

## Chapter 2

# Modeling and Propagation of Uncertainties

Paolo Emilio Pinto

**Abstract** The basic problem dealt within this chapter consists in finding the distribution of a scalar quantity  $Y$  which is a function of a vector  $\mathbf{X}$  of probabilistically qualified random quantities. The dependence of  $Y$  on  $\mathbf{X}$  is not explicit, requiring algorithms for its determination. An approach that facilitates the solution of the problem was first adopted in problems of nuclear safety, and consists in expressing the probability of  $Y$  exceeding a given value (if  $Y$  has a monetary connotation), or a given structural Limit-State of a whole system, as a multiple integral of a Markov chain of conditional probability functions. This approach is described in this chapter with reference to buildings, but the approach can be equally applied to various industrial systems. A crucial step in the chain is the passage from the intensity of the action, in the present case the intensity of the ground motion, to the vector of the structural response variables on which the state of the system, and ultimately  $Y$ , depend. This passage involves consideration of the variability of the action, in the form of different samples of ground motion together with all the uncertainties inherent in the numerical determination of the response, which include those related to the selection of the structural model and the uncertainties of its parameters. A number of approximate techniques for dealing with this problem are presented, starting from the simple but inadequate FOSM method, to the approach based on the use of a Response Surface in the space of the structural variables, and concluding with the potentially more accurate Latin Hypercube Sampling technique, underlining however its practical limits due to the computational effort required for large number of uncertain quantities.

---

P.E. Pinto (✉)

Department of Structural Engineering and Geotechnics, University of Rome 'La Sapienza',  
Via Gramsci 53, 00197 Rome, Italy  
e-mail: [pinto@uniroma1.it](mailto:pinto@uniroma1.it)

## 2.1 On the Nature of the Uncertainties

It is usual to separate the uncertainties affecting all human activities into two categories, namely, *aleatoric* or *epistemic*. A common definition of the first category includes the uncertainties related to events whose outcome cannot be known before an experiment is made, as for example the outcome of a fair throwing of a fair dice, even if the associated phenomenon is completely understood. The definition above is, however, not exhaustive: there exist in nature phenomena that are, or are considered to be, inherently random, and others for which the idea of experiments is not applicable.

The second category includes uncertainties deriving from our incomplete knowledge of the corresponding phenomenon. Uncertainties of this nature are in principle reducible concurrently with improved understanding of the relevant process.

The distinction between the two categories appears at first quite clear and well founded. To blur somewhat the picture, one might argue that in a deterministic view of the world where every fact occurs as a consequence of precise laws, all uncertainties would reflect ignorance, hence, they would all belong to the epistemic class. At the opposite end, in a view of the world represented as a huge stochastic process, all the uncertainties would be represented in the model, hence they would be classified as aleatoric, while the parameters of the process, in case they would be only partially known, would be epistemic in nature.

This brief introduction is meant to establish a terminological basis. In probabilistic applications to actual problems, like for example the risk analyses that have been carried out within SYNER-G, the two types of uncertainty coexist, and there is no conceptual difference between the two, except for the following one. The aleatoric uncertainties are describable, in the majority of cases, by means of continuous probability distributions; on the other hand, the epistemic ones are often of the discrete type, and the associated probabilities are to be assigned subjectively on the basis of experience. Exceptions occur in both cases with discrete aleatoric variables, e.g. the variable describing which seismic source is generating an event, among the finite set of sources affecting the area, or continuous epistemic variables, such as those describing the uncertainty in model parameters.

## 2.2 The “Propagation” of the Uncertainties in Complex Systems

### 2.2.1 General

The SYNER-G project includes a number of individual systems (buildings, power stations, bridges, electric networks, etc.) connected to each other so as to form what could be called a “live” super-system (since the functioning of the whole depends

on the proper interaction amongst all the functioning components), and aims at evaluating the risk of different levels of reduced performance of the whole super-system in consideration of a spatially distributed seismic hazard.

All components of the super-system are in themselves complex systems which are made up in general of a large number of elements, each one performing a specific function and exposed to being damaged by a seismic event.

In general, the functional logic of complex systems cannot be described by means of an explicit functional relationship linking the response of the system to the state of its elements. Further, the states of the elements are generally of the continuous type, evolving from complete functionality to complete loss of it, and also the particular state of any component has an influence on that of the others. Under these conditions, the classical system theory based on elements having only binary behaviour is of no use for a probabilistic analysis of the types of complex systems.

A further issue is relevant regarding the use of the two different types of uncertainties, aleatoric and epistemic. It is customary to proceed by first assigning values to, or making choices for, the epistemic variables, evaluate then the risk conditional on all the possible combinations of values/choices of these variables, and then to make use of the total probability theorem to de-condition the previously determined conditional probabilities. This way of proceeding lends itself to a graphical illustration called “logic tree”, of which a simple example is given in Sect. 2.2 (Fig. 2.1). Different techniques can be employed to evaluate the conditional probabilities in each branch of the logic tree.

