Chapter 2
Introduction to Engineering Risk Analysis

Overview of Risk Analysis for Engineered Systems

Can contemporary organizations and societies design, build, and operate complex engineering systems safely and reliably for long periods? Being able to do so is crucial if nuclear power is to be a viable option, if infrastructure such as advanced transportation systems and energy distribution networks is to be trustworthy, if minerals and petroleum are to be discovered and extracted safely, and if hazardous manufacturing and chemical storage facilities are to be located in convenient proximity to transportation hubs and population centers. This chapter, which is an update and extension of Bier and Cox (2007), discusses methods for quantifying the extent to which complex engineering systems can be designed and operated safely.

Opinions about the answer are divided. One school of thought, sometimes called *Normal Accident Theory* after the book that articulated it (Perrow, 1984), holds that engineered systems with high "interactive complexity" (presenting unexpected and surprising sequences of events that are difficult to detect or comprehend at first) and "tight coupling" of interdependent components or subsystems (so that changes propagate quickly among them) are inherently unpredictable and uncontrollable by human operators. Resulting accidents and catastrophic failures in such high-risk technological systems are seen as inevitable and unavoidable: in this sense, they are "normal." In this pessimistic view, adding redundancy to complex systems to reduce accidents makes them even more complex and prone to unpredictable failures. Case studies of accidents and near-accidents at chemical plants, nuclear reactors, airports, and other complex industrial facilities well illustrate Normal Accident Theory.

A different view, popularized in the catchphrase "Failure is not an option" (made famous by Oscar-nominated actor Ed Harris playing NASA Flight Director Gene Kranz in the movie *Apollo 13*), is that complex engineering systems *can* be built and operated safety by sufficiently disciplined, creative, well-organized, and well-trained teams and organizations. Sociologists, psychologists, and other researchers have sought common features of "high-reliability organizations" (HROs), meaning organizations with significantly fewer accidents and failures than normally expected. They have proposed lists such as preoccupation with failure, reluctance

L.A. Cox, Jr., *Risk Analysis of Complex and Uncertain Systems*,
International Series in Operations Research & Management Science 129,
DOI 10.1007/978-0-387-89014-2_2,

to (over-)simplify interpretations or to prematurely reject unexpected interpretations of observations, sensitivity to operations, commitment to resilience, and appropriate deference to expertise (as opposed to rank, seniority, or power) (Weick and Sutcliffe, 2001). Such habits can help to create the vigilant "mindfulness" needed to operate safely and to catch and correct potential problems before they cascade out of control. Routinely safe operations on aircraft carriers and in other high-stress, risky, and complex environments vividly illustrate that well-trained teams in well-designed organizations and environments can manage risks successfully. Similar principles might be applied in different settings, such as operating rooms and intensive care units. Eliminating "mindlessness" (e.g., blind rule following or deference) in the implementation of federally and locally funded programs to reduce infant mortality and preterm birth has been proposed as a way to reduce the frequency and severity of poor outcomes (Issel and Narasimha, 2007).

Probabilistic risk assessment (PRA) of engineered systems gives engineers and risk managers practical tools to understand, predict, and manage risks for a variety of complex engineered systems. It identifies how systems might fail, the likely (and not-so-likely) potential adverse consequences of failures, and how best to prevent failures and mitigate adverse consequences while meeting other goals, such as the continued productive operation of a hazardous facility. PRA methods include probability modeling techniques (both analytic and simulation-based) for quantifying engineering risks, typically expressed as the probabilities of adverse events and as the frequencies and severities of their adverse consequences over a stated period of time. PRA also includes optimization methods from operations research and safety and reliability engineering that can identify cost-effective ways to improve safety and reliability while satisfying other constraints (e.g., on system cost, weight, or performance).

Examples of complex engineering systems to which PRA has been successfully applied include nuclear power plants (beginning with the Reactor Safety Study in 1975, and continuing to the present day); the space shuttle (both before and especially after the Challenger disaster); dam and reservoir planning and operations; highway, bridge, and transportation infrastructure; emergency planning; liquefied natural gas (LNG) terminals and storage facilities; other hazardous chemical plants and operations; and electric power generation and distribution planning. The common elements in such systems is that they all involve (1) a *designed system* intended to withstand different levels of stress, with the option of incorporating different levels of backup and fail-safe design, (2) a *system operator/risk manager* faced with decisions about how to inspect, maintain, and use the system (e.g., when to launch, when to shut down, and generally what level of precaution to adopt), and (3) an uncertain *environment* that generates stresses and adverse conditions that the system should ideally be able to withstand. Uncertainties from the environment may involve random events, such as equipment failures or unexpectedly high or stressful transient loads (as in the case of the Tacoma Narrows bridge collapse); natural disasters such as earthquakes, floods, or hurricanes; terrorist attacks; or operator errors, perhaps arising from miscommunication, miscoordination, or misunderstanding of systems behavior among those running it. Unexpected behaviors of interacting software modules or other subsystems may also cause a system to fail, even if each component performs as it was designed to (Leveson, 2004).

Example: Unreliable Communication with Reliable Components

Setting: Suppose that two people call each other on their cell phones at the same time, so that each receives a busy signal. If each caller can reattempt a call at the beginning of each new time period (say, every 20 seconds), then what retry strategy minimizes the average time to connect?

Solution: If each caller tries to call with probability p at the start of a period, then the probability that they conflict again is p^2, the probability that neither calls is $(1-p)^2$, and the probability that they connect is $2p(1-p)$, which is maximized for $p = 0.5$. Thus, each should call with probability 0.5 at the start of each period, until they connect. The probability of connecting at the start of a period is then $2^*(0.5)^*(1-0.5) = 0.5$ and the expected number of periods until connection is established is therefore $1/0.5 = 2$. Thus, even if all parts of the system work perfectly and the two callers behave optimally (given what they know), the effective availability of a connection to the callers is less than 100%. This illustrates a *coordination failure* in the use of the system. Of course, such coordination failures are usually only minor annoyances. . . unless the safety of a system depends on being able to establish contact promptly when something goes wrong!

A system's designer and its operator/risk manager usually both want to make decisions so that the system operates as planned, given engineering and cost constraints and the uncertain environment. Of course, the decisions and trade-offs faced by the operator/risk manager typically reflect the decisions made by the system designer. PRA can help to quantify the trade-offs between cost and safety at the design stage and can help to identify policies and schedules for cost-effective inspection and testing, preventive maintenance, spare parts provisioning, redundancy allocation, and replacement of working parts to keep complex systems operating as intended throughout their design lives.

Example: Optimal Number of Redundant Components

Setting: Suppose that an airplane can have one, two, or four engines. Each engine independently has a probability $1-p$ of failing during the course of a mission. (Equivalently, it has probability p of surviving). A plane fails (crashes) if more than half of its engines fail.

Problem: What number of engines should a plane have, to maximize the probability of completing its mission?

Solution: A plane with one engine has success probability p. A plane with two engines has success probability $1-(1-p)^2$, the probability that both engines do not fail. Since $1-(1-p)^2 = 1-(1-2p+p^2) = p(2-p)$, this success probability is greater than p if and only if $2-p > 1$ and $p > 0$; in other words, for $0 < p < 1$. Thus, *a twin-engine plane is at least as likely to survive as a single-engine plane*, with equality only if $p = 1$ or $p = 0$. For a four-engine plane, the probability of success is one minus the probability of losing more than two engines: $1 - [\Pr(\text{lose 4 engines}) + \Pr(\text{lose 3 engines})] = 1 - [(1-p)^4 + 3p(1-p)^3] = 1-(1-p)^3\,[(1-p)+3p]$.

This is greater than the success probability for two engines if and only if the following inequalities hold:

$$
\begin{aligned}
&1 - (1-p)^3[(1-p) + 3p] > 1 - (1-p)^2,\\
&(1-p)^2 > (1-p)^3[(1-p) + 3p],\\
&1 > (1-p)[(1-p) + 3p],\\
&1 > (1-p)(1+2p),\\
&1 > 1 + p - 2p^2,\\
&2p^2 > p,\\
&2p > 1,\\
&p > 0.5.
\end{aligned}
$$

Thus, a four-engine plane is more likely to survive than a two-engine plane if and only if the individual engines are more likely than not to survive the mission ($p > 0.5$).

Example: Optimal Scheduling of Risky Inspections

Setting: Suppose that, in the absence of intervention, a component (perhaps an engine in the previous example) of age t has probability $1 - e^{-ht}$ of a defect that will increase the risk of failure when the component is next used. h is called the *hazard rate* for occurrence of the defect, and ht is the *cumulative hazard* accumulated by age t in the absence of intervention. At any time, an expensive inspection may be performed, and, if a defect is present, it will be found and repaired, effectively setting the age of the component back to 0. However, careless inspection may itself introduce an uncorrected defect that would not otherwise have occurred. The probability of this is $p \geq 0$ for each inspection.

Problem: What time between inspections minimizes the expected number of uncorrected defects per unit time?

Solution: If inspections take place every T time units, then each inspection removes hT expected defects and adds p expected defects. The optimal time between inspections makes the marginal "cost" (here meaning loss of reliability) from an inspection – that is, the expected new defects created, p – equal to its marginal benefit (that is, the expected effects removed, hT). Thus, the optimal time between inspections, denoted by T^*, satisfies $hT^* = p$, and so $T^* = p/h$ (for $h > 0$). More frequent inspections than this, with $T < T^*$, are expected to create more problems than they solve ($p > hT$). Less frequent inspections, with $T > T^*$, let the expected costs of not intervening sooner exceed the costs of doing so ($hT > p$).

PRA is usually applied to rare and catastrophic events for which it may be difficult to estimate risks directly due to the lack of empirical data, the possibility of unobserved changes (e.g., deterioration) in the system, and changes in the system's environment or use. Risk assessment can also be applied to predict routine

(e.g., occupational accident) risks, although in such cases it may be possible to rely primarily on empirical data, reducing the need for modeling. In general, PRA is used to estimate, predict, and find ways to reduce the risks to facility or system owners, employees, and the public. This chapter focuses on methodological advances in engineering risk analysis, with selected applications (including some applications of PRA methods and insights to fields other than engineering) to illustrate the methodology.

Using Risk Analysis to Improve Decisions

Risk analysis can help to inform design decisions (e.g., trade-offs among safety and performance, cost, etc.) as well as operational decisions (e.g., when to shut down a facility). It can be useful regardless of who makes the decisions – for example, facility owners and operators, regulators, or multiple stakeholders interacting through a participatory risk management and conflict-resolution process. Key technical challenges that PRA must address include: how to predict the probable performance and quantify the behaviors – both probable and improbable – of a complex system, given a design and the operator's decisions, in the face of inadequate data; how to optimize the joint decisions faced by the system designer and owner/operator (which can involve NP-hard combinatorial optimization problems, as well as problems of coordination and communication between different organizations); how to most effectively model interdependencies and uncertainties about the system's current state; the development of cost-effective "screening"-type methods for addressing the myriad possible risks in "open" systems (such as the risk of terrorist attack); and scale-up problems for extremely complex systems, such as infrastructure networks. Also, there is still room to benefit more fully from adaptation of methods developed in other fields, including decision analysis and related fields (such as Bayesian statistics).

Hazard Identification: What Should We Worry About?

