

Chapter 18

Accounting for Unknown Unknowns in Managing Multi-hazard Risks

Robert B. Gilbert, Mahdi Habibi, and Farrokh Nadim

Abstract A significant challenge in managing multi-hazard risks is accounting for the possibility of events beyond our range of experience. Classical statistical approaches are of limited value because there are no data to analyze. Judgment or subjective assessments are also of limited value because they are derived from within our range of experience. This chapter proposes a new framework, Decision Entropy Theory, to assess probabilities and manage risks for possibilities in the face of limited information. The theory postulates a starting point for assessing probabilities that reflect having no information in making a risk management decision. From this non-informative starting point, all available information (if any) can be incorporated through Bayes' theorem. From a practical perspective, this theory highlights the importance of considering how possibilities for natural hazards could impact the preferred alternatives for managing risks. It emphasizes the role for science and engineering to advance understanding about natural hazards and managing their risk. It ultimately underscores the importance of developing adaptable approaches to manage multi-hazard risks in the face of limited information.

18.1 Introduction

A significant challenge in managing multi-hazard risks is accounting for the possibility of events beyond our range of experience. A landslide in Oso, Washington, caused a debris runout that destroyed an entire community, including 43 lives, dammed a river creating a flood hazard upstream and then downstream when the dam was breached, and severed a transportation and utility corridor (GEER 2014). While this slope had failed multiple times in the past century, the debris never ran out far enough to impact the community until the event in 2014. The storm surge in Hurricane Katrina breached the levee system protecting neighborhoods below

R.B. Gilbert (✉) • M. Habibi
The University of Texas at Austin, Austin, TX, USA
e-mail: bob_gilbert@mail.utexas.edu

F. Nadim
Norwegian Geotechnical Institute, Oslo, Norway

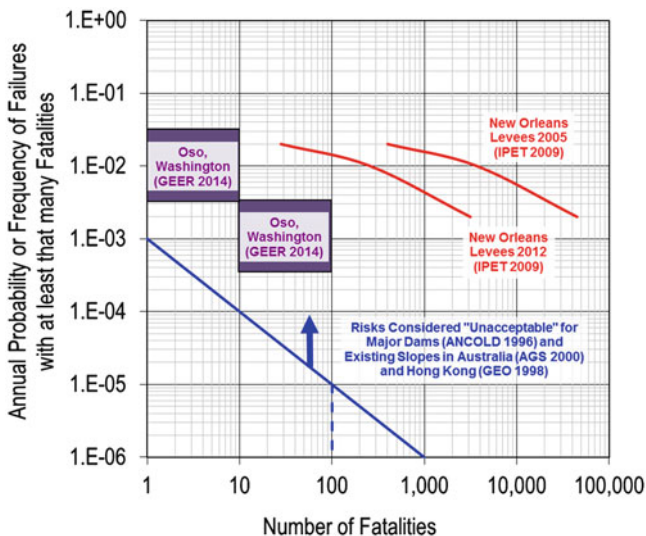


Fig. 18.1 Example comparisons of risks for natural hazards related to slopes and floods (AGS 2000; ANCOLD 1996; GEO 1998)

sea level, flooded the neighborhoods by connecting them to the ocean, and caused nearly 2000 deaths from drowning and exposure days after the storm. While storm surges had occurred in New Orleans before, the surge in Katrina was several meters higher than the maximum surge recorded previously in many locations (IPET 2009).

While rare and beyond our range of experience, these extreme hazards and multi-hazards are significant in terms of the consequences and the need to manage the associated risks (Fig. 18.1). Managing these risks requires first assessing probabilities of rare events (e.g., Liu and Nadim 2014; Nadim and Sparrevik 2013; Nadim 2011 and Lacasse and Nadim 2009). Probability assessments are typically based on historical data, observations, experience, and engineering judgment. However, historical data are sparse, and there is little or no experience with these extreme events because of their nature. In addition, models of combinations of highly complex events are inevitably simplified and uncertain. The lack of actual information with which to assess probabilities in these situations could lead to optimistic assessments that underrepresent risks or pessimistic assessments that overrepresent risks. It is therefore difficult to defend these assessments and rely on them to effectively manage the risks.

Taleb (2007) refers to possibilities beyond our experience as “black swan” events: “Before the discovery of Australia, people in the Old World were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence . . . [The sighting of the first black swan] illustrates a severe limitation to our learning from observations or experience and the fragility of our knowledge. One single observation can invalidate a general statement derived from millennia of confirmatory sightings of millions of white swans.” The United States Secretary of Defense, Donald Rumsfeld, infamously referred to possibilities beyond

our experience as “unknown unknowns” (DOD 2002): “There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don’t know. But there are also unknown unknowns. There are things we don’t know we don’t know.” The challenge is to logically and defensibly account for “black swans” or “unknown unknowns” in managing risk.

This paper proposes a framework, Decision Entropy Theory, to assess probabilities and manage risks for possibilities beyond our experience. First, the challenge of assessing probabilities with limited data is described. Next, the mathematical basis for the proposed theory is presented and illustrated. Finally, practical insights for risk management are drawn from the theory.

18.2 The Challenge with Prior Probabilities

Probability theory is used to represent uncertainty. This theory is based on starting with a comprehensive set of all possible events, known as the sample space and denoted S . Probabilities for events are assessed based on available information via Bayes’ theorem:

$$P(\text{Event } i | \text{Information} \cap S) = \frac{P(\text{Information} | \text{Event } i \cap S) P(\text{Event } i | S)}{\sum_{\text{all } j} P(\text{Information} | \text{Event } j \cap S) P(\text{Event } i | S)}$$

where $P(\text{Event } i | \text{Information} \cap S)$ is the probability for Event i given available information, which is the probability of interest and referred to as the posterior or updated probability; $P(\text{Information} | \text{Event } i \cap S)$ is the probability of obtaining the available information given that Event i occurs, which is referred to as the likelihood function; and $P(\text{Event } i | S)$ is the probability for Event i in the initial sample space, which is referred to as the prior or initial probability.

Consider the annual chance of a hazard, F . The available information is that this hazard has occurred x times in n years of experience. If we assume that occurrences follow a Bernoulli sequence,¹ then the likelihood function is given by the binomial distribution:

$$P(x \text{ occurrences in } n \text{ years} | F = f_i) = \left[\frac{n!}{x! (n-x)!} \right] f_i^x (1-f_i)^{n-x}$$

If there are no occurrences in the experience, then the likelihood function does not distinguish between small chances of occurrence that are about an order of magnitude less than the inverse of the length of experience (Fig. 18.2).

¹A Bernoulli sequence assumes that occurrences are independent from year to year and that the chance of occurrence each year is a constant.

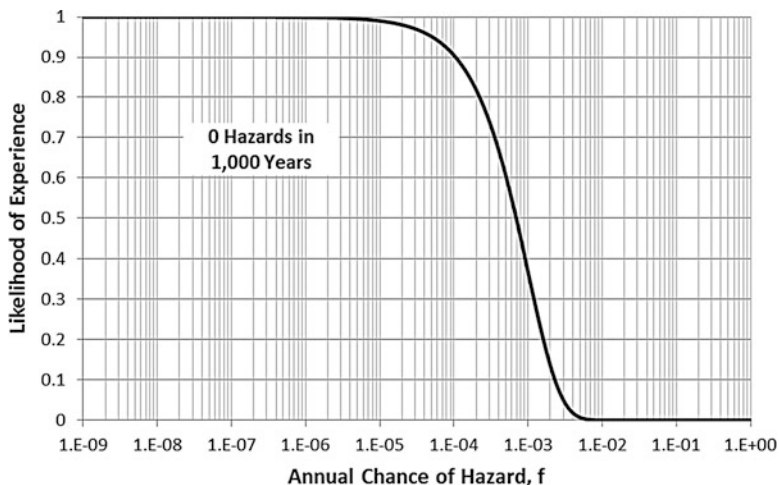


Fig. 18.2 Example likelihood function for no occurrences of a hazard in a period of experience versus the annual chance of occurrence

The challenge is the prior probability in Bayes' theorem, $P(F = f_i | S)$, because it is not conditioned or based on information. This prior probability represents “complete ignorance” in the words of Luce and Raiffa (1957) and is commonly referred to as a “non-informative prior” probability.² The non-informative prior sample space is important because it sets the stage for assessing or conditioning probabilities on any available information (whether subjective or objective). Consider a uniform prior distribution for the annual chance of occurrence between 0 and 1 (Fig. 18.3). The updated probability distribution for the annual chance of occurrence (Fig. 18.4) is the prior distribution filtered through the likelihood function; wherever the likelihood function is flat, such as for small chance of occurrence (Fig. 18.2), the updated distribution is entirely a reflection of the prior distribution.

The significance of the prior probability distribution is demonstrated using three different distributions: a uniform distribution on the annual chance of occurrence,³ a uniform distribution on the logarithm of the annual chance of occurrence,³ and a uniform distribution on the return period or the inverse of the annual chance of occurrence⁴ (Fig. 18.5). The updated probability distributions for the annual chance of occurrence are significantly different between the three possible prior

²The prior probability is non-informative because it does not depend on information: $P(S | \text{Information}) = P(S)$.

³Note that the logarithm of the annual chance of occurrence approaches negative infinity and the return period approaches positive infinity. A lower bound of 1×10^{-9} was used for the annual chance of occurrence (or an upper bound of 1×10^9 on the return period). Since the likelihood function is flat approaching this lower bound, the choice of a lower bound will affect the results and underscores the significance of the shape of the prior probability distribution.