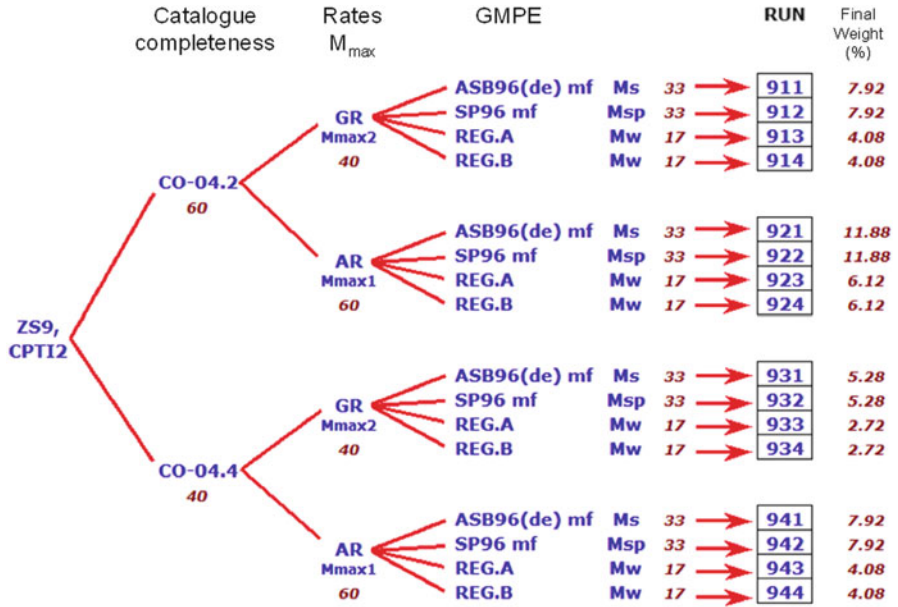
A full blown Monte Carlo (MC) approach would be perfectly suited for the purpose, but even enhanced with the modern variance reduction techniques it would result in an unrealistic computational effort.

The few applications that can be found in the structural engineering literature use a mixture of MC simulations for some steps of the procedure and First Order-Second Moment (FOSM) methods for others. While the MC procedure is asymptotically exact, application of FOSM leads to acceptable approximations in cases of linear or approximately linear relationships among the variables involved (Lee and Mosalam 2005; Baker and Cornell 2008). If the relation is not originally linear, a first-order Taylor expansion can be used:

$$Y = g(\mathbf{X}) \cong g(\mathbf{X}_0) + \nabla g|_{\mathbf{X}_0}(\mathbf{X} - \mathbf{X}_0) \quad (2.1)$$

where  $Y$  is a scalar quantity dependent on the vector  $\mathbf{X}$  collecting the  $N$  variables  $X_i$ ,  $g(\bullet)$  is a generic function of  $\mathbf{X}$  and  $\nabla g$  is its gradient with respect to  $\mathbf{X}$ , that can be obtained either analytically or numerically (through perturbation) depending on whether  $g(\mathbf{X})$  is known in explicit or algorithmic form. If the expansion is centred at the mean value of  $\mathbf{X}$ , owing to linearity one gets for the mean and the variance of  $Y$ :

$$\mu_Y = g(\mu_{\mathbf{X}}) \quad (2.2)$$



**Fig. 2.1** Logic tree employed in the latest Italian seismic hazard map derivation to handle epistemic uncertainty. The weights are the product of the weights of the branches leading to the RUN (e.g. RUN 911 has weight  $0.6 \cdot 0.4 \cdot 0.33 = 0.0792$ ). RUN numbers are composed of Seismic Zonation number (9), catalogue- $M_{max}$  combination number (1–4) and employed GMPE (1–4)

$$\sigma_Y^2 = \nabla g|_{\mu_X} \mathbf{C}_{XX} \nabla g^T|_{\mu_X} \quad (2.3)$$

where  $\mathbf{C}_{XX}$  is the known covariance matrix of  $\mathbf{X}$ . The expression above shows an example of how the variability of  $N$  variables can be condensed into that of a single one.

An approach that has been shown to offer practical advantages in probabilistic risk assessments dates back to early studies of seismic safety of nuclear plants in the 1950s of the last century. It consists in expressing the risk (defined, for example, as the annual rate of exceedance of a chosen measure of loss, functional or economic) in the form of a multiple integral of a Markovian chain of conditional probability functions. This approach has been adopted in recent years by the Pacific Earthquake Engineering Research (PEER) Center (Yang et al. 2009; fib 2012) with reference to buildings, and its presentation given in the following uses quantities that are meaningful for this class of structures, but the approach is general, and it can be adapted to other types of systems, such as industrial plants, hospitals, electric transformation/distribution systems, etc. Hospital systems, in particular, have been treated in SYNER-G with an approach having common points with the PEER approach.

## 2.2.2 The PEER Approach

### 2.2.2.1 Nomenclature

List of the terms used in this chapter:

- $IM$  = Intensity Measure: a parameter expressing the intensity of seismic action, such as the peak ground acceleration, and the spectral acceleration at the fundamental period of the system.
- $\lambda(im)$  = the hazard function, or mean annual frequency of exceedance of  $IM = im$
- **EDP** = Engineering Demand Parameter: a vector of response variables used for assessing the degree of damage to all structural and non-structural elements. The full characterization of the vector in probabilistic terms, i.e. in terms of a joint distribution function, is obtained from a number of dynamic response analyses.
- **DM** = Damage Measure: a vector having as many components as the number of structural and non structural elements. The passage from **EDP** = **edp** to **DM** = **dm** is obtained using “fragility curves” relative to different states of damage.
- $DV$  = decision variable: a scalar quantity, monetary, as in the case of PEER, or a suitable performance index in the case of a generic system. The passage from **DM** = **dm** to  $DV = dv$  is obtained through the so-called “loss curves” giving the probability of loss  $dv$  as function of the damage level **dm**. The number of the loss curves is given by the product of the number of damage levels times the different groups of damageable elements.
- $\lambda(dv)$  = the mean annual exceedance rate of  $DV = dv$ , unconditional from all previous variables.

