Comparing Probabilistic and Fuzzy Set Approaches for Design in the Presence of Uncertainty

by

Sophie Qinghong Chen

A dissertation submitted to the faculty of

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Aerospace Engineering

Dr. Efstratios Nikolaidis, Chairman

Dr. Harley H. Cudney, Co-chairman

Dr. Raphael T.Haftka

Dr. Hugh F. Vanlandingham

Dr. Mahendra P. Singh

August 7, 2000
Blacksburg, Virginia

Keywords: Probability, fuzzy set, design under uncertainty

COMPARING PROBABILISTIC AND FUZZY SET APPROACHES FOR DESIGNING IN THE PRESENCE OF UNCERTAINTY

By

Sophie Chen

Efstratios Nikolaidis & Harley H.Cudney
Aerospace and Ocean Engineering

(ABSTRACT)

Probabilistic models and fuzzy set models describe different aspects of uncertainty. Probabilistic models primarily describe random variability in parameters. In engineering system safety, examples are variability in material properties, geometrical dimensions, or wind loads. In contrast, fuzzy set models of uncertainty primarily describe vagueness, such as vagueness in the definition of safety.

When there is only limited information about variability, it is possible to use probabilistic models by making suitable assumptions on the statistics of the variability. However, it has been repeatedly shown that this can entail serious errors. Fuzzy set models, which require little data, appear to be well suited to use with designing for uncertainty, when little is known about the uncertainty. Several studies have compared fuzzy set and probabilistic methods in analysis of safety of systems under uncertainty. However, no study has compared the two approaches systematically as a function of the amount of available information. Such a comparison, in the context of design against failure, is the objective of this dissertation.

First, the theoretical foundations of probability and possibility theories are compared. We show that a major difference between probability and possibility is in the axioms about the union of events. Because of this difference, probability and possibility calculi are fundamentally different and one cannot simulate possibility calculus using probabilistic models. We also show that possibility-based methods tend to be more conservative than probability-based methods in systems that fail only if many unfavorable events occur simultaneously.

Based on these theoretical observations, two design problems are formulated to demonstrate the strength and weakness of probabilistic and fuzzy set methods. We consider the design of tuned damper system and the design and construction of domino stacks. These problems contain narrow failure zones in their uncertain variables and are tailored to demonstrate the pitfalls of probabilistic methods when little information is available for uncertain variables.

Using these design problems we demonstrate that probabilistic methods are better than possibility-based methods if sufficient information is available. Just as importantly, we show possibility-based methods can be better if little information is available. Our conclusion is that when there is little information available about uncertainties, a hybrid method should be used to ensure a safe design.

Acknowledgements

Foremost I would like to express my sincere gratitude to my advisor and committee chairman, Dr. Efstratios Nikolaidis for his guidance and encouragement throughout my days at Virginia Tech. Without his patient mentoring, I could not have finished. I am also grateful to my committee cochairman, Dr. Harley Cudney, for his humorous and insightful remarks on my research, devilish strictness in my education in every aspect. I would like to thank Dr. Raphael Haftka for a lot of invaluable advice during our discussions. His incredibly deep perception into the nature of problems has always been the source of inspiration and breakthrough. I also express my warmest appreciation to the rest of my committee, Dr. Mahendra Singh, Dr. Hugh Vanlandingham and Dr. Eric Johnson who devoted their time to the review of my work.

I would like to thank Dr. Alfred Wicks for his classes and instruction in modal analysis. Also, I must thank Dr. Keying Ye for his friendship and instructions. His classes in Bayesian statistics brought new ideas to our project. Next, I would like to thank Dr. Frederick Lutze for kindly helping me go through all required administrative procedures.

I would like to thank Raluca Rosca, for her cooperation and valuable suggestions in my part of this project. I would also like to thank Jeff Barry, Jonathan Abbott, for their time and assistance in experiments.

Special thanks are due to my friends in Blacksburg. Their companion has made my life here full of enjoyable and unforgettable memories.

Finally, my deepest appreciation goes to my parents for their support and love in everything of my life.

Contents

Lists of Tables

List of Figures

Nomenclature

A	The event that the construction cost of the tune-damper system exceeds the budget
B	The event that the vibration of the system exceeds safety level
D	The length of a domino block
E_1	Tower *A* is stable but too short
E_2	Tower *A* collapses but misses tower *B*
E_3	Tower *A* collapses and knocks *B* down.
E_C	The event that tower *A* collapse
E^c_C	The complementary event of E_C that tower A is stable
E_O	Tower *A* and *B* overlap
E^c_O	The complementary event of E_O
$F_X(x)$	Probability distribution for random variable X
$f_X(x)$	Probability density function for random variable X
$g(\cdot)$	Fuzzy measure
H	Target height of tower A
Hc	Collapse height of tower A
h	Value of target height of tower A
h_0	Minimum acceptable height
h_1	The lower limit of minimum acceptable height
h_2	The upper limit of minimum acceptable height
ic	Construction cost of the tuned-damper system
m	Mass of the absorber
M	Mass of the original system
ML1~ML3	Monetary losses for the E_1 ~ E_3
N	Sample size
NH1, *NH2*	the lower and upper limits of the minimum acceptable height

p_1	The probability that a tower collapses at the target height
p_2	The probability the target height is shorter than the minimum allowable height
p_3	The probability that two towers overlap in the x-coordinate
$P(\cdot)$	Probability of an event
R	Mass ratio of the absorber to the original system, $R=m/M$
R_l	The lower range of mass ratio
R_u	The upper range of mass ratio
S	Failure due to shortness
x_0, x_1	The lower and upper ranges for x-coordinates for the center of tower B
x_0', x_1'	The lower and upper ranges for x-coordinates for the center of tower A
x_A, x_B	x-coordinates for centers of tower A and B
xAB	the difference in the mean values of the x-coordinates of the two towers
β_1	Ratio of the natural frequency of the original system to the excitation frequency
β_2	Ratio of the natural frequency of the absorber to the excitation frequency
$\gamma_1,\gamma_2,\gamma_3$	Loss factors for the $E_1 \sim E_3$
δ^2	Mean square error of a sample
μ	Population mean
$\mu_1 \sim \mu_3$	Membership of losses for $E_1 \sim E_3$
μ_{Hc}	The mean of probability distribution of collapse height
$\pi_{1\sim3}$	Possibility derived from the least conservative possibility distribution of $p_1 \sim p_3$
$\pi_{E1\sim E3}$	Possibility of occurrence of $E_1 \sim E_3$