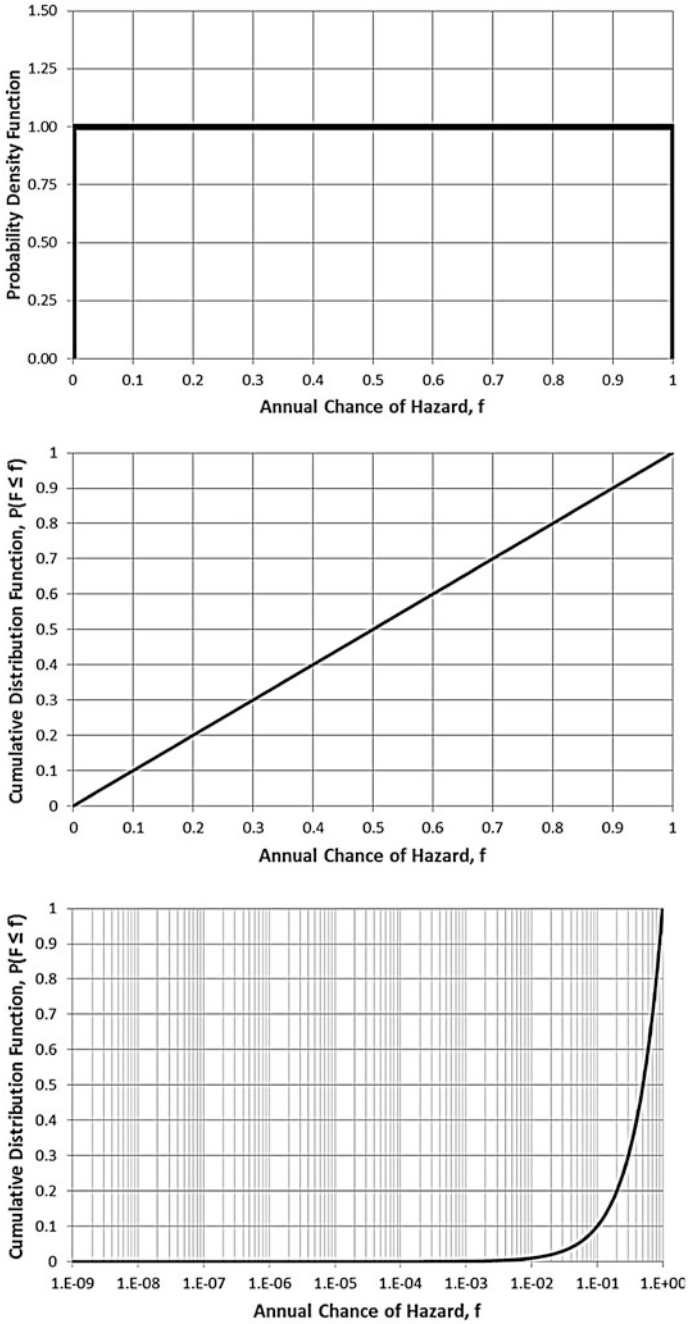


Fig. 18.3 A uniform prior probability distribution for the annual chance of occurrence. (Three different representations of the probability distribution are shown. Since we will typically be dealing with order-of-magnitude ranges for the chance of occurrence (Fig. 18.1), the representation with the cumulative distribution function on a logarithmic scale for the chance of occurrence will be used throughout this chapter.)

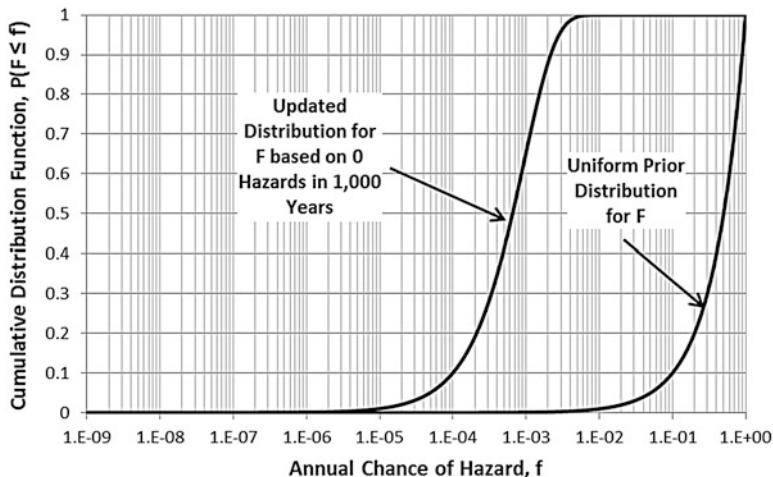


Fig. 18.4 Prior and updated probability distributions for the annual chance of occurrence

probability distributions (Fig. 18.6 and Table 18.1). The significance of the prior probability distribution for the annual chance of occurrence is further highlighted by considering the expected value for the updated probability distribution, which is the probability of an occurrence in 1 year (Fig. 18.7):

$$P(\text{Occurrence in One Year}) = E(F) = \sum_{\text{all } f_i} f_i p_F(f_i | x \text{ occurrences in } n \text{ years})$$

where $p_F(f_i | x \text{ occurrences in } n \text{ years})$ is the updated probability distribution for the annual chance of occurrence based on the experience.

The challenge of establishing a non-informative prior probability distribution has been the subject of theorists for centuries. Bernoulli (1738) postulated the principle of insufficient reason, which is paraphrased as: *If one is completely ignorant as to which state will occur, then the states should be treated as if they are equally probable.* Jaynes (1957, 1968) expressed the principle of insufficient reason as maximizing the entropy of information, H, where

$$H = \sum_{j=1}^{n \text{ States}} - \ln [P(\text{State } j | S)] P(\text{State } j | S)$$

The greatest “lack of information” is the set of prior probabilities that lead to the maximum entropy, which is obtained when these probabilities are equal for all states. In practice, this approach has been applied to a variety of problems (e.g., Tribus 1969; Box and Tiao 1973).

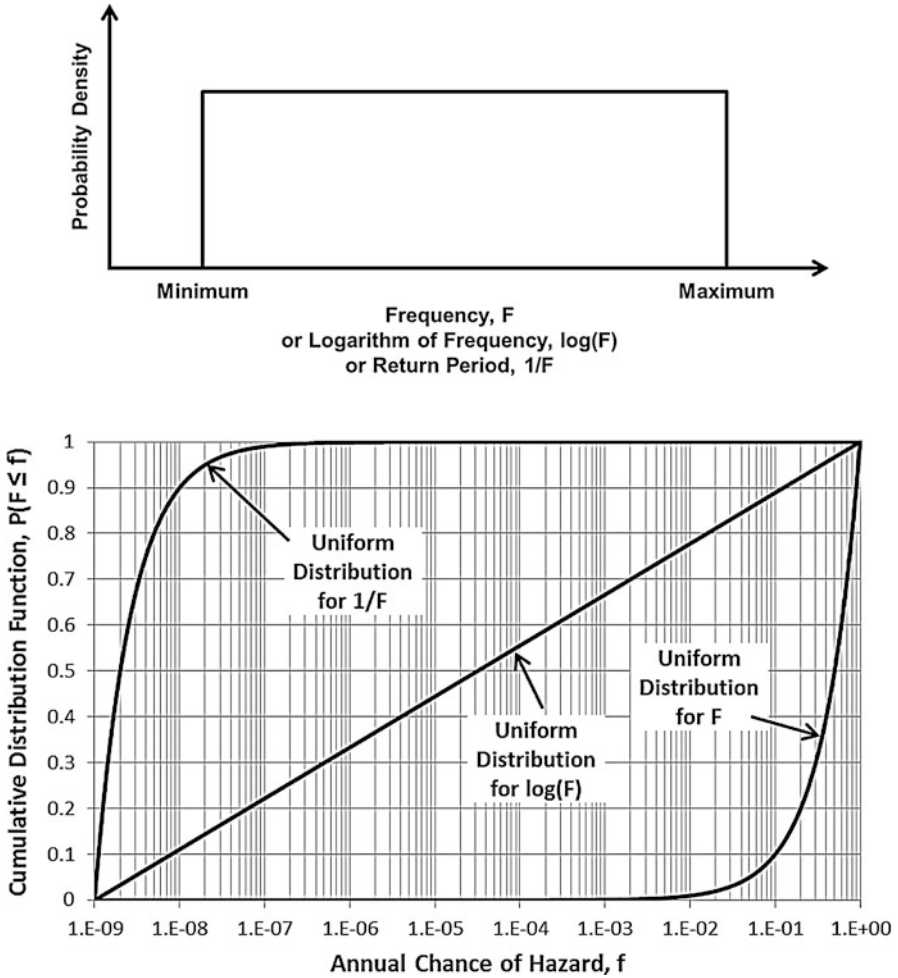


Fig. 18.5 Alternative prior probability distributions for annual chance of occurrence

While mathematically convenient, the principle of insufficient reason has consistently been criticized. Keynes (1921), who renamed it the principle of indifference, contends that this principle is ambiguous and can lead to paradoxical or contradictory conclusions. One of his examples was the specific volume of a substance, where the state of nature is bounded between 1 and 3 L^3/M . Based on the principle of insufficient reason, it would be equally probable for the specific volume to be between 1 and 2 or between 2 and 3 L^3/M . The density, the inverse of the specific volume, will be bounded between $1/3$ and $1 M/L^3$. In this case, the principle of insufficient reason suggests it would be equally probable that the density be between $1/3$ and $2/3 M/L^3$ or $2/3$ and $1 M/L^3$, meaning that it would be equally probable for the specific volume to be between 1 and $1.5 L^3/M$ or 1.5 and

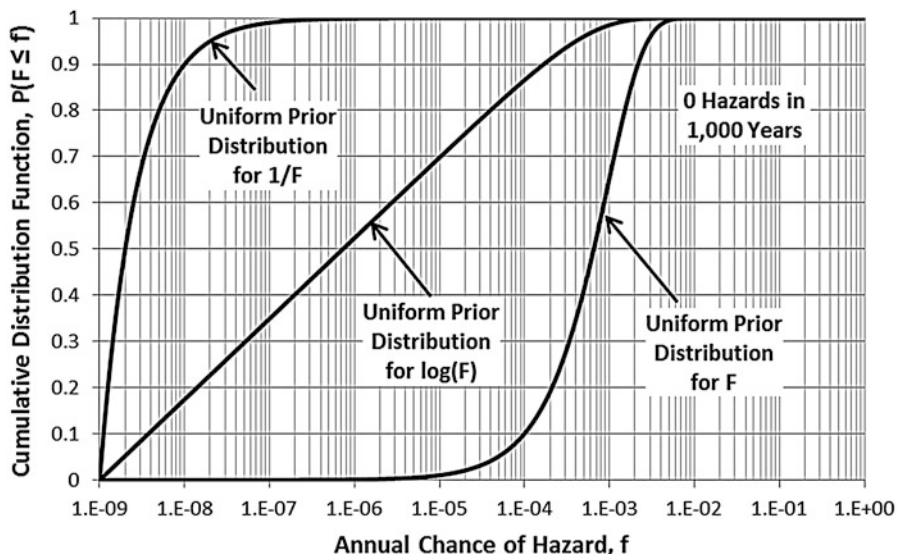


Fig. 18.6 Three different updated distributions for annual chance of occurrence

Table 18.1 Descriptors for updated distributions in Fig. 18.6

Descriptor	Uniform prior distribution for F	Uniform prior distribution for $\log(F)$	Uniform prior distribution for $1/F$
10th percentile	1×10^{-4}	4×10^{-9}	1.1×10^{-9}
50th percentile	7×10^{-4}	8×10^{-7}	2×10^{-9}
90th percentile	2×10^{-3}	2×10^{-4}	1×10^{-8}
Expected value	1×10^{-3}	8×10^{-5}	1×10^{-8}

$3 L^3/M$. Therefore, these two perspectives give different prior probabilities for the same states of nature.

