Special Workshop on Risk Acceptance and Risk Communication
March 26-27, 2007, Stanford University

Aleatory or epistemic? Does it matter?

Armen Der Kiureghian

University of California, Berkeley
723 Davis Hall, University of California, Berkeley, CA 94720-1710, USA, adk@ce.Berkeley.edu

and

Ove Ditlevsen

Technical University of Denmark
Nils Koppels Allé, Building 403, DK-2800 Kgs. Lyngby, Denmark, od@mek.dtu.dk

Abstract

The sources and characterization of uncertainties in engineering modeling for risk and reliability analyses are discussed. While many sources of uncertainty may exist, they are generally categorized as either aleatory or epistemic. Uncertainties are characterized as epistemic, if the modeler sees a possibility to reduce them by gathering more data or by refining models. Uncertainties are categorized as aleatory if the modeler does not foresee the possibility of reducing them. From a pragmatic standpoint, it is useful to categorize the uncertainties within a model, since it then becomes clear as to which uncertainties have the potential of being reduced. More importantly, epistemic uncertainties may introduce dependence between events, which may not be properly noted if their character is not correctly modeled. Influences of the two types of uncertainties in reliability assessment, codified design, performance-based engineering and risk-based decision-making are discussed. Two simple examples demonstrate the influence of statistical dependence arising from epistemic uncertainties on systems and time-variant reliability problems.

1. Introduction

The nature of uncertainties and the manner of dealing with them has been a topic of discussion by statisticians, engineers and other specialists for a long time (see, e.g., Paté-Cornell 1996, Vrouwenvelder 2003, Faber 2005). This paper attempts to reopen that discussion in the context of structural reliability and risk analysis one more time. It is unlikely that this paper will bring a closure to that discussion. Yet, we hope that it will shed some light on the topic as it relates to such issues as assessment of structural reliability, codified design, performance-based design and risk-based decision-making. In particular, we will consider systems reliability and time-variant reliability problems, for which proper treatment of uncertainties is more crucial than for time-invariant component reliability problems. We argue that the nature of uncertainties and how one deals with them depends on the context and the application.

Engineering problems, including reliability, risk and decision problems, without exception, are solved within the confines of a model universe. This universe contains the set of physical and probabilistic models (or sub-models), which are employed as mathematical idealizations of reality to render a solu-

tion for the problem at hand. The model universe may contain inherently uncertain quantities; furthermore, the sub-models are invariably imperfect giving rise to additional uncertainties. Therefore, an important part of building the model universe is the modeling of these uncertainties. Any discussion on the nature and character of uncertainties should be stated within the confines of the model universe.

While there can be many sources of uncertainty, in the context of modeling, it is convenient to categorize the character of uncertainties as either aleatory or epistemic. The word aleatory derives from the Latin *alea*, which means the rolling of dice. Thus, an aleatoric uncertainty is one that is presumed to be the intrinsic randomness of a phenomenon. Interestingly, the word is also used in the context of music, film and other arts, where a randomness or improvisation in the performance is implied. The word epistemic derives from the Greek $\varepsilon\pi\iota\sigma\tau\eta\mu\eta$ (*episteme*), which means knowledge. Thus, an epistemic uncertainty is one that is presumed as being caused by lack of knowledge (or data). The reason that it is convenient to have this distinction within an engineering analysis model is that the lack-of-knowledge-part of the uncertainty can be represented in the model by introducing auxiliary non-physical variables. These variables capture information obtained through the gathering of more data or use of more advanced scientific principles. An uttermost important point is that these auxiliary variables define statistical dependencies (correlations) in a clear and transparent way.

Most problems of engineering interest involve both types of uncertainties. In the modeling phase, sometimes it may be difficult to determine whether a particular uncertainty should be put in the aleatory category or the epistemic category. It is the job of the model builder to make the distinction. The choice the model builder makes is, of course, conditioned on the general state of scientific knowledge, but much more on the practical need for limiting the sophistication of the model to a level of significant engineering importance for the decisions yielding from the model.

To provide a context for the following discussion, we consider the model universe for a structural reliability or risk analysis problem that involves a set of input variables $\mathbf{x} = (x_1, \ldots, x_n)$ that take values as outcomes of a corresponding set of basic random variables $\mathbf{X} = (X_1, \ldots, X_n)$, a parameterized probabilistic sub-model $f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\Theta}_f)$ describing the distribution of the random vector $\mathbf{X}$, and a set of parameterized physical sub-models $y_i = g_i(\mathbf{x}, \boldsymbol{\Theta}_g)$, $i = 1, 2, \ldots, m$, describing relations between the quantities $\mathbf{x}$ and m derived quantities $\mathbf{y} = (y_1, \ldots, y_m)$, which are employed in modeling the reliability or risk problem under study. The random variables $\mathbf{X}$ are called basic because we assume they are directly observable and, hence, empirical data is available for them. They may represent such quantities as material properties (strength, ductility, toughness, fatigue life, etc.), load characteristics (e.g., earthquake magnitude, wind velocity, wave height), other environmental effects (e.g., temperature, concentration of toxins, amount of pollution), and geometric dimensions (e.g., cross sectional sizes, location of supports, out-of-straightness). The derived variables $\mathbf{y}$ usually are not directly observable, except in laboratory or field studies aimed at model development. Engineering performance criteria usually are described in terms of such derived quantities, e.g., stresses, deformations, stability limits, measures of damage, loss, downtime, concentration of toxins in downstream waters. The sub-models $f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\Theta}_f)$ and $g_i(\mathbf{x}, \boldsymbol{\Theta}_g)$, $i = 1, 2, \ldots, m$, are invariably imperfect mathematical idealizations of reality and contain uncertain errors. Their parameters, $\boldsymbol{\Theta}_f$ and $\boldsymbol{\Theta}_g$, are usually assessed through a process of "fitting" these sub-models to observed data.