Probabilistic risk assessment typically begins by defining a system to be analyzed and identifying undesired outcomes that might occur when it is operated. *Hazard identification* methods have been developed to identify the potential adverse consequences of system operation. Structured qualitative techniques include hazard and operability (HAZOP) studies and failure modes and effects analysis (FMEA), which describes potential failure modes, causes, effects, safeguards, and recommendations for reducing risks.

Fault trees and event trees can be used in a qualitative mode for hazard identification but can also be quantified to estimate the likelihood of adverse events. *Fault tree analysis* (Barlow, 1998) begins with an undesired outcome, called the "top event," and reasons backward to identify which combinations of more basic events

(e.g., component failures) could bring about the top event (e.g., failure of the system). The result is a tree that represents those sets of basic events that would be sufficient to cause the top event using "AND" and "OR" logic (and possibly more complicated logic gates as well). The tree generally goes down to the level of basic events whose probabilities can be reliably estimated from experience, judgment, and.or data.

Example: Fault Tree Calculations for Car Accidents at an Intersection

Setting: Suppose that a car accident (the top event) occurs at an intersection if and only if (two cars approach the intersection at the same time from different directions) AND (both cars proceed). The event "both cars proceed" can be further decomposed into a logical subtree, as follows: (both cars proceed) if and only if [(the signal is broken AND both cars proceed) OR (the signal is not broken AND both cars proceed)]. Reliable statistics show that the first event (sometimes called the *imitating event*), namely, "Two cars approach the intersection at the same time from different directions," occurs with an average annual frequency of 100 times per year. The signal is broken on 0.1% of these occasions (independently of traffic) and the conditional probability that both cars will proceed, following the initiating event, is 0.1 if the signal is broken and 0.01 if it is not broken.

Problem: (a) What is the average annual frequency of accidents at the intersection, given these numbers? (b) What fraction of accidents would be prevented if the signal never failed?

Solution: (a) The conditional probability of an accident, given the initiating event, is Pr(signal is broken)*Pr(both cars proceed | signal is broken) + Pr(signal is not broken)*Pr(both cars proceed | signal is not broken) = (0.1%)*(0.1) + (1 – 0.1%)∗(0.01) = 0.0001 + 0.9999*0.01 = 0.0101 (to four significant digits). (Here "|" is read as "given" or "conditioned on.") The average annual frequency of accidents is this conditional probability times the average annual frequency of initiating events: 0.0101*100 = 1.01 accidents per year. (b) The contribution of accidents with a broken signal to the total average annual frequency of accidents is only (0.1%)*(0.1)*100 = 0.01 accidents per year. If the signal were never broken, then the average frequency of accidents per year would still be 100*0.01 = 1 accident per year.

Comments: (a) *Dominant contributors.* In this example, accidents with the traffic signal working constitute a *dominant contributor* to the average annual accident frequency. This means that ignoring other, rarer events (namely, accidents with the signal broken) yields the same calculated risk number (about one expected accident per year), to one significant digit. One way to simplify fault tree calculations is to focus on dominant contributors, neglecting events that are rare enough that they do not change the numerical answer (within some desired level of precision, such as one or two significant digits). (b) *Poisson probabilities.* The calcu-

lated risk of about one accident per year can be viewed as the mean of a rare-event (approximately Poisson) process. This allows the probabilities for any number of accidents per year (under current conditions) to be estimated: It is Pr(x accidents in a year) $= \lambda\exp(-\lambda)/x!$ for $x = 0, 1, 2, \ldots$, where λ is the mean number of accidents per year (approximately 1, in this example). For example, the probability of zero accidents at this intersection in a year, if the accident process is Poisson with mean 1 accident per year, is $e^{-1} = 1/2.718 = 0.368$. (c) *Obtaining probabilities for basic events.* If reliable statistics were not available for the probabilities that both cars proceed when the signal is working and when it is broken, they might be estimated from experiments (e.g., using driving simulator results), models of driver behavior, or expert judgment. Uncertainty and sensitivity analyses would then typically be used to determine by how much the calculated risk might change if different plausible estimates or better future information about these inputs were to be used in the analysis. (d) *Recursive deepening of a tree.* Each event in a model, such as "both cars proceed," can potentially be expressed as a subtree consisting of a logical combination of more refined event descriptions, e.g., "(both cars proceed and weather is good) or (both cars proceed and weather is not good)." Infinite recursion is prevented by stopping further decomposition when the current description allows basic event probabilities to be quantified accurately enough to support risk management decisions.

Event tree analysis begins with an "initiating event" and works forward to identify its potential consequences. In essence, an event tree is a decision tree without decision nodes. It shows potential sequences of events, with the probability of each branch leaving an event node (representing the possible resolution of an uncertainty, often modeled as a possible value of a random variable) being conditionally independent of earlier information, given that the branch point (i.e., that event node) has been reached. The frequency of a given event sequence is then just the product of the conditional branch probabilities along that path multiplied by the frequency of the initiating event. Both fault trees and event trees can be represented as logically equivalent influence diagrams. They can be solved by more general-purpose influence diagram algorithms (Barlow, 1998; Bobbio et al., 2001).

Structuring Risk Quantification and Displaying Results: Models for Accident Probabilities and Consequences

A *quantitative risk model* typically consists of a formal mathematical and or simulation model of the system of interest, together with one or more consequence attributes of interest and one or more alternative risk management decisions to be evaluated or decision variables to be optimized. The model is used to predict the probable consequences of alternative decisions. Preferred decisions are those that yield preferred probability distributions (or, more generally, preferred stochastic processes) for the consequences of interest.

Risk modeling typically involves some or all of the following components.

- *System representation* (Barlow, 1998; Smith, 2005). An engineered system is often represented mathematically in one of the following forms: (a) A "black-box" *statistical model* (e.g., a lifetime hazard function quantifying the conditional failure rate of a system for different ages or elapsed times, given that it has not failed so far): (b) component failure rates combined via a *coherent structure function* (such as a fault tree or an event tree) mapping the states of system components to the states of the system. (A coherent structure function must be monotonically increasing, going from a system failure probability of zero if all components work to a system failure probability of one if all components fail.); (c) a stochastic *state-transition* model (e.g., a Markov or semi-Markov model for transitions among working and failed components, representing component failure and repair rates); (d) a discrete-event *simulation model* (Smith, 2005).
- *Environment representation*. Like a system model, a model of the environment may be a statistical black-box model (e.g., a function describing the frequency and intensity of stresses to the system's components), a stochastic process, or a simulation model. Plausible worst-case or bounding scenario analyses are sometimes used when probabilistic descriptions of uncertainty are unavailable or are difficult to obtain. The model of the environment is often incorporated directly into the system model, as with traffic levels and weather conditions in a traffic accident model.
- *Decision-rule representation.* A *decision rule* for managing an engineered system maps observed information about the system into a resulting action or intervention. For example, a component may be replaced based on the observed history of failures and repairs for its components. Optimization methods, including recently developed simulation-optimization techniques (see, for example, Ólafsson and Kim, 2002), can help to identify "good" or "best" decision rules, given a system model, an objective function (e.g., a multiattribute utility function), and a model of the environment. Of course, many decisions in the real world (even when informed by PRA) are made without a formal decision rule, either because the PRA results themselves make the best decision clear or because of the need to address the concerns of multiple stakeholders.

Example: Bug-Counting Models of Software Reliability

An example of a simple black-box risk model for software reliability is a "bug-counting" model in which the (unknown) initial number of bugs in a piece of code is represented by a random variable N with a prior distribution. As the code is tested and debugged, the remaining number of bugs presumably decreases, and the random times between successive bug discoveries stochastically increase. (Relatively sophisticated models also allow for the possibilities that detection and repair are imperfect processes and that debugging activities may introduce new bugs.) The

empirical record of bug discoveries can be used to trigger a decision rule such as "If no bugs have been discovered within M tester-hours, then release the software." Simulation optimization can then be used to numerically optimize the parameter M. For analytic alternatives, see Singpurwalla and Wilson (1999) and Wilson and Samaniego (2002).

Example: Risk Management Decision Rules for Dams and Reservoirs

Wurbs (2005) describes the use of decision rules to manage water releases for dams and reservoirs as follows:

> Release decisions depend upon whether or not the flood control storage capacity is exceeded . . . federal reservoirs are typically sized to contain at least a 50-year recurrence interval . . . flood and, for many projects, design floods greater than the 100-year flood . . . , perhaps much greater. A specified set of rules, based on downstream flow rates, are followed as long as sufficient storage capacity is available to handle the flood without having to deal with the water surface rising above the top of the flood control pool. . . . For extreme flood events which would exceed the reservoir storage capacity, moderately high damaging discharge rates beginning before the flood control pool is full are considered preferable to waiting until a full reservoir necessitates much higher release rates.

The outputs from quantitative risk models are often summarized as F–N curves (also sometimes called exceedance probability curves, or complementary cumulative frequency distributions), showing the expected annual frequency F of fatalities or damages exceeding any given level, N, for $N > 0$. (Technically, as discussed in Chapter 5, such diagrams make sense only for compound Poisson processes, not for more general renewal processes. However, F–N curves are often used to summarize the results of PRA calculations, which typically use compound-Poisson approximations to risk in any case.) F–N curves are not perfect summaries of the distribution of risk within a population, however – largely because they do not describe individual risks, which may differ substantially. Other risk displays show how risk varies by location, over time, and with other covariates. For example, it is common practice to plot "risk contours" showing risks to individuals at different locations around a potentially hazardous installation or transportation route.

Example: Different Individual Risks for the Same Exceedance Probability Curve

Suppose that three people, 1, 2, and 3, live near two hazardous facilities, A and B. Facility A can have any of three accidents: A small accident that kills individual 1 only; a medium-sized accident that kills individuals 1 and 2; or a large accident that kills individuals 1, 2, and 3. If an accident occurs at facility A, it is equally likely to be small, medium, or large. By contrast, an accident at facility B is equally likely to kill individual 3 only, kill individuals 1 and 2, or kill all three. Accidents at facilities A and B are equally frequent. Then A and B have identical F-N curves, since each

accident (at either facility) has probability 1/3 of causing one fatality, 1/3 of causing two fatalities, and 1/3 of causing three fatalities. But the individual risks from the two facilities are very different. An accident at facility *A* has a 100% probability of killing individual 1, a 2/3 probability of killing individual 2, and only a 1/3 probability of killing individual 3; but an accident at facility *B* has a 2/3 probability of killing each individual. This difference in the distribution of individual risks is not captured in an *F-N* curve, but could be shown in a risk contour plot if the three individuals are positioned at different locations.

Major technical challenges for developing PRA results include

1. *Constructing and validating models* of the system and its environment. Statistical analysis of accident precursors uses data on "near-misses" to validate and refine model-based predictions (Yi and Bier, 1998; Borgonovo et al., 2000; Phimister et al., 2004). Powerful model-building and model-checking methods have also been developed in the areas of *system identification*, which attempts to identify dynamic system descriptions of input-output relations from observed time course data (see Chapter 11), and *data mining and machine learning*, which seek to learn correct models (or at least subsets of especially plausible models) directly from data (see Chapters 6 and 7).
2. *Calculating, simulating, or estimating probabilities of rare events.* Methods for addressing this challenge, such as importance sampling, adaptive importance sampling, cross-entropy, and Markov chain Monte Carlo (MCMC) methods with carefully designed transition kernels, have advanced significantly in recent years (e.g., Bucklew, 2004; Rubinstein et al., 2004).
3. *Treatment of dependencies* among failure events and system components. Methods for treatment of dependencies presently include common-cause failure analysis (to show dependence in the failure rates of similar components due to a common underlying cause), dependency matrices and event trees (to show the dependence of some systems on "support" systems such as electric power), and external-events analysis (to capture the fact that events such as earthquakes, fires, and floods can affect multiple components of a system).

