TERJE AVEN | PIERO BARALDI | ROGER FLAGE | ENRICO ZIO

Uncertainty in Risk Assessment

The Representation and Treatment of Uncertainties by Probabilistic and Non-Probabilistic Methods

WILEY

Contents

Uncertainty in Risk Assessment

The Representation and Treatment of Uncertainties by Probabilistic and Non-Probabilistic Methods

Terje Aven

University of Stavanger, Norway

Piero Baraldi

Politecnico di Milano, Italy

Roger Flage

University of Stavanger, Norway

Enrico Zio

Politecnico di Milano, Italy
Ecole Centrale Paris and Supelec, France

WILEY

This edition first published 2014

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Aven, Terje, author.

Uncertainty in risk assessment : the representation and treatment of uncertainties by probabilistic and non-probabilistic methods / Terje Aven, Piero Baraldi, Roger Flage, Enrico Zio.

pages cm

Includes bibliographical references and index.

ISBN 978-1-118-48958-1 (hardback)

1. Risk assessment–Statistical methods. 2. Probabilities. I. Title.

HD61.A947 2014

338.5–dc23

2013034152

A catalogue record for this book is available from the British Library.

ISBN: 978-1-118-48958-1

Preface

The aim of this book is to critically present the state of knowledge on the treatment of uncertainties in risk assessment for practical decision-making situations concerning high-consequence technologies, for example, nuclear, oil and gas, transport, and so on, and the methods for the representation and characterization of such uncertainties. For more than 30 years, probabilistic frameworks and methods have been used as the basis for risk assessment and uncertainty analysis, but there is a growing concern, partly motivated by newly emerging risks like those related to security, that extensions and advancements are needed to effectively treat the different sources of uncertainty and related forms of information. Alternative approaches for representing uncertainty have been proposed, for example, those based on interval probability, possibility, and evidence theory. It is argued that these approaches provide a more adequate treatment of uncertainty in situations of poor knowledge of the phenomena and scenarios studied in the risk assessment. However, many questions concerning the foundations of these approaches and their use remain unanswered.

In this book, we present a critical review and discussion of methods for the representation and characterization of the uncertainties in risk assessment. Using examples, we demonstrate the applicability of the various methods and point to their strengths and weaknesses in relation to the situation addressed. Today, no authoritative guidance exists on when to use probability and when to use an alternative representation of uncertainty, and we hope that the present book can provide a platform for the development of such guidance. The areas of potential application of the theories and methods studied in the book are broad, ranging from engineering and medicine to environmental impacts and natural disasters, security, and financial risk management. Our main focus, however is, on engineering applications.

The topic of uncertainty representation and characterization is conceptually and mathematically challenging, and much of the existing literature in the field is not easily accessible to engineers and risk analysts. One aim of the present book is to provide a relatively comprehensive state of knowledge, with strong requirements for rigor and precision, while striving for readability by a broad audience of professionals in the field, including researchers and graduate students.

Readers will require some fundamental background in risk assessment, as well as basic knowledge of probability theory and statistics. The goal, however, has been to reduce the dependency on extensive prior knowledge, and key probabilistic and statistical concepts will be introduced and discussed thoroughly in the book.

It is with sincere appreciation that we thank all those who have contributed to the preparation of this book. In particular, we are grateful to Drs. Francesco Cadini, Michele Compare, Jan Terje Kvaløy, Giovanni Lonati, Irina Crenguza Popescu, Ortwin Renn, and Giovanna Ripamonti for contributing the research that has provided the material for many parts of the book, and to Andrea Prestigiacomo for his careful editing work. We also acknowledge the editing and production staff at Wiley for their careful and effective work.

Terje Aven
Roger Flage *Stavanger*
Piero Baraldi *Milano*
Enrico Zio *Paris*

June 2013

Part I

Introduction

1

Introduction

Risk assessment is a methodological framework for determining the nature and extent of the risk associated with an activity. It comprises the following three main steps:

- Identification of relevant sources of risk (threats, hazards, opportunities)
- Cause and consequence analysis, including assessments of exposures and vulnerabilities
- Risk description.