### 2.2.2.2 Formulation

Using the symbols defined in the previous section the annual rate of exceedance of  $DV = dv$  can be written as:

$$\lambda(dv) = \int_{\mathbf{DM}} \int_{\mathbf{EDP}} \int_{IM} G(dv|\mathbf{dm})f(\mathbf{dm}|\mathbf{edp})d\mathbf{dm}f(\mathbf{edp}|im)d\mathbf{edp}|d\lambda(im)| \quad (2.4)$$

where  $G(\bullet)$  indicates the complementary distribution function of the argument and  $|d\lambda(im)|$  is the absolute value of the derivative of the hazard function.

The above equation is valid in the assumption of a Markovian dependence between the successive functions in the chain: for example, it implies that  $f(\mathbf{dm}|\mathbf{edp})$  is the same as  $f(\mathbf{dm}|\mathbf{edp}, im)$ , that is, **dm** is only dependent on **edp** and not on  $im$ .

Equation (2.4) identifies four stages of the performance assessment: hazard analysis, response analysis, damage analysis and consequence analysis. This arrangement is convenient since it subdivides the total task into subtasks each one requiring a specific field of competence, starting from seismology to cost analysis.

### 2.2.2.3 Development

#### 2.2.2.3.1 Hazard Analysis

This stage requires availability of the *hazard curve* appropriate for the site together with a *set of accelerograms* needed for the structural response analysis.

The hazard curve, which is definitely the most important element in risk analysis, is the result of a process disseminated with epistemic uncertainties, such as the subdivision in homogeneous seismic sources, their spatial dimension, their activity in terms of magnitudes and frequencies, the functional form of the attenuation laws, etc.

These uncertainties are accounted for by having recourse to the “logic tree”, whose branches have as many nodes as the uncertainties considered, and at each node two or more choices are made for a particular uncertainty, with a subjective probability attached to each of them. Each branch is then characterized by a probability value, which is the product of the probabilities assigned at all choices defining the branch.

These probabilities are associated with the hazard evaluated using the choices along the branches, so that a discrete probability distribution of the hazard curves is obtained.

In order to reduce the burden of calculating the system risk using the different hazard curves and then convolving the risks so obtained with the probability of the hazard, in order to de-condition the system risk from the hazard uncertainty, current practice adopts the simplification of using the mean hazard obtained from the logic tree.

Figure 2.1 shows an example of the logic tree used to handle the epistemic uncertainty in the evaluation of the latest version of the seismic hazard of Italy (Stucchi et al. 2004) using a spatial grid of sides of  $5 \times 5$  km. Discrete variables were used to describe uncertainty in the following issues: seismic catalogue completeness (two levels), upper bound magnitude in the G&R recurrence law (two levels), and attenuation laws (four different laws), resulting in 16 different combinations. Subjectively assigned probabilities are given in percent in the figure.

A number of accelerograms, typically between 20 and 40, is generally adequate, depending on the structure and in particular on the number of significant modes of vibration.

The selection can be made using the hazard de-aggregation procedure, by which one gets the triplet of  $M$ ,  $r$  and  $\epsilon$  that gives the major contribution to a selected value of the local exceedance rate of the intensity measure utilized. Records are then selected based on this triplet, more frequently on the values of  $M$  and  $r$  alone. Whether the records should also be compatible with a uniform hazard spectrum (characterized by a specified mean return period), or the choice should be made according to other, finer criteria, is still debated.

### 2.2.2.3.2 From *im* to **edp**

The response of the structure to each individual accelerogram is obtained, as already indicated, through non-linear time history analysis (NLTHA). The resulting **edp** vectors (one vector for each accelerogram and for each value of its intensity) collect all the response variables that will be needed in the next stage for passing from response to damage. These normally include the (maximum) values of the nodal displacements, the relative displacements and distortions of the elements, the accelerations at various points, and internal forces for brittle elements.

Once the NLTHAs are completed for the whole set of accelerograms at a given intensity level, statistical analysis is carried out on the response variables, and estimates of the parameters of their joint density function conditional on the intensity; i.e. mean values and covariance matrix, are obtained.

As the dynamic analysis phase is computationally the more demanding task of the whole risk determination procedure, often an artifice that involves the following steps is used to increase the number of correlated response values in the **edp** vector beyond those directly obtained through the NLTHAs.

In the first place the common assumption is accepted that the **edp** variables are jointly lognormally distributed, so that their logarithms are jointly normal.

Denoting by **X** the initial **edp** vectors, their logarithms are taken, denoted by **Y**, and the mean  $\mu_Y$  and covariance matrix  $C_{YY} = \mathbf{D}_{YY}\mathbf{R}_{YY}\mathbf{D}_{YY}$  are constructed, where  $\mathbf{D}_{YY}$  indicates the diagonal matrix of the standard deviations of **Y** and  $\mathbf{R}_{YY}$  its correlation matrix.