Quantifying Model Components and Inputs

A model typically expresses risk (e.g., the probability of failure by a certain time) as a function of the performance of model components and or input parameters. These must be quantified from available data, perhaps using a combination of expert judgment and Bayesian statistics (due to the sparseness of directly relevant data). In Bayesian statistics, a prior distribution is updated by conditioning on observed data to yield a posterior probability distribution for the quantities of interest (Lee, 2004). Such methods include hierarchical Bayesian methods (in which partially relevant data are used to help construct the prior distribution) as well as empirical Bayesian methods (in which the actual data for the problem at hand are used to help construct the prior distribution); see Carlin and Louis (2000).

Although Bayesian approaches to quantifying risk models are frequently applied in practice, advances are still being made in numerous areas. These include designing more flexible and tractable models for treating probabilistic dependence in risk models, alternatives to relying on subjective prior distributions (which can be problematic if plausible differences in subjective priors significantly affect risk results), and treatment of model uncertainty.

Modeling Interdependent Inputs and Events

If the state of a system is described by a coherent structure function, and each component independently undergoes stochastic transitions over time (e.g., from "working" to "failed" to "repaired" or "replaced"), then the probability distribution for the system's state (i.e., the probability that it will be working rather than failed at any time) can be obtained relatively easily. Stochastic simulation of the behaviors of the components, or the routine application of combinatorial reliability models and algorithms, such as fault tree analysis or event tree analysis, is practical even for large systems. However, if component behaviors are interdependent (e.g., if each component failure increases the stress on those components that have not yet failed), then it becomes more complex to calculate the risk that the system will have failed by any given time. Simulating interdependent behaviors may be straightforward in principle, but, in practice, it requires specifying how events depend on each other – a potential combinatorial nightmare.

Dependence can also be a problem for uncertainty analysis. In particular, the failure rates (or probabilities) of the various components can be uncertain and statistically dependent on each other, even if their behaviors are conditionally independent given their failure rates. For example, learning that one component had a higher failure rate than expected may cause one to increase estimates of the failure rates of other similar components. The failure to take such dependence into account can result in substantial underestimation of the uncertainty about the overall system failure rate (or probability), and in some cases also underestimation of the mean failure probability of the system (e.g., if the components whose failure probabilities are dependent are functionally in parallel with each other); see Apostolakis and Kaplan (1981), Burmaster and Anderson (1994), and Kraan and Cooke (1997).

Historically, for reasons of computational tractability (among others), dependencies among random variables have often been either ignored, or else treated using unrealistic and simplistic assumptions such as perfect correlation. Fortunately, substantial progress is being made in modeling dependencies among components (and/or in the information about components). Two techniques, copulas and Bayesian networks, have become popular for specifying dependency relations. Bayesian networks are directed acyclic graphs (influence diagrams without decision nodes) in which nodes represent events and directed arcs ("arrows") between nodes show probabilistic dependencies. Each node's value has a conditional probability distribution that depends only on the values of the variables that point into

it. (Equally important, absent arrows indicate the conditional independence of each variable from those that do not point into it, given the values of those that do.) Sampling from the conditional distribution of each variable in turn, given the sampled values of its predecessors (after sorting the variables so that each appears only after those that point into it, if any), and repeating many times provides a way to sample from the joint distribution of the variables without having to explicitly specify it. [Such "Gibbs sampling" is a simple form of Markov chain Monte Carlo (MCMC) sampling that is well suited for Bayesian networks. In effect, the joint distribution is factored as a product of marginal distributions (for the input variables, meaning those with no predecessors) and conditional distributions (for all other nodes), thus allowing the potentially large size of a full joint distribution to be tamed by the relative sparseness of significant dependencies among variables in most real-world systems.] Free Windows software for Bayesian inference using Gibbs sampling ("WinBUGS"), called from the free statistical computing environment R, can be obtained by Googling on R2WinBUGS.

The use of copulas (functions that link a multivariate cumulative distribution to its one-dimensional cumulative marginal distributions; see, for example, Nelsen, 1999) has also become increasingly common in both financial and engineering risk analysis. Copulas have been applied, for example, to model dependencies between opinions from different experts (Jouini and Clemen, 1996; Lacke, 1998) and between system failure rates during normal and accident conditions (Yi and Bier, 1998). They are used extensively in financial risk analysis (e.g., in the Gaussian CreditMetrics or Basel II model) to describe correlated credit portfolio risks and interdependent risks of default (Frey et al., 2001).

Of course, copulas are not always the most convenient way to represent dependencies; see Joe (1997) for a compendium of multivariate distributions. Recently, Merrick et al. (2005) used an inverted Wishart distribution to model uncertainty about the dependencies among experts in assessing risks to the Washington State Ferries system while allowing the analyst to "learn about the dependencies between the experts from their responses." This is achieved by asking the experts to provide multiple different assessments of maritime risk under differing circumstances.

Cooke and colleagues (Bedford and Cooke, 2001; Kurowicka and Cooke, 2004) developed a practical method for specifying a joint distribution over n continuous random variables with specified rank correlations, using only $n(n-1))2$ assessments of conditional correlations. Kurowicka and Cooke (2004) point out that use of continuous multivariate distributions for a Bayesian belief net (a Bayesian network) allows for more tractable Bayesian updating than the commonly used discrete distributions (Lauritzen and Spiegelhalter, 1998).

Example: Analysis of Accident Precursors

Consider a risk analyst attempting to estimate the failure probabilities of critical safety systems in a nuclear power plant in the event of an accident. Fortunately, few if any accidents will have been observed on plants of that type, suggesting the

analyst may use data regarding failure probabilities of those systems during routine testing. However, this data will clearly be only partially relevant to the probabilities to be assessed; for example, one might expect that many systems will have higher failure probabilities under accident conditions than during routine testing.

Yi and Bier (1998) show how copulas can be used to represent dependency between the system failure probabilities under normal versus accident conditions. This makes it possible to perform a Bayesian update showing the effect of data collected under normal conditions on the system failure probabilities under accident conditions. Thus, for example, if routine testing showed a particular system to be much less reliable than was previously believed, this information could be used to update the expected failure probability of the system in the event of an accident. However, Yi and Bier's model is not sufficiently general to account for all relevant prior assumptions about dependencies. Thus, further work is needed to enhance ability to model dependencies.

Example: Flight-Crew Alertness

A challenge in modeling flight-crew alertness (Roelen et al., 2003) is that various predictive variables are correlated not only with crew alertness, but also with each other. For example, the crew's workload on a given flight is likely to be a function of both the length of the flight (with longer flights having higher total workload) and how much the crew members rest during the flight (with more rest being associated with a lower workload). However, assessing the combined impact of these variables on crew alertness may be difficult if longer flights also allow more rest time during flight.

Kurowicka and Cooke (2004) develop a continuous Bayesian belief net for this situation to allow airline managers to identify ways to compensate for known causes of poor alertness (such as long flights, or insufficient sleep prior to flight time). By allowing the variables in their model to have continuous distributions (rather than discrete distributions, which are more common in applications of Bayesian belief nets), they were able to achieve a highly parsimonious model requiring the assessment of only eight conditional rank correlations, compared to the many more assessments that would have been required for a discrete model.

Some Alternatives to Subjective Prior Distributions

Unlike classical statistical procedures, Bayesian analysis can be used in situations of sparse data, because subjective judgments and other nonstatistical types of evidence can be used in Bayesian estimation, inference, and decision processes. However, with sparse data, the results of Bayesian analyses are often sensitive to the analyst's choice of prior probabilities for models and parameters. Hence, Bayesian methods can be more subjective and less readily accepted when data are sparse.

Maximum-entropy distributions have sometimes been proposed to help solve this problem. They use whatever information is available about the uncertain quantity of interest (e.g., mean, median, or mean and variance) to constrain the assumed distribution for that quantity but presuppose as little additional information as possible beyond that, to avoid inadvertently assuming more than is actually known. A maximum-entropy distribution is defined to be the least informative distribution (in a precise technical sense) that satisfies the specified constraints (Jaynes, 2003). The resulting distribution can then be used either as a prior distribution for Bayesian analysis (if additional data become available) or as a partially informative distribution without updating. For example, Meeuwissen and Bedford (1997) use maximum entropy to identify the minimally informative distribution with a given set of rank correlation coefficients, using a piecewise constant numerical approximation (a so-called chessboard distribution).

However, maximum entropy and related approaches (such as "noninformative prior" distributions) lead to significant problems even in some relatively simple examples. For example, if all we know about a random variable X is that it is bounded by 0 and 1, then a maximum-entropy distribution for it would be uniform between these limits. Of course, exactly the same reasoning presumably applies to X^2, but X and X^2 cannot both be uniformly distributed between 0 and 1. Such lack of invariance to transformations of variables (e.g., from half-life to decay rate) means that maximum-entropy distributions may depend on essentially arbitrary choices of scale, or of how to represent the same physical situation. In addition, the maximum-entropy distribution can be difficult to compute in some cases (especially when quite a bit is known about the quantity of interest, so that the maximum-entropy distribution must satisfy numerous constraints).

Such limitations have raised interest in "robust" Bayesian methods and other bounding approaches. Robust Bayesian methods (Rios Insua and Ruggeri, 2000) update an entire class, family, or set (usually convex) of prior distributions with observed data, rather than just a single prior distribution. If the class is chosen carefully, the computational effort required to update all distributions in the class need not be substantially greater than for a single distribution. If all (or most) prior distributions in a suitably broad class give similar results, this can lead to greatly improved confidence in the results of the analysis.

In a similar spirit, probability bounds analysis (Ferson and Donald, 1998) propagates uncertainties (rather than choosing a prior distribution for Bayesian updating). The analyst specifies bounds on the cumulative distribution functions of the various input parameters to a model, rather than selecting specific cumulative distributions. These bounds are then propagated through the model. The uncertainty propagation process, which again can be quite computationally efficient, yields valid bounds on the cumulative distribution function for the final result of the model (e.g., a risk level). This approach can take into account not only uncertainty about the probability distributions of the model inputs, but also uncertainty about their correlations and dependence structure. This is valuable, because correlations will often be more difficult to assess accurately than marginal distributions, and correlations of 1 or –1 among the input variables do not necessarily produce the most extreme possible

distributions for the output variable(s) of interest; see, for example, Ferson and Hajagos (2006).

