About Knowledge and Responsibility in Probabilistic Seismic Risk Management

by Pasquale Erto, Massimiliano Giorgio, and Iunio Iervolino

ABSTRACT

In seismic risk management, apart from assessment, it is also necessary to develop decision-making strategies that are based on evaluating the probability of the occurrence of earthquakes, their consequences, and the effect of possible mitigation actions. There is a close connection between the knowledge level of the analyst and the resulting probabilistic evaluations, which should be considered correct if they incorporate all available information. Knowledge, probability, and loss are the key concepts in a state-of-the-art approach to risk estimation and management. This article, after reviewing the concepts of risk-based decision making and the possible interpretations of probability, discusses the importance of entrusting seismic risk assessment to experts on the phenomenon who are also knowledgeable regarding probability theory, and how these experts should combine their expert knowledge with probability theory, which is related to the responsibility of the seismic risk analyst.

NATURAL PHENOMENA, PROBABILITY, AND LOSS

The term "random experiment" refers to any operation, activity, or phenomenon whose outcome cannot be predicted with certainty. A classic example is the toss of a coin, whose outcome (i.e., head or tail) is uncertain par excellence. The same concept of a random experiment (in a broad sense) is also applicable when referring to the observation of a natural phenomenon; for example, an earthquake and/or its effect, whose outcome will remain uncertain until it actually occurs.

In almost all cases in which the outcome is uncertain, any loss (i.e., potentially adverse consequence) that the phenomenon may induce will also be characterized by uncertainty, either because losses will normally depend on the actual outcome or because it is often impossible to assign to each possible outcome a definitive value for the associated loss. Lack of complete knowledge on the mechanisms governing the experiment and the exact conditions under which the experiment takes place, as well as lack of knowledge of the exact value of loss associated with each outcome of the said experiment, prohibit *a priori* knowledge of any outcome or potential loss level. As a matter of fact, the outcome and/or the loss given by the outcome may vary if the experiment is repeated under nominally identical conditions (i.e., undistinguishable in the eyes of the experimenter).

Successfully dealing with this type of uncertainty is important. Probabilistic calculations provide the analytical tools necessary to exploit all available knowledge (incomplete as it may be) in order to express in a consistent manner the probability that a certain outcome, rather than a different one, is realized (or, more generally, that a given event occurs). Better yet, from a risk estimation point of view, the mathematical tools of probabilistic calculus can be used to directly express the probability that the given experiment/phenomenon produces a particular loss value (e.g., Marzocchi *et al.*, 2015).

Risk Estimation: Expected Loss

Under conditions of uncertainty, once all loss values L_j (let us say n in number) possibly resulting from a random experiment have been identified, and the probability $P[L_j]$ associated with observing each of them has been defined, it is possible to obtain a quantitative estimate of risk according to equation (1), which gives what is called "expected loss" $E[L]$, that is, the (weighted) average of all possible loss values producible by the considered experiment. As shown in equation (1), the weight attributed to each loss value is its probability:

$$\text{Risk} = E[L] = \sum_{j=1}^{n} L_j \cdot P[L_j]. \tag{1}$$

Evaluating risk in terms of expected loss is extremely useful for two reasons: first, it provides whomever undertakes the evaluation with a consistent way of treating the available information; second, it permits comparisons with risks of a completely different nature by expressing them in terms of a single measurement unit (e.g., euros) used to quantify the loss. The latter is essential when deciding risk management strategies (e.g., Benjamin and Cornell, 1970).