This difficulty of consistency is illustrated with the example of assessing the annual chance of occurrence for a hazard (Figs. 18.5, 18.6, and 18.7). There is no theoretical basis for defining the states of nature in terms of F , $\log(F)$, or $1/F$. One could justify $\log(F)$ on the basis that order-of-magnitude ranges are typically of interest for the chance of hazards (e.g., Fig. 18.1). One could justify $1/F$ on the basis that the hazards are typically represented by their return period.

Luce and Raiffa (1957) criticized the principle of insufficient reason in the context of decision making. They illustrated their criticism with an example decision problem where the consequences of different decision alternatives depend on uncertain states of nature. They showed they could affect which decision alternative was preferred by arbitrarily dividing states of nature into substates. Luce and Raiffa (1957) contended that this result is irrational; if there is “complete ignorance,” then how the states of nature are labeled should not affect the preferred decision alternative.

Journel and Deutsch (1993) also criticized the principle of insufficient reason for producing irrational results. Their example is an oil reservoir in which the permeability of the reservoir varies spatially. They showed that maximizing the entropy of information in the input (the permeability field) generally minimizes entropy in the output of interest for making decisions (the water breakthrough time).

In summary, the challenge with assessing probabilities in practice is in establishing a non-informative prior probability distribution as a starting point. This non-informative prior probability distribution can significantly affect the assessed probabilities, particularly when there is limited information available. Existing approaches to establish non-informative prior probabilities based on the principle of insufficient reason do not provide a rational or consistent basis.

18.3 Decision Entropy Theory

We postulate a theory for establishing a non-informative sample space that characterizes the information in terms of making a decision between two alternatives. The greatest lack of information for the decision, i.e., the non-informative starting point, is defined by the following three principles:

1. An alternative compared to another alternative is equally probable to be preferred or not to be preferred.
2. The possible gains or losses for one alternative compared to another alternative are equally probable.
3. The possibilities of learning about the preference of one alternative compared to another with new information are equally probable.

The premise of this theory is that probabilities provide input to decision making (i.e., managing risk); therefore, non-informative probabilities are probabilities that do not inform the decision. The mathematical framework for implementing these principles is presented in Appendix 1.

The intent of this theory is to overcome the shortcomings associated with the principle of insufficient reason:

1. It is consistent: It always produces the same non-informative distribution of the possible outcomes of a decision no matter how those outcomes formulated in terms of states of nature. This non-informative distribution is what will affect the preferred decision alternative and the value of obtaining additional information about the decision.
2. It is rational: It follows the premise that the purpose of assessing probabilities is to support decision making. The non-informative distribution about the possible outcomes of a decision represents the maximum uncertainty in making a decision.

To illustrate this theory, consider a basic decision in risk management either eliminating a risk with mitigation or accepting the risk with an uncertain annual

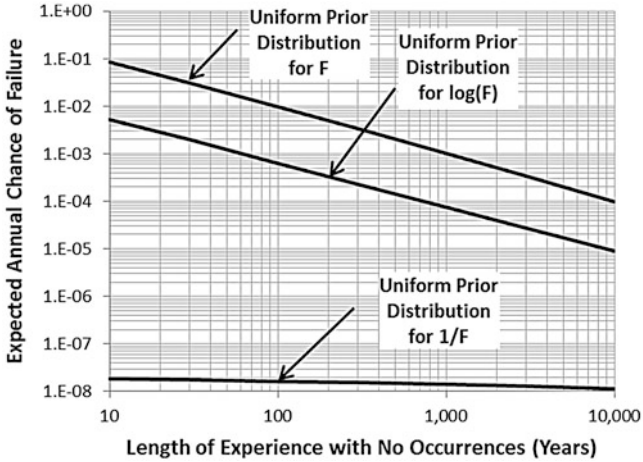


Fig. 18.7 Expected annual chance of failure of three different prior probability distributions for annual chance of occurrence

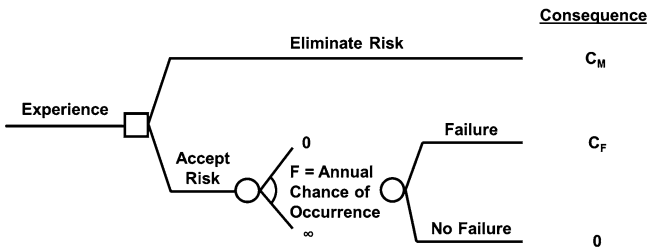


Fig. 18.8 Basic decision tree for risk management with uncertain chance of failure

chance of failure (Fig. 18.8). The alternative to accept the risk will be preferred when its expected annual consequence is smaller⁴ than that for eliminating the risk:

$$\text{Prefer Accept Risk if } c_F \times F \leq c_M \text{ or } F \leq \frac{c_M}{c_F} = \frac{1}{c_F/c_M}$$

where c_M is the annual cost of eliminating the risk, c_F is the cost of a failure due to the occurrence of the hazard, and F is the annual chance of occurrence for the hazard. Therefore, the ratio $\frac{1}{c_F/c_M}$ is indicative of the threshold for “tolerable” risk in Fig. 18.1.

⁴Consequence will be considered here as a positive cost. The degree of preference for an outcome increases as the cost of that outcome decreases.

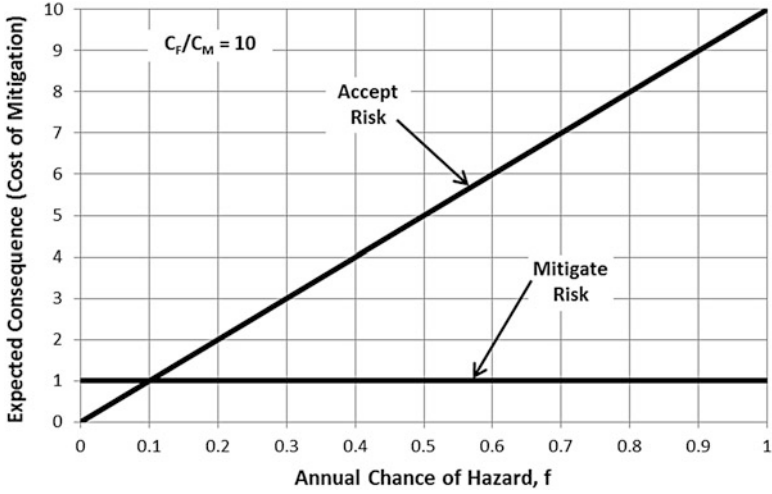


Fig. 18.9 Expected consequences versus annual chance of hazard for accepting risk and mitigating risk

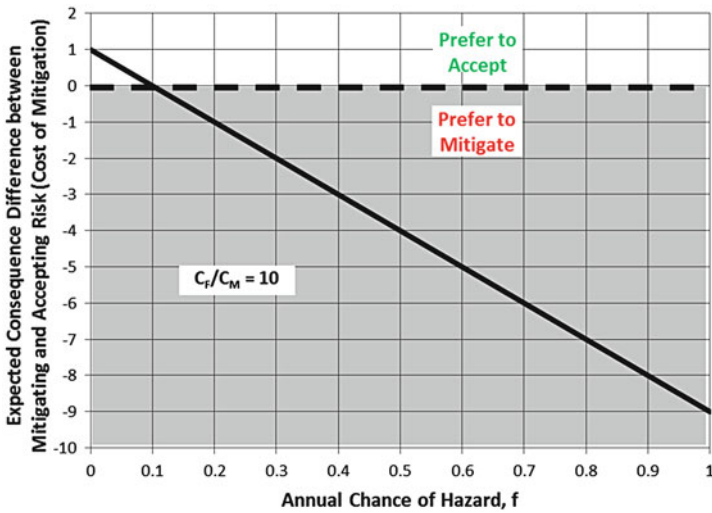


Fig. 18.10 Difference in expected consequences between mitigating risk and accepting risk versus annual chance of hazard

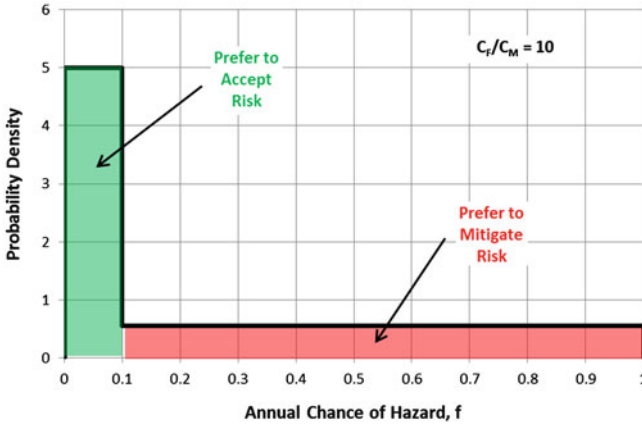


Fig. 18.11 Non-informative prior probability distribution for annual chance of hazard in decision between mitigating risk and accepting risk

The expected consequences for each alternative and the difference in expected consequences between alternatives are linear functions of the uncertain annual chance of occurrence for the hazard (Figs. 18.9 and 18.10). Subsequently, the non-informative prior probability distribution for the annual chance of occurrence based on the principle of Decision Entropy Theory is a bi-uniform distribution (Fig. 18.11). If a decision were made on the basis entirely of the prior probability distribution, the expected value for the annual chance of occurrence is 0.32 and the expected difference in the costs between risk mitigation and risk acceptance is $-2.2c_M$. Since this expected cost difference is negative, the risk mitigation alternative would be preferred (i.e., it has the smallest expected cost).