σ^2	Population variance
σ_1	Lower boundary for σ in Bayesian analysis
σ_e	The estimated standard deviation of the probability distribution of locations
σ_{Hc}	The standard deviation of probability distribution of collapse height
σ_t	The true standard deviation of the probability distribution of locations
σ_x/D	the normalized standard deviations of the x-coordinates of the two towers
$\sigma_{\delta 2}$	Standard error of δ^2
ω_e	Exitation frequency
ω_{n1}	Natural frequency of the original system
ω_{n2}	Natural frequency of the absorber
ζ	Damping ratio of the original system
$\Pi(\cdot)$	Possibility of an event
$\Pi_{1\sim3}(\cdot)$	The least conservative possibility distributions for probability density $p_1 \sim p_3$
Ω	Universal set

CHAPTER 1 INTRODUCTION TO DESIGN METHODS IN THE PRESENCE OF UNCERTAINTIES

This Chapter is a summary of historical development of design methods in the presence of uncertainties and an introduction to the significance of this dissertation. We first present the motivation for this research. Next, we present a taxomonomy of engineering design decision problems and focus on the importance of design in the presence of uncertainty. We review the literature on methods for design under uncertainty. We explain the significance of comparing probabilistic methods and fuzzy set methods. Next, we describe the objectives of this research and the approach to achieve these objectives. Finally, we describe the outline of this dissertation.

1.1 Motivation for studying design methods in the presence of uncertainties

The term uncertainty reflects lack of knowledge about the world or lack of ability to predict the outcome of a process. When we flip a coin we do not know if we will get heads or tails; when we have a batch of bulbs, we do not know the service life of each bulb. Unfortunately, there is no universal interpretation of uncertainty. For a long period uncertainty was understood merely as the impossibility to predict occurrences of events, and probability was the unique agent to represent uncertainty. When new areas such as artificial intelligence are explored, we discover that uncertainty is also associated with imprecise statements, such as "there is a good chance that the price of oil ten years from now will be pretty high". This is another form of uncertainty, which can not be represented by probability. One way to represent imprecise statements mathematically is by using fuzzy set theory.

We have reducible and irreducible uncertainties. Reducible uncertainties are due to lack of knowledge. They can be reduced if we collect more information. Reducible uncertainties can be categorized as either modeling uncertainties (uncertainties in the errors of models of physical phenomena) or statistical uncertainties (uncertainties in probabilistic models). Irreducible uncertainties cannot be reduced even if we collect more information. For example, even if we flip a coin many times and record the outcomes, we still cannot predict the outcome the next time we flip a coin. However, the statistical uncertainty in the probabilistic model of predicting the chance of "head" or "tail" reduces as we collect more observations.

The success of engineering design depends significantly on whether the most appropriate model of uncertainty has been identified and used. A model used for design that overlooks uncertainty might cause severe consequences such as loss of life or property. On the other hand, a design based on a model that overestimates uncertainty will waste resources. Also, a model that wrongly predicts the behavior of customers might cause a company to lose market share. In a highly competitive engineering world, properly addressing uncertainty is required for designs to be successful.

The simplest and still widely used approach to safeguard a design from uncertainty is the safety factor method, where safety factors derived from previous engineering experiences are used to consider uncertainty. The worst-case-scenario method and the Taguchi method are popular among design engineers (Otto and Antonsson, 1993). Probabilistic and fuzzy set methods are the most general and can be applied to model uncertainties in all types of problems.

Ever since the emergence of the fuzzy set theory, there have been disagreements between proponents of probabilistic and fuzzy set-based methods over their domains of applicability and robustness. Many researchers have carried out studies to resolve these controversies by comparing these two methods (Wood and Antonsson, 1990). Unfortunately, these studies compared the two methods in calculating the safety of the same system using probability and fuzzy set models. They considered simple structures. None of these studies conducted both theoretical and analytical comparisons. It is our objective in this dissertation to compare probabilistic and fuzzy methods thoroughly from theoretical and analytical perspectives. Moreover, the study described in this dissertation uses design rather than analysis to compare the methods. Specifically, the study tries to address directly the following question: "given the same information and the same resources, which design method leads to the best design?" The author believes that this is the proper way to compare methods since the ultimate test of a method is how well its designs fare in the field.

1.2 Taxonomy of design problems

Engineering design can be regarded as a decision-making process (Hazelrigg, 1996). In a decision-making problem, a decision-maker's task is to select an act from a set of alternatives that will result in the best outcome. To do this, he/she has to find out a set of alternative acts. He/she has to determine the set of all possible states of nature. The outcome of the design process depends on both the act chosen and the state of the world. Since the state of the world is uncertain the decision-maker cannot predict the outcome of an act.

Decision-making problems can be classified according to certain criteria. Here, we will categorize decision-making problems in terms of the following characteristics:

- Number of attributes needed to describe an outcome;
- Ability to predict the outcome of an act;
- Definition of success and failure (precise versus vague).

