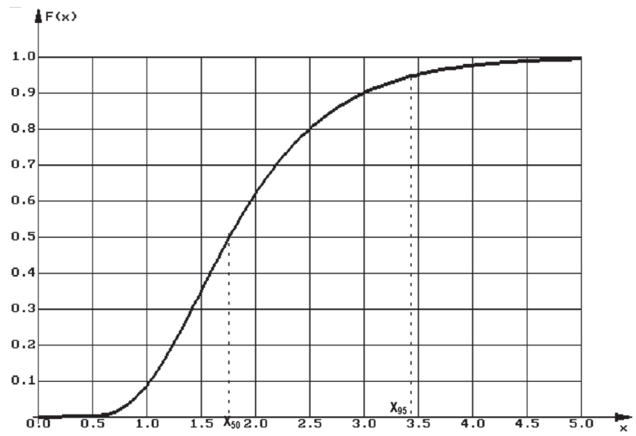

Figure 1. Lognormal distribution with the quantiles 50% and 95%

Let us find two values of the variable for quantiles with the probabilities 50% and 95%. They can be expressed with the corresponding values of the variable of the standard normal distribution and parameters (equation (8)):

$$x_{50} = \exp\left\{ \sigma Z_{50} + b \right\},$$

$$x_{95} = \exp\left\{ \sigma Z_{95} + b \right\}.$$

The value $\ln x_{95} - \ln x_{50}$ must characterize the internal scatter of the lognormally distributed random variable, and, therefore, its standard deviation as well. For tentative estimation of the latter let us assume that

$$\frac{x_{95}}{x_{50}} \approx 2,$$

i.e.

$$2 \approx \frac{\exp\left\{ \sigma Z_{95} + b \right\}}{\exp\left\{ \sigma Z_{50} + b \right\}} = \exp\left\{ \sigma \left(Z_{95} - Z_{50} \right) \right\},$$

out of which the desired “reference” standard deviation is:

$$\sigma = \frac{\ln 2}{Z_{95} - Z_{50}} \approx \frac{\ln 2}{1,645} \approx 0,421$$

(out of tables we know that $Z_{50} = 0$, $Z_{95} = 1.645$).

This value will be used in dispersions calculation:

$$\sigma_i^2 = \frac{0,421^2}{w_i} = \frac{0,177}{w_i}.$$

Identification of the latent risk of expert assessment

As it is known, the risk is defined by two factors: the probability of a certain adverse event and the expected losses caused by its realization. As part of the approach under consideration, a discrete model of a posteriori situation is adopted: the building has either collapsed or not, i.e. there are no intermediate options. Therefore, the expected magnitude of losses is the same, which allows disregarding it completely and equating the risk with the probability of destruction. Of course, that is just an approximation, and in more complex models the degree of destruction, among other things, can be considered as well, yet that is beyond the above described method.

The main point of the proposed method of evaluation of the latent risk of expert assessment consists in the following. For each value of the parameter that characterizes external effects, the local, “differential” risk is calculated, after which it is summed over this parameter subject to its distribution (in other words, the mathematical expectation is calculated). Naturally, that requires knowing this distribution (in the subject area under consideration that is the prediction of seismic situation that reflects the dependence of the probability of earthquake from its strength). Let us assume that it is known and designate it $f_0(x)$. The problem now is to obtain the expression for evaluation of the local risk given x.

The source of the latent risk of expert assessment associated with the involvement of experts is the probabilistic nature of the parameters estimation of the fragility curve that causes the possibility of its deviation from the actual situation, which ultimately leads to incorrect assessment of the initial risk that defined by the fragility curve itself. The probability of structure destruction, i.e. the value of the fragility curve in point x, should be considered the differential measure of the initial risk in that point. The local estimation of the latent risk of expert assessment, due to its nature, should be based on the value and probability of curve deviation (defined by the obtained distribution of lognormal distribution parameters) from the true value in point x.

It must be understood that unlike the estimation of the initial risk, the consequences of deviation in different directions are essentially different from each other. The fact that the curve obtained as the result of expert assessment is below the true one means that the experts underestimated the risk in this point. That is fraught with the destruction of the building with a higher probability, i.e. the latent risk of expert assessment is of the same type as the initial one. In the opposite situation, when the experts overestimate the risk, in terms of permission, there seems to be no negative consequences, yet if preventive measures are taken in order to reduce the residual risk to the acceptable level, overexpenditure may occur, which is also undesirable. Obviously, the consequences in different cases must be taken into consideration differently. However, sufficiently approximate estimation of the latent risk of expert assessment can be done uniformly, while it is most convenient to perform calculations using the first procedure, which will be done below (Figure 2).