Example: Effects of Exposure to Contaminated Soil

Ecological and environmental risk models frequently involve a high degree of uncertainty, because some important parameters in the model may not be readily measurable. Consider the problem of attempting to estimate the effect of soil contamination on predator species (Hope, 1999), which may be exposed to contamination both directly (through ingestion of soil) and indirectly (by ingestion of a variety of prey species). Estimating the exposure to the predator species requires estimating the concentration of the contaminant in the flesh of all prey species, some of which may themselves be predators. This requires estimating the overall food and water intake and diet composition for each relevant species, as well as the uptake of the contaminant. Good data or expert opinion may be available for some parameters, but for others (such as the fraction of a particular predator's diet made up of a particular prey species), experts may feel uncomfortable assessing an informative probability distribution and may prefer simply to state, for example, that the fraction must be between 0 and 1. Standard practice would either press the experts to provide informative distributions, or simply assume a uniform distribution between 0 and 1, but this may not always conform to the experts' judgments. Correlations between the fractions of the diet made up of differing foods can also obviously be difficult to estimate reliably.

Regan et al. (2002) compare a traditional two-dimensional Monte Carlo analysis of this problem to the results obtained using probability bounds. Even using bounds of 0 and 1 for some parameters, the qualitative conclusions of the analysis (e.g., that the predator species of interest was "potentially at risk" from exposure to soil contamination) remained essentially unchanged between the two-dimensional Monte Carlo analysis and the probability bounds analysis. Thus, bounding analysis can help support a particular decision if it shows that the qualitative results and recommendations resulting from the analysis are not highly sensitive to the specific choices of probability distributions used in the simulation.

The use of subjective prior probabilities and judgment-based probability models can also be simplified or avoided in many situations where probability theory provides the required forms of distributions and/or useful bounds on the probable values of uncertain quantities. Table 2.1 summarizes some important classes of situations where probability theory prescribes distributions and bounds. [Table 2.1 assumes familiarity with the various distributions mentioned, such as Poisson, Weibull, exponential, gamma, Gumbel, normal, and lognormal. See Ross (1996) and the hyperlinks in the table for technical details of these distributions and topics. Googling on the distribution names and italicized topics in Table 2.1 will provide a host of web resources and authoritative references, even if these specific links become obsolete.]

Many of these results can be applied even when the correct probability distributions are unknown or are only partly known, perhaps from statistical sampling or

Table 2.1 Selected asymptotic distributions and bounds from probability theory

Situation	Key results and references	Examples
Random occurrences, with the numbers of occurrences in disjoint time intervals being statistically independent count random variables, and with the expected number of occurrences (counts) in any interval being proportional to its length.	The random number of occurrences in an interval has a Poisson probability distribution with mean (and variance) λt, where t = length of interval and λ = mean number of occurrences per unit time (the "intensity" of the arrival process). If each occurrence has a random consequence, and consequences are independent identically distributed (i.i.d.) random variables, the process is a *compound Poisson* distribution process (Ross, 1996). If λ depends on a linear combination of explanatory variables, a *Poisson regression model* results.	Annual numbers of fires in a city, sporadic food poisoning cases, auto thefts, bank robberies, car accidents on a stretch of highway, etc. Number of typos per page, defects per square foot of material, per mile of pipe, etc.
Waiting times between events in a Poisson process.	The random time between consecutive events in a Poisson process with intensity λ is an *exponential distribution* with mean $1/\lambda$. The random time for k consecutive events to occur has a *gamma distribution* (Ross, 1996).	Time-to-failure in a process where failure occurs after k "hits" (or other event arrivals).
Rare events, independent or weakly dependent numbers of occurrences in disjoint time intervals.	The number of events in a time interval is approximately Poisson or compound Poisson. Bounds are available for approximation error (Barbour et al., 1995; Pekoz, 2006).	Reliability systems; sequences with positive values tending to be clumped in time.
Smallest value among a large number N of i.i.d. nonnegative random variables, e.g., failure times of N independent components, or first time for one of N competing processes to reach a certain stage.	• For large N, the smallest value has approximately a *Weibull distribution*. • If the N variables are unbounded (instead of being non-negative or, more generally, bounded from below), then the distribution of the smallest value approaches a *Gumbel distribution for minima* [an Extreme Value Type I distribution, having hazard function $h(x) = e^x$ when location = 0 and scale = 1]. www.itl.nist.gov/div898/handbook/apr/section1/apr163.htm www.itl.nist.gov/div898/handbook/eda/section3/eda366g.htm mathworld.wolfram.com/ExtremeValueDistribution.html	Failure time of a system with N components in series or with N competing sources or causes of failure.

Table 2.1 (continued)

Situation	Key results and references	Examples
Maximum value in a large set or sequence of N i.i.d. random variables.	The largest value has a probability distribution that approaches a *Gumbel distribution for maxima* for large N, provided that the variables have finite moments and unbounded tails that decrease at least as fast as an exponential (e.g., Normal distributions). More generally, maxima and minima in many situations have approximately a *generalized extreme value* (GEV) distribution. [This is a three-parameter family that includes Weibull (or reversed Weibull), Gumbel, and Fréchet distributions as special cases.] www.itl.nist.gov/div898/handbook/eda/section3/eda366g.htm http://rss.acs.unt.edu/Rdoc/library/evd/html/00Index.html	Maximum floods, rainfalls, traffic loads, insurance claims, bank deposits/withdrawals. Failure time of a reliability system with N components in parallel.
Sum of N i.i.d. random variables with finite mean μ and variance σ^2. *Sum* of N random variables with finite means and variances, *not necessarily i.i.d.,* satisfying the Lindeberg condition (each variance is small compared to the sum of all the variances).	The sum approaches a *Normal distribution*, with mean $N\mu$ and variance $N\sigma^2$ for large N. (This is a *central limit theorem*. Explicit bounds are available for the rate of convergence and closeness of the approximation.) For large N, the sum has approximately a *Normal distribution*, with mean = sum of the N means and variance = sum of the N variances. (This is the Lindeberg-Feller central limit theorem.) http://mathworld.wolfram.com/LindebergCondition.html http://mathworld.wolfram.com/Lindeberg-FellerCentralLimitTheorem.html	Total losses from N consecutive insurance claims; total damage accumulated in a population of N units exposed to a source of random damage; total time to use up N spare parts having i.i.d. random lifetimes.
Product of many positive well-behaved (square-integrable) i.i.d. random variables. Roughly, this may be interpreted as a product of many individually "small" random factors.	The product is asymptotically *lognormal*. Distributions of chemicals, particles, or microbes in the environment, and other *exposure variables*, are often lognormally distributed. Corrosion, diffusion or migration of ions, crack growth, and other material and chemical processes often lead to lognormal degradation processes and *failure times.* http://stat.ethz.ch/~stahel/lognormal/bioscience.pdf http://www.itl.nist.gov/div898/handbook/apr/section1/apr164.htm	Probability that all of many conditions hold; survival times for cancer patients; failure times of processes where further degradation occurs at a rate proportional to current degradation.

Table 2.1 (continued)

Situation	Key results and references	Examples
Any *monotone graph property P,* i.e., any graph property (such as connectivity) that does not depend on the labeling of graph vertices and is not destroyed by adding edges to a graph. *Percolation processes*, in which an effect propagates from site to neighboring site in a lattice if they are linked (and there is probability p that any two neighboring sites are linked).	The probability that property P holds in a random graph with N vertices and with edge probability p is close to 0 for p below some threshold value and is close to 1 for p above a second, larger threshold value. For large N, the two threshold values approach each other, forming a "sharp transition threshold," t, such that that the graph almost surely has property P if $p > t + \varepsilon$ and almost surely does not have property P if $p < t - \varepsilon$, where ε approaches 0 for large N. The system is said to undergo a *phase transition* for property P at the sharp transition threshold, t (Friedgut, 2005). Sharp transition thresholds also hold for *percolation processes* and for *random geometric graphs* (in which nodes are uniformly distributed in a spatial region and any two nodes within a specified distance d of each other are linked). http://www.stanford.edu/~ashishg/papers/gnrn.pdf	Spreading of forest fires, infectious diseases, invasive species, etc. have been modeled as percolation processes. Phase transitions from unreliable to reliable wireless networks as power and/or density of wireless nodes in an area increases. http://portal.acm.org/citation.cfm?id=970847
X is a possibly unknown nonnegative random variable with mean $E(X)$.	$\Pr[X \geq k^*E(X)] < 1/k$, for $k \geq 1$. This is *Markov's inequality*. It implies Chebyshev's inequality (next). http://mathworld.wolfram.com/MarkovsInequality.html	Random lifetime with unknown probability distribution, mean estimated from data.
X is a possibly unknown random variable with finite mean $E(X)$ and finite variance σ^2.	For any $k > 0$, $\Pr(\lvert X - E(X)\rvert \geq k\sigma) \leq 1/k^2$. Interpretively, values of a random variable are unlikely to be many standard deviations away from the mean. This is *Chebyshev's inequality* (Ross, 1996; http://en.wikipedia.org/wiki/Chebyshev's_inequality).	Probability of gain or loss exceeding a given limit. Financial risk analysis, banking risks.
Martingales: $X_0, X_1, \ldots$ is a sequence of random variables with finite means and bounded increments such that $E(X_{t+1} \mid X_0, \ldots, X_t) = X_t$ and $\lvert X_k - X_{k-1}\rvert \leq a_k$ almost surely. The variables need not be i.i.d.	$\Pr(\lvert X_t - X_0\rvert \geq d) \leq 2\exp(-d^2/[2(a_1{}^2 + a_2{}^2 + \ldots + a_t{}^2)]$. This is a form of *Azuma's inequality* for martingales. For example, if each step size $\lvert X_k - X_{k-1}\rvert \leq 1$, then the probability that the process will have increased by $\geq d$ after t steps satisfies $\Pr(X_t \geq X_0 + d) \leq \exp(-d^2/2t)$. Interpretively, the probability of a "large deviation" (of size d or more) for cumulative winnings in a sequence of fair gambles is exponentially small (Ross, 1996; http://en.wikipedia.org/wiki/Azuma's_inequality).	Probability of gain or loss exceeding a given limit. Machine learning with limited error probabilities; fault detection and change detection in martingale processes.

Table 2.1 (continued)

Situation	Key results and references	Examples
Binomial trials with unknown success probability, *p on each trial*.	If data consist of N trials with no successes, then there is approximately 95% confidence that p does not exceed $3/N$. This is the *rule of 3* (Chen and McGee, 2008; www.sinica.edu.tw/~jds/JDS-401.pdf).	Upper bound for probability of a never-observed event.
Coherent structure reliability system, with unknown structure and unknown independent component failure rates.	If each component has a failure rate (or hazard function) that is *increasing on average* [i.e., $H(t)/t$ is increasing in t, where $H(t) =$ cumulative failure rate], then so does the whole system. Many other similar results and bounds are available (Barlow, 1998).	Reliability system with uncertain structure and/or component failure rates.

simulation modeling that provides estimates of means and variances. For example, the sums, maxima, and minima of repeated random samples from most distributions encountered in practice have asymptotic distributions (as the number of samples becomes large) that do not depend on the specific distribution being sampled. Thus, it is unnecessary to know the underlying "parent distribution" to quantify the distribution of these statistics, all of which are of interest in various risk analysis applications. Similarly, a variety of inequalities quantify how unlikely it is that a value sampled from a distribution will fall far from its expected value. Again, these bounds do not require detailed knowledge of the parent distribution. As a result, empirical data that give only limited information about a risky process may still be adequate to obtain useful quantitative bounds on risks of interest.

