Quantifying uncertainty under a predictive, epistemic approach to risk analysis

S.Apeland
Stavanger University College, Norway

T.Aven
Stavanger University College, Norway

T.Nilsen
RF – Rogaland Research, Stavanger, Norway

July 2001

Abstract

Risk analysis is a tool for investigating and reducing uncertainty related to outcomes of future activities. Probabilities are key elements in risk analysis, but confusion about interpretation and use of probabilities often weakens the message from the analyses. Under the *predictive*, *epistemic* approach to risk analysis, probabilities are used to express uncertainty related to future values of observable quantities like the number of fatalities or monetary loss in a period of time. The procedure for quantifying this uncertainty in terms of probabilities is, however, not obvious. Examples of topics from the literature relevant in this discussion are use of expert judgement, the effect of so-called heuristics and biases, application of historical data, dependency and updating of probabilities. The purpose of this paper is to discuss and give guidelines on how to quantify uncertainty in the perspective of these topics. Emphasis is on the use of models and assessment of uncertainties of similar quantities.

Key words; risk analysis, Bayesian paradigm, predictive approach, uncertainty quantification, expert judgements, modelling

1 Introduction

Generally, in risk analyses with little or no relevant historical "hard data", engineering judgement needs to be used. How to use such judgement depends, however, on which probabilistic approach to risk and risk analysis that is applied. In order to outline a fully subjective approach, which is the main issue of this paper, we also refer to two other principally different approaches. These are the classical, statistical (frequentist) approach and the so-called combined classical and Bayesian approach, cf. Aven (2000a) and Apeland &

Aven (2000). Under both these approaches, focus in the risk analysis is on establishing estimates of presumed true statistical quantities, like probabilities and failure rates. If sufficient experience data are available, the estimation can be based purely on analysis of these under the classical, statistical approach. If data are scarce the combined classical and Bayesian approach allows use of engineering judgement to establish subjective uncertainty measures related to what the true values of the statistical quantities are. Thus, in both frameworks, uncertainty is related to two levels, i.e. the occurrence of future events *and* the true values of the probabilities and failure rates. The combined classical and Bayesian approach is also referred to as the probability of frequency framework, cf. Apostolakis & Wu (1993) and Kaplan (1992). In this framework the name probability is used for the subjective probability and frequency for the "objective", relative frequency based probability.

We find the above two frameworks for dealing with risk not very suitable for supporting decision-making, cf. Aven (2000a) and Apeland & Aven (2000). We are more attracted by a fully subjective approach to risk and risk analysis. Frameworks for such analysis have been developed recently, see Aven (2000a,b), cf. also Aven & Rettedal (1998), Apeland & Aven (2000). This approach is based on theory of subjective probability developed by de Finetti (1972, 1974), Savage (1962), Lindley (2000) and others. It is referred to as the *predictive*, *epistemic* approach. In this framework, probability is defined as a measure of uncertainty related to predictions of observable quantities, like production loss or volume of an oil spill in a given period of time. The uncertainty is purely epistemic, i.e. a result of lack of knowledge only. Both engineering judgement and historical "hard data" are used to quantify uncertainty related to the predictions.

It is, however, not obvious how one should proceed to quantify the uncertainty. Earlier research has discussed the general problem of establishing subjective probabilities, see for example Cooke (1991), Kahneman et al. (1982), von Winterfeldt & Edwards (1986). The purpose of this paper is to show how uncertainty can be quantified within the predictive, epistemic approach to risk and risk analysis. In particular some aspects related to modelling, the problem of interpreting and expressing dependency and how parametric distribution classes can be used in a practical setting , are discussed. In addition to addressing the contrasts between the predictive, epistemic approach and the two classical approaches, some characteristic differences to what seems to be the prevailing thinking when adopting a Bayesian view are discussed.

The remainder of this paper is organized as follows: In Section 2, we present the predictive, epistemic approach to risk and risk analysis. We discuss the problem of quantifying uncertainty within this framework in Section 3. Attention is drawn to the use of expert judgement, historical data and models. We use Section 4 to focus on one, in our view, particularly important issue in this context; the problem of quantifying uncertainty related to predictions of functions of so-called "similar" quantities, like the lifetimes of a system where repair brings the system to "as good as new" condition.

2 The predictive, epistemic approach to risk and risk analysis

In the predictive, epistemic approach to risk and risk analysis, focus is on predicting observable quantities, like occurrence, or not, of an accidental event, or the number of fatalities or the magnitude of financial loss in a period of time. Observable quantities express states of the "world", i.e. quantities of the physical reality or the nature; they are unknown at the time of the analysis, but become (or could become) known in the future. Let Y denote an unknown observable quantity and g the relationship between Y and a vector of unknown observable quantities on a more detailed level, $\boldsymbol{X} = (X_1, X_2, ..., X_n)$, such that:

$$Y = g(\boldsymbol{X}).$$

The function g, which we denote model, is deterministic. Thus if $\boldsymbol{X}$ were known, Y could be predicted with certainty, given the assumptions underpinning g. However, in most practical cases, such information is not available, and uncertainty related to the predictions has to be taken into account. In the predictive, epistemic approach uncertainty related to the future values of $\boldsymbol{X}$ is described through an uncertainty distribution $P(\boldsymbol{X} \leq \boldsymbol{x})$, $\boldsymbol{x} = (x_1, x_2, ..., x_n)$. This uncertainty is epistemic, i.e. a result of lack of knowledge. Then, uncertainty related to Y can be described through a distribution given by:

$$P(Y \leq y) = \int_{\{x: g(x) \leq y\}} dP(\boldsymbol{X} \leq \boldsymbol{x}).$$

The model g is a simplified representation of the world, and it is a tool used, allowing uncertainty of Y to be expressed through the distribution $P(\boldsymbol{X} \leq \boldsymbol{x})$, reflecting more or less of the available knowledge about the relationship between Y and $\boldsymbol{X}$. The selection of model is a choice made by the analyst, based on an overall consideration of the level of information, resources and analysis objectives.