To estimate the local risk in point x, let us compare the true (unknown) fragility curve, the lognormal distribution function $p(x;\theta',\omega')$, shown in Figure 2 with a dotted line, with the curve obtained with a certain probability as the result of expert assessment (solid line in Figure 2). According to the above described method, the probability of one or another position of the expert assessment of the fragility curve is defined by a posteriori distribution of parameters $\pi(\theta,\omega \mid E)$, and the solid line reflects one of these possible positions. In this case the experts underestimated the risk of destruction in point x: instead of the real probability $p(x;\theta',\omega')$ they predicted a lower one, $p(x;\theta,\omega)$. In this context it appears to be quite logical to examine the risk associated with the involvement of experts as a share of the total risk.

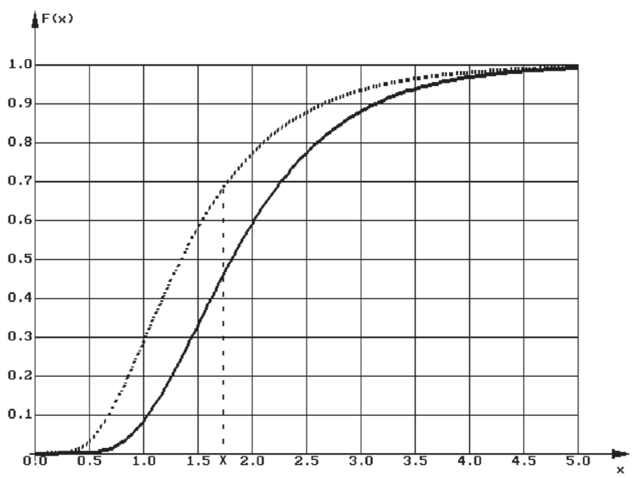

Figure 2. Identification of local risk

Thus, its quantity can be directly expressed as the difference between these probabilities. In Accordance with the chosen symmetrical approach to the assessment of deviations of unlike signs, in general, the module of this difference should be considered. Further, taking into consideration the probabilistic nature of the curve obtained as the result of analysis of the expert opinions, the local infinitely small risk in point x can be calculated as the mean module of difference with respect to the parameters:

$$d\tilde{R}=\iint\limits_{\theta,\omega}\left|p\left(x;\theta,\omega\right)-p\left(x;\theta',\omega'\right)\right|\,\pi\left(\theta,\omega\,|\,E\right)d\theta\,d\omega\,dx. \quad (21)$$

The tilde in this expression means that this is not the final formula. Its presence here has two reasons. First, as it was mentioned above, the parameters by the true fragility curve are unknown. However, taking into consideration the assumption of the absence of a systematic shift in the expert assessments (according to the error model), its good approximation for the purpose of averaging of the difference module is the most probable distribution with the parameters defined in (15) that was obtained above. Second, expression (21) does not reflected the fact that x is also a random variable with the distribution density $f_0(x)$. These two factors taken into consideration, the local risk writes as follows:

$$dR=\left\{\iint\limits_{\theta,\omega}\left|p\left(x;\theta,\omega\right)-p\left(x;\theta_m,\omega_m\right)\right|\pi\left(\theta,\omega|E\right)d\theta\,d\omega\right\}f_0(x)dx,$$

and, respectively, the risk associated with the involvement of experts in the analysis of the fragility curve, is defined by the expression:

$$R=\int\limits_x\left\{\iint\limits_{\theta,\omega}\left|p\left(x;\theta,\omega\right)-p\left(x;\theta_m,\omega_m\right)\right|\pi\left(\theta,\omega|E\right)d\theta\,d\omega\right\}f_0(x)dx.$$

This value should be compared with the total risk of structure destruction predicted by the experts. It is defined by the formula:

$$R_0=\int\limits_x p\left(x;\theta_m,\omega_m\right)f_0\left(x\right)dx.$$

If $R<<R_0$, the quality of the expert assessment must satisfy the analyst. With a sufficient certainty, he/she can analyze the various possible developments and make grounded decisions, e.g. as regards the advisability of facility construction or the development of preventive measures for reduction in the possible losses. But if value of the latent risk of expert assessment is comparable with the total risk (say, differs less than three to five times), the quality of the expert assessment does not allow considering its results as sufficient grounds for any decision making. This means that more competent experts must be involved or other methods be used for additional analysis.