Risk assessment is now widely used in the context of various types of activities as a tool to support decision making in the selection of appropriate protective and mitigating arrangements and measures, as well as in ensuring compliance with requirements set by, for example, regulatory agencies. The basis of risk assessment is the systematic use of analytical methods whose quantification is largely probability based. Common methods used to systematically analyze the causes and consequences of failure configurations and accident scenarios are fault trees and event trees, Markov models, and Bayesian belief networks; statistical methods are used to process the numerical data and make inferences. These modeling methods have been developed to gain knowledge about cause–effect relationships, express the strength of these relationships, characterize the remaining uncertainties, and describe, in quantitative or qualitative form, other properties relevant for risk management (IAEA, 1995; IEC, 1993). In short, risk assessments specify what is at stake, assess the uncertainties of relevant quantities, and produce a risk description which provides information useful for the decision-making process of risk management.

In this book we put the main focus on quantitative risk assessment (QRA), where risk is expressed using an adequate representation of the uncertainties involved. To further develop the methodological framework of risk assessment, we will need to explain in more detail what we mean by risk.

This introductory chapter is organized as follows. Following Section 1.1, which addresses the risk concept, we present in Section 1.2 the main features of probabilistic risk assessment (PRA), which is a QRA based on the use of probability to characterize and represent the uncertainties. Then, in Section 1.3, we discuss the use of risk assessment in decision-making contexts. Section 1.4 considers the issue of uncertainties in risk assessment, motivated by the thesis that if uncertainty cannot be properly treated in risk assessment, the risk assessment tool fails to perform as intended (Aven and Zio, 2011). This section is followed by a discussion on the main challenges of the probability-based approaches to risk assessment, and the associated uncertainty analysis. Alternative approaches for dealing with uncertainty are briefly discussed.

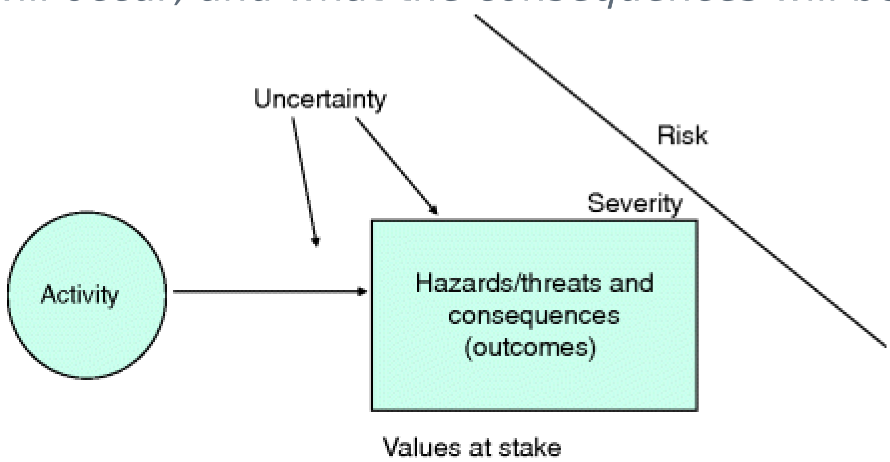
1.1 Risk

1.1.1 The Concept of Risk

In all generality, risk arises wherever there exists a potential source of damage or loss, that is, a hazard (threat), to a target, for example, people, industrial assets, or the environment. Under these conditions, safeguards are typically devised to prevent the occurrence of the hazardous conditions, and protection is put in place to counter and mitigate the associated undesired consequences. The presence of a hazard does not in itself suffice to define a condition of risk; indeed, inherent in the latter there is the uncertainty that the hazard translates from potential to actual damage, bypassing safeguards and protection. In synthesis, the notion of risk involves some kind of loss or damage that might be received by a target and the uncertainty of its transformation in actual loss or damage, see Figure 1.1. Schematically we can write (Kaplan and Garrick, 1981; Zio, 2007; Aven, 2012b)

$$\text{Risk} = \text{Hazards/Threats and Consequences (damage)} + \text{Uncertainty}. \quad (1.1)$$

Figure 1.1 *The concept of risk reflecting hazards/threats and consequences and associated uncertainties (what events will occur, and what the consequences will be).*



Normally, the consequence dimension relates to some type of undesirable outcome (damage, loss, harm). Note that by centering the risk definition around undesirable outcomes, we need to define what is undesirable, and for whom. An outcome could be positive for some stakeholders and negative for others: discussing whether an outcome is classified in the right category may not be worth the effort, and most of the general definitions of risk today allow for both positive and negative outcomes (Aven and Renn, 2009).