Most problems in reliability or risk analysis involve the above elements. Throughout this paper we will use these elements to discuss the modeling of uncertainties and to assess their relevance to risk and reliability evaluation in different application contexts.

2. Sources of uncertainty

In the context of the problem described above, one can identify the following sources of uncertainty:

1. Uncertainty inherent in the basic random variables $\mathbf{X}$, such as the uncertainty inherent in material property constants and load values, which can be directly measured.
2. Uncertain model error resulting from selection of the *form* of the probabilistic sub-model $f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\Theta}_f)$ used to describe the distribution of basic variables.
3. Uncertain modeling errors resulting from selection of the physical sub-models $g_i(\mathbf{x}, \boldsymbol{\Theta}_g)$, $i = 1, 2, \ldots, m$, used to describe the derived variables.
4. Statistical uncertainty in the estimation of the parameters $\boldsymbol{\Theta}_f$ of the probabilistic sub-model.
5. Statistical uncertainty in the estimation of the parameters $\boldsymbol{\Theta}_g$ of the physical sub-models.
6. Uncertain errors involved in measuring of observations, based on which the parameters $\boldsymbol{\Theta}_f$ and $\boldsymbol{\Theta}_g$ are estimated. These include errors involved in indirect measurement, e.g., the measurement of a quantity through a proxy, as in nondestructive testing of material strength.
7. Uncertainty modeled by the random variables $\mathbf{Y}$ corresponding to the derived variables $\mathbf{y}$, which may include, in addition to all the above uncertainties, uncertain errors resulting from computational errors, numerical approximations or truncations. For example, the computation of load effects in a nonlinear structure by a finite element procedure employs iterative calculations, which invariably involve convergence tolerances and truncation errors.

3. Categorization of uncertainties

In this section, we discuss the categorization of each of the uncertainty sources described above.

3.1. Uncertainty in basic variables

Consider a basic random variable X describing a material property constant, such as the compressive strength of concrete. Should the uncertainty in X be categorized as aleatory or epistemic? The answer depends on the circumstances. If the desired strength is that of the concrete in an existing building, then the uncertainty should be categorized as epistemic if it is decided that specimens taken from the building can be tested, yielding information about the strength. The testing may, of course, involve random errors of measurement, particularly if non-destructive methods are used. This measurement uncertainty should also be categorized as epistemic, if there is possibility of considering alternative methods of measurement. On the other hand, the uncertainty in the strength of concrete in a future building should be categorized as aleatory, if there will be no attempts to make more detailed modeling related to the control of the concrete production, for example. Until the building has been realized, no amount of testing will reduce the variability inherent in the strength of concrete of the future building.

The situation with demand (load) variables is somewhat different, as in assessing the reliability of both existing and future buildings, one is usually interested in future realizations of demand values. Hence, in this context, the uncertainty in basic demand variables is usually categorized as aleatory.

It is important to reiterate the difference between basic and derived variables. This is a choice made by the modeler, usually following standard engineering practice. Consider, for example, the annual maximum wind velocity, which may be of interest in designing a tower. The modeler may choose to consider this

quantity as a basic variable, in which case he/she would fit a probabilistic sub-model, possibly selected from some standard recommendation, to empirically obtained annual maximum wind velocity data. Alternatively, if such data are not available, the analyst may choose to use a predictive sub-model for the wind velocity derived from more basic meteorological data. In that case, the annual wind velocity is a derived variable of the form $y = g(\mathbf{x}, \boldsymbol{\Theta}_g)$, where $\mathbf{x}$ denotes the input meteorological variables and $g(\mathbf{x}, \boldsymbol{\Theta}_g)$ denotes the predictive sub-model of the wind velocity. The categorization of uncertainties in a derived variable is described below as a part of model uncertainty. As we will see, the uncertainty in a derived variable may be categorized as a combination of aleatory and epistemic uncertainties.

The arbitrariness in the choice of variables as basic or derived when building the analysis model suggests that the categorization of uncertainties in a problem depends on our choice of sub-models. By use of sub-models, we rely on empirical data on further basic variables, or sometimes on a priori probability assignments. A good example arises in seismic hazard analysis. Here, the interest is in the intensity of potential earthquake ground motions at a site, a demand variable. Since empirical data on the intensities of ground motions experienced at a specific site are hard to get, the common practice is to relate the intensity measure to the earthquake magnitude, for which empirical data is available, and to distance, for which an a priori sub-model can be used, e.g., the earthquake can be assumed to be equally likely to occur anywhere along an active fault. This is done through an "attenuation" law, which can be viewed as a predictive sub-model of ground motion intensity. In this formulation the ground motion intensity becomes a derived variable, whereas the basic variables are the earthquake magnitude and distance. In making this sub-model choice, we introduce additional uncertainties, which can have both aleatory and epistemic components, as described in the following section.