Risk related to Y is described through the entire uncertainty distribution $P(Y \leq y)$. Summarizing measures such as the mean, the variance and quantiles are risk measures that can give more or less information about the risk. The background information used is reported to the decision makers along with the presentation of the risk measures. All probabilities are conditional on this background information.

An observable quantity represents a state of the world. Thus it includes also quantities that would have been better described as potentially observable. Consider for example the number of injuries. Provided that a precise definition of an injury has been made, there exists a correct value. The fact that there could be measuring problems in this case – some injuries are not reported – does not change this. The point is that the true number exists and according to the definition of an injury and if sufficient resources were made available, that number could be found. Another example that makes this point clear is the following. A production company produces units, say mobile telephones, and say that we focus on the portion of units that fail during a certain period of time and according to a certain definition of failure, among all produced units in one year for one particular type of mobile telephones. This portion is potentially observable, since it can be measured exactly if sufficient resources were made available. Of course in practice, that would not normally be done. Yet we classify it as observable.

Now, what about a relative frequency probability? Is such a probability observable? Well, the answer is both no and yes. As an example, consider a case where the system is a production facility and we focus on the occurrence of an accidental event (suitably defined) for a one year period. Then we can define a relative frequency probability by the portion of similar production facilities where this event occurs. If this population of similar production facilities is just a thought experiment, it is fictional, then this relative frequency probability is not observable. We will not be able to observe the probability in the future. If, however, such a population can be specified, the probability can be viewed as observable - it is possible to measure it in the future. However, such a population is difficult to imagine in this case unless that we extend the meaning of similar to include “every” type of production facility. Then we would be able to obtain a value of the portion of facilities where this event occurs, but that portion would not be very relevant for the system we study. What is a real population and what is a fictional population needs to be determined in each application. As a general rule we would say that real populations naturally exist when we deal with repeatable games, controlled experiments, mass produced units and large physical populations like human beings, etc. For the mobile telephone example above, a real population can be defined and the relative frequency probability, i.e., the portion of failed units, is an observable quantity. Focus in this paper is however on other types of applications, where the system is rather unique in the sense that we cannot find reasonably similar systems without doing thought constructions. We are particularly interested in complex man-machine systems for which operational and maintenance aspects affect performance of the systems.

As a final remark related to a quantity being observable, consider the volume produced in a gas production facility, measured in some unit during a certain period of time, say one year. For all practical purposes, saying that this volume is for example 2.5 would be sufficiently accurate. If you go into the details, the exact production volume could be somewhat difficult to define and measure, but thinking practically, and using the conventions made for this kind of measurements, the correctness of the measurement, for example 2.5, is not an issue. If it were, more precise measurements would have been implemented.

Quantifying the probability distribution $P(\boldsymbol{X} \leq \boldsymbol{x})$ means that we have the same degree of belief in the event $\boldsymbol{X} \leq \boldsymbol{x}$ as we have in drawing a favorable ball from an urn with $P(\boldsymbol{X} \leq \boldsymbol{x}) \cdot 100\%$ favorable balls under standard experimental conditions. To specify the distribution, alternative approaches can be used; direct use of historical data, analyst judgements using all sources of information, formal expert elicitation, as well as Bayesian analysis. We refer to Sections 3 and 4 for brief discussions on the use of historical data and the Bayesian approach. It is beyond the scope of this paper to discuss all these approaches in detail.

The predictive, epistemic approach is of course linked to the Bayesian paradigm, as for example described by Bernando and Smith (1994), Lindley (2000) and Barlow (1998), but it cannot be directly derived from this paradigm. In fact, our approach represents a rethinking on how to approach risk and uncertainty within a subjectivistic, Bayesian framework, to obtain a successful implementation in a practical setting. This relates to the use of parametric probability models; when to introduce such models and how to interpret the various elements of the models, cf. Section 4. In a risk analysis comprising hundreds of observable quantities, of which uncertainty must be assessed, a pragmatic view to the Bayesian approach is required in order to be able to conduct the analysis. A number of direct probability assignments needs to be carried out – the introduction of probability models where it is needed to specify prior (posterior) distributions of parameters, makes the analysis extremely complicated. A Bayesian updating procedure may be used for expressing uncertainty related to observable quantities, but its applicability is in many cases rather limited. In most practical cases we would not perform a formal Bayesian updating to incorporate new observations – rethinking of the whole information basis and modelling is required when we conduct the analysis at a particular point in time, for example in the prestudy or concept specification phases of a project. Furthermore, we make a sharp distinction between probability and utility in contrast to the Bayesian school, which sees these two concepts as inseparable.

The framework or basis for risk analysis that is referred to as the predictive, epistemic approach is much more than subjective probabilities and Bayesian inference. It relates to fundamental issues like how to express risk and uncertainty, how to understand and use models, and how to understand and use parametric distribution classes and parameters in a risk analysis setting. Compared to the more traditional approaches for risk analysis in the engineering community, the predictive, epistemic approach gives different answers to all these issues.