Let A denote a hazard/threat, C the associated consequences, and U the uncertainties (will A occur, and what will C be?). The consequences relate to something that humans value (health, the environment, assets, etc.). Using these symbols we can write (1.1) as

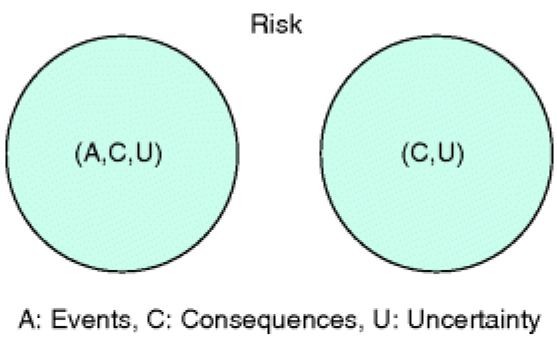
(1.2) $$\text{Risk} = (A, C, U),$$

or simply

(1.3) $$\text{Risk} = (C, U),$$

where C in (C, U) expresses all consequences of the given activity, including the hazardous/threatful events A. These two risk representations are shown in Figure 1.2.

Figure 1.2 *The main components of the concept of risk used in this book.*



Obviously, the concept of risk cannot be limited to one particular measuring device (e.g., probability) if we seek a general risk concept. For the measure introduced, we have to explain precisely what it actually expresses. We also have to clarify the limitations with respect to its ability to measure the uncertainties: is there a need for a supplement to fully describe the risk? We will thoroughly discuss these issues throughout the book.

A concept closely related to risk is vulnerability (given the occurrence of an event A). Conceptually vulnerability is the same as risk, but conditional on the occurrence of an event A:

(1.4)
$$\text{Vulnerability} \mid A = \text{Consequences} + \text{Uncertainty} \mid \text{the occurrence of the event } A,$$

where the symbol | indicates “given” or “conditional.” For short we write

(1.5) $\text{Vulnerability} \mid A = (C, U \mid A).$

1.1.2 Describing/Measuring Risk

The risk concept has been defined above. However, this concept does not give us a tool for assessing and managing risk. For this purpose we must have a way of describing or measuring risk, and the issue is how.

As we have seen, risk has two main dimensions, consequences and uncertainty, and a risk description is obtained by specifying the consequences C and using a description (measure) of the uncertainty, Q. The most common tool is probability P, but others exist and these also will be given due attention in the book. Specifying the consequences means identifying a set of quantities of interest C' that represent the consequences C, for example, the number of fatalities.

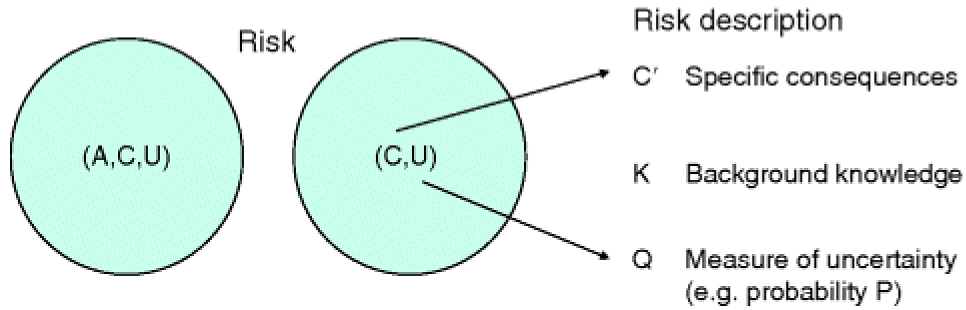
Now, depending on the principles laid down for specifying C' and the choice of [illegible], we obtain different perspectives on how to describe/measure risk. As a general description of risk we can write

$$\text{Risk description} = (C', Q, K),$$

(1.6) $$(\text{or, alternatively}, (A', C', Q, K)),$$

where K is the background knowledge (models and data used, assumptions made, etc.) that and the specification C' are based on, see Figure 1.3. On the basis of the relation between vulnerability and risk previously introduced, the vulnerability given an event is analogously described by $(C', Q, K|A)$.

Figure 1.3 *Illustration of how the risk description is derived from the concept of risk.*



1.1.3 Examples

1.1.3.1 Offshore Oil and Gas Installation

Consider the future operation of an offshore installation for oil and gas processing. We all agree that there is some “risk” associated with this operation. For example, fires and explosions could occur leading to fatalities, oil spills, economic losses, and so on. Today we do not know if these events will occur and what the specific consequences will be: we are faced with uncertainties and, thus, risk. Risk is two dimensional, comprising events and consequences, and associated uncertainties (i.e., the events and consequences being unknown, the occurrences of the events are not known and the consequences are not known).