This prior probability distribution for the annual chance of occurrence is updated with whatever experience is available, such as no occurrences of the hazard in 1000 years of experience (Fig. 18.12). In this case, the update expected value for the annual chance of occurrence is 1×10^{-3} and the expected cost difference between risk mitigation and risk acceptance is $+0.99$, meaning that risk acceptance is now preferred based on the available experience.

18.4 Practical Insights for Risk Management from Decision Entropy Theory

The Decision Entropy Theory provides practical insights into managing risk in the face of uncertainty. The significance of uncertainty depends on its impact in risk management decisions. The value of additional information in a decision is

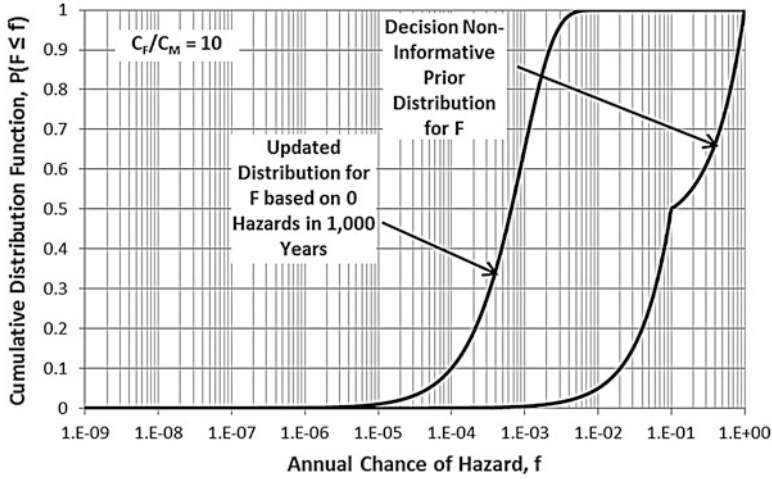


Fig. 18.12 Updated probability distribution for annual chance of hazard in decision between mitigating risk and accepting risk with no occurrences of hazard in 1000 years of experience

accentuated when a non-informative prior sample space is established. Finally, the possibility that available data are not relevant can have a significant impact on a risk management decision.

18.4.1 Significance of Uncertainty in Natural Hazards Depends on Risk Management Decisions

For the basic risk management decision (Fig. 18.8), the prior probability distribution for the annual chance of the hazard depends on the ratio of the cost of failure to the annual cost of mitigation, C_F/C_M . The larger the ratio, the smaller the threshold value for the annual chance of the hazard between preferring to accept versus mitigate the risk, $\frac{1}{C_F/C_M}$, and the more pronounced the left hand tail of the probability distribution for the annual chance of the hazard (Fig. 18.13). This manifestation of a “lack of information” at the start allows for the possibility that the actual chance of the hazard could be greater or smaller than the threshold (the decision point); note that the 50th percentile in the prior probability distribution is at $\frac{1}{C_F/C_M}$ (Fig. 18.13).

The result of starting open to the possibility of the decision going either way without any information is that subsequent information can be more influential in changing the decision (Fig. 18.14). To illustrate this point, consider the decision with $C_F/C_M = 1 \times 10^3$, meaning that the threshold value for the chance of the hazard is 1×10^{-3} . With the non-informative prior probability distribution, the preferred alternative is to mitigate the risk. However, an experience of no hazards occurring in a period of greater than 50 years is enough to change the preferred alternative to

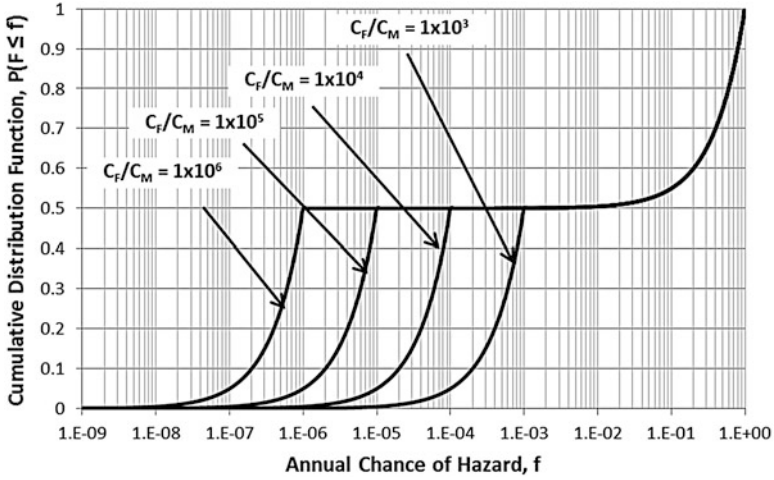


Fig. 18.13 Prior probability distributions for annual chance of hazard in decision between mitigating risk and accepting risk

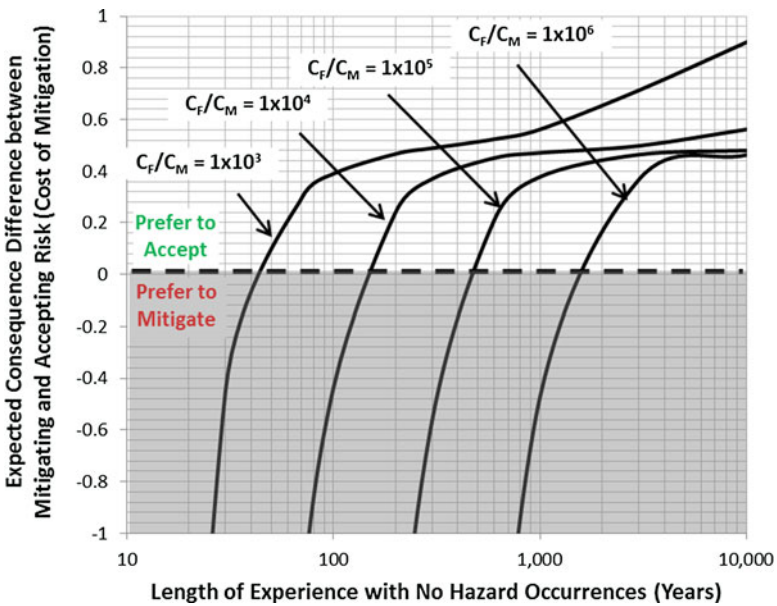


Fig. 18.14 Difference in expected consequences between mitigating risk and accepting risk versus length of experience with no hazard occurrences

accepting the risk (Fig. 18.14). For comparison, here is how other prior probability distributions would affect this decision:

- With a uniform prior probability distribution for the annual chance of the hazard, a period of greater than 1000 years with no hazards occurring would be required to change the preferred alternative to accepting the risk (the updated expected chance of occurrence is less than 1×10^{-3} for lengths of experience with no occurrences greater than 1000 years in Fig. 18.7).
- With a uniform prior probability distribution for the logarithm of the annual chance of the hazard, a period of greater than about years with no hazards occurring would be required to change the preferred alternative to accepting the risk (Fig. 18.7).
- With a uniform prior probability distribution for the return period or the inverse of the annual chance of the hazard, the preferred alternative would be accepting the risk without any information (Fig. 18.7).

Therefore, the greatest lack of information at the start (the non-informative prior probability distribution from Decision Entropy Theory) is not conservative or un-conservative in the context of risk management; it is intended to provide an unbiased starting point before incorporating any available information. In this particular example, the “black swan” that decision entropy is accommodating is the possibility that the annual chance of hazard may actually be “small” as opposed to assuming that it is “large” in the absence of information. While potentially counterintuitive (particularly for “conservative” engineers), this concept is of practical significance and used often implicitly in real-world decisions. We would essentially have never built a major dam (c_F/c_M of about 1×10^6 based on Fig. 18.1) if we needed to wait more than 1,000,000 years to assess the chance of extreme hazard occurrences. Decision entropy provides a rational and consistent basis to support making risk management decisions in the face of the inevitable lack of information about extreme hazard occurrences.

18.4.2 Value of Information Emphasized by a Non-informative Starting Point

The value of obtaining additional information (beyond the available experience) about the chance of hazard depends on how probable it is that additional information will change the risk management approach (Fig. 18.15). Quantitatively, the value of information is the maximum amount the decision maker would be willing to spend to pursue additional information. The value of information for the basic risk management decision is bounded by the cost required to eliminate the risk (c_M).