We define probability as a measure of uncertainty, which means reference to a certain standard such a drawing a ball from an urn. We do not link the definition to gambling situations with prizes and decision making, as is often done in the literature: When a person says that a probability of occurrence of an event *A*, is *P*(*A*), he implies that he is willing to pay (assuming linear utility for money) *P*(*A*) now in exchange for 1$ later if event *A* occurs. To us such a definition of probability is not so attractive, it complicates the assignments as it introduces more dimensions; decision making involving money.

3 Review and discussion of some basic elements in the quantification process

In this section we focus on three basic elements in the quantification process, which in practice are closely interrelated, namely: expert judgement of probabilities, use of historical data and application of models. The elements are discussed in the perspective of the predictive, epistemic approach, referred in Section 2. The review is not exhaustive, but in our view it captures some essential issues.

3.1 Expert judgement

Traditionally, risk analysts have used experience data extensively as information basis for frequency and probability assessments. A large number of databases are established for this purpose and a major subtask of practical analysis is gathering of data from these and from other sources. An important factor often causing relevant data to be scarce is that risk analyses typically deal with rare events, i.e. long system observation time is necessary. Since the systems studied in addition often represent new concepts and arrangements, little or no experience exists, and use of expert judgement is enforced.

The objection among many risk analysts regarding the introduction of subjectivity in risk analysis in terms of probabilities provided by experts, is mainly related to the association of such judgements with superficiality and imprecision. Experts' ability to express their uncertainty in terms of probabilities and whether such statements constitute a sufficiently credible basis in a decision-making context is discussed extensively in the literature. A large portion of this work is centered on so-called heuristics and biases.

The term "heuristics" is normally referred to as cognitive principles used by experts to evaluate quantities associated with uncertainty. As Kahneman et al. (1982) argue, probability statements from experts is also governed by heuristic principles. The "biases" are specific mechanisms causing systematic errors in judgements based on the various heuristics. Other examples of literature discussing heuristics and biases are Morgan & Henrion (1990), Cooke (1991), Wright & Ayton (1994) and Watson & Buede (1987).

In general, let us assume that the expert decides on a probability number based on a weighing of the circumstances that influence the generation of the event in question, and that the description of heuristics in the literature provides a partial explanation of this process. In practice we may then interpret the biases as mechanisms that:

1. lead to inconsistency between the expert's system knowledge and his evaluation of uncertainty (perceived uncertainty) or

2. introduce a disparity between the uncertainty as perceived by the expert and the probability figure he eventually puts down on paper (quantified uncertainty).

Under a combined classical and Bayesian approach to risk, quantification of probabilities means estimation of underlying, true probability numbers, and expert judgement is used in this process to express the expert's best evaluation of what the true values are. In this context the biases must be seen as a source to uncertainty in the probability estimates. Different methods have been developed for dealing with the variation among judgements from several experts about a given quantity, giving a best-fit statement from these, see e.g. Apostolakis & Mosleh (1986) and for calibration of experts, ref. e.g. Cooke (1991). Such methods may capture and reduce some effects of biases, but they do not offer a full uncertainty treatment of subjective probability statements. We have to acknowledge that not all bias effects will be reflected consistently in the uncertainty distributions specified for the underlying true probabilities. The results of the analysis, its goodness and importance, have to be evaluated in this perspective. If the decision-makers believe that the analyst has based his analysis on biased expert judgements, this will weaken the message from the analysis.

We face the same problem when considering expert judgement of probabilities related to observable quantities under the predictive, epistemic approach, but the setting is somewhat different, since uncertainty does not apply to the probabilities to be judged. In a decision-making context, the goodness or the weight of the probability statements is ruled by the state of information taken into account in the actual evaluations behind these. The effect of both types of mechanisms, (1) and (2) above, can be considered as a misrepresentation of this basis information, i.e. the information is not fully and consistently reflected in the probability number, and thus the goodness or weight it will be ascribed in a decision-making setting will be reduced. Consequently, if one believes that one or more biases might influence an expert probability statement, one should attempt arranging the judgement process in a manner that reduces the effect of the biases, for example calibration, involving more than one expert or some form of iterative elicitation procedure. Such arrangements should increase the confidence in the quantification process as a whole from a decision-maker's point of view.

In this relation it should also be mentioned that others argue that human beings are good at quantifying uncertainty on basis of experience and question the existence and influence of cognitive biases, see for example Cosmides & Tooby (1992).

Another basic problem related to expert judgement is related to assignment of small probabilities, see for example Lindley (1985), Lichtenstein et. al (1978). Since many applications of risk analysis deal with catastrophic events, causal factors that are considered only theoretically possible or limiting to the unlikely may be of interest. In general the problem of superficial and imprecise judgements becomes more distinctive when dealing with little likely events, with probabilities typically $<10^{-3}$. To some extent the situation may be improved by applying assessment aids such as a set of standardized reference events with commonly agreed probabilities, for example drawing balls from an urn (ref. Section 2), that may be compared with the event under consideration, or graphical tools like the so-called "probability wheel", see De Groot (1970). However, faced with rare events, the expert simply has difficulties to relate his uncertainty to low probability levels and differ cognitively between numbers such as 10^{-5} and 10^{-6}. For a given expert one might then question the meaning of such numbers directly assigned in a subjective approach under which probabilities express the expert’s uncertainty related to future outcomes.

Although all experts seem to have a limit probability level under which expressing uncertainty in numbers becomes difficult, practice shows that this limit can be lowered by training, see Berg Andersen et. al (1997). Through repeatedly facing the problem of assigning probabilities to rare but known events, discussing the causal factors and comparing their likelihood, the expert familiarizes with this to him often unusual way of thinking. The expert will gradually feel more comfortable with applying smaller numbers, but still training alone will hardly solve this problem. It seems like we must accept that the area of application of probability judgement has a boundary at the lower and upper ends of the probability scale.