It is worth noting that the different categorization of uncertainties in an existing versus a future building dictates a fundamental difference in the methods used for assessing their reliabilities. For an existing building, the reliability assessment should aim at evaluating the reliability conditioned on the known history of the building. For example, the knowledge that the building has survived an earthquake of known intensity can be used to truncate the lower tail of the strength distribution. As more information is gathered, the uncertainty in the assessment decreases. In essence, this is a problem of information updating, for which Bayesian techniques are ideally suited. On the other hand, the problem of assessing the reliability of a future building, say during the design process, is one of determining the state of a random sample taken from a population. After all reasonable control measures have been taken into account, no updating with direct information can be performed until the building has been realized. This distinction in assessing the reliability of an existing versus a future structure has often been missed in the literature on structural reliability.

3.2. Model uncertainty

Consider a physical quantity y, which is uniquely determined in terms of two sets of basic variables **x** and **z**. We wish to develop a mathematical model (or sub-model) to predict *y*. Very often the exact form of the relationship between *y* and $(\mathbf{x}, \mathbf{z})$ is unknown. Furthermore, the modeler may not be aware of the dependence of *y* on **z**, or for reasons of pragmatism he/she may not wish to include these variables in a predictive model of *y*. For example, it may be practically impossible to gather data on the variables **z** and, therefore, including them in the model would not be useful.

As a specific example, consider the ground motion intensity attenuation model described above. We are well aware that the intensity at a site is dependent, in addition to the earthquake magnitude and distance,

on such variables as the propagation velocity of the fault rupture, the mechanical characteristics of the path of propagation of seismic waves, the geologic features surrounding the site, and so on. However, from a pragmatic standpoint, it is difficult if not impossible to measure these variables for a given site. Therefore, we exclude them in the attenuation model. These variables, as well as others of which we may not be aware, constitute the missing variables **z** in the ground motion attenuation model that is expressed only in terms of the earthquake magnitude and distance, which constitute the vector of basic variables **x** for the model.

The predictive model of y may be written in the form

$$y = \hat{g}(\mathbf{x}, \boldsymbol{\Theta}_g) + \varepsilon \tag{1}$$

where $\hat{g}(\mathbf{x}, \boldsymbol{\Theta}_g)$ is an idealized mathematical model involving the basic variables $\mathbf{x}$ with $\boldsymbol{\Theta}_g$ as its parameters, and $\varepsilon = y - \hat{g}(\mathbf{x}, \boldsymbol{\Theta}_g)$ is the model error (the residual). The parameters $\boldsymbol{\Theta}_g$ are usually estimated through statistical analysis of the model against observed data on y and $\mathbf{x}$. It is noted that, while it may be difficult to observe y for the particular risk analysis problem of interest, paired observations of y and $\mathbf{x}$ are necessary to assess the model in (1). These observations are usually conducted under special laboratory or field studies aimed at model development.

The model error ε has two components: (a) the effect of the missing variables $\mathbf{z}$, which are absent in the model, and (b) the effect of the potentially inaccurate form of the model. For example, the relationship between y and $\mathbf{x}$ could be nonlinear, while the model may use a linear form. Since these effects are uncertain, ε is modeled as a random variable. Usually, one is interested in an unbiased model. In that case, parameters $\boldsymbol{\Theta}_g$ are determined by setting the mean of ε equal to zero. Furthermore, by a proper transformation of the model, it is often possible for ε to have a normal distribution with its standard deviation σ_ε – a measure of the inaccuracy of the model – being independent of $\mathbf{x}$. This is known as the homoskedastic form of the model (Box and Tiao 1992). Thus, in order to completely define the model, the set of model parameters to be estimated is $(\boldsymbol{\Theta}_g, \sigma_\varepsilon)$. When more than one sub-model is involved, in addition to all the parameters $\boldsymbol{\Theta}_g$ and the standard deviation σ_ε for each sub-model, one will need to also determine the correlation coefficients between the error terms for different sub-models.

We now examine the nature of the uncertainties in the model of the form in (1). As explained above, ε accounts for the uncertain effects of the missing variables $\mathbf{z}$ as well as the potentially inaccurate form of the model. Both these uncertainties can be reduced if the model is refined to include one or more of the missing variables and/or mathematical expressions (analytical or algorithmic), which provide a better approximation to the correct form. In this sense, the uncertainty in ε is categorized as at least partly epistemic. However, our limited state of scientific knowledge may not allow us to further refine the model form and our inability to measure the missing variables may preclude the possibility of expanding the model. In such cases, at least a portion of the uncertainty in ε is categorized as aleatory. In particular, the part of the uncertainty in ε that arises from the effect of the missing variables is reasonably categorized as aleatory if these variables, though unknown, are characterized as being aleatory random variables.