Example: The "Rule of Three" for Negative Evidence

Setting: People sometimes worry about events that *might* happen in theory, even though they *have not* (yet) happened in practice. How reassuring should one consider such "negative evidence" (i.e., the absence of occurrences of a feared event, despite past opportunities for occurrence), bearing in mind the adage that "Absence of proof [of a hazard] is not proof of absence"? This can be an important topic when new technologies or poorly understood systems are involved, ranging from the Large Hadron Collider particle accelerator at CERN, which some feared might destroy the world by producing micro black holes, to the systems of interlocking safeguards that countries establish to try to protect against diseases such as bovine spongiform encephalitis (BSE, or "mad cow" disease). We will use the latter example to illustrate how negative evidence (i.e., the observation that a feared event has not yet been observed) can be used to bound risk.

Problem: Supposed that a country concerned about the possibility that its domestic cattle might be infected with BSE tests 1,000,000 randomly selected cattle and finds no cases. How confident can one be, based on this data, that the true prevalence proportion of BSE in the sampled population is not large? Assume that how BSE originates and spreads among cattle is not understood well enough to simulate or model with high confidence and that the effectiveness of any safeguards against BSE is not yet known. Thus, we want an upper-bound risk estimate based on the empirical "negative evidence" of no observed cases among a million animals tested, since calculations based on a well-validated understanding of the BSE disease process are not available.

Solution: A useful nonparametric confidence bound is based on the following "rule of 3" (Chen and McGee, 2008): If an event that has the same probability p (which may be unknown) of occurring on each trial has not occurred in any of N independent trials (e.g., in a simple random sample of size N), then, with at least 95% confidence, its occurrence probability on each trial satisfies $p \leq 3/N$. Thus, the unknown prevalence proportion of detectable BSE in this example would satisfy $p \leq 3/1{,}000{,}000 = 0.000003$. This bound does not require or assume any specific prior distribution for p, or any knowledge of the (probably complex) processes by which BSE might enter the country and spread domestically.

Example: A Sharp Transition in a Symmetric Multistage Model of Carcinogenesis

Setting: This example illustrates how probability laws can be used to model complex processes such as cancer, even if the molecular-level details of causal pathways are unknown. As a simplified illustration, consider the following *symmetric multistage model* of carcinogenesis. A cell line gradually accumulates transformations (e.g., somatically heritable mutations) from a set of K possible transformations. Transformations occur randomly and independently over time. The K transformations arrive according to independent Poisson processes, with (at least approximately) equal intensities, given by λ average occurrences per unit time. (Transformations with occurrence rates much less than this common value are not rate-limiting and thus may be disregarded.) Once any of the K transformations has occurred, we assume that it is permanent and irreversible. If a specific transformation occurs more than once, the occurrences after the first one are wasted, i.e., the cell genotype has already acquired that transformation and does not reach malignancy any faster if it occurs again. The cell line survives for a finite lifetime of duration T. If all K distinct transformations occur before time T, then the cell line becomes malignant.

Problem: Under these conditions, what is the probability that the cell line will become malignant before death at time T? If it does become malignant before time T, then what can be said about the (random) time at which the first malignant cell is formed?

Solution: The somewhat surprising answer is that, for sufficiently large K, there is a "sharp transition" time such that the first malignant cell is very unlikely to be formed much sooner or much later than that time. In other words, a nearly deterministic occurrence time for the first malignant cell emerges simply as a consequence of there being many stages in this simple stochastic transition model.

Result: In this completely symmetric multistage model, there is a "sharp transition" time $T^* \approx (1/\lambda)[(\ln(K) + \gamma]$, where λ is the expected number of transformations events per unit time, i.e., their average occurrence rate, and γ = Euler's constant = 0.57721.... In particular, the expected time until the first malignant cell is formed is T^*; moreover, the coefficient of variation of the actual (random) time of formation of the first malignant cell (i.e., the ratio of its standard deviation to T^*) approaches 0 for large K.

Proof: The expected *number* of transformation occurrences, including wasted (i.e., repeated) ones, until a malignant cell is formed (i.e., until all K transformations have occurred at least once) is given by the harmonic sum: $E(n^*) = K(1 + 1/2 + 1/3 + \ldots + 1/K) \approx K[(\ln(K) + \gamma]$, where n^* denotes the random number of the transformation occurrence event at which all K transformations are first completed and γ is Euler's constant, $\gamma = 0.57721\ldots$. This follows from previously known results for the "Coupon Collector's Problem" with equal probabilities (e.g., Ross, 1996, p. 414; Motwani and Raghavan, 1995) or for the maximum of K independent exponential random variables (e.g., Nelson, 1995, p. 173). [Intuitively, this result is motivated by the fact that any of the K transformations can occur first and be nonredundant, after which the probability that the next transformation is nonredundant drops to $(K-1)/K$, then to $(K-2)/K$, ..., and finally, for the last transformation, to $1/K$.] The

expected *time* until a malignant cell is formed is therefore $T^* = E(t^*) = E(n^*)/(K\lambda) \approx (1/\lambda)[(\ln(K) + \gamma]$, where T^* denotes the random time at which all K transformations are first completed, and $K\lambda$ is the rate at which transformation events arrive (since each of the K types independently arrives at rate λ). This proves part (a) of the theorem. The fact that the probability distribution of n^* has a sharp concentration around $E(n^*)$ is proved in Motwani and Raghavan (1995). Given this key result, hold n^* fixed. The time until n^* transformations (including redundant ones) have occurred has a gamma distribution with mean $n^*/(K\lambda)$ and variance $n^*/(K^2\lambda^2)$, by standard results for waiting times in Poisson arrival processes and for the mean and variance of the gamma distribution (e.g., Ross, 1996, p. 18). The ratio of the standard deviation to the mean of this waiting time is therefore $(n^*)^{-1/2} \approx [K((\ln(K) + \gamma)]^{-1/2}$, which goes to 0 as K increases.

Discussion: An interesting, and perhaps unexpected, aspect of this result is that it establishes a form of nearly deterministic behavior for a stochastic system: If the sharp transition time T^* is smaller than the death time T, then formation of a malignant cell by time T is almost certain; otherwise, it is very unlikely. (This qualitative behavior is typical of what is sometimes called a 0–1 law in stochastic processes.)

If K is not large enough to guarantee a sharp transition at time T^*, then the qualitative behavior can be generalized as follows: For any $\varepsilon > 0$, no matter how small, there is an interval of times $[T^-, T^+]$ such that the probability of a malignant cell being formed before T^- or after T^+ is less than ϵ. The cumulative probability distribution for the occurrence time of the first malignant cell increases from almost 0 to almost 1 over this interval. As K increases, the width of this interval shrinks toward zero, with T^- and T^+ approaching a common value, T^*.

Realistic models of carcinogenesis are more complex than this example (see, for example, Chapters 11 and 12), but this simplified illustration shows that sometimes the behaviors of complex stochastic systems can be described well by phase transitions and probability laws, even if the details of the systems (such as which specific events occur along different causal pathways leading to cancer) are unknown. (Chapter 16 describes a similar phase-transition result for the ability of telecommunications networks to recover from deliberate coordinated attacks at multiple locations.)

Dealing with Model Uncertainty: Bayesian Model Averaging (BMA) and Alternatives

Copulas and maximum-entropy methods are mainly used to deal with uncertainties about the *parameters* and *input distributions* for particular models. However, *model uncertainties* about (a) which variables to include in a model when many potential predictors (including some possibly irrelevant ones) have been measured and (b) the most appropriate *functional form* for a model – or, more generally, how to calculate or predict a model's outputs from its inputs – are even more impor-

tant in practice than input and parameter uncertainties, in applications ranging from dose-response models in toxicology to the reliability modeling of complex systems. Some researchers have suggested assessing a probability distribution over multiple plausible models by evaluating the consistency of the various models with the observed data (in much the same way as the likelihood function in Bayesian updating evaluates the consistency of various parameter values with observed data) and determining how much weight to put on each model based on its consistency with the data. Failing to consider model uncertainties can lead to spuriously narrow statistical confidence intervals for parameter estimates and to spuriously high confidence in model-based predictions (Hoeting et al., 1999).

However, it is frequently not reasonable to attempt to estimate the probability that a given model is "correct," because, as Box (1979) pointed out, "All models are wrong, some models are useful." For example, it seems highly implausible that any of the current models for estimating the probability of human error on a given task is close to being "correct" (because all are gross oversimplifications of the real world), nor can the current models be considered a collectively exhaustive set of possible models of human error. Bayesian updating of probability distributions over such partial subspaces of possible models may not always work well in practice. Some models may be intentionally conservative (e.g., for regulatory and/or screening purposes) or intentionally simplified (e.g., for computational tractability, or to yield qualitative insights). That such models may be inconsistent with observed data does not necessarily invalidate their use for their intended purposes.

Finally, of course, more complex models, with larger numbers of parameters, will often fit the observed data well in many situations (subject to the possible limitations of overfitting), but may not always be preferable, if only for reasons of parsimony and/or generalizability. Thus, standard approaches for dealing with uncertainty probabilistically are often not well suited for handling model uncertainty. Bayesian model averaging (BMA) (see Chapter 7) was motivated largely by these challenges. BMA avoids basing all of one's conclusions on any single model if multiple models are about equally plausible. It avoids giving high weight to models that are excessively complex if simpler ones give comparably good (or better) descriptions of the data, as measured by the likelihood of the data given a model. BMA generally performs reasonably well in practice, e.g., as evaluated by its ability to give well-calibrated uncertainty interval estimates for uncertain outputs, taking into account model uncertainty (Hoeting et al., 1999; Raftery and Zheng, 2003).

An alternative that avoids assigning probabilities to individual models, "comprehensive uncertainty evaluation" (Brown, 1999), involves subjectively adjusting the probability distributions resulting from a particular model to try to take into account known weaknesses of the model (such as conservatisms, or risks that are not adequately modeled). This is consistent with subjective utility theory and avoids some of the theoretical conundrums associated with assigning probabilities to models. Brown has applied this method (for example, to support regulatory decision making for nuclear power plants), but it has not yet seen widespread application by other analysts in practice.

In many applications of Bayesian analysis to situations involving model uncertainties, the input parameters are assumed to be known, and the model results are used to update the prior distribution over model outputs (see, for example, Chick, 1997). However, observing the output of a model could also cause one to revise the prior distribution over model inputs if the true values of the model outputs were known reasonably well (e.g., from empirical data). Thus, for example, Bayesian analysis could be used to estimate which values for the rate of disease progression are most consistent with the observed data on disease prevalence and severity (Andradóttir and Bier, 2000).

Risk Characterization

The output of a PRA to support risk management decision making is a characterization of the risk for each decision option being evaluated. Occasionally, the decision task is to identify an optimal risk management policy from a large set of possibilities, rather than to explicitly characterize the risks for each of a small number of alternatives. Then, simulation-optimization algorithms or special-purpose techniques such as Markov decision processes or stochastic optimal control theory may be required (see Tables 2.2 and 2.3). However, explicit comparison of risks from a few options is more usual, and is the main focus of this section.