When performing a risk assessment we describe and/or quantify risk, that is, we specify (C', Q, K). For this purpose we need quantities representing C' and a measure of uncertainty; for the latter, probability is introduced. Then, in the example discussed, C' is represented by the number of fatalities, $Q=P$, and the background knowledge covers a number of assumptions that the assessment is based on, for example, related to the number of people working on the installation, as well as the models and data used for quantification of the accident probabilities and consequences. On this basis, several risk indices or metrics are defined, such as the expected number of fatalities (e.g., potential loss of lives, PLL, typically defined for a one-year period) and the fatal accident rate (FAR, associated with 100 million exposed hours), the probability that a specific person will be killed in an accident (individual risk, IR), and frequency–consequence (*f*–*n*) curves expressing the expected number of accidents (frequency f) with at least *n* fatalities.

1.1.3.2 Health Risk

Consider a person's life and focus on the condition of his/her health. Suppose that the person is 40 years old and we are concerned about the "health risk" for this person for a predetermined period of time or for the rest of his/her life. The consequences of interest in this case arise from "scenarios" of possible specific diseases (known or unknown types) and other illnesses, their times of development, and their effects on the person (will he/she die, suffer, etc.).

To describe risk in this case we introduce the frequentist probability p that the person gets a specific disease (interpreted as the fraction of persons that get the disease in an infinite population of "similar persons"), and use data from a sample of "similar persons" to infer an estimate p^* of p. The probability p can be considered a parameter of a binomial probability model.

For the consequent characterization, C', we look at the occurrence or not of a disease for the specific person considered, and the time of occurrence of the disease, if it occurs. In addition, we have introduced a probability model with a parameter p and this p also should be viewed as a quantity of interest C'. We seek to determine p, but there are uncertainties about p and we may use confidence intervals to describe this uncertainty, that is, to describe the stochastic variation in the data.

The uncertainty measure in this case is limited to frequentist probabilities. It is based on a traditional statistical approach. Alternatively, we could have used a Bayesian analysis based on subjective (judgmental, knowledge-based) probabilities P (we will return to the meaning of these probabilities in Chapter 2). The uncertainty description in this case may include a probability distribution of p, for example, expressed by the cumulative distribution function $F(p') = P(p \leq p')$. Using to measure the uncertainties (i.e., $Q = P$), we obtain a risk description (C', P, K), where p is a part of C'. From the distribution $F(p')$ we can derive the unconditional probability $P(A)$ (more precisely, $P(A|K)$) of the event *A* that the person gets the disease, by conditioning on the true value of p (see also Section 2.4):

(1.7) $$P(A) = \int P(A \mid p') \, \mathrm{d}F(p') = \int p' \, \mathrm{d}F(p').$$

This probability is a subjective probability, based on the probability distribution of the frequentist probability p. We see that $P(A)$ is given by the center of gravity (the expected value) of the distribution F.

Alternatively, we could have made a direct subjective probability assignment for $P(A) = P(A \mid K)$, without introducing the probability model and the parameter p.

1.2 Probabilistic Risk Assessment

Since the mid-1970s, the framework of probability theory has been the basis for the analytic process of risk assessment (NRC, 1975); see the reviews by Rechard (1999, 2000). A probabilistic risk assessment (PRA) systematizes the knowledge and uncertainties about the phenomena studied: what are the possible hazards and threats, their causes and consequences? The knowledge and uncertainties are characterized and described using various probability-based metrics, as illustrated in Section 1.1.3; see also Jonkman, van Gelder, and Vrijling (2003) for a comprehensive overview of risk metrics (indices) for loss of life and economic damage. Additional examples will be provided in Chapter 3, in association with some of the detailed modeling and tools typical of PRA.

A total PRA for a system comprises the following stages:

1. *Identification of threats/hazards.* As a basis for this activity an analysis of the system is carried out in order to understand how the system works, so that departures from normal, successful operation can be identified. A first list of hazards/threats is normally identified based on this system analysis, as well as on experience from similar types of analyses, statistics, brainstorming activities, and specific tools such as failure mode and effect analysis (FMEA) and hazards and operability (HAZOP) studies.

2. *Cause analysis.* In cause analysis, we study the system to identify the conditions needed for the hazards/threats to occur. What are the causal factors? Several techniques exist for this purpose, from brainstorming sessions to the use of fault tree analyses and Bayesian networks.