Use is then made of the fact that a vector **Z**, having mean value  $\mu_Y$ , standard deviation matrix  $\mathbf{D}_{YY}$ , and correlation matrix  $\mathbf{R}_{YY}$  can be obtained from the expression:

$$\mathbf{Z} = \mu_Y + \mathbf{D}_{YY}\mathbf{L}_{YY}\mathbf{U} \quad (2.5)$$

with **U** being a standard normal vector and  $\mathbf{L}_{YY}$  the lower triangular decomposition of the correlation matrix ( $\mathbf{R}_{YY} = \mathbf{L}_{YY}\mathbf{L}_{YY}^T$ ).

Based on the relationship above one can generate samples of **U**, obtain the vector **Z**, and finally compute a vector **X** by taking the exponential of **Z**.

This simulation procedure of creating additional **edp** vectors having the proper statistical structure is very efficient, so that large numbers of these vectors can be generated with a minimal computational effort.

It is important to recall that in order to follow these steps one needs random vectors of **edp** for several levels of the intensity.

In the discussion thus far, related to the passage from *im* to **edp**, attention has focused on the variability of the response due to the variability of the ground motion, given a measure of its intensity. Though it is generally recognized that this variability is quite possibly the main contributor to the total uncertainty, this does not allow ignoring further sources of uncertainty of different nature whose relevance varies from case to case, and sometimes can be of significant importance.

As discussed in the following, the further sources are two, and are different in nature.

In order to determine **edp** from *im* a model of the structure has to be set up. Here the term model is used in a wide sense; it includes the structural modelling, e.g., whether and how certain elements have been included and their behaviour described, as for example the beam-column joints, or whether account has been taken of the shear–flexure interaction, whether beam-column elements are of the distributed or concentrated plasticity type, etc.

Various combinations of the above mentioned features are clearly possible. None of the existing models can be said to be perfect, and it is not guaranteed that the most “sophisticated” ones give the most accurate response.

One clearly recognizes the above situation as one where the uncertainty is of the epistemic type, whose solution could consist of the following steps: selecting two (or more) models, all of them considered in principle as valid candidates, running the whole procedure of risk analysis, and assigning the weight (degree of confidence) attributed to each of the models to the risks values computed. In this particular instance of application of the logic tree all models would have the same weight, it would make in fact little sense to use a model considered as less reliable than the others. The final value of the risk would clearly be the weighted average of the different risk values.

The second source of uncertainty can be of both aleatoric and epistemic type, and is related to the parameters of the models, for example the mechanical properties of the reinforced concrete components (if there are doubts on other aspects, such as the exact layout of the reinforcement, they would belong to the previous category). The mechanical properties of the materials are random variables whose distributions must be assumed as known. Correlations exist between some of them (as for ex strength and ultimate deformability of concrete), as well as spatial correlations of the same variable at different locations.

It has been observed that for structures designed according to modern codes, and for not extreme ranges of the response, the variability of these quantities has normally a reduced effect on the variability of the dynamic response, and its relevance becomes quite modest in consideration of the other major uncertainties affecting the whole procedure (see following steps). In view of their reduced importance their effect is treated in an approximate way, consisting in sampling all the variables from their distributions as many times as is the number of the dynamic runs, and in associating to each run (i.e. to each accelerogram) a different model having the properties of the corresponding sample.

In the last few years, however, much research has been devoted to obtain models suitable for describing the degrading behaviour of existing, non code-conforming structural elements, down to their total loss of vertical load carrying capacity. These models are of completely empirical derivation, with a rather restricted experimental base. As a consequence, their parameters are characterized by large dispersions, inducing a variability of the response of the same order of magnitude of that due to ground motion variability. The developing state of such models leads to consider the variability of the parameters as belonging to the epistemic class.



The simplified PEER approach is not suited for these cases, and a review of the techniques that have been recently proposed to deal with them is presented at the end of this section.

Continuing with the PEER approach, the stage of progress reached thus far is the one characterized by the largest number of variables: a number of 100/200 edp vectors (each one having the dimension necessary to describe the state of the structure), necessary for the need of the simulation procedure to follow, times the number of levels of the intensity to be considered (usually around ten). The subsequent steps will need to gradually condense this large vector down to the final one, where a single variable (or possibly a small number of alternative variables in the form of performance indexes) will express the risk of the system.

### 2.2.2.3.3 From edp to dm

Real systems are made up of a very large number of individual components that can be broadly classified into two categories: the *structural* components, into which one may include, for convenience, both *load-resisting elements* (beams, columns, floors, . . .), and the so-called *architectural components* like partition walls, ceilings, glazing, etc., and *functional* components, i.e. those allowing each particular system to operate.

The focus of risk analysis varies with the type of the system. For ordinary buildings the risk is normally defined in terms of the total economic loss, while for a hospital the definition is in terms of its continued operability, and for an electric generation/transformation station the risk is in terms of the number of lines that remain active and of the quantity and quality of the power that can be delivered, etc.

Whatever the adopted definition of risk, each particular **edp** vector determines a particular state of the system, involving both structural and functional components according to the functional logic of the system.