The value of perfect information about the annual chance of occurrence for the hazard depends both on the decision (i.e., c_F/c_M) and the available experience (Fig. 18.16). For relatively limited experience (small lengths of experience with no

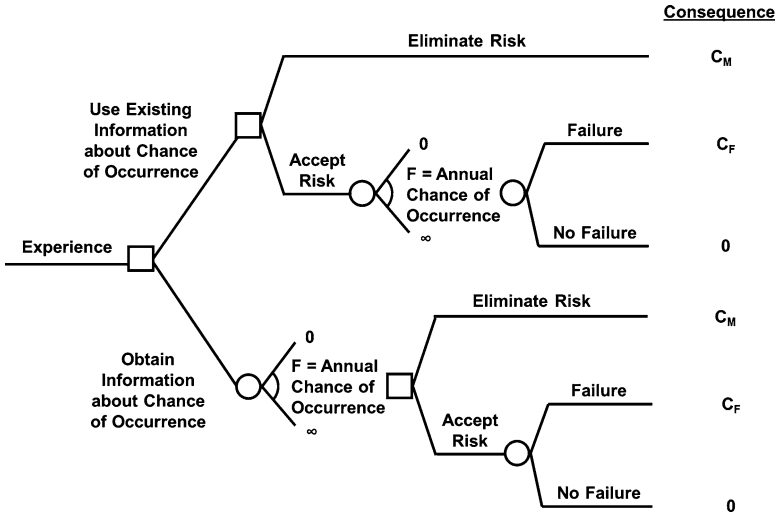


Fig. 18.15 Decision tree to assess value of information in managing risk

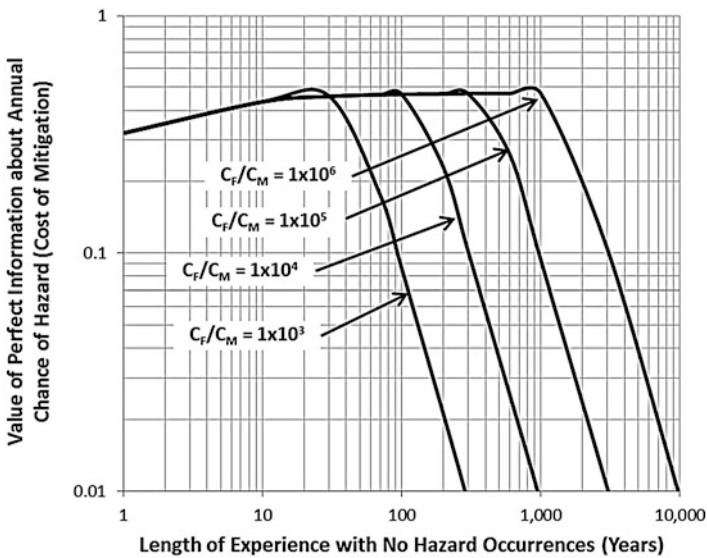


Fig. 18.16 Value of perfect information about annual chance of hazard versus length of available experience with no hazard occurrences

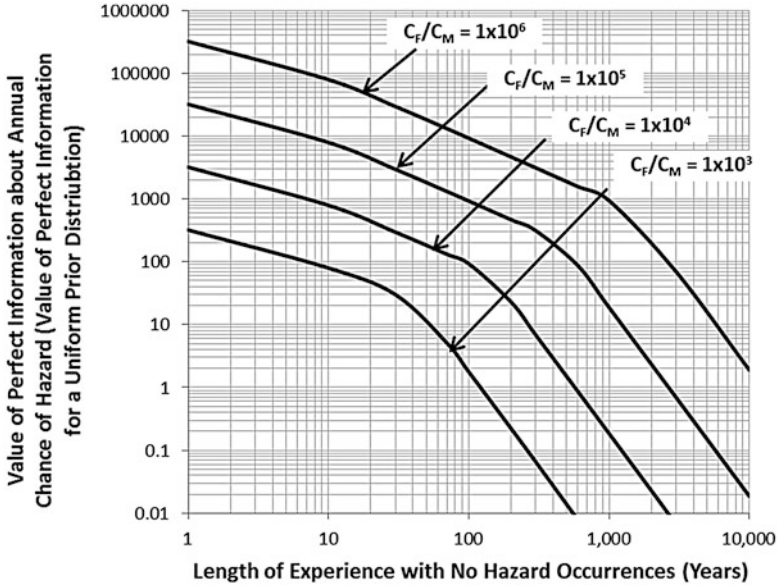


Fig. 18.17 Value of perfect information about annual chance of hazard relative to that obtained with a uniform prior probability distribution for the annual chance of hazard versus length of available experience with no hazard occurrences

hazard occurrences), the value of perfect information is between 25 and 50 % of the cost of risk mitigation. It is interesting that the value of perfect information increases initially with small amounts of experience (Fig. 18.16); this result reflects that potential to change this particular decision (i.e., mitigate the risk) based on additional information at first increases as length of experience with no occurrences increases. As the length of experience without an occurrence exceeds the threshold where the preferred alternative changes from risk mitigation to risk acceptance (Fig. 18.14), the value of perfect information decreases because it becomes less probable that new information beyond the available experience will change the risk management approach (Fig. 18.16). In practice, we will typically have experience bases less than 100 years, and the value of perfect information will be the greatest. Also, the value of perfect information is typically orders of magnitude greater than that obtained using a uniform prior probability distribution for the annual chance of the hazard (Fig. 18.17).

The value of obtaining additional information about the chance of hazard is of practical significance. The greater the value of information, the more important the role is of science and engineering to advance our understanding about natural hazards and managing their risks. For example, greater knowledge about the causes of long debris runouts from landslides could provide information about the chances of them occurring beyond simply waiting for a long record of experience. In addition, the greater the value of information, the greater the value of an adaptable

approach to managing risk. For example, a means of mitigating the risk from debris runout that could readily be modified if additional information indicates that accepting the risk is preferred will be more effective than one that cannot easily be changed once it is implemented.

18.4.3 Possibility that Available Experience Is Not Relevant Can Have Significant Impact on Risk Management

In many cases, there is a possibility that the available experience about the annual chance of hazard occurrence may not be directly relevant to predicting what it will be in the future for the purpose managing risk. For example, the annual chance that the storm surge at a location on the Gulf of Mexico coast exceeds a particular height may be different in the next 100 years compared to what it was in the last 100 years due to changes in hurricane frequencies and intensities caused by climate changes. The limiting cases are that the experience is relevant or that it is not relevant.

The relevancy of the data affects the risk management decision because the relevancy is uncertain in making the decision (Fig. 18.18), where F_A is the annual chance of occurrence in the period of available experience, Data Set A, and F_B is the annual chance of occurrence in the future for purposes of managing risk, Data Set B.

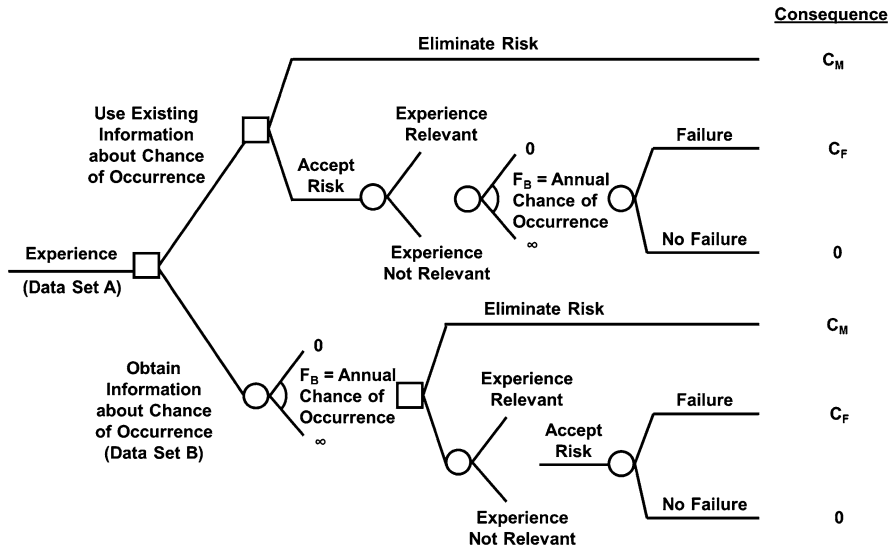


Fig. 18.18 Decision tree to assess impact of relevance of available experience in managing risk

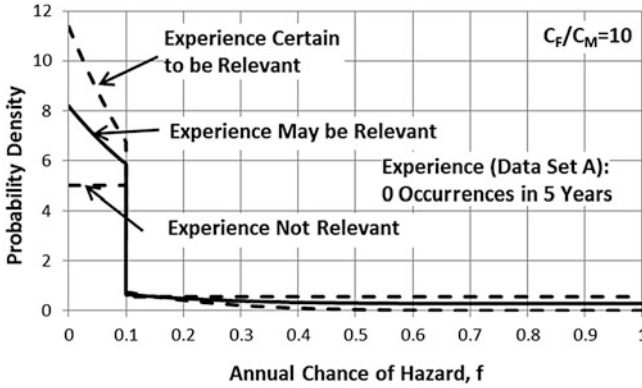


Fig. 18.19 Example prior probability distribution considering the possibility that the experience may be relevant to the risk management decision

The third principle of Decision Entropy Theory establishes the prior probability that the experience (Data Set A) is relevant. If the experience is relevant, then the greatest possible is learned from the experience. If the experience is not relevant, then the least possible (nothing) is learned from the experience. Therefore, the third principle establishes that the non-informative prior probability that the experience is relevant is equal to one-half (Appendix 2), and the prior probability distribution for the annual chance of the hazard in the decision (F_B) is in the middle of the two extremes where the experience is or is not relevant (Fig. 18.19)

To illustrate the significance of data relevancy, consider the value of obtaining additional information about the annual chance of occurrence in the future for purposes of managing risk, F_B (Fig. 18.18); see Appendix 2 for details. If the new data are not consistent with the experience, then the updated distribution reflects more weight on the new data versus the experience and the probability that the experience is relevant decreases (Fig. 18.20). Conversely, if the new data are consistent with the experience, then the updated distribution reflects the combined data and the probability that the experience is relevant increases (Fig. 18.20). Therefore, allowing for the possibility that experience may not be relevant allows for the preferred risk management decision to change more quickly with additional information.

The value of perfect information about the annual chance of occurrence for the purposes of making a risk management decision balances the two extremes where the experience is or is not relevant (Fig. 18.21). Keeping the possibility open that the experience may not be relevant can significantly increase the value of new information when the available experience is seemingly extensive (Fig. 18.21).