3. *Consequence analysis.* For each identified hazard/threat, an analysis is carried out addressing the possible consequences the event can lead to. Consequence analysis deals to a large extent with the understanding of physical phenomena, for example, fires and explosions, and various types of models of the phenomena are used. These models may for instance be used for answering questions like: How will a fire develop? What will be the heat at various distances? What will the explosive pressure be in case an explosion takes place? And so on. Event tree analysis is a common method for analyzing the scenarios that can develop in the different consequences. The number of steps in the sequence of events that form a scenario is mainly dependent on the number of protective barriers set up in the system to counteract the initiating event of that sequence. The aim of the consequence-reducing barriers is to prevent the initiating events from resulting in serious consequences. For each of these barriers, we can carry out failure analysis to study their reliability and effectiveness. Fault tree analysis is a technique often used for this purpose.

4. *Probabilistic analysis.* The previous stages of analysis provide a set of sequences of events (scenarios), which lead to different consequences. This specification of scenarios does not address the question of how likely the different scenarios and the associated consequences are. Some scenarios could be very serious, should they occur, but if the likelihood of their occurrence is low, they are not so critical. Using probability models to reflect variation in the phenomena studied and assigning probabilities for the occurrence of the various events identified and analyzed in steps 2 and 3, overall probability values and expected consequence values can be computed.

5. *Risk description.* Based on the cause analysis, consequence analysis, and probabilistic analysis, risk descriptions can be obtained using various metrics, for example, risk matrices showing the computed/assigned probability of a hazard/threat and the expected consequences given that this event has occurred, as well as IR, PLL, and FAR values.

6. *Risk evaluation.* The results of the risk analysis are compared to predefined criteria, for example, risk tolerability limits or risk acceptance criteria.

PRA methodology is nowadays used extensively in industries such as nuclear power generation (e.g., Vesely and Apostolakis, 1999; Apostolakis, 2004), offshore petroleum activities (e.g., Falck, Skramstad, and Berg, 2000;

Vinnem, 2007), and air transport (e.g., Netjasov and Janic, 2008).

The current default approach to a comprehensive quantitative PRA is based on the so-called set of triplets definition of risk, introduced by Kaplan and Garrick (1981); see also Kaplan (1992, 1997). In this approach, risk is defined as the combination of possible scenarios , resulting consequences x, and the associated likelihoods l. Loosely speaking: What can happen (go wrong)? How likely is it? What are the consequences? Within this conceptual framework, three main likelihood settings are often defined (Kaplan, 1997): repetitive situation with known frequency ($l=f$, where f is a known frequentist probability), unique situation ($l=p$, where p is a subjective probability), and repetitive situation with unknown frequency ($l=H(f)$, where H is a subjective probability distribution on an unknown/uncertain frequentist probability f). Of course, the first case is a special case of the third. The last-mentioned setting is typically dealt with using the so-called probability of frequency approach, where all potentially occurring events involved are assumed to have uncertain frequency probabilities of occurrence, and the epistemic uncertainties about the true values of frequency probabilities are described using subjective probabilities. For the sake of simplicity, in the following we will often use the short term “frequency” instead of “frequentist probability.”

The probability of frequency approach is in line with the standard Bayesian approach (Aven, 2012a) as will be described below. It is also considered “the most general and by far the most powerful and useful idea” by Kaplan (1997, p. 409), and corresponds to the highest level of sophistication in the treatment of uncertainties in risk analysis according to the classification by Paté-Cornell (1996).

In this book, however, we adopt a broader perspective of risk by which the set of triplets is not *risk* per se but a *risk description*. In this view, the outcome of a risk assessment is a list of scenarios quantified in terms of probabilities and consequences, which collectively describe the risk. As we will thoroughly discuss throughout the book, this risk description will be shown to be more or less adequate for describing the risk and uncertainties in different situations.

Numerous textbooks deal with methods and models for PRA, for example, Andrews and Moss (2002), Aven (2008), Cox (2002), Vinnem (2007), Vose (2008), and Zio (2007, 2009). Some also deal specifically with foundational issues, in particular with the concepts of uncertainty and probability, for example, Aven (2012a), Bedford and Cooke (2001), and Singpurwalla (2006).