3.2 Use of historical data

In the predictive, epistemic approach to risk and risk analysis focus is on quantifying epistemic uncertainty related to observable quantities. Thus, uncertainty only exists on one level in this approach; related to what is going to happen in the future. To quantify this uncertainty, all available information, i.e. available historical data and engineering judgements, is used. If sufficient relevant data is available, we can use this information directly in the probability assignment process. To illustrate, consider the following example. An analyst group wishes to express uncertainty related to the occurrence of an event in a fault tree. Suppose that the observations show three "successes" out of 10. Then we obtain a probability of 0.3. This is our (i.e. the analyst's) assessment of uncertainty related to the occurrence of the event.

This method is appropriate when the analyst judges the observational data to be relevant for the uncertainty assessment of the event, and the number of observations is large. What is considered sufficiently large, depends on the setting. As a general guidance, we find that about 10 observations is typically enough to specify the probability probabilities at this level using this method, provided that not all observations are either "successes" or "failures". In this case the classical statistical procedure would give a probability equal to 1 or 0, which we would normally not find adequate for expressing our uncertainty about the event. Other procedures then have to be used, expert judgements or Bayesian analysis, see Section 4.

Note that the probability assigned following this procedure, is not an estimate of an underlying true probability as in the classical setting.

Next, suppose that we consider the problem of assigning the distribution of a quantity Y taking values in the set of the real numbers. Suppose observational data are available and that the analyst judges these data to be relevant for the uncertainty assessment of Y. Furthermore, suppose the number of observations is large. Then we can proceed along the same lines as for the event case discussed above also for specification of $P(Y \leq y)$ as the portion of observations having values less than or equal to y. Thus $P(Y \leq y)$ is given by the empirical distribution function in the classical statistical set-up. In most cases we would prefer to use a continuous function for $P(Y \leq y)$ - as this is mathematically convenient. Such a function is obtained by a fitting procedure where the empirical distribution is approximated by a continuous function, for example a normal distribution function. Classical statistical methods for fitting a distribution function to observed data, is the natural candidate for such a procedure. As for the event example above, note that we use classical inference just as a tool for producing our uncertainty distribution for Y, not for estimating an underlying true distribution function for Y.

The same type of procedures can be used for more complicated situations involving censoring, for example when the interesting quantities are time to failures of a unit and we do not exactly know all the failure times.

As for the event case this procedure works with a high number of observations. What if the number of observations is not large, say 6, or what if most of the observations are zero, say, and we are most concerned about a possible large value of Y, i.e. the tail of our uncertainty distribution of Y? Clearly, in these cases it is problematic to use the above procedure, as the information given by the data is limited. Other procedures should then be adopted, expert judgements and Bayesian analyses, cf. Section 4.

3.3 Modelling

A model is a simplified representation of a real world system. The general objective of developing and applying models in this context is to arrive at risk measures based on information about related quantities and a simplified representation of real world phenomena. Detailed modelling is required to be able to identify critical factors contributing to risk and evaluate the effect of risk reducing measures. The simplified form makes models suitable for analysis, and in model construction this property is traded off against the needs for complexity required for producing sufficiently credible results. Typical factors governing the selection of models are the form in which and at what level of detail system information is available, the resources available in the specific study and whether focus is on the overall risk level or comparing decision alternatives, see for example Kaplan & Burmaster (1999). In general the advances seen within computer technology have improved the conditions for analyzing complex models.

It is beyond the scope of this paper to give a detailed review of the great variety of modelling approaches used in practical risk analysis, which range from traditional tools such as fault and event trees and methods based on limit-state functions to less formal and improvised approaches. In order to clarify the basis for the quantification process, we primarily, call attention to the distinction between models of the actual system in question, i.e. representation of cause-coherence relationships, and so-called stochastic models (probabilistic models or uncertainty models), reflecting uncertainty related to the system. Both categories are frequently denoted models among risk analysts and also discussions in the literature are often unclear with respect to this distinction. However, if the latter category, which includes the probability figures and the mathematical expressions behind these, is considered to be part of the model structure, the real world system is ascribed inherent stochastic elements. Consequently we may argue that in this case the risk analysis must be seen as a tool for estimating aleatory uncertainty under a classical interpretation of the risk concept.

From the point of view of the predictive, epistemic approach, as stated in Section 2, all uncertainty is epistemic. Thus, since models are used to reflect the real world, they only include descriptions of relationships between observable quantities, i.e. the former of the two categories mentioned above, while probabilistic expressions reflect uncertainty or lack of knowledge related to the values of such quantities. In the risk analysis we assess the uncertainty related to selected quantities. Modelling is a tool allowing us to express the uncertainty in the format found most appropriate to fulfil the objectives of performing the analysis.

Risk analysis is a tool for comparing decision alternatives at some level, i.e. to aid the selection between alternative system arrangements, operational set-ups or strategies. Still, within many areas of application, the industry is reluctant to fully utilize the potential of risk analysis in this type of decision processes. Examples are planning of wells and planning against process leaks, ref. Nilsen et al. (2000). We believe that this reluctance is partly explained by the widespread confusion among risk analysts and decision-makers with respect to the theoretical foundation for risk analysis, including modelling and use of expert judgement as information basis. In our view the selection of models to a great extent governs how far the quantification process will satisfy the objective of the risk analysis. Below we discuss important issues related to selection of modelling approach.