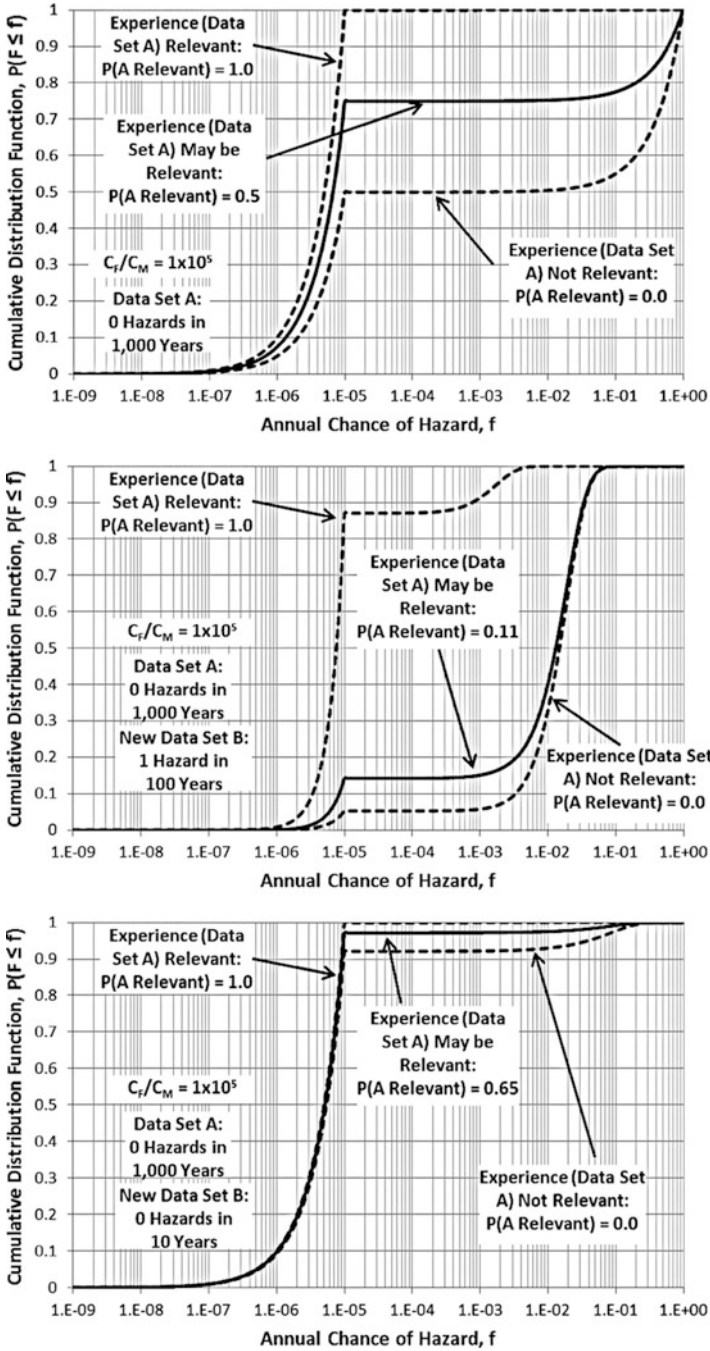


Fig. 18.20 Example prior (top graph) and updated probability distributions (lower graphs) considering possibility that experience may be relevant to the risk management decision

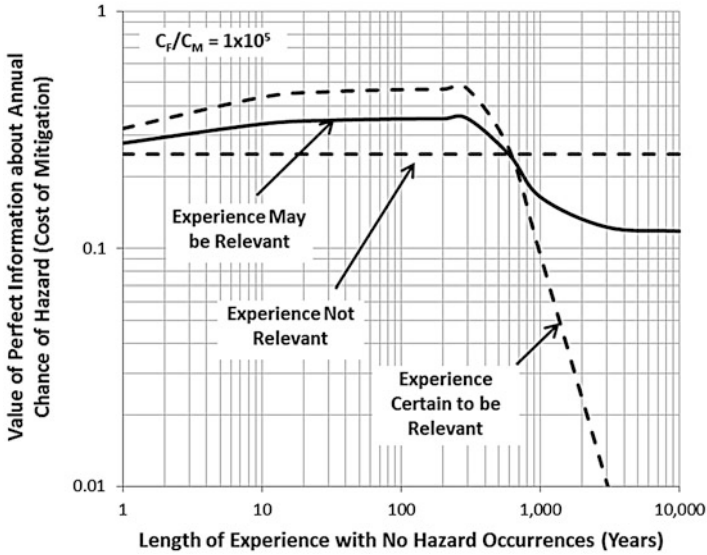


Fig. 18.21 Value of example updated probability distributions considering the possibility that the experience may not be relevant to the risk management decision

18.5 Summary and Conclusions

This chapter proposes a new framework, Decision Entropy Theory, to assess probabilities and manage risks for natural hazard possibilities that are beyond our experience. The theory postulates a starting point for assessing probabilities that reflect having no information in making a risk management decision between two alternatives:

1. An alternative compared to another alternative is equally probable to be preferred or not to be preferred.
2. The possible gains or losses for one alternative compared to another alternative are equally probable.
3. The possibilities of learning about the preference of one alternative compared to another with new information are equally probable.

From this non-informative starting point, all available information (if any) can be incorporated through Bayes' theorem. Decision Entropy Theory attempts to provide for consistency and rationality that is lacking in available approaches for assessing probabilities with limited information.

From a practical perspective, Decision Entropy Theory highlights the importance of considering how possibilities for natural hazards could impact the preferred alternatives for managing risks. A lack of information at the start means the greatest uncertainty in the preferred alternative, not the greatest uncertainty in the

hazard. Limited information does not necessarily justify a conservative approach for managing risk; additional information could lead to either less expensive or more expensive means of risk management being preferred. The value of obtaining new information for managing risk is accentuated when limited information is available; therefore, the role science and engineering to advance our understanding about natural hazards and managing their risk is emphasized. Ultimately, this framework underscores the importance of developing adaptable approaches to manage multi-hazard risks in the face of limited information.

A.1 Appendix 1: Mathematical Formulation for Principles of Decision Entropy

The following appendix provides the mathematical formulation characterizing the entropy of a decision.

Principle Number 1 In a non-informative sample space, a selected alternative is equally probable to be or not to be preferred compared to another alternative.

Given that alternative A_i is selected and compared to A_j , the maximum lack of information for the decision corresponds to maximizing the relative entropy for the events that an alternative is and is not preferred:

$$\begin{aligned} & \text{Maximize } H_{\text{rel}} \left(\text{Preference Outcome} \mid A_i \text{ Selected and Compared to } A_j \right) \\ & = -P[A_i \text{ Preference to } A_j] \ln \{P[A_i \text{ Preference to } A_j]\} \\ & \quad - P[A_i \overline{\text{Preference to } A_j}] \ln \{P[A_i \overline{\text{Preference to } A_j}]\} - \ln(2) \end{aligned}$$

where $H_{\text{rel}}(\text{Preference Outcome} \mid A_i \text{ Selected and Compared to } A_j)$ is the relative entropy of the decision preference, $P[A_i \text{ Preferred to } A_j]$ is the probability that alternative A_i is preferred compared to alternative A_j , and $P[A_i \overline{\text{Preferred to } A_j}]$ is the probability that alternative A_i is not preferred compared to alternative A_j or $1 - P[A_i \text{ Preferred to } A_j]$. If the preference for one alternative versus another is characterized by the difference in the utility values for each alternative, then the relative entropy of the decision preference can be expressed as follows:

$$\begin{aligned} & \text{Maximize } H_{\text{rel}} \left(\text{Preference Outcome} \mid A_i \text{ Selected and Compared to } A_j \right) \\ & = -P \left[u(A_i) - u(A_j) > 0 \mid A_i \text{ Compared to } A_j \right] \\ & \quad \times \ln \left\{ P \left[u(A_i) - u(A_j) > 0 \mid A_i \text{ Compared to } A_j \right] \right\} \\ & \quad - \left\{ 1 - P \left[u(A_i) - u(A_j) > 0 \mid A_i \text{ Compared to } A_j \right] \right\} \\ & \quad \times \ln \left\{ 1 - P \left[u(A_i) - u(A_j) > 0 \mid A_i \text{ Compared to } A_j \right] \right\} - \ln(2) \end{aligned}$$

where $u(A_i)$ is the utility for alternative A_i .

The entropy is a measure of the frequencies or probabilities of possible outcomes divided between the two states: A_i Preferred to A_j and A_i Preferred to A_j . The natural logarithm of the entropy is used for mathematical convenience by invoking Stirling’s approximation for factorials; note that maximizing the logarithm of the frequencies of possible outcomes is the same as maximizing the frequencies of possible outcomes since the logarithm is a one-to-one function of the argument. The relative entropy normalizes the entropy by the number of possible states (in this case two); the maximum value of the relative entropy is zero.

For two outcomes, A_i Preferred to A_j and A_i Preferred to A_j , the relative entropy $H_{rel}(\text{Preference Outcome} | A_i \text{ Selected and Compared to } A_j)$ is maximized in the ideal case where it is equally probable that one or the other alternative is preferred, or

$$P [u(A_i) - u(A_j) > 0 | A_i \text{ Compared to } A_j] \\ = 1 - P [u(A_i) - u(A_j) > 0 | A_i \text{ Compared to } A_j].$$

When $i = j$ (i.e., an alternative is compared to itself), the relative entropy becomes equal to its minimum possible value, $-\ln(2)$, because there is no uncertainty in the preference [$p \ln(p) \rightarrow 0$ as $p \rightarrow 0$ or $p \rightarrow 1$].

Principle Number 2 In a non-informative sample space, possible degrees of preference between a selected alternative and another alternative are equally probable.