To better clarify the usefulness of these concepts in the context of seismic risk, let us assume, for instance, that a seismic swarm of moderate intensity has just occurred in a zone near a site of interest, and let us further assume that whomever has been charged with evaluating that risk believes that the said swarm might herald the occurrence of a strong earthquake in the following 6 months, an event indicated by Q_S, with a probability $P[Q_S] = 0.01$. Having assigned this probability, and by assuming that the loss due to the earthquake corresponds to a billion euros, whereas the loss in the case the earthquake does not occur $\overline{Q_S}$ is 0€, then using equation (1) one obtains the risk value of the following equation:

$$E[L] = E[L|\overline{Q_S}] \cdot (1 - P[Q_S]) + E[L|Q_S] \cdot P[Q_S] \\ = (0€) \cdot 0.99 + (1 \cdot 10^9 €) \cdot 0.01 = 1 \cdot 10^7 €. \quad (2)$$

Note that the implied risk, while expressed in euros, does not coincide with any actual value that loss may assume. In fact, the risk as defined in equation (1) should not be interpreted as the effective loss value one may expect to observe; it should rather be understood as a value representing the average of losses that are observable in a large number of nominally identical situations (to follow).

Expected Loss and Optimal Decisions

In order to complete the process of risk management, it is necessary, apart from evaluation, to develop decision-making strategies to control it, which means strategies for distinguishing which action (A^*) among those available ($A_i, i = 0, 1, \ldots, l$) is most rational to undertake in order to minimize the loss that could incur from the phenomenon under consideration. To this end, the definition of risk in equation (1) comes in handy. It seems natural to define the optimal decision as the one which results in selecting, among all possible actions, action A^*, corresponding to the smallest expected loss $E[L(A^*)]$

$$E[L(A^*)] \le E[L(A_i)] \quad \forall i = 0, 1, \ldots, l. \quad (3)$$

Note that among available actions one should always include action A_0, which represents taking no action. To clarify these concepts, let us assume that—referring to the previous example—those responsible for managing the seismic risk must decide which action would be the best to take among the following alternatives: (A_0) undertake no action; (A_1) do not evacuate the site, but prepare an enhanced plan of civil protection for the period during which it is feared that an earthquake could follow the swarm, in order to better cope with an eventual emergency; (A_2) evacuate the site for 6 months. It is assumed that relocating the residents of the entire site for 6 months would cost 500 million euros and that such an action would nullify losses from an eventual earthquake (in general, evacuation does not completely avoid losses, as damage to assets may still occur; however, this kind of loss is independent of the decisions considered in the example, then it would not change the optimal one, as such it is neglected for simplicity). On the other hand, it is assumed that action A_1 would cost 12 million euros, but would only succeed in halving losses due to an earthquake. Expected losses are calculated as follows:

$$E[L(A_i)] = CA_i + E[L(A_i)|\overline{Q_S}] \cdot (1 - P[Q_S]) \\ + E[L(A_i)|Q_S] \cdot P[Q_S] \;\; \forall i = 0, 1, \ldots, l, \quad (4)$$

in which $E[L(A_i)|Q_S]$ and $E[L(A_i)|\overline{Q_S}]$ represent the losses in the cases the action is taken whether the earthquake occurs or not, respectively, and CA_i is the cost of the action A_i.

The results are given in equation (5). In this example, if no mitigating action is taken, $E[L(A_0)]$ would amount to 10 million euros; that is, equation (2). The expected loss in case of an evacuation $E[L(A_2)]$ is equal to about half a billion euros, whereas adopting enhanced planning reduces it to around 17 million euros, that is $E[L(A_1)]$. It is evident that in this case (purely academic and only hypothetical), the optimal decision would be to take no mitigating action:

$$\begin{cases} E[L(A_0)] = (0€) \cdot 0.99 + (1 \cdot 10^9 €) \cdot 0.01 = 1 \cdot 10^7 € \\ E[L(A_1)] = 12 \cdot 10^6 € + (0€) \cdot 0.99 + (5 \cdot 10^8 €) \cdot 0.01 = 1.7 \cdot 10^7 € \\ E[L(A_2)] = 5 \cdot 10^8 € + (0€) \cdot 0.99 + (0€) \cdot 0.01 = 50 \cdot 10^7 € \end{cases} . \quad (5)$$

The expected loss is not to be interpreted as the most likely result of the next random experiment (the next seismic swarm, in this example), yet it coincides with the limit of the average loss computed over a virtually infinite number of replications of the considered experiment (i.e., independent repetitions). In this sense, the decision based on the minimal expected loss is the optimal one in the long run; that is, it may be the rational one for authorities that may have to deal with several of these crises over a period of time.