In some problems, the outcomes may have only one attribute. For example, we might design steel plates with the minimum weight. Hence, weight is the only attribute to describe our outcomes. In another case, we may design plates to achieve two goals: minimize cost and minimize weight. Cost and weight may be conflicting attributes. Also, in many problems, there is uncertainty in the occurrence of each outcome. For example, due to fluctuation in the steel prices, we may not know the price of steel. Thus, a problem can be categorized in terms of how many attributes are needed for each outcome and whether there is uncertainty in predicting the occurrence of each outcome. Table 1.1 lists categories of decision problems and the corresponding approaches that a decision-maker can use to solve these problems. The transition between survival and failure is gradual (i.e. there is imprecision in the decision-maker's preferences).

Problems in which there is no uncertainty and involve a single attribute (Type I in Table 1.1) can be solved using optimization. The decision-maker needs only to determine the optimum act that maximizes a function measuring the value of each outcome to the decision-maker. Multiple objective optimization (Pareto, 1906), utility theory or fuzzy set theory can be employed to solve problems with many attributes (Type III). These approaches find multiple optima (efficient solutions) or use the utility function (Thurston, 1991) or the membership function (Diaz, 1988, Wood and Antonsson, 1990, Carnahan et al. 1994) to express the worth of each outcome to the decision-maker so they can select the best action.

Table 1.1 Taxonomy of decision problems

	Certainty about the outcomes of actions	**Uncertainty about the outcomes of actions**
One attribute is sufficient for describing an outcome	Type I problems Approach: Deterministic optimization	Type II problems Approaches: Utility theory, fuzzy set theory
Multiple attributes are needed for describing an outcome	Type III problems Approaches: Multiple objective optimization (i.e. Pareto optimal method), utility theory, fuzzy set theory	Type IV problems Approaches: Utility theory, fuzzy set theory

In problems involving uncertainty (Types II and IV), we can express the likelihood of an event using probability or possibility. Utility theory addresses problems where uncertainty in outcomes is modeled by a probability distribution. In utility theory, a decision-maker selects the best act by maximizing the expected utility (Thurston, 1994). Wood et al. (1990) proposed the *method of imprecision* to solve design problems involving uncertainty. They used membership functions to express the decision-maker's preferences about the values of the design variables (actions) as well as the preferences about the outcomes. Probability was employed to quantify the likelihood of events associated with uncontrolled stochastic uncertainties, such as manufacturing variability. Possibility was used to model non-stochastic uncertainties, such as modeling errors.

In many decision problems, preferences are imprecise, i.e. there is no clear, sharp boundary between success and failure. In other problems, this boundary is clear and crisp. An auto manufacturer who wants to build a "quiet" car faces a problem with imprecise preferences, in that it is insensible for him/her to call a design that exceeds the noise limit by 0.01db a failure.

On the other hand, a construction company who wants to design a building or a bridge that should not collapse in the next hundred years faces a problem with a clear, sharp boundary between success and failure.

Finally, problems involving uncertainty can be also partitioned into those in which there is sufficient information to accurately determine the probabilities of all the states of nature and those in which there is insufficient information. Most real life problems fall under the latter category.

1.3 Review of the probabilistic approach for designing in the presence of uncertainties

The probabilistic approach to design accounts for the uncertainties existing in the parameters of engineering design problems. This method uses probability theory to combine the effects of uncertainties to create a prediction about the reliability of the resulting design. In a simple yet typical probabilistic design problem, a designer might be asked to find the optimal section dimensions for a beam to minimize the probability of structural failure, when the beam is subjected to an external load which has known probability distribution.

Mathematically, probability is defined as a number assigned to events of a universal set (the set of all possible events or outcomes of an experiment). Probability satisfies the three axioms of Kolmogorov (Papoulis, 1965), which dictate that: a) the probability of any single event occurring is greater or equal to zero; b) the probability of the universal set is one (since the universal set includes all possible outcomes, we are certain that an experiment will create an outcome) and c) the probability of the union of mutually exclusive events is equal to the sum of the probabilities of these events. This last axiom is called the "additivity axiom".

There are two principal interpretations of probability, the objective and the subjective. According to the objective interpretation, probability is *the relative frequency of occurrence of an event* (Siddall, 1983). In this objective sense, probability must be estimated from a large number of observations. However, many events only occur once or rarely occur so it is impossible to estimate their probabilities using data. For example, we could not objectively predict the probability of the outburst of the Y2K problem in 1999 based on observations because we did not have any data.

The second interpretation of probability is the subjective (Bayesian) interpretation (Savage, 1972). According to this interpretation, probability is *the degree of belief that an event will occur*. Whether or not a large number of observations are available, many people still use judgment to estimate probabilities of events. For example, although there is abundant data about the number of candidates admitted to universities, many people may still believe that their child will PROBABLY be admitted based on his/her high school record, and not because of the large number of observations. Subjective probabilists (also called *Bayesians*) maintain that probabilistic methods are also useful for problems where there are only a few observations and probability is based on one's experience or intuition that an event will occur. With these two definitions, probabilistic methods can model uncertainties of a stochastic nature as well as of a subjective nature, either when the available information is based on judgment or when it is based on measurements. Indeed, in practice, subjective perception is always present because we are rarely fortunate to have sufficient data to create an accurate probabilistic model that reflects the real random nature of a phenomenon or an event. In this sense, every probabilistic model is a combination of the two interpretations.

The probabilistic approach is applied in reliability-based structural engineering design by introducing the probability of failure as a measure of unreliability. The lower the probability of failure is, the higher the reliability of a structure will be. The probability of failure can either be the objective of a design (e.g., minimize the probability of failure) or be the constraint of a design. In the latter case, the weight or the cost of the structure may be the objective (e.g., minimize the weight of the design subject to the probability of failure being less than, say, 0.01%).