In spite of the maturity reached by the methodologies used in PRA, a number of new and improved methods have been developed in recent years to meet the needs of the analysis brought about by the increasing complexity of the systems and processes studied, and to respond to the introduction of new technological systems. Many of the methods introduced allow for increased levels of detail and precision in the modeling of phenomena and processes within an integrated framework of analysis covering physical phenomena, human and organizational factors, and software dynamics (e.g., Mohaghegh, Kazemi, and Mosleh, 2009). Other methods are devoted to the improved representation and assessment of risk and uncertainty. Examples of more recently developed methods are Bayesian belief networks, binary digit diagrams, multi-state reliability analysis, and advanced Monte Carlo simulation tools. For a summary and discussion of some of these models and techniques, see Bedford and Cooke (2001) and Zio (2009).

The probabilistic analysis underpinning PRA is based on one or the other of two alternative conceptual foundations:

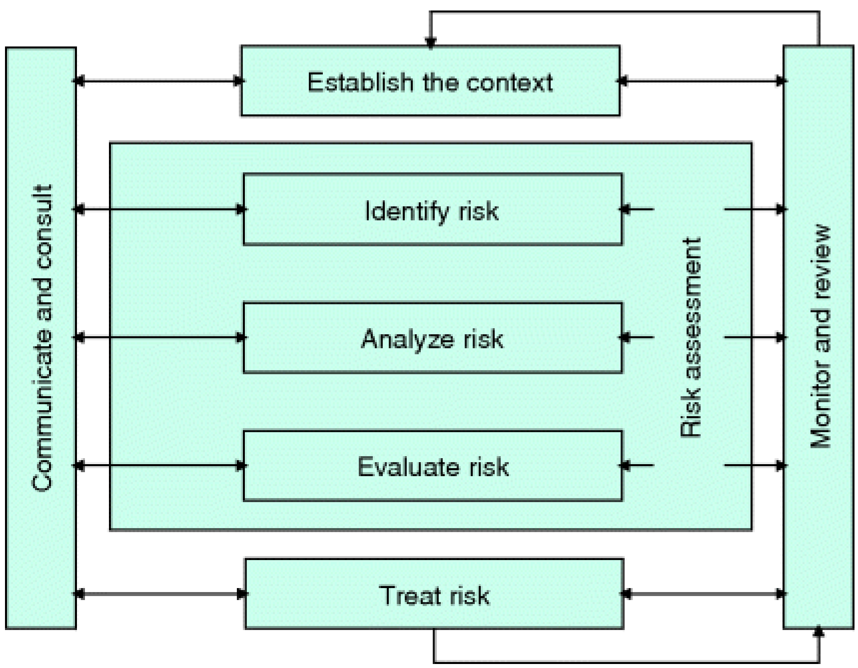
the traditional frequentist approach and the Bayesian approach (Bedford and Cooke, 2001; Aven, 2012a). The former is typically applied in situations in which there exists a large amount of relevant data; it is founded on well-known principles of statistical inference, the use of probability models, the interpretation of probabilities as relative frequencies, point estimates, confidence interval estimation, and hypothesis testing.

By contrast, the Bayesian approach is based on the concept of subjective (judgmental, knowledge-based) probabilities and is applied in situations in which there exists only a limited amount of data (e.g., Guikema and Paté-Cornell, 2004). The idea is to first establish probability models that adequately represent the aleatory uncertainties, that is, the inherent variability of the phenomena studied, such as the distribution of lifetimes of a type of system. The epistemic uncertainties, reflecting incomplete knowledge or lack of knowledge about the values of the parameters of the models, are then represented by prior subjective probability distributions. When new data on the phenomena studied becomes available, Bayes' formula is used to update the representation of the epistemic uncertainties in terms of the posterior distributions. Finally, the predictive distributions of the quantities of interest – the observables (e.g., the lifetime of new systems) – are derived by applying the law of total probability. The predictive distributions are epistemic statements, but they also reflect the inherent variability of the phenomena being studied, that is, the aleatory uncertainties.

1.3 Use of Risk Assessment: The Risk Management and Decision-Making Context

Risk management can be defined as the coordinated activities to direct and control an organization with regard to risk (ISO, 2009). As illustrated in Figure 1.4, the main central steps of the risk management process are: establishment of the context, risk assessment, and risk treatment. Context here refer to the internal and external environment of the organization, the interface of these environments, the purpose of the risk management activity, and suitable risk criteria. Risk treatment is the process of modifying risk, which may involve avoiding, modifying, sharing or retaining risk (ISO, 2009).

Figure 1.4 *The risk management process (based on ISO, 2009).*



Note that, according to ISO (2009), source (hazard/threat/opportunity) identification is not included as part of risk analysis. Many analysts and researchers do