To represent the maximum lack of information for the decision between two alternatives, maximize the relative entropy for the possible positive and negative differences in utility values between alternative A_i and A_j , $\Delta u_{i,j} = u(A_i) - u(A_j)$. The sample space where A_i is Selected and Compared to A_j is divided into two states, A_i Preferred to A_j and A_i Preferred to A_j , which are respectively subdivided into $n_{\Delta u_{i,j} | A_i > A_j}$ and $n_{\Delta u_{i,j} | A_i \preccurlyeq A_j}$ substates for each interval of $\Delta u_{i,j}$, where the operator $>$ means “preferred to” and the operator \preccurlyeq means “not preferred to” or the complement of $>$. Maximizing the relative entropy for degree of preference is then given by the following:

$$\text{Maximize } H_{rel} \left(\text{Preference Degrees} \middle| A_i \text{ Selected and Compared to } A_j \right) \\ = H_{rel} \left(\text{Preference Degrees, } A_i \text{ Preferred to } A_j \right) \\ + H_{rel} \left(\text{Preference Degrees, } A_i \overline{\text{Preferred to}} A_j \right)$$

where

$$\begin{aligned}
 &H_{\text{rel}}(\text{Preference Degrees, } A_i \text{ Preferred to } A_j) \\
 &= \sum_n \sum_{\Delta u_{i,j} | A_i \succ A_j} -P(\Delta u_{i,j} | A_i \text{ Preferred to } A_j) P(A_i \text{ Preferred to } A_j) \\
 &\quad \times \ln \left[P(\Delta u_{i,j} | A_i \text{ Preferred to } A_j) P(A_i \text{ Preferred to } A_j) \right] \\
 &\quad - \left(\frac{1}{2} \right) \ln \left(n_{\Delta u_{i,j} | A_i \succ A_j} \times 2 \right)
 \end{aligned}$$

and

$$\begin{aligned}
 &H_{\text{rel}}(\text{Preference Degrees, } A_i \overline{\text{Preferred to } A_j}) \\
 &= \sum_n \sum_{\Delta u_{i,j} | A_i \preceq A_j} -P(\Delta u_{i,j} | A_i \overline{\text{Preferred to } A_j}) P(A_i \overline{\text{Preferred to } A_j}) \\
 &\quad \times \ln \left[P(\Delta u_{i,j} | A_i \overline{\text{Preferred to } A_j}) P(A_i \overline{\text{Preferred to } A_j}) \right] \\
 &\quad - \left(\frac{1}{2} \right) \ln \left(n_{\Delta u_{i,j} | A_i \preceq A_j} \times 2 \right)
 \end{aligned}$$

The maximum value for $H_{\text{rel}}(\text{Preference Degrees} | A_i \text{ Selected and Compared to } A_j)$ is equal to zero, and it is achieved when $P(A_i \text{ Preferred to } A_j) = P(A_i \overline{\text{Preferred to } A_j}) = 1/2$, all $P(\Delta u_{i,j} | A_i \text{ Preferred to } A_j) = 1/n_{\Delta u_{i,j} | A_i \succ A_j}$ and

all $P(\Delta u_{i,j} | A_i \overline{\text{Preferred to } A_j}) = 1/n_{\Delta u_{i,j} | A_i \preceq A_j}$. Note that the number of substates

for intervals of $\Delta u_{i,j}$ in $A_i \text{ Preferred to } A_j$ or in $A_i \overline{\text{Preferred to } A_j}$ is not important in maximizing the relative entropy; the entropy is maximized when the possible intervals (however many there are) are as equally probable as possible.

Principle Number 3 In a non-informative sample space, possible expected degrees of information value for the preference between a selected alternative and another alternative are equally probable.

To represent the maximum lack of information for a value of information assessment for the decision between two alternatives, maximize the relative entropy for the possible positive and nonpositive expected changes with information in the differences in expected utility values between alternative A_i and A_j . These possible expected changes are termed the “information value” and denoted by $\Omega_{E_{k,l}}$:

$$\begin{aligned}
 (\Omega_{E_{k,l}})_{i,j} &= E(\Delta u_{i,j} | E_k \cap S_{i,j}) - E(\Delta u_{i,j} | E_l \cap S_{i,j}) \\
 &= E[u(A_i) - u(\overline{A_i}) | E_k \cap S_{i,j}] - E[u(A_i) - u(\overline{A_i}) | E_l \cap S_{i,j}]
 \end{aligned}$$

where E_k and E_l are two sets of possible information about the preference between A_i and A_j , $\Delta u_{i,j} = u(A_i) - u(A_j)$. The sample space for $E_{k,l}$ is divided into two subsets, an expected positive information value (i.e., $\Omega_{E_{k,l}} > 0$) with $n_{\Omega_{E_{k,l}} > 0}$ states and an expected nonpositive information value (i.e., $\Omega_{E_{k,l}} \leq 0$) with $n_{\Omega_{E_{k,l}} \leq 0}$ states. Maximizing the relative entropy for possible information values is then given by the following:

$$\begin{aligned} \text{Maximize } H_{\text{rel}} \left(\text{Information Value Degrees} \middle| \text{Information about } A_i \text{ Selected and} \right. \\ \left. \text{Compared to } A_j \right) = H_{\text{rel}} \left(\text{Information Value Degrees, Information has Positive} \right. \\ \left. \text{Value for } A_i \text{ Compared to } A_j \right) + H_{\text{rel}} \left(\text{Information Value Degrees, Information} \right. \\ \left. \text{has Non-Positive Value for } A_i \text{ Compared } A_j \right) \end{aligned}$$

where

$$\begin{aligned} H_{\text{rel}} \left(\text{Information Value Degrees, Information has Positive Value for } A_i \text{ Compared} \right. \\ \left. \text{to } A_j \right) = \sum_{n_{\Omega_{E_{k,l}} > 0}} \frac{-P \left(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} > 0 \right) P \left(\Omega_{E_{k,l}} > 0 \right)}{\times \ln \left[P \left(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} > 0 \right) P \left(\Omega_{E_{k,l}} > 0 \right) \right]} - \left(\frac{1}{2} \right) \ln \left(n_{\Omega_{E_{k,l}} > 0} \times 2 \right) \end{aligned}$$

and

$$\begin{aligned} H_{\text{rel}} \left(\text{Information Value Degrees, Information has Non-Positive Value for } A_i \right. \\ \left. \text{Compared to } A_j \right) = \sum_{n_{\Omega_{E_{k,l}} \leq 0}} \frac{-P \left(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} \leq 0 \right) P \left(\Omega_{E_{k,l}} \leq 0 \right)}{\times \ln \left[P \left(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} \leq 0 \right) P \left(\Omega_{E_{k,l}} \leq 0 \right) \right]} \\ - \left(\frac{1}{2} \right) \ln \left(n_{\Omega_{E_{k,l}} \leq 0} \times 2 \right) \end{aligned}$$

The maximum value for the relative entropy for possible information value is equal to zero and achieved when $P(\Omega_{E_{k,l}} > 0) = P(\Omega_{E_{k,l}} \leq 0) = 1/2$, all $P(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} > 0) = 1/n_{\Omega_{E_{k,l}} > 0}$ and all $P(\Omega_{E_{k,l}} \middle| \Omega_{E_{k,l}} \leq 0) = 1/n_{\Omega_{E_{k,l}} \leq 0}$.

This principle is consistent with the first two principles where the alternative of obtaining information, $E_{k,l}$, is compared with the alternative of not obtaining information for a given preference comparison (i.e., $E_l = E_0 = \emptyset$): when the relative entropy of the information value is maximized, there is an equal probability that obtaining the information is preferred (i.e., has positive information value or $\Omega_{E_{k,l}} > 0$) and is not preferred (i.e., has nonpositive information value or $\Omega_{E_{k,l}} \leq 0$) and the possible positive and nonpositive degrees of information value are equally probable.

Fig. 18.22 Sample space for decision with three alternatives

		Comparison Alternative, A_j		
		A_1	A_2	A_3
Selected Alternative, A_i	A_1			
	A_2			
	A_3			

A.1.1 Multiple Pairs of Alternatives

The principles of decision entropy establish a sample space for the comparison of any two decision alternatives, A_i and A_j . The sample space for the set of all possibilities of comparison for a given decision problem is denoted the decision sample space. In this sample space, each possible combination of an alternative that is selected, A_i , and an alternative that could have been selected, A_j , is equally probable (Fig. 18.22). For n_A alternatives, there are n_A^2 pairs of i, j and $P(A_i \text{ Selected and Compared to } A_j) = 1/n_A^2$ and $P(A_i \text{ Compared to } A_j | A_i \text{ Selected}) = 1/n_A$. The preferred decision alternative has the maximum expected degree of preference compared to all other alternatives:

$$E[u(A_i) - u(\bar{A}_i) | A_i \text{ Selected}] = \sum_{\text{all } j} E(\Delta u_{i,j} | A_i \text{ Selected and Compared to } A_j) \times P(A_i \text{ Compared to } A_j | A_i \text{ Selected})$$

where

$$E(\Delta u_{i,j} | A_i \text{ Selected and Compared to } A_j) = \sum_{\text{all } \Delta u_{i,j}} \Delta u_{i,j} P(\Delta u_{i,j} | A_i \text{ Selected and Compared to } A_j)$$

The use of the expected degree of preference (or utility difference) as a measure of preference is consistent with utility theory analysis (e.g., Von Neumann and Morgenstern 1944; Hurwicz 1951; Savage 1951; Hodges and Lehmann 1952; Luce and Raiffa 1957; Raiffa and Schlaifer 1961; Benjamin and Cornell 1970; Ang and Tang 1984, etc.). In a conventional decision analysis, the sample space for utility values is not conditioned on a particular alternative being selected, A_i , meaning that the expected utility for a selected alternative can be calculated without considering the alternatives to which it is being compared. However, the absolute magnitude

of expected utility is irrelevant¹; its relevance depends on comparing it with the expected utility values for other alternatives. Therefore, it is the differences between utility values that are of interest.

Mathematically, comparing the expected degrees of preference, $E[u(A_i) - u(\bar{A}_i) | A_i \text{ Selected}]$, is the same as comparing expected utility values in a conventional decision analysis. In a conventional analysis where the probabilities for utility values given that an alternative has been selected do not depend on the alternative to which it is being compared, the expected degree of preference for an alternative can be expressed as follows:

$$\begin{aligned} E[u(A_i) - u(\bar{A}_i) | A_i \text{ Selected}] &= \sum_{\text{all } j} E(\Delta u_{i,j} | A_i \text{ Selected and Compared to } A_j) \\ &\quad \times P(A_i \text{ Compared to } A_j | A_i \text{ Selected}) \\ &= \sum_{\text{all } j} [E(u_i | A_i \text{ Selected}) - E(u_j | A_j \text{ Selected})] \\ &\quad \times P(A_i \text{ Compared to } A_j | A_i \text{ Selected}) \\ &= E(u_i | A_i \text{ Selected}) - \sum_{j=1 \text{ to } n_A} E(u_j | A_j \text{ Selected}) \times \left(\frac{1}{n_A}\right) \end{aligned}$$

Therefore, the expected degree of preference for an alternative in a conventional decision analysis is equal to the expected utility for that alternative minus a constant (the average expected utility for all alternatives). Therefore, the order of comparisons is the same whether the expected utility values, $E(u_i | A_i \text{ Selected})$, or the expected degree of preference values, $E[u(A_i) - u(\bar{A}_i) | A_i \text{ Selected}]$, are used in comparisons.