Sundararajan (1995) collected probabilistic design applications in a wide spectrum of industries. These applications include:

- Analysis of the limit state (i.e. performance requirements for a structure to function properly, the violation of which will result in the structural failure) for the design of concrete structures in nuclear plants;

- Reliability-based aircraft design considering variability in external loads, material properties and finite element methods;
- Reliability-based fatigue design for ship structures, taking into account variations in the quality of workmanship and welding.

Recent applications of probabilistic design include probabilistic structural design of thin shells (Torng and Lin et al., 1998), and optimum design of composite laminates (Mahadaven and Liu, 1998).

It has been repeatedly shown that designs obtained using the same safety factor actually have significantly different failure probabilities (Marek, 1999, and Frangopol, 1985). The latter reference showed that the use of deterministic optimization may give designs that have either an unacceptably low or an unnecessarily high level of reliability as a result of inefficient use of resources. Probabilistic methods yield designs with reliability levels that are consistent with the cost because these methods explicitly account for the probability of failure. For example, in a structure consisting of bolts and a beam, the probabilistic method will find out that, to achieve the same reliability increment, it is cheaper strengthening the bolts than strengthening the beam.

Safety factors in deterministic models are not appropriate for problems involving innovative designs because safety factors are based on experience and there is no experience for these problems. Probabilistic design works better for this type of problems because it is based on first principles rather than on experience. For example, new materials like composites may have large scatter in their properties and no safety factors properly consider this variability. However, we can develop probabilistic models from sample tests to represent uncertainties in their properties.

One of the reasons that deterministic methods are still popular is that probabilistic methods involve prohibitively expensive calculations. With reliability-based method, we need to estimate the reliability of tens of hundreds of alternative designs. Specifically, we can calculate the probability of failure of a design using either Monte Carlo simulation or second moment methods. Monte-Carlo simulation requires a number of deterministic analyses ranging between few hundred to tens of thousands. Second moment methods (see for example, Madsen, Krenk and Lind, 1986) require fewer deterministic analyses (from 10 to 100) provided that we have analytical, closed form expressions for the sensitivity derivatives of the performance function with respect to the random variables. However, second moment methods do not always converge and they can yield erroneous estimates of the failure probability of structures for which there are multiple most-probable failure points. Moreover, probabilistic data is difficult to get. Based on the above, we conclude that the cost of reliability-based design often presents a formidable challenge (Maglaras and Nikolaidis, 1990).

To mitigate this high cost, Hasselman et al. (1996) proposed a probabilistic safety margin (PSM) method to approximate probabilistic design optimization. However, this method is inaccurate for large uncertainty or situation where a high degree of reliability is required. Yu and Choi (1997) suggested a mixed interactive design approach, which combines deterministic design, design trade-off analyses and what-if studies to achieve efficiency in computation. However, using a purely probabilistic design approach is still impractical for design problems involving large numbers of design variable uncertainties and constraint uncertainties.

The validity of the probabilistic method is based on the accuracy of the probabilistic model. It has been shown that small uncertainties in the probability distributions can cause large errors in the computed probability of failure (Ben-Haim and Elishakoff, 1990, Fox and Safie, 1992). Although Fox and Safie (1992) proposed guidelines for proper choice of probabilistic distributions when little numerical information is available, these guidelines are based on experience. It has not been proven that they always lead to safe and efficient designs. Therefore, it is imperative that one should examine the efficacy of alternative methods, such as fuzzy set methods, for these types of problems.

Bayesian interpretation of probability is criticized for the additive axiom it has to abide and for the requirement that probabilities of all elementary events should be precisely expressed by real numbers. In real world, imprecise probabilities exist because of little information. These imprecise probabilities violate the additive axiom. For example, if we flip a bent coin, we can assess the upper and lower probabilities of the event "heads" to account for the lack of information. If one does not have experience with this game, he/she will assume a large difference between these two probabilities to account for the lack of information. When he/she has sufficient observations, these two probabilities will converge (Klir, 1994 and Walley, 1991).

For example, we might assign upper probabilities 0.8 for the events "head" and "tail" respectively. The sum of these two probabilities will be greater than one in this case.

In engineering practice, it is often the case that we have few or no data to build a probability distribution. For example, when we design new equipment to manufacture steel plates, we are in a situation where we have no data to estimate the variability for the thickness of the processed plates. In this event, subjective judgements based on experts' experience with similar equipment are used to derive a probabilistic model. Fox and Safie (1992) illustrated how to create a conservative probabilistic model that best represents the current knowledge for the underlying probability distribution by using subjective judgements.

If one uses Bayesian probability, one can describe the uncertainty in probabilistic models by considering that the distribution parameters (e.g., the mean value or the standard deviation of a random variable) are random variables. Suppose we flip a bent coin. We do not know the probability of the event “head” because the coin is bent and because we do not have data to estimate this probability. In Bayesian statistics, the probability of the event "head", which is unknown, can be considered a random variable with its own probability distribution function. This distribution function is called a prior probability distribution function. We can predict outcomes of flipping based on our observations and the prior distribution of the parameter. However, the selection of the prior probability distribution function is still a controversial issue (Berger, 1985). Therefore, Bayesian probability might not be appropriate to account for the ignorance about the true probability destribution of a random variable because of these restrictions.

To summarize, we have reviewed the probabilistic approach and also shown that it is a widely accepted and effective method for design in the presence of uncertainties. However, for problems with a large number of variables, or with very little data from which to model uncertainties, the probabilistic approach can be either too expensive or not effective. A method based on fuzzy sets will be discussed in the next section, which addresses some of the shortcomings of the probabilistic approach.