"Risk" is usually defined in engineering risk assessments and PRA as the frequency and severity of losses arising from operation of the designed system in its uncertain environment, including a specification of losses (i.e., which adverse consequences matter, and to whom). An effective display of risk shows how it is affected by different actions (e.g., different risk management decisions) and allows "drill-down" to view the risks to particular subpopulations, as well as the contributions of various different causes to the overall level of risk. For example, seeing how risk curves shift when risk-reducing measures are implemented would help managers identify the most effective measures. Uncertainty and sensitivity analysis are also essential to risk characterization, because they support estimates of the value of information.

Engineering vs. Financial Characterizations of "Risk": Why Risk Is Not Variance

The variance (or standard deviation) of the return on investment is widely used as a measure of risk in financial risk analysis, where mean-variance analysis is applied to calculate "efficient" frontiers and undominated portfolios, defined as those having maximum expected return for a given variance. Why, then, do health, safety, environmental, and reliability risk analysts insist on defining risk more flexibly, as being determined by probabilities *and* consequences, rather than simply by variances (or, for that matter, semivariances, value-at-risk, or modern coherent risk mea-

Table 2.2 Some decision modeling frameworks

Framework	Act	State	Consequence	Optimization algorithms
Normal form • $\Pr(s)$ = state probability model • $u(c)$ = utility model $c \varepsilon C$, a set of possible consequences	a = act = choice of *controlled input* to system, or of a *decision rule* mapping observations to controlled input values	s = state (random or uncontrolled inputs) $s \varepsilon S$, a set of possible states	• $c(a, s)$ = deterministic consequence model • $\Pr(c \mid a, s)$ = stochastic consequence model • $a \varepsilon A$, a set of feasible acts	Choose a from A to maximize expected utility (EU): $\max_{a \in A} \Sigma_c u(c) p(c \mid a)$ s.t. $p(c \mid a) = \Sigma_s \Pr(c \mid a, s) \Pr(s)$
Utility table	a = row	s = column	expected utility for each cell (i.e., for each row-and-column pair)	Eliminate dominated rows, then choose act to maximize EU
Decision tree	Choice of an act at each decision node	Outcome at each chance node	Utilities at tips of tree	Backward dynamic programming
Influence diagrams and Bayesian networks	Choice of an act at each decision node	Outcome at each chance node	Value at value node	• Gibbs sampling • Bucket elimination • Graph algorithms (e.g., arc reversal)
Markov decision process (MDP); semi-Markov decision process	Choice of act in each state	Transition time and next state	Reward per unit time in states and/or at transitions; transition rates among states	• Value iteration • Policy iteration • Neurodynamic programming • Linear programming

Table 2.2 (continued)

Framework	Act	State	Consequence	Optimization algorithms
Partially observable MDP	Decision rule mapping observations to acts	Underlying state of process	Function of uncertain state and choice of acts, and/or transitions	Similar to MDP (A POMDP is an MDP with probabilities of states for its state.)
Stochastic optimal control and robust control of uncertain systems	Decision rule (maps observed history to choice of controlled input values)	Probability of next state, given history to date (and inputs)	Function of state and/or control trajectory	Stochastic dynamic programming; adaptive optimization, robust control, reinforcement learning algorithms
Simulation model	Controllable inputs	States of model components	Function of component states and/or input sequence	Simulation-optimization heuristics and algorithms
Unknown model	Sequence of acts		Sequence of rewards	Adaptive learning, online, and minimal-regret algorithms

Table 2.3 Selected literature on decision optimization frameworks and algorithms

- Decision trees:
  - Game Trees for Decision Analysis – Shenoy (1996), http://citeseer.ist.psu.edu/shenoy96game.html
- Influence diagrams and Bayesian networks:
  - Sampling Methods for Action Selection in Influence Diagrams – Ortiz, Kaelbling (2000), http://citeseer.ist.psu.edu/ortiz00sampling.html
  - A Forward Monte Carlo Method for Solving Influence Diagrams. – Charnes, Shenoy (2000), http://citeseer.ist.psu.edu/charnes00forward.html
  - A Simple Method to Evaluate Influence Diagrams – Xiang, Ye (2001), http://citeseer.ist.psu.edu/ye01simple.html
  - Learning Bayesian Networks with R, http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/BottcherDethlefsen.pdf; see also http://www.cs.ubc.ca/∼murphyk/Software/bnsoft.html
- Markov decision processes (MDPs) and partially observable MDPs (POMDPs):
  - Reinforcement Learning for Factored Markov Decision Processes – Sallans (2002), http://citeseer.ist.psu.edu/sallans02reinforcement.html
  - Symbolic Dynamic Programming for First-Order MDPs – Boutilier, Reiter, Price (2001), http://citeseer.ist.psu.edu/boutilier01symbolic.html
  - Speeding Up the Convergence of Value Iteration in POMDPs – Zhang, Zhang (2001), http://citeseer.ist.psu.edu/zhang01speeding.html
  - Solving POMDP by On-Policy Linear Approximate Learning Algorithm – He (1999), http://citeseer.ist.psu.edu/335710.html
- Optimal and robust control and reinforcement learning for uncertain and nonlinear systems:
  - Feedback Control Methodologies for Nonlinear Systems – Beeler, Tran, Banks (2000), http://citeseer.ist.psu.edu/Beeler 00feedback.html (for deterministic nonlinear systems)
  - An Overview of Industrial Model Predictive Control Technology – Qin, Badgwell (1997), http://citeseer.ist.psu.edu/qin97overview.html
  - http://citeseer.ist.psu.edu/kaelbling96reinforcement.html
  - http://citeseer.ist.psu.edu/sutton98reinforcement.html
  - http://www.princeton.edu/∼noahw/palgrave2.pdf (introduces robust control)
- Simulation-optimization:
  - A Survey of Simulation Optimization Techniques and Procedures – Swisher, Jacobson et al. (2000), http://citeseer.ist.psu.edu/517471.html
  - Simulation Optimization of Stochastic Systems with Integer Variables by Sequential Linearization – Abspoel et al. (2000), http://citeseer.ist.psu.edu/516176.html
  - Simulation Optimization: Methods and Applications – Carson, Maria (1997), http://citeseer.ist.psu.edu/carson97simulation.html
  - http://opttek.com/simulation.html (overview and link to commercial software)
- Minimal-regret, online, and adaptive learning algorithms:
  - Minimizing Regret: The General Case – Rustichini (1999), http://citeseer.ist.psu.edu/rustichini98minimizing.html
  - Adaptive Strategies and Regret Minimization in Arbitrarily Varying Markov Environments – Mannor, Shimkin (2001), http://citeseer.ist.psu.edu/467490.html
  - Nearly Optimal Exploration-Exploitation Decision Thresholds – Dimitrakakis (2006), http://citeseer.ist.psu.edu/dimitrakakis06nearly.html
  - Combinatorial Online Optimization in Real Time – Grötschel, Krumke, Rambau (2001), http://citeseer.ist.psu.edu/448491.html; see also http://citeseer.ist.psu.edu/foster97regret.html

sures developed in financial risk theory)? This section suggests a partial answer by providing a simple proof that mean-variance decision making violates the principle that a rational decision maker should prefer higher to lower probabilities of receiving a fixed gain, all else being equal. Indeed, simply hypothesizing a continuous increasing indifference curve for mean-variance combinations at the origin is enough to imply that a decision maker must find unacceptable some prospects that offer a positive probability of gain and zero probability of loss. Unlike some previous analyses of the limitations of variance as a risk metric, this section does not require the additional framework of von Neumann-Morgenstern utility theory.

Incompatibility of Two Suggested Principles for Financial Risk Analysis

Two plausible principles for managing financial investment risks are the following:

1. *Rule 1: Make dominating choices.* Other things being equal, given a choice between a smaller probability of gain and a larger probability of gain, a decision maker should always choose the larger probability of gain. For example, given a choice between winning $100 with probability 0.1 and winning $100 with probability 0.2, rational decision makers who prefer more dollars to fewer should choose the option that gives a 0.2 probability of winning the $100.
2. *Rule 2: Seek mean-variance efficiency (higher variance requires higher mean return).* Given a choice among risky prospects, an investor should require more expected return to accept a prospect with more variance than to accept a prospect with less variance. For example, a 0.2 chance of winning $100 (else nothing) has a higher variance than a 0.1 chance of winning $100, but it also has a higher mean.

Rule 1 is implied by the decision-analytic principle of first-order stochastic dominance (Sheldon and Sproule, 1997): Prospects that give higher probabilities of preferred outcomes (and lower probabilities of less preferred outcomes) should be preferred. Rule 2 provides the basis for many current efficient portfolio and mathematical optimization (e.g., quadratic programming) approaches to optimal investment (http://en.wikipedia.org/wiki/Modern_portfolio_theory). Although theorists have noted that some risk-averse decision makers may prefer some mean-preserving increases in variance (ibid.), the idea that volatility in returns, as measured by variance or standard deviation, is generally undesirable to risk-averse investors, and that it should be avoided or compensated by higher expected returns, is still widely taught and practiced.

However, *Rules 1 and 2 are incompatible* in general. Simply hypothesizing that a decision maker has continuous upward-sloping indifference curves for mean-variance combinations (so that increasing the variance in the random return from an investment prospect or portfolio requires increasing its mean return in order to

leave the investor equally well off) violates Rule 1 for some simple prospects, as demonstrated next.

Following the literature on mean-variance decision making, suppose that a decision maker has positively sloped continuous indifference curves in mean-variance space (e.g., Wong, 2006). To any mean-variance pair (m, v) (a point in the mean-variance space) there corresponds a *certainty equivalent*: namely, the point at which the indifference curve through (m, v) reaches the horizontal (mean) axis. The indifference curve through the origin (0, 0) separates *acceptable risks* (those with positive certainty equivalents, lying below and to the right of the curve, if return is desirable) from *unacceptable risks* (those with negative certainty equivalents, lying above and to the left of it). To make an unacceptable risk acceptable in this framework, one must either increase its mean return or reduce its variance. (A risk-neutral decision maker who cares only about means and not about variances would have vertical indifference curves, but we will focus on the case, implied by Rule 2, of positively sloped indifference curves.)

The hypothesis that upward-sloping mean-variance indifference curves exist has some surprising consequences.

Theorem 1 *If the indifference curve through the origin slopes upward, then the decision maker finds unacceptable some prospects with positive expected values and no possibility of loss.*

Proof The proof is constructive. Let the slope of the indifference curve through the origin be s at the origin. By hypothesis, $0 < s \infty$. Now, consider a Bernoulli random variable $X(p)$ that gives a positive return of $2s$ with probability p (the "win probability") and no return (\$0) with probability $(1 - p)$. For a given value of p between 0 and 1, inclusive, $X(p)$ has mean $2ps$ and variance $4s^2p(1 - p)$ (since it is a scaled version of a Bernoulli random variable). Therefore, as p ranges from 0 to 1, $X(p)$ traces out a parabola in mean-variance space, with variance $= 0$ at $p = 0$ and at $p = 1$, and with a positive maximum variance of s^2 at $p = 0.5$ (see Fig. 2.1). A line from (0, 0) to the point on this parabola corresponding to a particular value of p has slope $4s^2p(1 - p)/2ps = 2s(1 - p)$. As p approaches 0, this slope approaches $2s$. Hence, the parabola traced out by $X(p)$ as p ranges from 0 to 1 starts above and to the left of the indifference curve through the origin (since it is constructed in such a way as to have twice the slope of the indifference curve at the origin), but it ends below and to the right of the indifference curve [since it is constructed to pass through the point $(2s, 0)$ when $p = 1$]. Therefore, the parabola must intersect the indifference curve somewhere above and to the right of the origin (since it starts above it and ends below it). Let p^* denote the value of the win probability for this intersection point. Then the decision maker prefers (0, 0) to all prospects $X(p)$ with $p < p^*$ since, by construction, these are unacceptable (i.e., above and to the left of the indifference curve through the origin). Hence, the decision maker finds unacceptable all such prospects giving probability p of $2s$ (else \$0) for $p < p^*$ even though he or she has positive win probabilities and even though none of them offers the possibility of a loss.