We now turn to the probabilistic model (or sub-model) $f_{\mathbf{X}}(\mathbf{x}, \boldsymbol{\Theta}_f)$. This model is normally selected by fitting a theoretical distribution to available data. Various methods for evaluating goodness of the fit are

available. However, when events with small probabilities are of interest, as is the case in most structural reliability and risk problems, the tail of the probability distribution becomes important. Unfortunately, standard goodness-of-fit tests do not guarantee accuracy of the fit in the tail. For example, Ditlevsen (1994) has shown that equally well-fitted distributions can lead to significantly different probability estimates. Therefore, in computing probabilities, particularly for rare events, an error of uncertain magnitude arises from the assumed distribution model. This error can be placed in the epistemic category, since gathering of more data would allow a better fit of the distribution and, therefore, a reduction in the model uncertainty. However, unlike the case of physical models described above, it is difficult to assess the magnitude of the error arising from the choice of a distribution model. A logical way to do this would be to compute the probability of interest for all viable distribution models and assess the variability in the computed probability values. A second approach, suggested in Der Kiureghian (1989), is to parameterize the choice of the distribution. The uncertainty in the distribution model is then represented by the uncertainty in the parameter. However, both these approaches are demanding of large amounts of analysis.

The arbitrariness in the choice of the distribution model and the "tail-sensitivity" of small probabilities has lead to the recommendation that probabilistic structural design codes standardize probability distributions for load and resistance quantities (Ditlevsen and Madsen, 1989). One point of view is that in such a construct the computed probabilities should be considered as notional values and that caution should be exercised in using them in an absolute sense, e.g., for computing the expected costs of rare events. However, this view assumes that absolute probability exists as a physical entity outside the mathematical model by which it is computed.

It is only in very special problems, where one can think of the probability of an event as the relative frequency of the physical occurrence of the event in a long series of independent repetitions of an unchanged experiment, in which the event can occur. In the field of structural safety, several highly important sources of uncertainty do not exhibit such a repetitive behavior under identical circumstances. One can safely state that the interpretation of the stable long run occurrence frequency as an absolute probability in the physical sense belongs to utopia. Consequently, the usefulness of the probability concept must rest on another rational foundation. However, as a mind construct (and a simulation tool) the relative frequency interpretation of a mathematical probability is decisive for its usefulness as a degree of belief regarding the occurrence of an event. To make a probabilistic degree-of-belief model subject to a pragmatic test of falsification (a concept by Matheron based on Popper), and thus defensible as an objective tool, it is necessary that some type of relative frequency behavior be associated with the probabilistic model. A detailed discussion of the philosophy of this objectivity issue is given in Ditlevsen and Madsen (1996).

The mentioned codification of selected probability distributions should be seen as a consensus of the structural reliability engineering profession concerning the model elements to which the calculated probabilities are sensitive. Otherwise the engineering practice becomes open to unjustified distribution tail choices made for competitive reasons. Moreover, a useful common knowledge bank is obtained in this way, not the least with respect to the choice of distributions of the epistemic uncertainties. Clearly, for distributions based on sample data, the knowledge bank should be subject to revision as more data and better quality data become available.

To overcome the problem of arbitrary distribution choice, probabilistic codes are developed by a process of calibration to accepted practice, whereby it becomes reasonable at least to use probability as a means of comparison and adjustment. It even makes sense to use the standardized distributions in models for optimal decision-making, provided the intangible utility values are also calibrated so that the accepted practice on the average is the optimal practice.

In recent years considerable attention has been paid to developing performance-based engineering, particularly with regard to design of buildings and other structures to resist earthquake forces (Cornell and Krawinkler 2000). Central to this approach is the promise of computing risk associated with various structural performance requirements, including those of rare events such as extreme damage and collapse. In the rapidly developing literature in this field, little attention is paid to such issues as the tail sensitivity problem or the characterization of uncertainties inherent in the modeling and estimation. While this paper may not contribute to solving this problem, it raises a concern and hopefully sheds some light on the underlying issues and problems of an approach that relies on values of probabilities for rare events.

3.3. Parameter uncertainty

The parameters $(\boldsymbol{\Theta}_g, \sigma_\varepsilon)$ of the physical sub-models and $\boldsymbol{\Theta}_f$ of the distribution sub-model are estimated by statistical analysis of observed data. Specifically, $(\boldsymbol{\Theta}_g, \sigma_\varepsilon)$ are estimated based on pair-wise observations of Y and $\mathbf{X}$, and $\boldsymbol{\Theta}_f$ are estimated based on observations of $\mathbf{X}$. The preferred approach is the Bayesian analysis, which allows incorporation of prior information on the parameters, possibly in the form subjective expert opinion. The uncertainty in the parameter estimates is directly related to the amount and quality of the available information. By amount, we refer to the size of the available samples of observations. By quality, we refer to the accuracy in the observations. Any measurement error present in the observations deteriorates the information content and, hence, the quality of the data. The quality also refers to the information content in the prior. This kind of analysis is now routine and we will not discuss further details.

Parameter uncertainties are strictly epistemic because the uncertainty in the estimation decreases and may asymptotically vanish with increasing quantity and quality of the available observational data.

3.4. Final remark

The above discussions may raise the philosophical question whether there is any aleatory uncertainty at all. Clearly this question does not make sense outside the model universe. From a linguistic point of view, all uncertainties are the same as lack of knowledge. However, as explained above, it is convenient within a probabilistic model (mathematical statistical model, in particular) to introduce the categorization of uncertainties into aleatory and epistemic. Thus, within the model universe, the word epistemic assumes a more narrow meaning than just lack of knowledge.