1.4 Review of the fuzzy set approach for designing in the presence of uncertainty

Fuzzy set is *"a class with a continuum of grades of membership"* (Zadeh, 1965). This concept is a generalization of a classical set where an element either belongs or does not belong to the set. We can claim that a specific element belongs to a fuzzy set to a certain degree. Fuzzy set theory was first introduced by Zadeh to model uncertainty in subjective information. In natural language, many expressions are ambiguous and imprecise, like "Mark is tall". Though one can theoretically assign a definite height to discriminate between "tall man" and "short man", in practice it is not feasible because most of us will agree Mark is tall even if he is 0.1 inch shorter than that marking height.

In fuzzy set theory, a membership function is used to represent the degree that elements belong to the set in question. For example, "the class of quiet cars" is a fuzzy set. A new Mercedes belongs to this fuzzy set to the degree of 1.0, while a 20-year-old pickup with no muffler belongs to it with the degree of 0.01. Mercedes and Lexus may both have a membership 1.0, however their technical specifications for noise level may be quite different.

Shackle (1961) first proposed the possibility theory as a non-probabilistic framework in which decision process was modeled in term of "possibility". By Shackle's definition, possibility is the degree to which it is easy for an event to occur. He stated that possibility of an event is equal to one minus a person's degree of surprise if the event occurs. He stated that possibility should be used instead of probability when the conditions under which we have to make a decision under uncertainty cannot be reproduced. For example, if a company is to decide whether to make a risky investment, which will lead to bankruptcy if it is not successful and there is no data from which to estimate the probability of success, then the company should use possibility instead of probability to model uncertainty.

Zadeh (1978) formulated possibility theory as an extension of fuzzy set theory. According to Zadeh, a possibility distribution is numerically equal to the membership function. Therefore, given that the **membership function** of a new Mercedes is 1.0 in the fuzzy set "quiet cars," the **possibility** that a particular car is a Mercedes is 1.0 if we know that this particular car is quiet.

Possibility distributions can also be derived from evidence theory without association with fuzzy set theory (Klir and Yuan, 1995). In evidence theory, a measure is a set function that assumes nonnegative values and satisfies underlying axioms. Evidence theory studies uncertainties using belief measures and plausibility measures. The belief measure represents the total evidence that an element belongs to a set and to its subsets. The plausibility measure represents the total evidence that an element belongs to a set and its subsets and the evidence that the element belongs to other sets that overlap with the set. When the evidence assigned to the set and its subsets does not conflict (i.e. all these sets can be ordered in a sequence without overlapping), plausibility measures will become possibility measures.

A possibility distribution is a mapping of the elements in the set to the unit interval. A possibility distribution function uniquely represents a possibility measure and vise versa.

In our research, we will adopt the possibility theory developed by Zadeh, and consider "fuzzy set" and "possibility based" interchangeable. Therefore, we accept that fuzzy set theory satisfies the axioms for possibility measures in evidence theory.

There are many approaches to determine a fuzzy set membership function (possibility distribution) from experts' judgement or from experimental data (Pedrycz and Gomide, 1998).

For uncertainties that are modeled simultaneously by probability distributions and possibility distributions, there exist some rules to ensure consistency between the two measures (Klir and Yuan, 1995). We use these rules when constructing a possibility distribution from the given probability distribution or vise versa. The most evident rule is simply common sense: something that is probable must be possible. For example, "It will probably snow" indicates stronger evidence of chance of snowing than "It will possibly snow". More specifically, the first rule of consistency between the two distributions is that the possibility of any event should be greater than or equal to the probability of the same event. This will be called the "weak consistency principle". A stronger consistency principle requires that any event with a non-zero probability has a possibility of 1.0. For a given probability distribution, there are infinite numbers of possibility distributions, which satisfy the weak consistency principle. To meet the strong consistency principle, there is only one possibility distribution, and that distribution is equal to one everywhere.

We choose the possibility distribution, which is *the least conservative* among all the consistent possibility distributions, which satisfy the requirement, that the possibility of an event is larger than the probability of the same event (Nikolaidis et al., 1998). This transformation principle guarantees that we add the smallest possible amount of information to our model of uncertainty when we transform a probability distribution into a possibility distribution.

Fuzzy set methods have been used in engineering design. Wood et al. (1990) introduced the Method of Imprecision (MoI), which used mathematics of fuzzy sets to model uncertainties in the form of poorly defined, incomplete design descriptions at the early design stage. Scott and Antonsson (1998) applied MoI in preliminary design of vehicle structures to account for styling preferences and engineering requirements. Carnahan and Thurston (1994) used fuzzy rating to identify the levels of attributes and their relative importance in a multiple attribute decision-making problem at the preliminary design stage. Butler, Rao and LeCair (1995) modeled a designer's judgement about spatial relationships in detailed layout design for process manufacturing facilities using fuzzy sets. Sawyer and Rao (1999) proposed a concept of the fuzzy safety factor for rating designs of mechanical and structural systems in terms of strength-based reliability and damage tolerance.

Fuzzy set methodology provides a mathematical framework for modeling linguistic imprecision and graduality in propositions. It provides a convenient approach when we need to model complex systems on the basis of vague pieces of knowledge from descriptions of human language (Dubois and Prade, 1994). This uncertainty can not be represented adequately by probability theory. For example, to say somebody is *probably* old is not equivalent to express that he/she is *very* old. Klir (1994) stated that the additive structure of probability is not appropriate for some uncertainties. For example because observations in close neighborhoods of the "crisp" boundaries of events may not be reliable, additivity can not be actually achieved.

Opponents of fuzzy set theory criticized the "reality hypothesis", the "subjective hypothesis", the "behaviorist hypothesis", the "probability as fiction hypothesis" and the "superset hypothesis" in the fuzzy set theory from a philosophical viewpoint (Laviolette and Seaman, 1994). They also criticized the lack of operational meaning for the membership function, and the lack of theory for inference from data (Almond, 1995). Laviolette and Seaman (1994) demonstrated an example where the fuzzy set decision model can be insensitive to radical changes in the underlying probabilistic model.