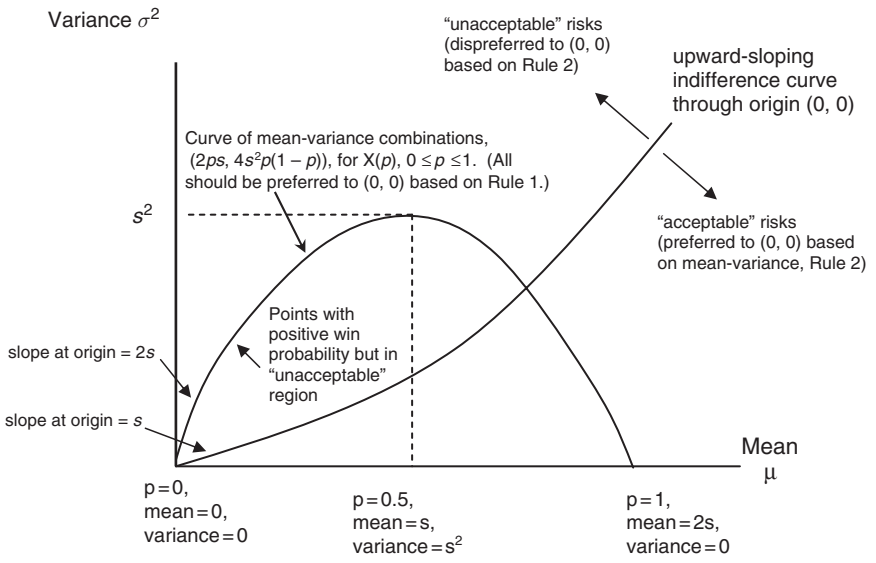

Fig. 2.1 Geometry of inconsistency between Rules 1 and 2

The proof of Theorem 1 implies that if the indifference curve through the origin has positive slope, then the decision maker prefers some prospects that give zero probability of winning a positive amount (namely, $2s$) to other prospects that give a positive probability of winning the positive amount (and otherwise nothing). Such a decision maker prefers the status quo or "nothing ventured, nothing gained" point (0, 0) to the possibility of winning a positive amount without any possibility of a loss, violating Rule 1. In this sense, Rules 1 and 2 are incompatible.

More generally, other parabolas can easily be constructed that intersect indifference curves twice, once for the ascending (positively sloped) portion of the parabola and once for its descending (negatively sloped) portion (Borch, 1969). In any such construction, the rightmost intersection represents a stochastically dominant prospect (which should be preferred, by Rule 1) compared to the leftmost intersection. That both points lie on the same indifference curve violates Rule 1.

In summary, although students of elementary finance are often taught that "risk" should be characterized by the variance or standard deviation of returns around an expected value, students of health, safety, environmental, and reliability risk analysis are usually taught instead that "risk" is determined by the probabilities of different consequences. Theorem 1 shows why the second approach (considering different specific consequences, such as \$0 and \$2s, and their probabilities) can be preferable to considering only means and variances.

The finding that variance is problematic as a measure of risk has a history at least several decades old in the financial and decision sciences literatures. A common critique in the theoretical decision analysis and financial economics literatures is that mean-variance analysis is compatible with von Neumann-Morgenstern expected

utility theory only under restrictive conditions (e.g., if all risky prospects have normal or location-scale distributions and utility functions are quadratic, implying that less money is preferred to more, for some amounts) (Markowitz, 1959; Baron, 1977). Mean-variance dominance and stochastic dominance relations for location-scale distributions do not coincide in general (Wong, 2006). Indeed, *expected utility theory is inconsistent with all possible moment-based preference models* (in which preferences are determined by mean, variance, skewness, kurtosis, etc.) for many utility functions (Brockett and Kahane, 1992). Variance is also inconsistent with proposed normative axioms for "coherent" financial risk measures (nonnegativity; homogeneity and subadditivity, which together imply that deterministic outcomes have zero risk; and shift-invariance, which implies that adding a constant to a random variable does not change its risk) (Pedersen and Satchell, 1999). Empirical studies since the 1960s have demonstrated that real decision makers pay attention to more than mean and variance in their choices among risky prospects (Jia et al., 1999).

Thus, Theorem 1 is consistent with a long line of previous research. However, in contrast to much previous work, it demonstrates a conflict between Rules 1 and 2 making only minimal assumptions (in particular, not requiring the framework of von Neumann-Morgenstern expected utility theory or other sets of normative axioms for risk measures) and using only elementary mathematics. It may therefore be useful for understanding why specifying the variances and expected returns from alternative investment choices (or other actions) does not adequately characterize risk or identify the choice with the most desirable probability distribution of consequences.

In fairness, it should be noted that financial risk analysts have developed much more sophisticated and satisfactory measures of risk than variance and that characterizing risk by frequency and severity is not problem-free (see Chapter 5). A recent triumph of financial risk theory has been the definition and analysis of *coherent risk measures* (Artzner et al., 1999). These provide formulas for assigning numbers to risky prospects so that normative axioms are satisfied, such as that risk remains unchanged if the same predictable constant is added to or subtracted from all possible consequences of a prospect; and that comparisons of risk should be logically consistent with each other over time. Financial risk theorists have shown that various sets of normative axioms imply intuitively pleasing quantitative representations of risk, such as that the "risk" of a financial prospect is its minimum (worst-case) expected net present value (ENPV), when ENPVs are calculated for each of a set of mutually consistent probability measures (Riedel, 2004). Older proposed measures of financial risk, including variance and Value-at-Risk (VaR), which reflects the probability of losing at least a specified amount, do not satisfy these axioms (Artzner et al., 1999; Pedersen and Satchell, 1999). Although coherent risk measures provide a substantial advance in methodologies for characterizing financial risks, they do not apply to other risks that cannot be traded, valued, or diversified away via financial markets. The characterization of health, safety, environmental, and reliability risks in terms of probabilities or frequencies of different consequences, having different magnitudes or severities, is still the norm. Chapter 5 discusses further the use and limitations of frequency in risk characterization.

Challenges in Communicating the Results of PRAs

Risk communication (including both presenting the results of risk analyses to stakeholders, decision makers, and other audiences, and listening to and actively eliciting their concerns so that they can be addressed in the risk analyses in the first place) facilitates the effective participation and interaction of technical experts, stakeholders, and decision makers in risk management decisions and deliberations. There is an extensive body of literature on risk communication (including guidelines, survey results, experiments, and web resources on risk communication).

Even more than PRA, risk communication is still an art rather than a science, but one that can be informed and improved by theory, experience, and experiments. Current challenges in risk communication include dealing with framing effects, communicating highly technical results to decision makers who may not be intimately familiar with some of the methods used in the risk analysis, and building trust among affected stakeholders and members of the public more generally. Adding to the difficulty is the fact that the communication and presentation styles that are most effective in accurately expressing the technical content of risk assessment findings may not always be those that invite and elicit public understanding, participation, and interaction.

Cullen and Frey (1999) discuss the distinctions between state-of-knowledge uncertainty and population variability (sometimes referred to simply as uncertainty and variability, respectively). State-of-knowledge uncertainty typically reflects uncertainties that affect all of the units being studied (e.g., can certain standardized industrial systems fail in a particular way? Can a certain chemical cause particular health effects?). These uncertainties could be reduced through further research. Variability refers to variations among the elements being studied (often assumed to be due to randomness in production processes, phenotypes, etc.). For example, differences in how different individuals in the population would respond to a chemical being studied, or strengths of different samples of a material, would reflect variability. Variability is often taken to be essentially irreducible through further study.

With the development and increased popularity of so-called second-order Monte Carlo analysis for quantifying uncertainty about risks, it is now common practice to distinguish between uncertainty and variability. This increases the value of the risk analysis for decision making, because different policy options may be appropriate for dealing with uncertainty rather than variability. For example, in situations of high population variability but low state-of-knowledge uncertainty, such as airbag effectiveness (Thompson, 2002), it may make sense to target risk-reducing efforts at those facilities or members of the population with the highest estimated risks (in this case, children and small adults). By contrast, situations of low variability but high uncertainty would tend to suggest that further research may be desirable before undertaking costly risk reduction actions. However, the widespread use of second-order Monte Carlo simulation does increase the challenges of effectively communicating ever more sophisticated and sometimes abstruse risk analysis methods and results to decision makers and members of the public in a way that clearly supports improved decision making (Bier, 2001a).

Of course, technically accurate risk communication by itself is not sufficient to achieve other key goals of risk communication, such as changing people's behavior (Blaine and Powell, 2001), gaining their trust in the results of the analysis, or even giving them the information they need to make improved decisions. Rather, effective and persuasive communication about risks generally requires a concerted effort to build trust, gain and maintain credibility and legitimacy, and summarize relevant information simply and clearly (Bier, 2001b). Brevity, clarity, focus, candor, the use of cogent examples, and avoiding negative stereotypes of risk communicators may be crucial for communicating technical risks to nonspecialist audiences in a way that ensures the message is heard and absorbed rather than tuned out or dismissed (e.g., Byrd and Cothern, 2000). As discussed in Chapter 1, audience members generally respond not only (and sometimes not primarily) to technical information about risks, but also to message framing, the source of the information, and the emotional style and assumed motives of the presenter in assessing the credibility of risk communication messages (Chartier and Gabler, 2001).

Methods for Risk Management Decision Making

Formal methods of decision analysis and optimization for uncertain systems have been extensively developed in operations research and systems engineering and applied to both the design and the operation of complex engineering and industrial systems. Table 2.2 sketches some of the best-known frameworks for decision making when a decision maker's choice of act is related only probabilistically to resulting consequences.

Although some of the methods and algorithms mentioned in Table 2.1 are quite sophisticated, most share a simple common structure. The risk manager must choose from a set of feasible *controllable inputs* that influence a system's behavior. There are other facts and inputs (sometimes thought of as being selected by "Nature" or "Chance," and referred to as the *state* of the world) that cannot be directly selected by the risk manager but that also influence the system's behavior. The risk manager's acts and the state of the world together determine probabilities for different *consequences* (and for the system's next state, in systems dynamics and optimal control formulations of decision problems). Finally, a *utility function* represents preferences for different consequences (or time streams of consequences) produced by the system.

Various optimization algorithms and heuristics can be used to identify optimal (i.e., expected utility-maximizing) or approximately optimal *acts* (i.e., values of controllable inputs), given available information, or to identify optimal or approximately optimal *decision rules* (also called *policies*) that prescribe what acts to take based on the information available when decisions are made. Optimization algorithms for solving decision problems are constantly being refined and improved by ongoing research. Thus, it is worth Googling the topics and solution methods in Table 2.2 (leftmost and rightmost columns, respectively) before selecting a

framework and solution methods for a particular problem. As of this writing (in 2008), the links in Table 2.3 provide points of entry to the technical literature and solution algorithms.