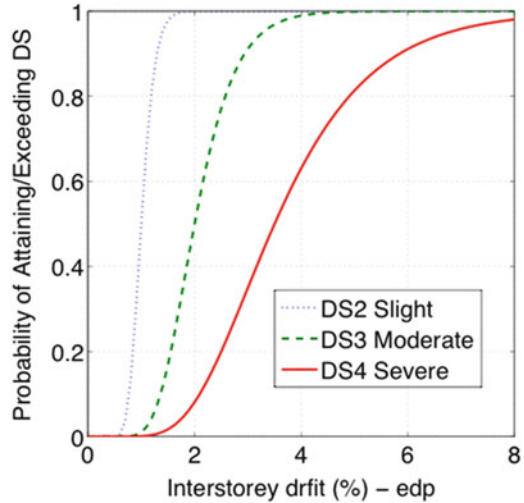
The PEER procedure, which is exemplified in this chapter, has been developed with reference to buildings, so the risk refers to the total monetary cost.

In order to reduce the variables of the problem, elements are divided into groups with the criterion that the members belonging to one group are (approximately) attributed to the same fragility function. Fragilities are available for discrete states of damage, frequently three for structural elements and two for architectural elements.

The states of damage for structural elements are described in qualitative terms, as for example light, moderate and severe, having in mind that to each term and for each type of element there should be associated (in probabilistic terms) a cost of repair.

The passage from edp to dm is illustrated with reference to Fig. 2.2, showing three fragility curves relative to interstorey drift ratio. Damage State 1 (DS1) corresponds to negligible damage and is not represented in the figure, DS2 corresponds to slight damage requiring superficial repair, DS3 represents severe damage

**Fig. 2.2** Fragility curves for increasing damage states



requiring substantial repair, and DS4 is for damage requiring a cost of repair equal to that of replacement.

It is recalled that each fragility curve expresses (in this case) the probability that the damage of the element (or group of elements) is equal or larger than that relative to the specified curve for a given value of *edp* (the component of the **edp** vector that influences the state of the element/group). The vertical distance between two curves (e.g. DS2 and DS3) provides the probability of being in the damage state corresponding to the higher curve (DS2).

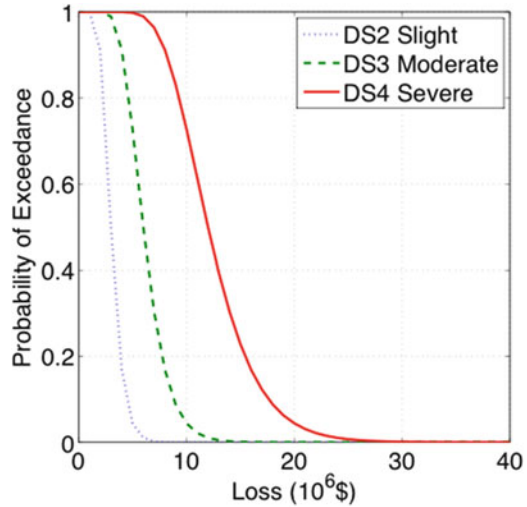
The passage from the probability of **edp** to that of **dm** is thus the following:

- an **edp** vector is chosen from the collection relative to a specific intensity level.
- each group of elements is considered in turn and its damage level is determined using the appropriate fragility curve. This is done by sampling a uniform random number (in the interval 0–1) and checking where it falls in the intervals defined by the fragility curves at the current *edp* value.
- the operation is repeated for all **edp** vectors, and the distribution of the **dm** vector (for a particular intensity level) is thus obtained.

#### 2.2.2.3.4 From **dm** to *dv*

The passage from **dm** to *dv* is accomplished by introducing the so-called “loss functions”, which give for each group of elements the probability of the cost associated with each level of damage. An example of such types of functions is given in Fig. 2.3.

**Fig. 2.3** Probability of exceedance of economic loss thresholds as a function of damage state



The passage from the probability of  $\mathbf{dm}$  to that of the total  $dV$  is as follows:

- for each group, the probability that the value  $DV = dV$  is exceeded is obtained as the sum of the products of the probability of  $dV$  conditional to the generic value of damage  $dm$  (loss curve), times the probability of that damage.
- the operation above is repeated for all the desired values of  $dV$  and for all levels of damage  $dm$ .
- since damage levels are mutually exclusive, the total probability (complementary distribution) of  $DV$  exceeding any given value  $dV$  is obtained as the sum of the probabilities over all groups and all damage levels.

#### 2.2.2.3.5 Consideration of Collapse

In the developments thus far, structural and non-structural elements have been considered to be susceptible to damages of various severity.

However, the integrity of the whole system, i.e. its ability to continue sustaining its own weight, has not been considered. Yet the total physical collapse of a building is an event that has a weight in the post-earthquake decisions, and for this reason the probability of its occurrence and the associated DV are often included in the overall procedure. This is simply done as follows.

Starting from a certain intensity level, in some of the simulations the integrity of the structure is so seriously endangered as to suggest the use of the term “collapse”.

These cases should be considered separately from the “non collapse” states and, for any given intensity, a probability of collapse should be approximately evaluated as the ratio between the number of simulations where collapse has occurred to the total number of simulations. The complement of this probability, namely, the “non

collapse probability”, is the one to be associated to the computation of damages as previously described.

It is important to note, however, that the question of how to define this extreme state of a structure is not yet completely settled, and different approaches are in use.