In this section, we have reviewed fuzzy sets, possibility theory and their relations with probability theory. Fuzzy set theory is useful for designs in the presence of uncertainties which involve modeling linguistic imprecision and uncertainties with little numerical data to develop a valid probabilistic model.

1.5 Review of other design methods in the presence of uncertainty

Besides probabilistic and fuzzy set methods, there are other methods effective for certain types of uncertainties.

Safety factor method is traditionally used for uncertainties that are not quantified but known to exist. For example, the assumptions we use to derive a mathematical model induce modeling uncertainty. The finite element method we use to solve boundary value problems involves many approximations such as the representation of the response of a system in terms of shape functions. Safety factors are determined from experience. Traditional design approaches use a single safety factor to account for all uncertainties. For example, the stress in a structure is scaled by a safety factor that is greater than one to account for all of the uncertainties.

Methods using a single safety factor consider the worst-case scenario. Therefore, usually they are overly conservative. Since it is impractical to determine safety factors on case-to-case basis, design codes recommend using the same safety factor in a wide range of applications. It has repeatedly shown that different structures designed using the same safety factor have significantly different reliability levels (Frangopol, 1984 and Marek, 1999). Most of these structures are over-designed but few are unsafe.

Recently, methods that apply multiple safety factors, called *partial safety factors*, to different load and strength parameters have replaced traditional methods using a single safety factor. The larger the uncertainty in a load parameter, the larger is the value of the applied safety factor. In design of offshore platforms, different safety factors are applied to the static load, the wave induced loads and the strength. This method is called, *load and resistance factor design* (LRFD) (American Institute of Steel Construction, 1986).

Methods using partial safety factors are very popular in design codes for structures such as buildings, bridges and ships. These factors are calibrated using probabilistic methods in a way that designs obtained using these partial safety factors have a target reliability level (or equivalent failure probability). Methods for calibrating partial safety factors can be found in Ditlevsen and Madsen (1996).

Taguchi method introduces statistical concepts into traditional engineering design. Its main objective is to improve quality by making robust designs against uncontrollable variability in manufacturing processes and the operating environment. The philosophy of the Taguchi method is that in many design problems it is more efficient to design a system in a way that it is insensitive to uncontrollable variability rather than to reduce variability. Taguchi method for robust design involves the following main steps: identify and distinguish the controllable and uncontrollable factors in design settings; arrange statistically designed experiments; determine combinations of controllable factors that reduce the performance variations caused by the uncontrollable factors (internal or external noise) while still maintaining the average performance (Logothetis and Wynn, 1989, Phadke, 1989).

Taguchi method is very popular in product design because it does not require a designer to develop complex models of uncertainties, yet it is very effective in increasing the quality of a design.

Grieve and Barton et al. (1998) combined FEA with the Taguchi method for design of a lightweight automotive brake disc that considers all the critical design and material factors. Sunar and Hyder (1998) applied Taguchi robust design methodology to a closed-loop structure containing a cantilever beam and a thermopiezoelectric actuator pair, where the size and location of actuators are considered as controllable factors.

Taguchi method is criticized because it does not form a model for the response, and does not model interactions among control or noise variables. Also Taguchi method uses array designs, which are more expensive than fractional factorial designs. Besides, the signal and noise ratio is not sufficient as a criterion of merit (Myers, 1992).

Ben-Haim (1996) proposed the robust reliability method, which uses convex-set models of uncertainties and uses the concept of robustness-to-uncertainty as a measure of reliability.

Convex sets contain functions or vectors that enclose all possible combinations of values of the uncertain variables. For example, when we design a pressure vessel subject to fluid pressure with unknown probability density, we can assign a convex set consisting of all possible probability densities that are consistent with available information. The size of the convex model is controlled by the uncertainty parameter. The objective of robust reliability design is to maximize the value of the uncertainty parameter that corresponds to the largest variation in the uncertain variables that the design can tolerate without failure. Elseifi, Gurdal and Nikolaidis (1998) applied convex models in design of stiffened composite panels with uncertainties in the geometric imperfections. Pantelides and Tzan (1996) used convex models for the seismic design of a structure.

Robust reliability design method is actually a worst scenario method since the convex models have to cover the uncertainty that causes the worst consequence. Although robust reliability may lead to overly conservative designs compared to probabilistic methods it is a logical alternative to probabilistic methods when the parameters needed to create probabilistic models cannot be precisely determined due to lack of data.

Finally, there are methods that consider the performance of a design under the worst-case scenario (Emch and Parkinson 1994) or interval methods (Mullen and Muhanna, 1999, Koyluoglu and Elishakoff, 1998).

1.6 Significance of comparing probabilistic and fuzzy set design methods

Probability theory used to dominate the domain of modeling uncertainty. Recently, alternative theories, such as fuzzy set theory and evidence theory have been introduced for modeling uncertainty. Proponents of fuzzy set theory argued that probability theory is not adequate for uncertainties related to the intrinsic meanings in human language. Proponents of probability theory defended that "probability is the only sensible description of uncertainly and is adequate for all problems involving uncertainty" (Lindley, 1987). With the concept of subjective probability, probability theory can model non-stochastic uncertainty as effectively as fuzzy sets.

Proponents of fuzzy sets and probability theory do not agree with the philosophy behind these two theories. Proponents of fuzzy sets insist on the necessity of modeling human thinking and emulating human behavior which is a descriptive perspective, whereas proponents of probability insist on the prescriptive standpoint which eliminates the incoherence in human thinking when modeling subjective behavior by imposing axioms (Bernardo and Smith, 1994). Besides the philosophical differences of probabilistic and fuzzy set methods, we are interested in the differences in their theoretical foundations. Because of these differences each theory is suitable for a specific kind of uncertainty.