Naturally, in order to use this approach it is essential to define the weights used in calculating expected loss, that is, the probabilities $P[L_j]$. In the following section, the discussion focuses on the approaches and mathematical tools which analysts may employ to obtain a consistent measure of the said probabilities by using all available data and knowledge. This, along with quantification of potential losses, constitutes the responsibility of those tasked with undertaking such probabilistic analyses.

DEFINITIONS OF PROBABILITY AND TRADITIONAL CALCULATION CRITERIA

The calculation of probabilities is a mathematical science developed from its governing postulates (i.e., rules that ensure consistency as formulated by Kolmogorov, 1933). Postulates are dictated without discussing their practical meaning; thus, all other rules of calculation (i.e., theorems) are derived from them. In other words, such a modern approach considers probability as a primitive concept, much like point and line are considered primitive concepts in geometry. In fact, it does not supply any interpretation of the term probability, it only addresses how to measure it. Consequently, any criterion that the analyst would decide to adopt in order to measure probability has to be considered mathematically legitimate, provided it satisfies the postulates.

It is important to note, however, that a mathematically legitimate evaluation might still not have practical validity, if the analyst did not take on the responsibility of using all and solely the available information. As a matter of fact, from a practical point of view, one should recognize that whoever formulates a probability is consequently burdened by responsibility, much in the same way as in any other professional or scientific field. For example, a building analyzed by correctly following the rules of structural engineering may end up designed incorrectly with regard to its intended purpose, if considerations regarding a potential source of stress—such as seismic demand—are omitted. In an analogous manner, a probability that is formulated following the rules of probability theory may end up incorrect if an essential element of knowledge related to the experiment is ignored. The responsibility associated with all evaluations of probability should lead to the purging of all arbitrariness from the analysis (including conservative assumptions) and should use all knowledge available at the time the calculation takes place.

The choice of which criterion should be used in formulating a probability is motivated by opportunity, professional evaluations, subjective knowledge, and so forth. Meanwhile, the mathematical rules that need to be observed say nothing about the choice of the formulation criterion, but merely guide towards a consistent development of the deductive calculations necessary for evaluating the risk. In turn, this situation determines a twofold responsibility/opportunity for the expert: (1) to choose the criterion based on the knowledge of the phenomenon, which allows probabilistic evaluations with practical value; and (2) to convincingly motivate the choice of the criterion, so as to gather consensus among others also interested in the same problem.

Even in situations when everyone operates under the same calculation choices (which thus become the norm), one is still obligated to explicitly cite the adopted hypotheses. This is necessary so as not to—unconsciously—preclude further possible refinements to the calculation, which may prove useful or even indispensable if hypotheses become untenable in light of any new information.

Although it has been discussed that probability is considered a primitive concept, it is nevertheless appropriate to briefly recall the three historical definitions that constitute useful tools for measuring probability in some typical experimental situations.

Classical Definition

This definition, formalized by Laplace (1812), refers to an experiment which may have N mutually exclusive, equally possible outcomes. In this context, assuming that N_A of these outcomes correspond to the occurrence of event A, the probability $P[A]$ can be defined as follows:

$$P[A] = N_A/N. \tag{6}$$

Although based on a calculation criterion valid for the specific situation, equation (6) is an unacceptable definition, as it would be circular. In fact, the concept of probability is implicitly employed prior to its definition. Furthermore, the classical criterion is not applicable when the possible outcomes of the experiment are virtually infinite and/or not equally probable. Then, for example, it cannot be used to formulate the probability of observing a specific value of macroseismic intensity for the next earthquake occurring at a certain location, as alternative intensity values are not equally likely.