Perhaps it is just a matter of time before it becomes sufficient to consider models that do not need the aleatory category, assuming that we learn about all missing variables and exact forms of models. Perhaps even basic variables can be explained through exact predictive models. In such a world, if uncertainty exists, it will only be epistemic. This utopian world, however, is too far from the reality of engineering practice today. The advantage of separating the uncertainties into aleatory and epistemic is that we thereby make clear which uncertainties can be reduced and which uncertainties are less prone to reduction, at least in the near-term. This categorization helps us in allocation of resources and in devel-

oping engineering models. Furthermore, better understanding of the categorization of uncertainties is essential in order to properly formulate risk or reliability problems. For example, epistemic uncertainties may introduce dependence among the estimated performances of the components of a system, and non-ergodic uncertainties may introduce dependence among a sequence of events in time or space. In practice, these dependences are often neglected due to improper treatment of the uncertainties. The examples in the following section demonstrate the influences of such effects.

4. Influence of uncertainties

In this section we present two examples to demonstrate the influence of uncertainties on reliability assessment. The first example demonstrates the influence of statistical dependence introduced by epistemic uncertainties among the components of a system. The second example demonstrates the influence of non-ergodic uncertainties, both epistemic and aleatory, in a time-variant reliability problem.

4.1. System reliability

Consider a *k*-out-of-*N* system. Such a system survives if at least *k* out of *N* components survive, where $1 \le k \le N$. The extreme values $k = 1$ and $k = N$ respectively define the special cases of parallel and series systems. For the sake of simplicity, we assume the components have statistically independent and identically distributed capacities represented by the random variable X_1, and also statistically independent and identically distributed demands represented by the random variable X_2. In essence, the component capacities and demands are random realizations from the distributions of X_1 and X_2, respectively. Thus, the components have identical limit-state functions defined by,

$$g(\mathbf{x}) = x_1 - x_2 \tag{2}$$

with $\{g(\mathbf{X}) \le 0\}$ indicating the failure event. We further assume that X_1 and X_2 are normal random variables with unknown means μ_1 and μ_2 and known standard deviations σ_1 and σ_2, respectively. According to the terminology introduced earlier in this paper, $\mathbf{X} = (X_1, X_2)$ are basic random variables and $\mathbf{\Theta}_f = (\mu_1, \mu_2)$ are distribution parameters to be estimated. Suppose the available information for estimating μ_1 and μ_2 are sample observations of size n of the capacity and demand values with respective sample means $\bar{x}_1$ and $\bar{x}_2$. It is convenient to adopt Bayesian modeling, in which μ_1 and μ_2 are considered as realizations of Bayesian random variables M_1 and M_2. Assuming independence of M_1 and M_2 and diffuse priors, these imply posterior distributions of M_1 and M_2, which are normal with means $\bar{x}_1$ and $\bar{x}_2$ and standard deviations $\sigma_1 / \sqrt{n}$ and $\sigma_2 / \sqrt{n}$, respectively.

As described earlier, the statistical uncertainty in the distribution parameters μ_1 and μ_2 is epistemic in nature. Since the component capacities and demands are identically distributed, this uncertainty is shared by all the components of the system. Hence, the statistical uncertainty inherent in the estimation of the distribution parameters introduces statistical dependence among the estimated states of the system components. To investigate this effect, we proceed as follows:

Observe that, for the case of linear limit-state function and normal random variables, the conditional reliability index of a typical component for given values of μ_1 and μ_2 is

$$\beta(\mu_1, \mu_2) = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{3}$$

Since M_1 and M_2 are normally distributed posterior to obtaining the data, it can be easily shown that $B = \beta(M_1, M_2)$ has the normal distribution with mean $\mu_B = (\bar{x}_1 - \bar{x}_2)/\sqrt{\sigma_1^2 + \sigma_2^2}$ and standard deviation $\sigma_B = 1/\sqrt{n}$. It is evident that the uncertainty in the reliability index in account of the statistical uncertainty is directly related to the size of observation samples. The probability of failure of the component conditioned on the distribution parameters is given by $p_f(\mu_1, \mu_2) = \Phi[-\beta(\mu_1, \mu_2)]$. Viewing this as a transformation between β and p_f, one easily determines that the distribution of the (Bayesian) random failure probability $P_f = p_f(M_1, M_2)$, reflecting the effect of the statistical uncertainty, is

$$f_{P_f}(p) = \sqrt{n} \exp\left[-\frac{n}{2}\left(\Phi^{-1}(1-p) - \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right)^2 + \frac{1}{2}\left[\Phi^{-1}(1-p)\right]^2 \right] \tag{4}$$

Figure 1 shows plots of the distributions of B and P_f for $\mu_B = 3$ and $n = 10$ and 30. The so-called predictive failure probability, denoted $\tilde{p}_f$, is the mean of this distribution. Its value can be obtained by either of the following formulas:

$$\begin{aligned} \tilde{p}_f &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_f(\mu_1, \mu_2) f_{M_1}(\mu_1) f_{M_2}(\mu_2)\, \mathrm{d}\mu_1\, \mathrm{d}\mu_2 \\ &= \int_0^1 p f_{p_f}(p)\, \mathrm{d}p \end{aligned} \tag{5}$$

The corresponding predictive reliability index is $\tilde{\beta} = \Phi(1 - \tilde{p}_f)$. Plots of $\tilde{\beta}$ and $\tilde{p}_f$ for $\mu_B = 3$ and as functions of the sample size n are shown in Figure 2. Note that with increasing n the predictive reliability index asymptotically approaches $\mu_B = 3$ and the predictive failure probability approaches $\Phi(-\mu_B) =$ 0.00135, the limiting values without statistical uncertainty. Increasing statistical uncertainty (decreasing sample size) tends to increase the predictive failure probability and decrease the predictive reliability index.