Investigating the differences in their theoretical foundations and the ensuing impacts on their efficacy of modeling uncertainties will help us identify the advantages and limitations in each theory. Also it will help us develop guidelines for choosing an appropriate method according to the available information.

Many researchers have conducted studies comparing fuzzy set theory and probability theory from different perspectives. Klir and Yuan (1995) compared the mathematical properties of possibility theory and probability theory. Wood and Antosson (1990) compared probability and fuzzy calculus for representing design imprecision (the uncertainty in choosing among alternatives). They found that the fuzzy calculus is well suited for the imprecision aspect of uncertainties, and that probability is best for stochastic uncertainty (the uncertainty which exists in processes a designer can not directly control, for example, the manufacturing tolerance). Chiang, Dong and Hasselman (1987) stated that possibility could be better than probability if there is little information about uncertainty. Vadde et al. (1994) obtained heavier structural designs using a fuzzy set based approach than a probabilistic approach. Maglaras et al. (1997) applied probabilistic and fuzzy set optimization to the design of a cantilevered truss structure and compared the respective optimum designs with experimentally. The results revealed that probabilistic methods yield significantly better designs than fuzzy set methods when there is sufficient information about uncertainties and a crisp definition of failure.

1.7 Objective of comparing probabilistic and fuzzy set design methods

Our objective in this dissertation is to understand the differences and similarities between probabilistic and fuzzy set methods in modeling uncertainties by comparing the theoretical foundations of probability and possibility theories and also by demonstrating the advantages and limitations of these two methods on design problems. In the end, we will try to develop guidelines for selecting appropriate design methods in the presence of uncertainty depending on how much information is available.

The study considers problems that involve random and modeling uncertainties but is limited to cases where there is only numerical information.

1.8 Approach to compare probabilistic and fuzzy set design methods

In this study, we first compare the theoretical foundations of probabilistic and fuzzy set methods. Fuzzy set methods use possibility to measure safety of a design; therefore, we compare the respective axioms that probability measures and possibility measures adhere to. Probability measures satisfy the three axioms of Kolmogorov, whereas possibility measures also satisfy three axioms as a special branch of fuzzy measures (Sugeno, 1977).

Also, we compare probabilistic and fuzzy set based models of uncertainties. We identify the properties of these models and their differences in modeling uncertainties when there is limited numerical information.

Next we compare the calculus of probabilistic and fuzzy set methods to calculate functions of uncertain variables. Based upon these comparisons, we want to show that probabilistic and fuzzy set methods are distinct and we cannot simulate one method with the other.

We exhibit cases where probabilistic or fuzzy set methods underestimate the risk of failure of a system. We use both methods to analyze the reliability of systems that are composed of parallel or serial components. Both of these two methods can underestimate risk of failure. Probabilistic methods underestimate the risk of systems for which there is a narrow failure zone in the space of the random variables.

As the last part of theoretical comparisons, we compare the algorithms the methods use to find out their optimums to minimize the chance of failure. These differences result in totally different designs for the same design problem.

As we have stated, we use design rather than analysis to evaluate the advantages and limitations of probabilistic and fuzzy set methods. In the design comparison, probabilistic and fuzzy set methods use the same amount of numerical data to create their models of uncertainties. They assess the safety of the design with different metrics, the probabilistic method by probability of failure, the fuzzy set method by possibility of failure. Accordingly each method finds out its own optimum design. The best design between the two optimum designs is the one that performs better under real life conditions. We use the design of a tuned damper system to compare probabilistic and fuzzy set methods analytically. This tuned damper system consists of an original system and a tuned damper. The tuned damper is designed to minimize the vibration of the original system as well as to minimize the construction budget. We choose this particular problem because there is a narrow failure zone in this problem and consequently the probabilistic method tends to underestimate risk when little information is available. Failure is defined as the excessive vibration of the original system or the construction cost overrun. In this problem, failure is catastrophic, which means that the boundary between failure and success is sharp. We use probabilistic and fuzzy set models to represent uncertainties in the design. The probabilistic method minimizes the probability of failure and the fuzzy set method minimizes the possibility of failure. We transform probabilistic models of uncertainty into fuzzy models using a transformation that minimizes the loss of information. When developing the probabilistic models from numerical data, we use standard statistics method and Bayesian method respectively to estimate distribution parameters and develop conservative probabilistic models. We compare these two design methods in terms of the amount of data used to create uncertainty models. For each sample data size, we also compare probabilistic and fuzzy set designs when the wrong type of the probability distribution of the random variables is used.

Finally, we consider a problem involving design and construction of stacks of dominoes to demonstrate advantages and limitations of probabilistic and fuzzy-set methods both analytically and experimentally.

1.9 Outline of the dissertation

In Chapter 2, we examine the theoretical foundations of probabilistic and fuzzy set methods. We compare the underlying axioms of probability measure and possibility measure. We also compare probabilistic and fuzzy calculus for functions of uncertain variables. Next, we demonstrate that both probabilistic and fuzzy set methods can underestimate chances of failure for a system with many components. Based on these investigations, we summarize observations about the theoretical differences of these two methods to model uncertainties.

In Chapter 3, we first propose a general approach for comparing probabilistic and fuzzy set methods for design in the presence of uncertainties. Then we formulate the problem of design for a tuned damper system. We describe the approaches of creating the probability and possibility distributions for the uncertain parameters from numerical data. The probabilistic method finds the optimum design to minimize the probability of failure of the system whereas the fuzzy set method finds its optimum design to minimize the possibility of failure of the system.

In Chapter 4, we explain that we can compare probabilistic and fuzzy set methods by calculating the true probabilities of failure of each optimum design. We describe the factors that contribute to the differences of these two design methods, and then we present the analytical results for each method considering all the factors. Based on these results, we summarize our observations of the advantages and limitations of these two methods.