An important principle that cuts across many solution techniques for complex decision models is that *adaptive random sampling of potential solutions* is computationally tractable and finds "good" (optimal or nearly optimal) solutions to many problems that are too hard (e.g., too computationally complex) to solve using exact methods. Monte Carlo methods and related meta-heuristics (such as genetic algorithms, simulated annealing, Tabu Search, or particle swarm optimization) can estimate and optimize the expected utility of different acts or decision rules even for large and complex stochastic systems. Much as the mean value of a variable in a population can be estimated accurately from a random sample, regardless of the uncertainties and complexities of processes that created the population distribution of the variable, so the expected utility of a decision rule, policy, or act (followed by future optimized acts, in dynamic settings) can often be estimated accurately using optimization algorithms that incorporate random sampling and adaptive improvement components.

Example: A Bounded-Regret Strategy for Replacing Unreliable Equipment

Setting: Suppose that a piece of machinery (such as a crane) that is being used in a major construction project breaks down frequently. The construction project must continue for 1,000 more days in order to meet a key deadline; after 1,000 more days, all activity (and further costs) on this effort will stop. Use of the current unreliable equipment costs \$1,000 per day in maintenance, repair, insurance, and overtime costs. The unreliable equipment will eventually break down completely; if this happens before the end of the project, it must then be replaced. The cost of replacement is \$1,200,000, and a new machine is highly reliable, costing \$0 per day, after it has been purchased, for the remaining duration of the project. Suppose that the current machine will last for an unknown number, T, of additional days before breaking down completely, where T is an unknown integer. Assume that not enough is known about the machine's remaining lifetime (e.g., from historical experience, accelerated life testing, reliability modeling, etc.) to assess a credible, well-calibrated probability distribution for T.

Problem: Assuming that the probability distribution for the remaining lifetime T is unknown, devise a decision rule for when to replace the machine (if at all) that is guaranteed to cost no more than twice as much as the least cost that could be achieved if T were known. (For simplicity, ignore discounting.)

Solution: If the lifetime T of the current machine were known, then the total cost of replacing the machine after $t < T$ days would be $\$1{,}000t + \$1{,}200{,}000$ if $T < 1{,}000$ days, in which case the optimal decision would be to replace the current machine immediately (set $t = 0$) and the minimized cost would be \$1,200,000; otherwise, if $T > 1{,}000$ days, then the unreliable machine should be used for the rest

of the project's duration, for a total cost of $1,000,000. Now consider the following myopic decision rule: *Wait until the machine fails completely, and then replace it.* The worst (most expensive) case is that the machine fails on day $T = 999$, in which case the total cost is $999,000 + $1,200,000 = $2,199,000. This is less than twice the optimal cost if T were known, $1,200,000 (for immediate replacement). If $T \geq$ 1,000 days, then the myopic decision rule yields the same optimal decision as if T were known. If $T <$ 1,000 days, then the myopic rule has less than twice the cost of the optimal decision if T were known.

Discussion: Although this example is trivial, it illustrates that analysis of decision rules is possible for some problems, even if uncertain quantities cannot be characterized by probability distributions. A number of nontrivial results on "online" decision and optimization procedures show that, in many sequential decision problems, it is possible to do almost as well on average, over the long run, using cleverly designed decision rules, as if the uncertain quantities (such as T in this example) were known.

Despite these advances in methods for decision analysis and optimization under uncertainty, in practice, formal decision analysis is seldom applied directly to make important risk management decisions. In part, this is because different participants may have different utility functions (which may be their own private information), different trade-offs among goals (e.g., minimizing average risk versus reducing inequities in the distribution of risks), and different tolerances for risk. In such cases, consensus utilities may not exist, and risk management decision making requires not only analysis and deliberation (Stern and Fineberg, 1996), but also negotiation and compromise.

Even when decision analysis is not directly applied, however, its conceptual framework is still useful for organizing analysis and deliberation (Apostolakis and Pickett, 1998), separating beliefs from preferences, and identifying and resolving relevant conflicts and/or uncertainties about facts and values. Byrd and Cothern (2000) and Cox (2001) further discuss individual and group decision-making processes and frameworks for risk management decision making.

Methods of Risk Management to Avoid

Well-informed and effective risk management (i.e., risk management that is likely to produce the desired consequences) requires considering *all* of the most important impacts – good and bad – that an intervention is likely to create. Unfortunately, many risk assessments exhibit a form of tunnel vision, focusing on one or a few narrowly defined issues (such as quantifying the reduction in risk that would be caused by contemplated actions) while ignoring other, possibly more important, ones, such as the risks that proposed risk management interventions might inadvertently *create* (Dowell and Hendershot, 1997; Bier, 1997). This represents a breakdown in sound risk assessment and risk management. Rational risk management requires considering and comparing the *total* consequences of the risk management decision options being evaluated. Risk characterization should therefore provide risk managers with

a balanced accounting of the adverse effects that a risk management intervention might *cause*, as well as of those that it might *prevent*.

Risk management recommendations that are based primarily on protecting the status quo or on beliefs about what might constitute "precautionary" actions should also be avoided if they do not explicitly identify and compare the probable consequences of alternative decision options. Decision analysis teaches that it is more effective to use quantitative information about the *probable consequences* of alternative interventions to eliminate dominated options, and to choose the best among those that remain. Heal and Kriström (2002) have argued on theoretical grounds that precautionary measures might make sense in situations where harm is irreversible, but their argument is based on, and consistent with, utility theory and real options theory.

Game-Theory Models for Risk Management Decision Making

Game theory has long been viewed by risk analysts as being of little relevance for practical risk management decision making. Several recent developments have started to change that view. These include not only increased interest in terrorism, homeland security, and critical infrastructure protection (which can be viewed as games between an attacker and a defender), but also increased interest in risk-informed regulation (which can be viewed as a game between a regulator and a regulated firm). As a result of such developments, game theory is becoming an important research tool in a variety of application areas related to risk.

Hausken (2002) has applied game theory to study the allocation of resources to ensuring component (and hence system) reliability in situations where different agents are responsible for the reliability of different components. In this situation, system reliability is viewed as a "public good." For example, agents responsible for the reliability of a component in a parallel system or subsystem might "free-ride" on investments in the reliability of other components in that system – e.g., postponing needed reliability enhancements in the hopes that some other agent will implement such improvements instead.

Recent work on reliability optimization (e.g., Levitin et al., 2001; Levitin and Lisnianski, 2003) attempts to identify cost-effective risk reduction strategies; for example, by optimizing physical separation of components that are functionally in parallel with each other, or by allocating physical protection to various hierarchies of a system (e.g., whether to harden the system as a whole, or individual components). However, the "threat" against which systems are to be hardened is generally taken to be static in this work.

Game-Theory Models for Security and Infrastructure Protection

Following September 11, 2001, there has been increasing interest in security, including the protection of public and commercial buildings, water supply systems, and

computer systems and software. Numerous researchers and practitioners have proposed the use of risk analysis in one form or another for homeland security (e.g., Paté-Cornell and Guikema, 2002; Garrick et al., 2004), especially for critical infrastructure (Haimes et al., 1998; Ezell et al., 2001; Apostolakis and Lemon, 2005). Most of this work is not formally game-theoretic. For instance, Paté-Cornell and Guikema discuss the need for periodic updating of the model and its input to account for the dynamic nature of counterterrorism but do not attempt to anticipate the effects of defensive investments on attacker strategies. Protection from intentional sabotage or terrorism differs from many other areas of risk management, because sabotage protection involves an intelligent adversary that can adapt in response to protective measures. Thus, reducing the vulnerability of some systems may cause adversaries to shift their attacks to other systems that have not yet been "hardened" to the same degree. Risk management in this context can be modeled as a game against an adversary or, conversely, as a game between defenders, because security investment by one defender can have either positive or negative externalities on the threats faced by other defenders (Kunreuther and Heal, 2003).

There is a large body of work on applications of game theory to security, much of it by economists (e.g., Frey and Luechinger, 2003; Arce et al., 2001; Enders and Sandler, 2004; Keohane and Zeckhauser, 2003; Lakdawalla and Zanjani, 2005). Much of this work is intended to inform policy-level decisions, e.g., by clarifying the relative merits of public versus private funding of defensive investments, or deterrence versus other protective measures. Recently, efforts have begun to focus more on operational risk management decisions, such as deciding how much defensive investment to allocate to particular assets (e.g., O'Hanlon et al., 2002), and have more of a risk analysis flavor (e.g., taking the success probabilities of potential attacks into account); see, for example, Bier et al. (2005) and Woo (2002).

Game-Theory Models of Risk-Informed Regulation

In health, safety, and environmental regulation, regulated parties often know more than regulators about the operations and risks of facilities. As a result, regulators may wish to provide incentives to encourage regulated parties to accurately disclose unfavorable information about their risks. Such situations can be modeled as games of asymmetric information between regulators and regulated parties. More widespread use of risk analysis results in regulatory decision making has the potential to both reduce risk and decrease compliance cost, by increasing management flexibility in determining how to achieve acceptable levels of safety (Bier and Jang, 1999). However, this approach has been slow to be adopted in practice, in part because of the inability of regulators to directly and accurately measure risk (Chinander et al., 1998) and because companies may have incentives not to disclose unfavorable risk information to regulators and or not to collect such information in the first place (Wagner, 1997).

Game-theoretic work in environmental economics to date (e.g., Heyes, 2000; Livernois and McKenna, 1999) has emphasized applications such as pollution

monitoring, in which a regulator can (with some effort) determine a firm's level of performance essentially with certainty, and firm performance can reasonably be modeled as binary (e.g., compliant with pollution-control regulations or not). Lin (2004) considers *risk-informed regulation*, in which regulators may not be certain to detect high risk levels even with substantial effort, and continuous risk levels may be more relevant than binary compliance status. Lin shows conditions under which it is still optimal (more efficient than traditional direct-monitoring regulation) for regulators to offer a loosened standard to firms that voluntarily disclose their risk levels.

Conclusions

This chapter has surveyed methods and concepts for PRA and decision making in engineered systems. Although the modeling of uncertain systems has been tremendously enabled by recent advances (such as Bayesian belief networks, with dependencies among inputs expressed via copulas), PRA still poses many challenges. Technical challenges remain in how best to construct useful (and at least approximately valid) models of systems and their environments from engineering knowledge and data, and in identifying optimal or near-optimal risk management policies. Communicating the results effectively and using them to guide improved decision making by multiple parties (e.g., teams of stakeholders) also poses practical questions that go beyond the framework of single-person decision theory. If the environment in which a system operates includes intelligent adversaries, then insights from novel methods (e.g., game-theoretic principles) may be needed to ensure that risk reduction strategies are effective and cost-effective (see Chapters 14 and 16). These challenges are likely to stimulate further advances in both the theory and practice of decision sciences for engineering risk analysis.

http://www.springer.com/978-0-387-89013-5

Risk Analysis of Complex and Uncertain Systems

Cox Jr., L.A.

2009, XXVIII, 436 p., Hardcover
ISBN: 978-0-387-89013-5