We now consider the system failure probability. Even though the component states are statistically independent, statistical dependence among their probability estimates is present due to sharing of the uncertain distribution parameters M_1 and M_2 by all components. Thus, to properly handle this dependence, we must first evaluate the conditional probability of the system failure given the distribution parameters, and then integrate over all possible values of the parameters. Using the complementary binomial cumulative probability function, the conditional probability of failure is

$$p_{sf}(\mu_1, \mu_2) = \sum_{j=N-k+1}^{N} \frac{N!}{j!(N-j)!} \left[p_f(\mu_1, \mu_2)\right]^j \left[1 - p_f(\mu_1, \mu_2)\right]^{N-j} \tag{6}$$

Using the distribution of $P_f = p_f(M_1, M_2)$ given in (5), the predictive system failure probability is obtained as

$$\tilde{p}_{sf} = \int_0^1 \sum_{j=N-k+1}^{N} \frac{N!}{j!(N-j)!} p^j (1-p)^{N-j} f_{P_f}(p)\, \mathrm{d}p \tag{7}$$

As mentioned earlier, when n approaches infinity, $\tilde{p}_f = \Phi(-\mu_B)$. The corresponding value of the system failure probability is obtained by substituting this value in place of $p_f(\mu_1, \mu_2)$ in (6). We denote this value as $\tilde{p}_{sf,n\to\infty}$. Figure 3 shows plots of the ratio $\tilde{p}_{sf} / \tilde{p}_{f,n\to\infty}$ for series systems $(k = N)$ with $\mu_B = 3$ and for parallel systems $(k = 1)$ with $\mu_B = 2$, both systems having 1 to 5 components. It can be seen that increasing statistical uncertainty (decreasing sample size n) increases the predictive failure probability for both systems. For series systems, the effect is relatively modest. For parallel systems, the increase in the failure probability can be by orders of magnitude. Figure 4 shows the same ratio for systems with $N = 5$ components with varying k for $n = 10$ and 30 and $\mu_B = 3$. As can be seen, the influence of the statistical uncertainty increases with increasing redundancy of the system (decreasing k). This is because the positive correlation arising from statistical uncertainty effectively reduces the redundancy of the system.

4.2. Time-variant reliability

Consider a structure subjected to repeated applications of earthquake loads. Each earthquake produces a stochastic ground motion at the site of the structure. Following common practice, we use a single measure (e.g., the peak ground acceleration) to characterize the intensity of the motion. Let S denote this intensity measure and assume, for a given earthquake of random characteristics, it has a lognormal distribution with parameters λ_S and ζ_S. Also let ν denote the mean rate of earthquakes per year. If the occurrence of earthquakes are assumed to follow a Poisson process, then $h(s) = 1 - \exp\{-\nu\Phi[-(\ln s - \lambda_S)/\zeta_S]\}$ is the annual seismic hazard function for the site. Let R denote the capacity of the structure (for any performance criterion of interest) expressed in terms of the ground motion intensity measure, and assume it has the lognormal distribution with parameters λ_R and ζ_R. In general, there are errors in modeling the structure. Furthermore, a single intensity measure cannot fully characterize the effect of a stochastic motion on the structure. To account for these errors, we adopt the limit-state function

$$g(r, s, \varepsilon) = \ln r + \varepsilon_1 - \ln s + \varepsilon_2 \tag{8}$$

where r and s are realizations of R and S, respectively, and ε_1 and ε_2 are model error terms, the former reflecting errors in modeling the structure (model form error) and the latter reflecting the effect of the stochastic ground motion (missing variables). We assume the model errors are normally distributed with zero means and standard deviation σ_1 and σ_2, respectively. Furthermore, we assume R, S, ε_1 and ε_2 are statistically independent. It follows from these assumptions that g has the normal distribution with mean $\lambda_R - \lambda_S$ and variance $\zeta_R^2 + \sigma_1^2 + \zeta_S^2 + \sigma_2^2$. The set of distribution parameters $\boldsymbol{\Theta} = (\nu, \lambda_R, \zeta_R, \lambda_S, \zeta_S, \sigma_1, \sigma_2)$ are, of course, subject to statistical uncertainty and so we let $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ denote their posterior joint distribution.

For a given earthquake, no distinction between the uncertainty types needs to be made, and the predictive failure probability is given by

$$\tilde{p}_f = \int_{\boldsymbol{\theta}} \Phi\left[-\frac{\lambda_R - \lambda_S}{\sqrt{\zeta_R^2 + \sigma_1^2 + \zeta_S^2 + \sigma_2^2}} \right] f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} \tag{9}$$

The product $\mu_\nu \tilde{p}_f$, where μ_ν denotes the mean of ν, represents the mean rate of earthquake-induced failures per year. This observation has lead many investigators (see Der Kiureghian 2005 for several references) to assume that the failure events are Poisson and therefore the following expression for the failure probability in $(0, t)$ has been used:

$$\widetilde{P}_{f,Psn} = 1 - \exp\left(-\mu_\nu \, \widetilde{p}_f \, t\right) \tag{10}$$