At an overall level the quantification process of risk analyses under the predictive, epistemic approach can be summarized in three steps:

1 Represent the outcomes of interest by conditioning on observable system and component states in a description of the system cause-coherence by modelling.
2 Express uncertainty with respect to future states in terms of probabilities based on available information, i.e. expert judgement and experience data.
3 Propagate the uncertainty statements in the model to obtain the resulting description of uncertainty related to the predicted quantities.

Presupposing that Step 3 is performed in accordance with the rules of probability calculus, the remainder of this subsection relates to Steps 1 and 2.

Experience data applied in risk analysis are often given on the form of the number of occurrences of an outcome y out of a number of trials n, registered during similar activity in the past. However, the "similar activity" often comprises a mix of experiences representing a non-existing average system. This makes it hard to differentiate with respect to the decision alternatives at hand. Especially it becomes hard to defend alternatives that involve new technology not represented in the data. In addition y often represents events at relatively rough causal level, which makes it difficult to differentiate through modelling. In other words, the data often fail to support a system a representation that is sufficiently detailed to capture the points of distinction in the fundamental properties of the alternatives that influence the performance measures to be considered.

Even if the data do differentiate between the main decision alternatives, they do not necessarily reflect system-specific information existing in other forms, e.g. related to local operating conditions or technical or organizational measures implemented, expected to impact the risk level and thus considered essential by the decision-makers. Such additional system information usually exists as a mix of detailed system specifications and expert knowledge. Typically it is associated with a set of factors $\boldsymbol{x}$ on a more detailed system level than the outcome y, for which the data exist, but influence the occurrence y, e.g. through a causal relation $f(\boldsymbol{x})$. This information is thus not directly applicable in the existing models. Attempts to transform the information to a relative frequency of y, which would allow application of the established model, in most cases involve great cognitive challenges. The result is usually experts restraining to provide specific numbers and an impression of superficiality among the decision-makers and risk analysts. In many decision processes for which risk analysis is not formally required as documentation, this situation leads to the risk analysis being neglected by the decision-makers, if performed at all.

To improve the situation described above, further system modelling to include the quantities $\boldsymbol{x}$, is necessary. Differentiation between the decision alternatives is made possible through a more detailed system representation. By utilizing knowledge of causal factors in system configuration and physics, the modelling process makes the risk analyst free to define the quantities to represent model input based on a total evaluation of the state of system information. Referring to the above generic example this implies including $f(\boldsymbol{x})$ in the model structure if the information available about $\boldsymbol{x}$ together with the system information in $f(\boldsymbol{x})$ itself is considered to represent the system better than the experience data available for y.

The freedom to specify the quantities, for which uncertainty statements are to be gathered, also improves the conditions for specifying subjective probabilities, ref. the discussion in Section 3.1. For example the problem related to small probability numbers can in many cases be avoided through modelling, see for example Raiffa (1968), Armstrong (1985). Let us assume that we are interested in quantifying uncertainty related to whether the event A will occur in a given period as input to a risk analysis. If A is judged to be little probable and the experts has difficulties in relating to it quantitatively, the problem may be handled by moving focus to observable quantities on a lower causal level, associated with a higher probability level. For example, if A is judged dependent on the occurrence of the conditions B and C, the expert may express his uncertainty with respect to these events instead, and a probability of A, $P(A)$, may be assigned e.g. by $P(A) = P(B) \cdot P(C|B)$. Another alternative is to formulate A by a limit-state function, i.e. A occurs if $f(\boldsymbol{X}) < 0$, see also structural reliability analysis (SRA) textbooks, for example Toft-Christensen & Baker (1982) or Krenk et al. (1986). The probability $P(A)$ can then assessed by expressing uncertainty about the event through the probability distributions of the observable quantities $\boldsymbol{X}$.

Another essential factor for describing uncertainty through expert probability judgement is the expert's interpretation of the quantities for which probability statements are to be gathered. The degree of correspondence between the definition of the quantities in the analysis and the expert's perception of these according to his experience, will govern the ability to relate quantitative uncertainty statements to the quantities. This is in line with the relation between state of information, goodness and credibility of probability statements described in Section 3.1. If keeping this in mind in the modelling process sufficient correspondence can be achieved.

In summary we can say that under the predictive, epistemic approach, modelling is a tool for identifying and expressing uncertainty and thus also a means for potentially reducing uncertainty. The uncertainty can be identified by including more system-specific information in the analyses, in terms of an expanded information basis for uncertainty statements and in terms of the model structure itself. Furthermore, modelling adds flexibility to the risk analyses since it allows us to express uncertainty in the format found most appropriate to obtain the objectives of the analysis.

A topic closely related to use of models, which is widely discussed in the literature, is model uncertainty. Several approaches to interpretation and quantification of model uncertainty are proposed, see for example Mosleh et. al. (1994). If the predictive, epistemic approach is adopted, a model $Y = f(\boldsymbol{X})$ is a purely deterministic representation of causal mechanisms judged essential by the analyst. It provides a framework for mapping uncertainty about the observable quantity of interest, Y, from expressions of epistemic uncertainty related to the observable quantities $\boldsymbol{X}$, and does not in itself introduce additional uncertainty. In this setting the model is merely a tool judged useful for expressing partial knowledge of the system. However, an approach based on weighting of confidence in candidate causal mechanisms, that are considered to possibly govern the future outcome of Y, is not in conflict with this basis. Furthermore, it is important and in line with the framework to evaluate the goodness or appropriateness of specific models to be used in a specific risk analysis and decision context. Modelling uncertainty related to the predictive, epistemic approach to risk analysis is further discussed in In Nilsen & Aven (2001).