Frequentist Definition

A second definition of probability, which overcomes some of the limitations of the classical one, is the frequentist (von Mises, 1928). Assuming that an experiment can be replicated an unlimited number of times, with all repetitions conducted under nominally (i.e., macroscopically) identical conditions, probability is defined by means of the limit given below:

$$P[B] = \lim_{n\to\infty} \frac{n_B}{n}, \tag{7}$$

in which n_B is the number of repetitions in which event B was realized and n is the number of total repetitions. This may be, for example, the probability that, given an earthquake will occur on a certain seismic source, the said earthquake has a magnitude which falls in an interval of interest.

Even equation (7), although again suggesting a valid calculation criterion for the case considered, is far from being an acceptable definition. In fact, one can never be in a position to guarantee that the considered limit exits, and even if it does, no one can guarantee that it will coincide with the intended measurement, which is the probability of B, whose very existence is an act of faith. Furthermore, even this definition does not possess general validity, because it presumes that the experiment is repeatable, which is not always the case. The repeatability issue arises, for example, in the case of very rare natural events such as earthquakes of magnitude never observed yet (as the 2011 Tohoku earthquake case; e.g., Kagan and Jackson, 2013). Finally, even when it can be assumed that repetition of the experiment is possible, one still has to clarify what exactly is meant by nominally identical experimental conditions, which is a vague concept, considering that different repetitions normally produce different outcomes because the experimenter cannot control the values or the effects of all the factors influencing the results. In other words, to apply the frequentist definition to the magnitude bin example above, it was necessary to implicitly assume that the earthquakes' generating process is such that each earthquake results from the repetition of the same random experiment; for example, the random rupturing phenomenon does not show any trend or any form of memory of the seismic history.

Subjectivist Definition

Because the two previous definitions of probability are permeated by a strong component of subjectivity (e.g., "equally possible" appearing in the classical definition and the existence of the limit in the frequentist case), a third line of thought, named "subjectivism" or "neo-Bayesianism", has developed a definition completely centered on the concept of probability as a degree of belief, without any particular prior assumption (Ramsey, 1926). This definition is often employed to assign a level of credibility to each of a set of models in the seismic risk assessment context with epistemic uncertainty (e.g., Der Kiureghian and Ditlevsen, 2009).

To better comprehend the subjectivist definition, imagine a hypothetical bet in which one is prepared to wager sum a on an experiment proving that event Q will occur, against sum b if Q will not occur. In this situation, one can consequently formulate, as a subjective estimate of the probability of Q, the ratio

$$P[Q] = a/(a + b). \qquad (8)$$

With this approach, one avoids the drawbacks of defining probability using one of its evaluation methods. Furthermore, it requires effort toward formulating the probabilities using all available knowledge (which is one of the responsibilities of the analyst, as discussed later on). In particular, to ensure fair evaluations, whoever is betting should be ready to exchange his role with the bookmaker; in other words, he should be in agreement with betting sum b, to receive the $a + b$ amount in case Q does not occur. The consistency and operative aspects of this definition are evident and bring to mind the contents of that wise rule-of-thumb according to which when a certain amount needs to be divided between two parties, one should be charged with dividing it and the other with choosing his share.

It is useful to delve deeper into the meaning of the last example. Assume that a sum of $a + b$ euros is the value for a certain item to be insured against earthquakes; that is, the amount of lost value in the case event Q (an earthquake) occurs. Because the occurrence of Q is uncertain during the period of insurance coverage, the premium that the owner of the item is predisposed to pay is not the full value of the loss ($a + b$ euros). He would be willing to pay a lower price instead (let us say a) which would be proportionate to the degree of belief in the occurrence of Q. Thus the ratio defined in equation (8) can be assumed to constitute a measure of the probability $P[Q]$.