B.1 Appendix 2: Implementation of Bayes' Theorem with Bernoulli Sequence for Two Possibly Related Sets of Data

Define F_A as the annual chance of occurrence in the period of available experience, Data Set A, and F_B as the annual chance of occurrence in the future for purposes of managing risk, Data Set B (Fig. 18.18). If the annual chance of occurrence is the same in the past and the future, then the likelihood of obtaining the available data from a Bernoulli sequence (i.e., x_A occurrences in n_A years) for a particular value of f_{Bi} is given by:

$$P(x_A \text{ in } n_A | f_{Bi} = f_{Ai}) = \left\{ \left[\frac{n_A!}{x_A! (n_A - x_A)!} \right] f_{Bi}^{x_A} (1 - f_{Bi})^{n_A - x_A} \right\}$$

¹Utility values can be scaled arbitrarily with linear transformations.

Conversely, if the annual chance of occurrence is not the same in the past and the future, then the likelihood of obtaining the available data (i.e., x_A occurrences in n_A years) is given by:

$$P(x_A \text{ in } n_A | f_{Bi} \neq f_{Ai}) = \sum_{\text{all } f_{Ai}} \left\{ \left[\frac{n_A!}{x_A! (n_A - x_A)!} \right] f_{Ai}^{x_A} (1 - f_{Ai})^{n_A - x_A} \right\} P(f_{Ai} | f_{Bi} \neq f_{Ai})$$

where $P(f_{Ai})$ is the prior probability for the annual chance of occurrence in the experience. If the experience is relevant, then the prior probability for F_A is the same as that for F_B : $P(f_{Ai} | f_{Bi} = f_{Ai}) = P(f_{Bi})$. Furthermore, $P(f_{Ai} | f_{Bi} \neq f_{Ai}) = P(f_{Ai} | f_{Bi} = f_{Ai}) = P(f_{Ai}) = P(f_{Bi})$ since the probability for F_A does not depend on whether or not the experience is relevant (i.e., the prior probability of obtaining a particular set of data from Set A is the same whether or not Sets A and B are the same, and the probability of obtaining a particular set of data from Set A does not depend on the chance of occurrence in Set B if the two sets are different). Therefore,

$$P(x_A \text{ in } n_A | f_{Bi} \neq f_{Ai}) = \sum_{\text{all } f_{Bi}} \left\{ \left[\frac{n_A!}{x_A! (n_A - x_A)!} \right] f_{Bi}^{x_A} (1 - f_{Bi})^{n_A - x_A} \right\} P(f_{Bi})$$

meaning that the likelihood of the information from the experience is a constant with respect to f_{Bi} (i.e., the updated probability distribution will be same as the prior probability distribution for F_B if the data are not relevant). The composite likelihood function considering the possibility that the data may or may not be relevant is given by:

$$\begin{aligned} P(x_A \text{ in } n_A | f_{Bi}) &= P(x_A \text{ in } n_A | f_{Bi} = f_{Ai}) P(f_{Bi} = f_{Ai}) + P(x_A \text{ in } n_A | f_{Bi} \neq f_{Ai}) P(f_{Bi} \neq f_{Ai}) \\ &= P(x_A \text{ in } n_A | f_{Bi} = f_{Ai}) P(f_{Bi} = f_{Ai}) + P(x_A \text{ in } n_A | f_{Bi} \neq f_{Ai}) [1 - P(f_{Bi} = f_{Ai})] \end{aligned}$$

where $P(f_{Bi} = f_{Ai})$ is the prior probability that the experience is relevant.

If the experience is relevant, then the greatest possible is learned from the experience. If the experience is not relevant, then the least possible (nothing) is learned from the experience. Therefore, the third principle of Decision Entropy Theory establishes that the probability that the experience is relevant is 0.5, or $P(f_{Bi} = f_{Ai}) = 0.5$.

If information could be obtained about the annual chance of hazard occurrence in the period of the decision (F_B) before making the decision (Fig. 18.18), then the likelihood of obtaining a particular set of data (i.e., x_A occurrences in n_A years and x_B occurrences in n_B years) is given by the following:

$$\begin{aligned} &P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi}) \\ &= P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} = f_{Ai}) P(f_{Bi} = f_{Ai}) \\ &+ P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} \neq f_{Ai}) [1 - P(f_{Bi} = f_{Ai})] \end{aligned}$$

where

$$P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} = f_{Ai}) \\ = \left\{ \left[\frac{n_A!}{x_A!(n_A - x_A!)} \right] f_{Bi}^{x_A} (1 - f_{Bi})^{n_A - x_A} \right\} \times \left\{ \left[\frac{n_B!}{x_B!(n_B - x_B!)} \right] f_{Bi}^{x_B} (1 - f_{Bi})^{n_B - x_B} \right\}$$

and

$$P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} \neq f_{Ai}) \\ = \left\{ \sum_{\text{all } f_{Bi}} \left\{ \left[\frac{n_A!}{x_A!(n_A - x_A!)} \right] f_{Bi}^{x_A} (1 - f_{Bi})^{n_A - x_A} \right\} P(f_{Bi}) \right\} \times \left\{ \left[\frac{n_B!}{x_B!(n_B - x_B!)} \right] f_{Bi}^{x_B} (1 - f_{Bi})^{n_B - x_B} \right\}$$

Hence, the updated probability distribution for the annual chance of occurrence for the hazard in the risk management decision (F_B) is obtained from Bayes' theorem as follows:

$$P(F_B = f_{Bi} | x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B) \\ = \frac{\left\{ P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} = f_{Ai}) P(f_{Bi} = f_{Ai}) \right. \\ \left. + P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} \neq f_{Ai}) [1 - P(f_{Bi} = f_{Ai})] \right\}}{\sum_{\text{all } f_{Bj}} \left\{ P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bj} = f_{Aj}) P(f_{Bj} = f_{Aj}) \right. \\ \left. + P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bj} \neq f_{Aj}) [1 - P(f_{Bj} = f_{Aj})] \right\}}$$

Likewise, the probability that the data from the experience (Data Set A) is relevant is updated with the data obtained from the period of the decision (Data Set B):

$$P(f_B = f_{Ai} | x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B) \\ = \frac{P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bi} = f_{Ai}) P(f_{Bi} = f_{Ai})}{\sum_{\text{all } f_{Bj}} \left\{ P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bj} = f_{Aj}) P(f_{Bj} = f_{Aj}) \right. \\ \left. + P(x_A \text{ in } n_A \text{ and } x_B \text{ in } n_B | f_{Bj} \neq f_{Aj}) [1 - P(f_{Bj} = f_{Aj})] \right\}}$$

References

- AGS. (2000). Landslide risk management concepts and guidelines. *Australian Geomechanics*, 35(1), 49–92.
- ANCOLD. (1996). *Commentary on ANCOLD guidelines on risk assessment*. Sydney, Australia: Australian National Committee on Large Dams.

- Ang, A. A-S. & Tang, W. H., (1984). *Probability Concepts in Engineering Planning and Design, Volume II - Decision, Risk and Reliability*, John Wiley & Sons, New York.
- Bernoulli, D. (1738). *Specimen Theoriae Novae de Mensura Sortis. Commentarii Academiae Scientiarum imperialis Petropolitanae, Tomus, V, 175–192* [Trans. (1954) Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1), 23–36].
- Benjamin, J. R., & Cornell, C. A. (1970). *Probability, statistics, and decision for civil engineers*. New York, NY: McGraw-Hill.
- Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading, MA: Addison-Wesley.
- DOD. (2002). News Transcript, Presenter: Secretary of Defense Donald H. Rumsfeld, Washington, DC: United States Department of Defense.
- GEO. (1998). *Landslides and Boulder falls from natural terrain: Interim risk guidelines*. Geotechnical Engineering Office Report 75, Government of Hong Kong.
- GEER. (2014). The 22 March 2014 Oso Landslide, Snohomish County, Washington. Contributing Authors: J. R. Keaton, J. Wartman, S. Anderson, J. Benoit, J. deLaChapelle, R. Gilbert, & D. R. Montgomery. Geotechnical Extreme Event Reconnaissance, National Science Foundation.
- Hodges, J. L., Jr., & Lehmann, E. L. (1952). The uses of previous experience in reaching statistical decisions. *Annals of Mathematical Statistics*, 23, 396–407.
- Hurwicz, L. (1951). Some specification problems and applications to econometric models (abstract). *Econometrica*, 19, 343–344.
- IPET. *Performance evaluation of the New Orleans and Southeast Louisiana Hurricane protection system*. (2009). Final Report, Interagency Performance Evaluation Task Force, U.S. Army Corps of Engineers.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4), 620–630.
- Jaynes, E. T. (1968). Prior probabilities. *IEEE Transactions on System Science and Cybernetics*, 4(3), 227–241.
- Journel, A. G., & Deutsch, C. V. (1993). Entropy and spatial disorder. *Mathematical Geology*, 25(3), 329–355.
- Keynes, J. M. (1921). *A treatise on probability*. London: The MacMillan and Company Limited.
- Lacasse, S., & Nadim, F. (2009). Landslide risk assessment and mitigation strategy, Chapter 3. In *Landslides – disaster risk reduction*. Berlin: Springer-Verlag.
- Liu, Z. Q., & Nadim, F. (2014). A three-level framework for multi-risk assessment. In *Geotechnical safety and risk IV* (pp. 493–498). London: Taylor & Francis Group.
- Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York, NY: Wiley.
- Nadim, F. (2011, October 24–26). Risk assessment for earthquake-induced submarine slides. In *5th International Symposium on Submarine Mass Movements and their Consequences, ISSMMTC*. Kyoto, Japan: Springer.
- Nadim, F., & Sparrevik, M. (2013, June 17–19). Managing unlikely risks posed by natural hazards to critical infrastructure. In *22nd SRA Europe Conference*. Trondheim, Norway.
- Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association*, 46(253), 55–67.
- Raiffa, H., & Schlaifer, R. (1961). *Applied statistical decision theory*. Boston, MA: Harvard University Graduate School of Business Administration.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. New York, NY: Random House, Inc.
- Tribus, M. (1969). *Rational descriptions, decisions and designs*. New York, NY: Pergamon Press.
- Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior* (3rd ed.). Princeton, NJ: Princeton University Press.