4 Expressing uncertainty related to similar observable quantities

The principles discussed in this section will be illustrated through the following case. Consider a system that, when failing, is repaired to a condition which is judged "as good as new" (alternatively we may think of the system being replaced by a new and identical one upon failure). We ignore the repair (replacement) times. An analyst group is established to predict the time to the nth system failure, Y. Let $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ denote the consecutive times between failures. Then we can link Y and $\boldsymbol{X}$ by writing

$$Y = g(\boldsymbol{X}),$$

where the function g is given by $g(\boldsymbol{X}) = X_1 + X_2 + ... + X_n$. The quantities X_i are "similar", meaning that the analyst group has no information that implies that the group's uncertainty related to the time between failure should be different from one X_i to the other. The aim is to express an uncertainty distribution $P(Y \leq y)$. This distribution can, of course, be established by specifying a simultaneous distribution $P(\boldsymbol{X} \leq \boldsymbol{x})$, $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, and normally, similar quantities will always be judged dependent, as the value of one or more observations give knowledge about the other quantities. In a practical setting, however, the analyst group needs to simplify, and the issue here is how to do this. In the following we will discuss alternative approaches relevant in this context.

4.1 Independence

The simplest way of establishing $P(Y \leq y)$ through a simultaneous distribution $P(\boldsymbol{X} \leq \boldsymbol{x})$ is to judge the X_is independent. For example, if the exponential distribution

$$P(X_i \leq x) = 1 - e^{-\lambda x}$$

with a fixed value of λ is used to express uncertainty related to the future values of each X_i, it follows directly that Y is gamma distributed with parameters n and λ. The use of the exponential distribution does not mean that the existence of a true failure rate λ is assumed, the exponential distribution with the quantity λ is simply a function in a mathematical class of functions that is considered suitable for expressing uncertainty related to X_i.

Judging the quantities independent simplifies the analysis, but when can such a simplification be justified? Before we enter this discussion, the more general Bayesian approach to dealing with the dependency problem will be discussed.

4.2 Parametric approach

The standard Bayesian alternative to a direct specification of a simultaneous distribution over the X_is would be a parametric probability model approach, where we specify a probability model $P(X_i \leq x \mid \lambda)$, for example the exponential, and assign a subjective uncertainty distribution H over the parameter λ (the prior or posterior distribution of λ). Combining these two elements we derive at the so-called predictive distribution of X_i, which is given by

$$P(X_i \leq x) = E[P(X_i \leq x|\lambda)] = \int P(X_i \leq x|\lambda) dH(\lambda). \tag{1}$$

The predictive distribution for Y, $P(Y \leq y)$, is given by

$$P(Y \leq y) = E[P(Y \leq y|\lambda)] = \int G_n(y, \lambda) dH(\lambda), \tag{2}$$

where $G_n(y, \lambda)$ is equal to the cumulative Gamma distribution function with parameter n and λ. Thus, the X_is are considered independent, *conditioned*, however, on the value of λ. The starting point in the Bayesian setting is that the X_is are judged to be exchangeable random quantities, i.e. their joint distribution function is invariant under permutations of coordinates, cf. Bernardo and Smith (1994). The concept of exchangeability gives a mathematically precise way of defining what is meant by similar units. The parameter λ is interpreted as the long run number of failures per unit of time when considering an infinite number of lifetimes X_i, and the uncertainty distribution H expresses our uncertainty about this value. In this way uncertainty about X_i is divided into two, the stochastic (aleatory) uncertainty expressed by

$P(X_i \leq x | \lambda)$ and $P(Y \leq y | \lambda)$, and the state-of-knowledge (epistemic) uncertainty expressed by H. The distribution $H(\lambda)$ is updated if new information becomes available, using Bayes formula.

Note that in this approach, fictional parameters, based on thought experiments, would be introduced and uncertainty of these assessed. Thus, from a practical point of view, there is not much difference between this framework and the "combined classical and Bayesian approach" discussed in the introduction. Of course, Bayesians would not speak about true, objective risks and probabilities, and the predictive form is seen as the most important one. However, although the role of parametric inference is typically that of an intermediate structural step, in practice, Bayesian analysis is often seen as an end-product of a statistical analysis. The use and understanding of probability models gives focus on limiting values of quantities constructed through a thought experiment, which are very close to the mental constructions of probability and risk used in the classical relative frequency approach.

A parametric procedure given by (1) and (2) can also be applied when adopting the predictive, epistemic framework. However, the interpretations of $P(X_i \leq x | \lambda)$ and $H(\lambda)$ would then be different. The probability distribution $P(X_i \leq x | \lambda)$ is a candidate for our subjective probability for $X_i \leq x$, such that if λ is chosen, we believe that about $(1 - e^{-\lambda x})100\%$ of the X_is will take a value equal to, or less than x. The probability distribution $H(\lambda)$ is a confidence measure, reflecting for a given value of λ, the confidence we have in $(1 - e^{-\lambda x})$ for being able to predict the number of (percentage) X_is taking this value. Following this interpretation, note that $P(X_i \leq x | \lambda)$ is not a model, and $H(\lambda)$ is not an uncertainty measure. We refer to this as the confidence interpretation.

If a suitable infinite (or large) population of "similar units" can be defined, in which X_i, $i = 1, 2, \ldots, n$, belong, then the above standard Bayesian framework applies as the parameter λ is an observable quantity as defined in Section 2. Then $H(\lambda)$ is a measure of uncertainty and $P(X_i \leq x | \lambda)$ truly is a model - a representation of the portion of lifetimes in the population having the property that they are less than or equal to x. We may refer to the variation in this population, modelled by $P(X_i \leq x | \lambda)$, as aleatory uncertainty, but still the uncertainty related to the values of the X_is is seen as a result of lack of knowledge, i.e., the uncertainty is epistemic. This nomenclature is in line with the basic thinking of e.g. Winkler (1996), but not with that commonly used in the standard Bayesian framework; see above.