PROBABILITY, STATE OF KNOWLEDGE, AND BAYES' THEOREM

As shown, any estimate of probability should be considered conditional to all knowledge that the analyst possesses at the time when the estimation takes place. Consequently, if one's state of knowledge changes then one's estimate of probability should be updated. It is obvious that the analyst is confronted with the problem of revising calculations in a consistent manner, more specifically, in a manner that produces results not in conflict with the estimate already derived based on the prior state of knowledge. Toward this end, Bayes' theorem can prove useful, because it represents the appropriate strategy for correctly performing such an update/revision in mathematical terms. In this context, using the Bayes theorem does not necessarily mean to adopt a subjectivist/Bayesian approach. Indeed, the probabilities in the Bayes' theorem can be indifferently assigned in a frequentistic and/or in a subjective manner (for example when dealing with epistemic/model uncertainty), without affecting the calculations and the way they are performed by any means.

Bayes' theorem is sometimes also referred to as the theorem of probability of the cause, as it allows computation of the probability of a certain event (A) given the occurrence of another event (B). Therefore, its result may be interpreted as the probability that A was causative for the occurrence of B, in those cases when a cause-effect relationship exists between A and B. This is the case, for example, of disaggregation of seismic hazard (e.g., Bazzurro and Cornell, 1999), in which the Bayes' theorem is used to compute how likely a specific variable involved in hazard assessment is causative for some feature of ground motion.

We introduce the expression for this theorem by means of an example, so as to demonstrate its potential and hint at its implications. Specifically, revisiting the example given in the first section, we shall see how Bayes' theorem permits us to estimate the probability that a strong earthquake (Q) follows within 6 months of a seismic swarm (S)—which in the example was given as 0.01. To this end, it is assumed that the probability of observing an earthquake in the region of interest within a 6-month span is one in a thousand, based on historical data (this is the probability estimate regardless of whether or not a seismic swarm has been previously observed). It is further assumed that the analysis of historical data also indicates that in 40% of the cases when a strong earthquake occurred in the region, it was preceded by a swarm. On the other hand, it is assumed that in 4% of the cases when no earthquake eventually occurred, a seismic swarm occurred anyway. Based on this information, the probabilities in the following equation were determined, in which the vertical bar (when present) indicates that it is assumed that the event to the right of the bar has occurred (i.e., a conditioning event):

$$\begin{cases} P[Q] = 0.001 \\ P[S|Q] = 0.4 \\ P[S|\overline{Q}] = 0.04 \end{cases} . \quad (9)$$

Having assigned these probabilities and assuming that a seismic swarm has occurred, we may calculate the probability $P[Q|S]$ that the observed swarm is a prelude to a strong earthquake, by means of Bayes' theorem, as shown in the below equation (note that this probability was indicated as $P[Q_S]$ in equations 2 and 4):

$$\begin{aligned} P[Q|S] &= \frac{P[Q \cap S]}{P[S]} \\ &= \frac{P[\text{Swarm and strong earthquake both occur}]}{P[\text{Swarm is observed}]} \\ &= \frac{P[S|Q] \cdot P[Q]}{P[S|Q] \cdot P[Q] + P[S|\overline{Q}] \cdot (1 - P[Q])} \\ &= \frac{0.4 \cdot 0.001}{0.4 \cdot 0.001 + 0.04 \cdot 0.999} = 0.01. \end{aligned} \quad (10)$$

As can be seen, solely on the basis of the added information of a swarm's occurrence, the probability of an earthquake taking place $P[Q|S]$ grew tenfold with respect to the value based exclusively on historical data (i.e., $P[Q]$).

PROBABILITY, EXPERTIZE, AND RESPONSIBILITY

Referring to the example in the Probability, State of Knowledge, and Bayes' Theorem section, one can actually imagine that if information related to other credible precursors were available, the estimate of earthquake probability could be further refined, again using Bayes' theorem. In fact, with ever-increasing knowledge of the phenomenon (i.e., knowledge on its mechanisms and determining factors), the outcome of the experiment will tend toward becoming predictable with certainty. This shows that the calculation of probabilities does not conflict with determinism.

In general, one could say that results of practical value can only be produced by those who are experts on both the phenomenon and the calculus of probability. It is actually for this reason that estimates of hazard (i.e., the probability of occurrence of hazardous events) and the associated risks are normally requested to the experts—the intended meaning of this term being those who are able to reduce the uncertainty about the phenomenon to a minimum, having studied said phenomenon more than anyone else.