The first and still widely adopted proposal consists in looking at collapse in a global way, and defining this state as the one when the calculated dynamic displacements tend to increase indefinitely for an infinitesimal increase of the intensity (so-called dynamic instability). This approach does not look directly at the state of the individual structural elements, some of which may well have exceeded their individual deformation capacity and be in the post-peak negative stiffness branch of response. Rather, the approach relies on the ability of the model to correctly reflect the effect of the local damage at the global level.

As in the passage from *im* to **edp**, models have been developed in the last few years to describe the degrading behaviour of reinforced concrete beam-column elements subjected to cyclic normal force, bending and shear down to the exhaustion of their vertical load bearing capacity.

Availability of these models, however, even when they will have become more accurate and robust, will not per se provide a unique solution to the problem of defining the collapse of an entire structure, since collapse can occur involving the failure of a variable number of elements depending on structural topology and robustness.

The frequent choice of considering a structure as a series system whose failure is made to depend on the first complete failure of a single element can be in many cases a rather conservative approach. Attempts to overcome this conservatism include for instance a floor-level comparison between gravity load demand and capacity (to account for vertical load redistribution), as done for instance in Baradaran Shoraka et al. (2013).

#### 2.2.2.3.6 Evaluation of Risk

The complementary distribution of  $DV$  calculated as above is a function of the intensity of the seismic action. As such, it may be already sufficiently informative for those in charge of taking decisions if, for instance, the interest is in knowing the loss associated with a given “design” action (e.g. earthquake with 1,000 years return period).

If, on the contrary, the interest is in knowing the total risk contributed by all possible intensities, one should simply integrate the product of the probability of  $DV$  conditional on  $IM = im$  times the mean annual rate of the  $IM$ . In performing this integration, collapse and non-collapse cases must be kept separate and weighted with the corresponding probabilities:

$$\lambda(dv) = \int_{IM} [G(dv|im, NC)P_{NC}(im) + G(dv|C)P_C(im)]d\lambda(im) \quad (2.6)$$

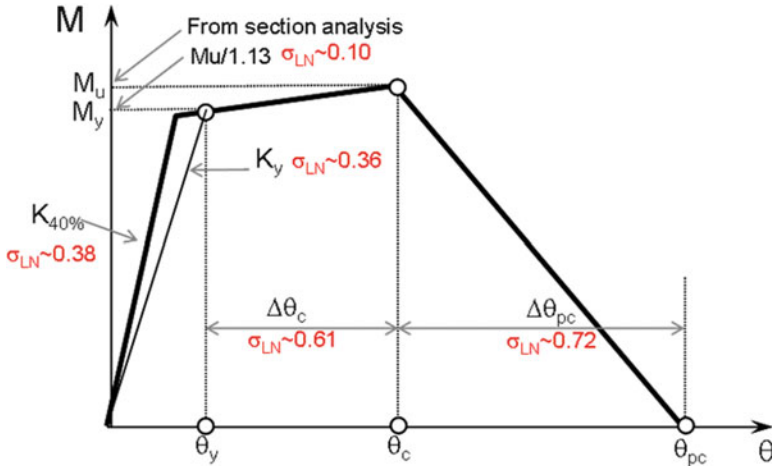


Fig. 2.4 Section degrading moment-rotation curve with indication of parameter dispersion (Models by Haselton and Deierlein 2007)

### 2.2.3 Treatment of “Epistemic” Model Uncertainties

As mentioned before, models have been recently developed which are capable of describing progressively degrading states of reinforced concrete elements down to their complete loss of bearing capacity. It has also been observed that the parameters of these models are affected by large variability whose final effect is that of a significant increase of the variance of the Limit State fragilities (Ibarra et al. 2005; Ibarra and Krawinkler 2005; Goulet et al. 2007).

Figure 2.4 shows a monotonic envelope of a moment-chord rotation relationship typical of this class of models used in a demonstration study by Liel et al. (2009). The logarithmic standard deviations attributed to the corner points of the diagram are 0.6 and 0.7 for the rotations up to and post peak, and 0.5 to a parameter which regulates the cyclic degradation. These values are much higher than, for example, the values adopted for the beam or column strength (0.19) or stiffness ( $0.36 \div 0.38$ ).

A standard procedure for introducing model uncertainties into a seismic fragility function does not exist. All methods available in the literature are approximate, and every improvement in approximation is paid with a rapidly increasing amount of additional computations.

The simplest and perhaps most widely used method is FOSM which was briefly described in Sect. 2.2.1. It is applicable to any type of LS fragility, from light damage to collapse, including all intermediate LSs. As explained earlier, the method consists of deriving a linear relationship  $g(\mathbf{X})$  between the LS of interest and the variables (in this case the epistemic quantities) whose influence on the LS is sought (an expression involving response and capacity, as a function of the

variables  $\mathbf{X}$ ). Linearity implies that the mean value of the distribution of the LS is unchanged with respect to that obtained by using the mean values of the epistemic variables,  $\mu_g = g(\mu_{\mathbf{X}})$ , while the variance of the fragility due to the variability of these variables is (approximately) evaluated.

FOSM has two limitations: the linear dependence can become grossly inadequate as the considered LS moves towards the collapse state and, second, the mean value of the fragility remains unchanged, a fact that is shown to be untrue by more elaborate approaches.