However, due to the presence of non-ergodic uncertainties, the failure events in time are not statistically independent and, therefore, cannot constitute Poisson events. More specifically, the aleatory or epistemic uncertainties in R and ε_1 as well as the epistemic uncertainties in $\boldsymbol{\Theta}$ are shared by all earthquake events, while the aleatory uncertainties in S and ε_2 are renewed at each earthquake. To overcome this dependence, we note that the conditional failure events given $R = r$, $\varepsilon_1 = e$ and $\boldsymbol{\Theta} = \boldsymbol{\theta}$ are Poisson with the mean rate $\nu\Phi[-(\ln r + e - \lambda_S)/\sqrt{\zeta_S^2 + \sigma_2^2}]$. Hence, the predictive failure probability over the time interval $(0,t)$ is given by

$$\widetilde{P}_f = 1 - \int_{r,e,\boldsymbol{\theta}} \exp\left[-\nu\Phi\left(-\frac{\ln r + e - \lambda_S}{\sqrt{\zeta_S^2 + \sigma_2^2}}\right) t\right] f_R(r) f_{\varepsilon_1}(e) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) \, \mathrm{d}r \, \mathrm{d}e \, \mathrm{d}\boldsymbol{\theta} \tag{11}$$

To investigate the difference between the approximation in (10) and the exact result in (11), we assume the following distributions and parameter values: ν is lognormal with varying mean μ_ν and 50% coefficient of variation, $\zeta_R = 0.294$, λ_R is normal with a zero mean and variance ζ_R^2/n, $\zeta_S = 0.472$, λ_S is normal with mean -1.0 and variance ζ_S^2/n, $\sigma_1 = 0.3$ and $\sigma_2 = 0.5$, where n is a measure of the quality of statistical information (analogous to sample size). Figure 5 shows plots of $\widetilde{P}_{f,\mathrm{Psn}}$ and $\widetilde{P}_f$ as a function of $\mu_\nu t$ for $n = 10$ and $n \to \infty$. It can be seen that neglecting the dependence between successive events due to the non-ergodic uncertainties results in an overestimation of the failure probability for large $\mu_\nu t$ values. Also, increasing statistical uncertainty (small n) increases the failure probability estimate. The differences between the two results, however, are relatively insignificant. This is because the considered time-variant problem is analogous to a series system problem with a random number of components.

5. Conclusions

The characterization of uncertainties into aleatory and epistemic in risk and reliability analysis and in codified or performance-based design is discussed. The distinction between aleatory and epistemic uncertainties is determined by our modeling choices. The distinction is useful for identifying sources of uncertainty that can be reduced, and in developing sound risk and reliability models. It is shown that for proper formulation of reliability, careful attention should be paid to the categorization (epistemic, aleatory, ergodic or non-ergodic) of uncertainties. Failure to do so may result in underestimation or overestimation of failure probability, which can be quite significant (orders of magnitude) in certain cases.

References

Cornell, C. A., and Krawinkler, H. (2000). Progress and challenges in seismic performance assessment. *PEER Center News*, Spring 2000. http://peer.berkeley.edu/news/2000spring/index.html

Box, G.E.P., and G.C. Tiao (1992). *Bayesian inference in statistical analysis*. Addison-Wesley, Reading, Mass.

Der Kiureghian, A. (1989). Measures of structural safety under imperfect states of knowledge. *J. Structural Engineering*, ASCE, **115:**1119-1140.

Der Kiureghian, A. (2005). Non-ergodicity and PEER's framework formula. *Earthquake Engineering and Structural Dynamics*, **34**:1643-1652.

Ditlevsen, O. (1994). Distribution arbitrariness in structural reliability. *Structural Safety & Reliability*, G. Schuëller, M. Shinozuka and J. Yao, Eds., Balkema, Rotterdam, The Netherlands, Proceedings of ICOSSAR'93, 1241-1247.

Ditlevsen, O. and H. O. Madsen (1989). Proposal for a code for the direct use of reliability methods in structural design. JCSS Working Document, 1989.
Ditlevsen, O., and H.O. Madsen (1996). *Structural reliability methods*. J. Wiley & Sons, New York, NY.
Faber, M. H. (2005). On the treatment of uncertainties and probabilities in engineering decision analysis. *J. Offshore Mechanics and Arctic Engineering*, 127:243-248.
Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: six levels of treatment. *Reliability Engineering and System Safety*, **54**:95-111.
Vrouwenvelder, A.C.W.M. (2003). Uncertainty analysis for flood defense systems in the Netherland. *Proceedings, ESREL*, 2003.