This method of calculation of the latent risk of expert assessment examination, certainly, is not universal. It has limitations caused by the assumptions that were used to ensure logical transitions as part of distribution estimation. They substantially simplified the computations, yet at the same time reduced the method's applicability. One must be fully aware of what conditions must be fulfilled in order to obtain satisfactory results.

The first hypothesis concerns the chosen model of the experts' behavior when assessing the quantiles of the distribution, according to which they do not make systematic errors, while the random value is defined by one parameter (dispersion) directly associated with the expert's rating. In more complex models, in case of many expert assessments, the expected value of systematic shift can be statistically estimated and subsequently taken into consideration. Here, it is assumed that the subject area experts are sufficiently experienced to not allow it.

Further, an assumption, possibly, the strongest of all, is made regarding the mutual independence of all expert assessments for all quantiles. It can be performed only with a certain degree of approximation. On the one hand, the information sources used by the experts are largely common, which leads to the correlation of the opinions of different experts. On the other hand, usually they look at the distribution as a whole, and as the result the assessments made by the same expert for different quantiles begin to depend on each other. This effect is most apparent with an increasing number of quantiles, therefore in order to guarantee at least an approximate fulfillment of conditions of independence one must restrict oneself with just a few.

Another important aspect of the estimation is the requirement of the true distribution belonging to the parametric family. Although there are no restrictions on the nature of the family and number of parameters, i.e. in this sense a very wide spectrum of problems is covered, this condition is necessary, and in cases where for some reasons it is not possible to indicate the only family, this method is not applicable.

As a whole, the proposed method of evaluation of an unknown distribution and calculation of risk is sufficiently universal and can be used in the context of mechanical stability of structures, but also a wide class of problems that involve the assessment of a certain probabilistic distribution on the basis of subjective data about it.

Conclusion

The paper suggests a risk calculation method associated with involving experts into the analysis of risk of destruction of various structures (buildings, railways, highways, etc.) in case of earthquakes. The source of this particular latent expert assessment risk is the imperfection of experts as sources of information that causes uncertainty in the obtained results. It determines the additional risk caused by inadequate ideas of the possible development of the situation in case of materialization of unfavorable factors.

The Bayesian approach was chosen in order to take account of this uncertainty as it is best suited for its description. Based on a number of works, in which it was developed, as well as subject matter research, a method was proposed for the estimation by an analyst of the probabilistic distribution (fragility curve) on the basis of the opinions of a group of experts that allows, using the obtained results, formalizing and explicitly expressing the latent risk of expert assessment. The described method is based on some additional assumptions given above, therefore this definition of the latent risk of expert assessment is not universal and can only be used only in the context of problems of the same type with the same limitations.

References

[1]. Pagni CA, Lowes LN. Fragility functions for older reinforced concrete beam column joints. Earthquake Spectra 2006;22:215-238.

[2]. Kruschke JK, Aguinis H, Joo H. The time has come: Bayesian methods for data analysis in the organizational sciences. Organizational Research Methods 2012;15: 722-752.

[3]. Morris PA. Combining Expert Judgements: A Bayesian Approach. Management Science 1977: 23.

[4]. Steel PDG, Kammeyer-Mueller J. Bayesian Variance Estimation for MetaAnalysis: Quantifying Our Uncertainty. Organizational Research Methods 2008;11:54-78.

[5]. Zhang Z, Lai K, Lu Z, Tong X. Bayesian inference and application of robust growth curve models using student's t distribution. Structural Equation Modeling 2013;20(1):47-78.

[6]. Vбzquez Z, Esther R, O'Hagan A, Soares Bastos L. Eliciting expert judgements about a set of proportions. Journal of Applied Statistics 2014;41(9):1919-1933.

[7]. Baker E, Chon H, Keisler J. Advanced Solar R&D: Applying expert elicitations to inform climate policy. Energy Economics 2009;31:37-49.

[8]. Bordley RF. Combining the opinions of experts who partition events differently. Decision Analysis 2009;6:38-46.

[9]. Wisse B, Bedford T, Quigley J. Expert judgement combination using moment methods. Reliability Engineering and System Safety 2008;93(5):675-686.

About the authors

Sergey K. Dulin, Doctor of Engineering, Professor, Chief Researcher, JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS); Lead Researcher, Federal Research Center Information Technology and Control of the Russian Academy of Sciences, e-mail: s.dulin@ccas.ru

Igor N. Rozenberg, Doctor of Engineering, Professor, Director General, JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), e-mail: i.rozenberg@vniias.ru

Vladimir I. Umansky, Doctor of Engineering, Deputy Director General, JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIAS), e-mail: v.umanskiy@vniias.ru

Received on: 01.03.2018