A second approach, described in Liel et al. (2009), makes use of a *response surface* in the space of model uncertainties, combined with a MC procedure. Response surfaces can be set up to give the log-mean and the log-standard deviation of the LS capacity (IM that induces the LS), in the form of complete second order polynomial functions of  $\mathbf{X}$ . This allows one to capture both direct effects of the variables, up to their squared values, and the interactions between any two of them on the quantity of interest.

With four random variables, each polynomial function contains 15 coefficients. This gives already an idea of the number of “experiments” to be carried out, since this number must be significantly larger than the number of the coefficients to be estimated in order to reduce the variance of the “error term”.

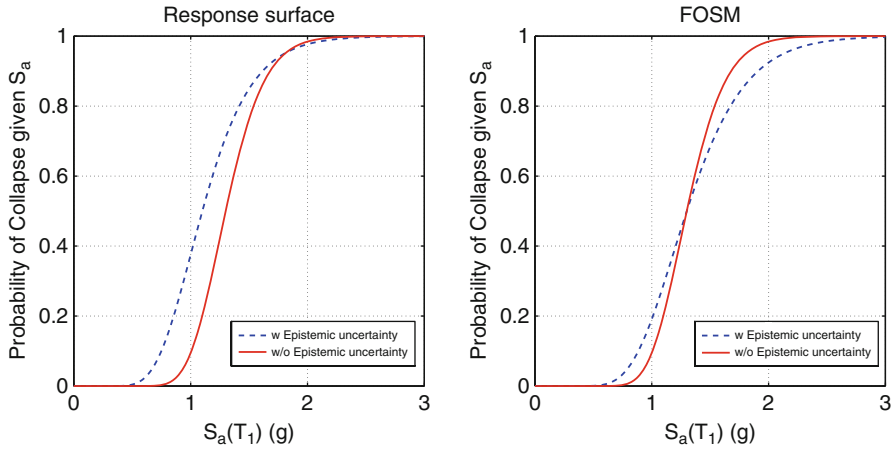
For a complete quadratic form of the function the so-called Central Composite Design of (numerical) experiments is appropriate (Pinto et al. 2004). This plan requires a complete two-level factorial design involving experiments for all the  $2^4 = 16$  combinations of the (4) variables complemented by the addition of two further “star” points located along each of the variable axes and a “centre” point, for an additional  $2 \times 4 + 1 = 9$  points, which makes a total of 25 “experiments”.

Each “experiment” consists in performing an incremental dynamic analysis on one particular model out of the 25 ones using an adequate set of accelerograms (in the order of 20–30), and in calculating log-mean and log-standard deviation values of the selected LS capacity. When all the experiments are concluded, standard Least Square method is used to obtain the numerical values of the coefficients of the two response surfaces  $\mu_{\ln IM}(\mathbf{X})$  and  $\beta(\mathbf{X}) = \sigma_{\ln IM}(\mathbf{X})$ .

Once the response surfaces are created, a MC procedure is used to sample a large number of sets of the modelling variables from their distributions, which yield the fragility parameters through the fitted surfaces. The unconditional fragility function, accounting both for record-to-record variability and epistemic model uncertainty, is obtained by averaging over all the samples. Results generally show a decrease in the median and an increase in the dispersion, more pronounced for LS closer to collapse due to the fact that the uncertainty affecting ultimate deformation is larger than that associated with elastic or low-ductility response.

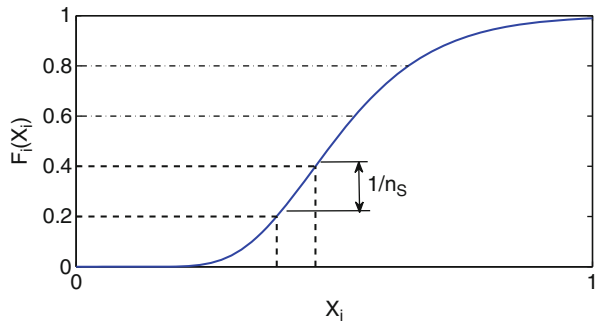
Figure 2.5 shows the collapse fragility obtained by both the response surface (left) and the FOSM approach (right), including or neglecting the modelling epistemic uncertainty. The curves show that the median collapse capacity is influenced by the epistemic uncertainty.

A further alternative to account for the effect of both the ground motion variability and all other types of uncertainty on the structural fragility has been



**Fig. 2.5** Fragility curves with and without epistemic uncertainty, as obtained by the response surface (*left*) and the FOSM (*right*) procedures (Adapted from Liel et al. 2009)

**Fig. 2.6** Equal probability values obtained by stratified sampling (Adapted from Vořechovský and Novák 2009)



adopted in a number of recent publications, e.g. Fragiadakis and Vamvatsikos (2010) and Dolšek (2012). The procedure is conceptually the same as described in Sect. 2.2.2.3 for the passage from *im* to **edp**, but made more efficient and potentially accurate through the use of the so-called Latin Hypercube Sampling (LHS) technique (Helton and Davis 2003) for the sampling of the random variables describing the modelling uncertainties. This technique is much more efficient than ordinary (random) MC. While random sampling produces standard errors that decrease with the number  $N$  of simulations according to  $\sqrt{N}$ , the error with LHS goes down at a much faster rate, approaching  $\sqrt[3]{N}$  for linear functions.