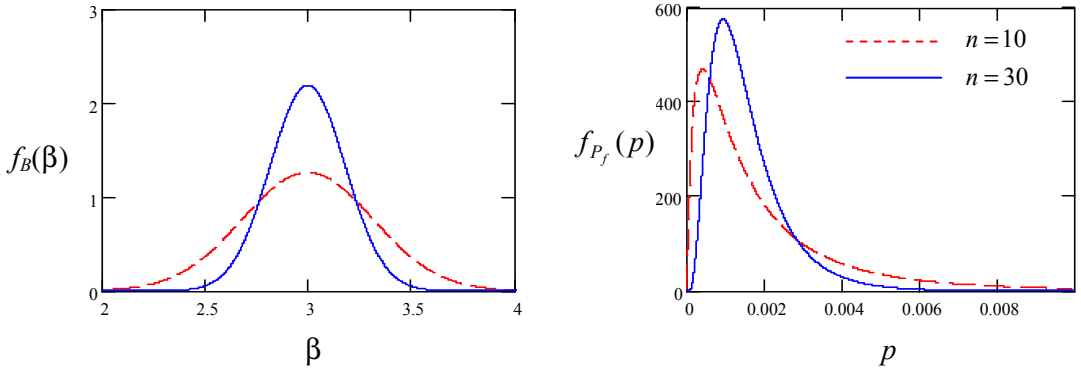

Figure 1: Distributions of reliability index (left) and failure probability (right).

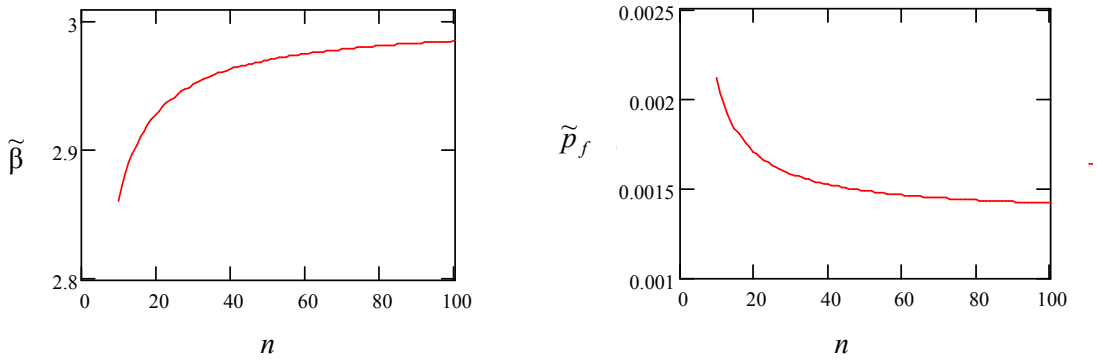

Figure 2: Predictive reliability index (left) and failure probability (right) as functions of sample size.

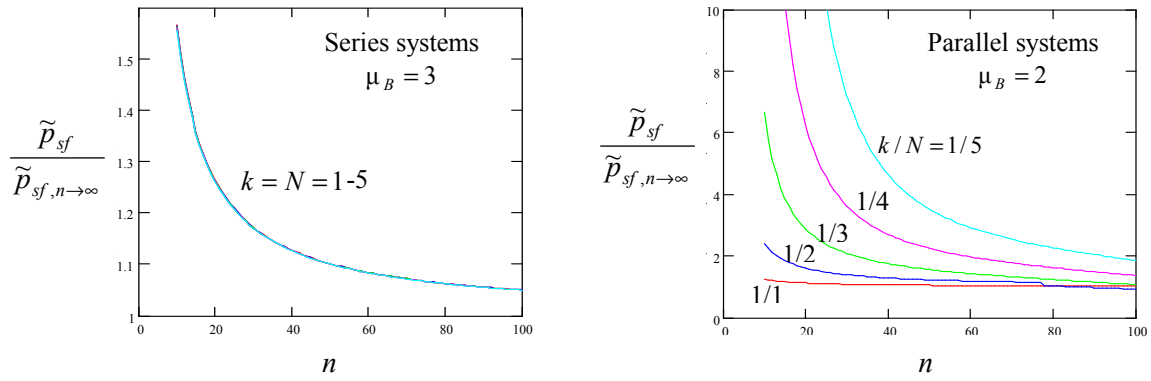

Figure 3: Influence of statistical uncertainty on series (left) and parallel (right) systems.

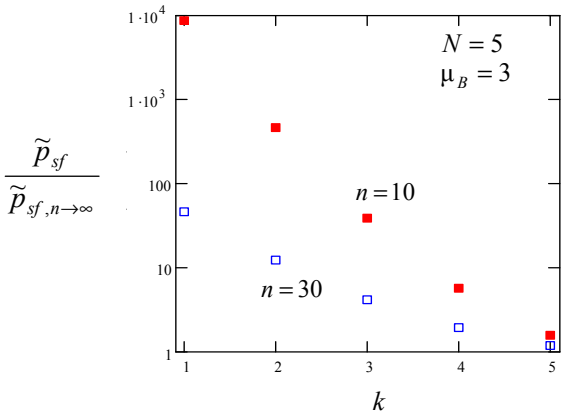

Figure 4: Influence of statistical uncertainty on k-out-of-N system.

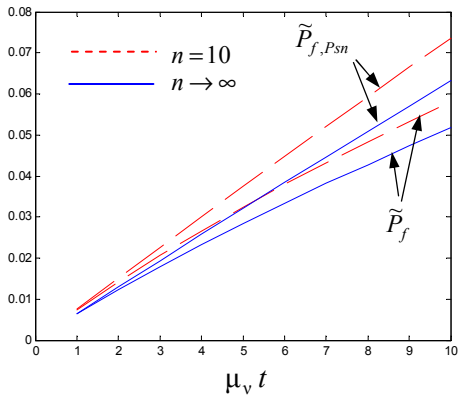

Figure 5: Influence of non-ergodic uncertainties on time-variant reliability