Consequently, it may be concluded that two of the principal responsibilities of the seismic risk analyst, when the analysis is framed in probability theory, are

1. to ensure that the risk assessment takes advantage of the fullest level of information, about the phenomenon of interest, at the time of the analysis;
2. and to employ such information in a (mathematically) coherent manner as prescribed by the rules of probability theory.

SUMMARY

The intrinsic link between the knowledge of the seismic risk analyst and the results of the probabilistic analysis was presented. More specifically, stating that any criterion to measure probability could be considered mathematically legitimate does not imply that said evaluation may be formulated in an arbitrary manner, but rather highlights the responsibility which burdens the analyst to produce an evaluation incorporating all effectively available information in a probabilistically consistent manner. This means that estimates of probability can be considered correct if and only if they correctly incorporate all available information.

This requires careful deliberation and also implies that seismic risk estimates will vary according to the analyst's level of information and the validity of the said estimates decays once new information is acquired and/or the event of interest occurs. Consequently, the importance of entrusting seismic risk assessment to experts on the phenomenon, who also know probability theory, was highlighted.

It was also shown how it is convenient to express risk in terms of the losses that the event of interest may produce and the associated probabilities. It is only by evaluating the consequences that one may compare risk management strategies that refer to different risks of different nature. This is where the possibility to direct resources for risk mitigation in an optimal manner, according to a utility criterion, derives from. In this context, the concepts of expected loss and optimal decision making were briefly discussed. In conclusion, knowledge, probability, and loss are the key factors for a state-of-the-art approach toward seismic risk evaluation and management.

DATA AND RESOURCES

All resources of this article came from the listed references.

ACKNOWLEDGMENTS

Associate Editor Alan Kafka and an anonymous reviewer, whose comments helped to improve the quality and readability of the article, are gratefully acknowledged.

REFERENCES

Bazzurro, P., and C. A. Cornell (1999). Disaggregation of seismic hazard, *Bull. Seismol. Soc. Am.* **89,** no. 2, 501–520.

Benjamin, J. R., and C. A. Cornell (1970). *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, New York, New York.

Der Kiureghian, A., and O. Ditlevsen (2009). Aleatory or epistemic? Does it matter? *Struct. Saf.* **31,** no. 2, 105–112.

Kagan, Y. Y., and D. D. Jackson (2013). Tohoku earthquake: A surprise? *Bull. Seismol. Soc. Am.* **103,** no. 2B, 1181–1194.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, Germany (in German).

Laplace, P. S. (1812). *Théorie analytique des probabilité*, Courcier, Paris, France (in French).

Marzocchi, W., I. Iervolino, M. Giorgio, and G. Falcone (2015). When is the probability of a large earthquake too small? *Seismol. Res. Lett.* **86,** no. 5, 1674–1678.

Ramsey, F. P. (1926). Truth and probability, in Ramsey, 1931, in *The Foundations of Mathematics and other Logical Essays*, R. B. Braithwaite (Editor), Chapter VII, Kegan, Paul, Trench, Trubner & Co.,/Harcourt, Brace and Company, New York/London, 156–198.

von Mises, R. (1928). *Wahrscheinlichkeitsrechnung, Statistik und Wahrheit*, Verlag von Julius Springer, Wien (in German).

Pasquale Erto
Dipartimento di Ingegneria Industriale
Università degli Studi di Napoli Federico II
p.le Tecchio 80
80125 Naples, Italy

Massimiliano Giorgio
Dipartimento di Ingegneria Industriale e dell'Informazione
Seconda Università di Napoli
via Roma 29
81031 Aversa (CE), Italy

Iunio Iervolino
Dipartimento di Strutture per l'Ingegneria e l'Architettura
Università degli Studi di Napoli Federico II
via Claudio 21
80125 Naples, Italy
iunio.iervolino@unina.it

Published Online 10 August 2016