According to Helton and Davis (2003) LHS operates in the following way to generate a sample of size  $n_s$  from a random vector  $\mathbf{X} = (X_1, X_2, \dots, X_{n_X})$ , consistently with the marginal distributions of the  $X_i$ 's. The range of each variable  $X_i$  is exhaustively divided into  $n_s$  intervals of equal probability content as shown in Fig. 2.6, and one value is sampled at random from each interval. The  $n_s$  values thus

obtained for  $X_1$  are paired at random without replacement with the  $n_S$  values obtained for  $X_2$ . These  $n_S$  pairs are combined in a random manner with the  $n_S$  values of  $X_3$ , and the process continues until a set of  $n_S$   $n_X$ -tuples is formed. The resulting matrix of  $n_S$  rows by  $n_X$  columns constitutes the LHS sample, and the values contained in each row represent a possible model for the structure.

The obtained LHS sample, however, is characterized by a correlation matrix that is not the one specified by the analyst according to the specific features of the structure, and hence needs to be modified. This can be achieved according to a procedure proposed in Helton and Davis (2003) and Vořechovský and Novák (2009), which is based on rearranging the values on the individual columns of the original matrix.

The LSH technique is in principle applicable for any size  $n_X$  of the vector of the modelling variables. However, the number of variables that can practically be treated is limited by computational considerations. Although there are no fixed rules, the sample size  $n_S$ , i.e., the number of different models, must be a multiple (of the order of, say, two) of the size  $n_X$  of the vector  $\mathbf{X}$ . Since each model must be subjected to dynamic analyses under the full set of the selected accelerograms, and for different intensity levels, the total number of models should not be exceedingly large.

To reduce the number of variables under consideration (the components of vector  $\mathbf{X}$ ), correlations can be assumed between the variables within each structural member and among the members in the structure. For example, within a RC member, variables related to strength, such as stiffness and yield moment could be assumed as perfectly correlated, and similarly for the deformation parameters (see Fig. 2.4). At the system level, all elements having the same properties could also be assumed as perfectly correlated. In the examples found in the literature, e.g. Fragiadakis and Vamvatsikos (2010) and Dolšek (2012), the dimension of vector  $\mathbf{X}$  does not exceed 12. It is noted that in the mentioned literature the components of the vector  $\mathbf{X}$  are all indicated as epistemic, though they include both the material properties such as, for ex., concrete and steel strength, which are usually categorized as aleatoric, as well as the other parameters of the model such as those shown in Fig. 2.4. All the components of  $\mathbf{X}$  are however represented by continuous variables, i.e. they do not include epistemic variables such as the consideration of alternative models, alternative methods of analysis, etc., for which the logic tree approach remains necessary and appropriate.

## References

- Baker JW, Cornell CA (2008) Uncertainty propagation in probabilistic seismic loss estimation. *Struct Saf* 30(3):236–252
- Baradaran Shoraka M, Yang TY, Elwood KJ (2013) Seismic loss estimation of non-ductile reinforced concrete buildings. *Earthq Eng Struct Dyn* 42(2):297–310
- Dolšek M (2012) Simplified method for seismic risk assessment of buildings with consideration of aleatory and epistemic uncertainty. *Struct Infrastruct Eng* 8(10):939–953



- Fragiadakis M, Vamvatsikos D (2010) Fast performance uncertainty estimation via pushover and approximate IDA. *Earthq Eng Struct Dyn* 39:683–703
- Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter KA, Stewart JP (2007) Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building—from seismic hazard to collapse safety and economic losses. *Earthq Eng Struct Dyn* 36(13):1973–1997
- Haselton CB, Deierlein GG (2007) Assessing seismic collapse safety of modern reinforced concrete moment frame buildings. John Blume Center technical report 156
- Helton JC, Davis FJ (2003) Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf* 81:23–69
- Ibarra LF, Krawinkler H (2005) Global collapse of frame structures under seismic excitations. John Blume Center technical report 152
- Ibarra LF, Medina RA, Krawinkler H (2005) Hysteretic models that incorporate strength and stiffness deterioration. *Earthq Eng Struct Dyn* 34(12):1489–1511
- International Federation of Structural Concrete (fib) (2012) Bulletin 68. Probabilistic performance-based seismic design
- Lee TH, Mosalam KM (2005) Seismic demand sensitivity of reinforced concrete shear-wall building using FOSM method. *Earthq Eng Struct Dyn* 34(14):1719–1736
- Liel AB, Haselton CB, Deierlein GG, Baker JW (2009) Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. *Struct Saf* 31:197–211
- Pinto PE, Giannini R, Franchin P (2004) Seismic reliability analysis of structures. IUSS Press, Pavia. ISBN 88-7358-017-3
- Stucchi M, Akinci A, Faccioli E, Gasperini P, Malagnini L, Meletti C, Montaldo V, Valensise G (2004) Derivation of the seismic hazard map for Italy – final report to the Department of Civil Protection (in Italian)
- Vořechovský M, Novák D (2009) Correlation control in small-sample Monte Carlo type simulations I: a simulated annealing approach. *Probab Eng Mech* 24:452–462
- Yang TY, Moehle J, Stojadinovic B, Der Kiureghian A (2009) Seismic performance evaluation of facilities: methodology and implementation. *ASCE J Struct Eng* 135(10):1146–1154