Note that the confidence interpretation can be used also if a population cannot be defined without performing mental constructions.

For the example considered, where a system is undergoing repair, it would in most cases be difficult to specify a large population of the world of similar times between failures. So we would use the confidence interpretation.

Adopting the confidence interpretation, we see that that the predictive, epistemic approach is based on the same assessment procedures, (1) and (2), as the standard Bayesian. So are the difference between these two approaches just semantics? No, we are discussing fundamental issues related to how to think when conducting risk analyses, and the predictive, epistemic approach gives a new structure for this thinking. In some cases, this structure would give the same mathematics as the standard Bayesian, in other cases it would prescribe other (simpler) procedures, for examples direct probability assignments and the use of independence where the standard approach introduce probability models with parameters and specification of uncertainty distributions over these. The mathematical procedures of the Bayesian paradigm, as such, are not being challenged by the predictive, epistemic approach, but the use – when we should adopt a certain procedure, for example (1) and (2) - and the interpretation of the elements of the framework.

4.3 Evaluation and recommendations

A simple approach is generally the most attractive one. Thus, we prefer independent marginal uncertainty distributions $P(X_i \leq x)$ (with fixed parameters) in situations where this can be justified. This justification is obtained if the analyst group's confidence in a particular marginal uncertainty distribution (parameter value) related to each X_i is considered so good that the group does not see the need for changing its view related to the uncertainty distributions based on possible new observations of X_is. This means, for example, that if uncertainty related to X_i is expressed through the exponential distribution with $1/\lambda = 50$, the analyst group would not change its uncertainty related to X_i, $i = 6, 7, \ldots, n$, even if the observed values of $X_1, X_2, \ldots, X_5$ are 17, 20, 25, 24, 28. The increased failure tendency that may be indicated by the observations will, in the case of assuming independence, not be reflected in the uncertainty distributions related to $X_6, X_7, \ldots, X_n$; the analyst group's initial information ("hard data" and knowledge) is considered so good that, even based on the observed values of $X_1, X_2, \ldots, X_5$, the group still has a strong belief in values of $X_6, X_7, \ldots, X_n$ according to the exponential distribution with $1/\lambda = 50$. Of course, observing $X_1, X_2, \ldots, X_5$ will give additional information when predicting $X_6, X_7, \ldots, X_n$, which thus implies dependency between $X_1, X_2, \ldots, X_5$ and $X_6, X_7, \ldots, X_n$. The question is whether the additional information is considered so significant that it justifies the increased degree of complexity obtained by treating the X_is as dependent in the analysis, or if the background information dominates the information when we add one or more observations of the X_is, the latter implying that an independent approach can be used.

The parametric approach of Section 4.2 is preferably used when the analyst group has rather vague information about the quantity of interest, and thus needs to take account for a range of possible characteristics, for example failure propagation processes. Note that also this approach means some type of simplification: if we use the exponential distribution together with $H(\lambda)$, to express uncertainty related to X_i, we ignore dependencies that are "in conflict with" the exponential distribution, for example introduced by observing increasing failure tendencies, as illustrated above. Independence as in Section 4.1 is, however, a stronger simplification.

Other procedures can also be formulated for establishing the simultaneous distributions of the X_is. An interesting example is the so-called operational Bayesian approach, see Barlow (1998), which uses indifference principles for exchangeable random quantities. The approach is parametric Bayesian as described above, but the parameter is expressed through a function of observations from a finite population of units. An expression for the joint distribution of the X_is is established, and the standard Bayesian setting is derived by letting the number of observations converge to infinity. The joint distribution in the finite population case, provides a reference for studying the "goodness" of the independence approximation.

Ad-hoc procedures can also be defined. Consider the following simple one. If the exponential distribution is used to express uncertainty related to the X_is, and minimal information is available, it can be evaluated as having the same value as observing 1 outcome of the X_is, having the value $M_1 = 1/\lambda$. Then, observing X_1 implies updating λ to $\lambda = [2 / (M_1 + X_1)]$. Similarly, if the initial information is considered good, it can be evaluated as having the same value as observing for example 10 outcomes of X_is, having a statistical mean $M_{10} = 1/\lambda$, and λ is updated to $\lambda = [11 / \{(M_{10} \cdot 10) + X_1\}]$. The difference in these two situations is, of course, the magnitude of influence the observed value of X_1 has on the updated uncertainty distributions. Such a process can be simulated, for example through a Monte Carlo simulation. We draw X_1 according to the initial uncertainty distribution $P(X_i \leq x)$, and update the distribution $P(X_i \leq x)$, $i = 2, 3, ..., n$, as described above. Then, X_2 is drawn according to the updated distribution, this distribution is updated again, and so on. In general we would not recommend the use of such an ad-hoc procedure, it is too "arbitrary". The parametric approach is preferred as it is well-established and has a solid basis. However, if the background information is strong, we would use independence. Since independence is a special case of the parametric approach, we could of course limit ourselves to this latter framework; if we have no uncertainty with respect to the true value of the parameter, we obtain independence. Our point is, however, that we should think of independence as a starting point; when can it be justified as a reasonable approximation? If it cannot be justified, apply the more complex parametric approach.

Acknowledgements

The authors are grateful to Bjørn Auestad, Jan Terje Kvaløy, and the referees for valuable comments and suggestions to an earlier version of this paper, and to the Norwegian Research Council for financial support.

References

Apeland, S., Aven, T. 2000. Risk based maintenance optimization: foundational issues. *Reliability Engineering and System Safety* **67** 285-292.

Apostolakis, G. 1990. The Concept of Probability in Safety Assessments of Technological Systems. *Science* **250** 1359-1364.

Apostolakis, G., and Mosleh, A. 1986. The assessment of probability distributions from expert opinions with an application to seismic fragility curves. *Risk analysis* **6** 447-461.

Apostolakis, G., and Wu, J.S. 1993. The interpretation of probability, De Finetti's representation theorem, and their implications to the use of expert opinions in safety assessment. In Reliability and Decision Making, Barlow, R.E. and Clarooti, C. A. (eds.), Chapman & Hill, London, 311-322.

Armstrong, J.S. 1985. *Long-range forecasting: From crystal ball to computer*. New York: Wiley.

Aven, T. 2000a. Risk analysis – A tool for expressing and communicating uncertainty. Proceedings of *European Safety and Reliability Conference* (*ESREL*), Edinburgh, 21-28.

Aven, T. 2000b. Reliability analysis as a tool for expressing and communicating uncertainty. In *Recent Advances in Reliability Theory: methodology, practice and inference* 23-28. Boston: Birkhauser.

Aven, T., Pörn, K. 1998. Expressing and interpreting the results of quantitative risk analyses; Review and discussion. *Reliability Engineering & Systems Safety* **62** 3-10.

Aven, T., Rettedal, W. 1998. Bayesian frameworks for integrating QRA and SRA methods. *Structural Safety* **20** 155-165.

Barlow, R.E., 1998, *Engineering Reliability*, SIAM, Philadephia.

Barlow, R.E. and Clarooti, C. A.,1993, Reliability and Decision Making, Preface, Chapman & Hill, London.

Berg Andersen, L., Nilsen, T, Aven, T., Guerneri, A. 1997. A practical case of assessing subjective probabilities – a discussion of concepts and evaluation of methods. Proceedings of *European Safety and Reliability Conference* (*ESREL*), Lisbon, 209-216.

Bernardo, J.M., Smith, A. 1994. *Bayesian Theory*. Chichester: Wiley & Sons.

Cooke, R.M. 1991. *Experts in uncertainty: Opinion and Subjective Probability in Science*. New York: Oxford University Press.

Cosmides, L., Tooby, J. 1992. *Cognitive Adaptions for Social Exchange*. Oxford: Oxford University Press.

de Finetti, B. 1972. *Probability, Induction and Statistics*. New York: Wiley & Sons.

de Finetti, B. 1974. *Theory of Probability*. New York: Wiley & Sons.

de Groot, M.H. 1970. *Optimal Statistical Decisions*. New York: McGraw-Hill.

Kahneman, D., Slovic, P. & Tversky, A. (Eds.) 1982. *Judgement under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kaplan, S., Burmaster, D. 1999. Foundations: How, when, why to use all of the evidence. Workshop proceedings of *When and how can you specify a probability distribution when you don't know much? Risk Analysis* **19**(1) 55-62.

Kaplan, S., 1992, Formalisms for handling phenomenological uncertainties: the concepts of probability, frequency, variability, and probability of frequency. *Nuclear Technology*, **102**, 137-142.

Kyburg, H.E. Jr., Smokler, H.E. 1980. *Studies in subjective probability*. New York: Robert E Krieger Publishing Company.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., Combs, B. 1978. Judged frequency of lethal events. *J. Exp. Psychol.: Human learn. memory* **4** 551-578.

Lindley, D.V. 1985. *Making Decisions*. New York: Wiley & Sons.

Lindley, D.V. 2000. The philosophy of statistics. *The Statistician* **49** 293-337.

Madsen, H.O., Krenk, S., Lind, N.C. 1986. *Methods of Structural Safety*. London: Prentice Hall.

Morgan, M.G., Henrion, M. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis*. New York: Cambridge University Press.

Mosleh, A., Siu, N., Smidts, C., and Lui, C. 1994. Model Uncertainty: its Characterization and Quantification. Proceedings of *Workshop 1, Advanced Topics in Risk and Reliability Analysis*, prepared for the Nuclear Regulatory Commission, NUREG/CP-0138.

Nilsen, T. and Aven, T. 2001. Models and model uncertainty in the context of risk analysis. Submitted for publication in *Reliability Engineering and System Safety*.

Nilsen, T., Aven, T, Jakobsen, G.S. 2000. Requirements and principles for cause analysis in QRA – With application. Proceedings of *European Safety and Reliability Conference* (*ESREL*), Edinburgh, 641-646.

Raiffa, H. 1968. *Decision Analysis*. Reading MA: Addison-Wesley.

Savage, L.J. 1962. Subjective probability and statistical practice. In *The Foundations of Statistical Inference*. New York: Wiley & Sons.

Singpurwalla, N.D. 1988. Foundational issues in reliability and risk analysis. *Society for Industrial and Applied Mathematics* **30** 264-281.

Toft-Christensen, P., Baker, M.J. 1982. *Structural Reliability Theory and its Applications*. New York: Springer-Verlag.

von Winterfeldt, D., Edwards, W. 1986. *Decision Analysis and Behavioral Research*. Cambridge: Cambridge University Press.

Watson, S.R., Buede, D.M. 1987. *Decision Synthesis: The Principles and Practice of Decision Analysis.* New York: Cambridge University Press.

Winkler, R.L. (1996) Uncertainty in probabilistic risk assessment, *Reliability Engineering and System Safety*, **54**, 127-132.

Wright, G., Ayton, P. 1994. *Subjective Probability*. New York: Wiley & Sons.