In Chapter 5, we compare both analytically and experimentally probabilistic and possibility-based methods on a problem involving design of domino towers. First, we formulate the design problem. Next, we explain the probabilistic and possibility-based approaches to the design problem. Also, we explain the method for comparing the designs from these approaches. Then, we present and compare these designs both analytically and experimentally. Finally, we present the conclusions about the effectiveness of the probabilistic and possibility-based methods.

In Chapter 6, we draw conclusions based on our observations from the comparison the theoretical foundations of the methods and the design comparisons. Then, we propose guidelines for how to choose between these two methods in terms of amount of numerical data available. Finally, we suggest future studies for investigating design methods in the presence of uncertainty.

In this dissertation, "we" refers to the author, Dr. E. Nikolaidis and Dr. H.H. Cudney.

1.10 References

Almond, R.G., 1995, “Fuzzy Logic: Better Science? Or Better Engineering?” *Technometrics*, Vol. 37, No. 3, pp. 267-270.

American Institute of Steel Construction, (AISC), 1986, *Load and Resistance Factor Design Specifications*, AISC, Chicago.

Ben-Haim, Y., and Elishakoff, I., 1990, *Convex Models of Uncertainty in Applied Mechanics*, Elsevier, Amsterdam.

Ben-Haim, Y., 1996, *Robust Reliability in the Mechanical Sciences*, Springer-Verlag, Berlin, Heidelberg.

Berger, J.O., 1985, *Statistical Decision Theory and Bayesian Analysis*, 2^{nd} edition, Springer-Verlag, New York.

Bernardo, J.M., and Smith, A.F.M., 1994, *Bayesian Theory*, John Wiley, New York.

Butler, A.C., Rao, S.S., and LeCair, S.R., 1995, “Fuzzy Computer Aided Design of Process Manufacturing Facilities,” *Research in Engineering Design*, Vol. 116, pp. 511-521.

Carnahan, J. V., Thurston, D. L., and Liu, T., 1994, “Fuzzying Ratings for Multiattribute Decision-Making,” *Journal of Mechanical Design*, Vol. 116, pp. 511-521.

Chiang, W.-L. and Dong, W.-M., 1987, “ Dynamic Response of Structures with Uncertain Parameters: A Comparative Study of Probabilistic and Fuzzy Set Models,” *Probabilistic Engineering Mechanics*, Vol. 2, No. 2, pp. 82-91.

Diaz, A., 1988, “Goal Aggregation in Design Optimization,” *Engineering Optimization*, Vol. 13, pp. 257-273.

Ditlevsen, O., and Madsen, H.O., 1996, *Structural Reliability Methods*, John Wiley & Sons, New York.

Dubois, D., and Prade, H., 1994, "Fuzzy Sets — A Convenient Fiction for Modeling Vagueness and Possibility", *IEEE Transactions on Fuzzy Systems*, Vol. 2, No.1, pp. 16-25.

Emch, G., and Parkinson, A., 1994, "Robust Optimal Design for Worst-Case Tolerances," *Journal of Mechanical Design*, Vol. 116, pp. 1019-1025.

Fox, E.P., and Safie, F., 1992, "Statistical Characterization of Life Drivers for a Probabilistic Design Analysis", *28th AIAA/SAE/ASME/ASEE Joint Propulsion Conference and Exhibit*, July 6-8, 1992, Nashville, TN.

Elseifi, M., Gurdal, Z., and Nikolaidis, E., 1998, "Convex and Probabilistic Models of Uncertainties in Geometric Imperfections of Stiffened Composite Panels," *39th AIAA/ASME/ASCE/ AHS/ASC SDM Conference*, April, 1998, Long Beach, CA.

Frangopol, D.M., 1985, "Structural Optimization Using Reliability Concepts", *ASCE Spring Convention, Symposium on Structural Safety Studies,* April 23- May 3, 1985, Denver, Colo.

Grieve, D.G., and Barton, D.C. et al., 1998, "Design of a Lightweight Automotive Brake Disc Using Finite Element and Taguchi Techniques," *Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering,* Vol. 212, No.4, pp. 245-254.

Hasselman, T.K., Chen, X., et al., 1996, "Enhanced Aircraft Design Capabiltiy for the Automated Structural Optimization System", WLTR-96-3105 prepared for the Flight Dynamics Directorate, Wright Laboratory, Air Force Material Command, Wright-Patterson AFB, OH, by ACTA Inc., Torrance, CA.

Hazelrigg, G. A., 1996, *Systems Engineering : An Approach to Information-based Design*, Prentice Hall, NJ

Klir, G., 1994, "On the Alleged Superiority of Probabilistic Representation of Uncertainty", *IEEE Transactions on Fuzzy Systems,* Vol. 2, No.1, pp. 27-31.

Klir, G., and Yuan, G., 1995, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey.

Koyluoglu, H.U., and Elishakoff, I., 1998, "A Comparison of Stochastic and Interval Finite Elements Applied to Shear Frames with Uncertain Stiffness Properties," *Computers and Structures*, Vol. 67, pp. 91-98.

Laviolette, M., and Seaman, J.W., 1994, "The Efficacy of Fuzzy Representations of Uncertainty", *IEEE Transactions on Fuzzy Systems*, Vol. 2, No.1, pp. 4-15.

Lindley, D.V., 1987, "The Probability Approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Ssytems," *Statistics Science*, Vol. 2, pp. 17-24.

Logothetis, N., and Wynn, H.P., 1989, *Quality through Design: Experimental Design, Off-line Quality Control and Taguchi's Contributions*, Clarendon Press, Oxford.

Maglaras, G., Nikolaidis, E., 1990, "Integrated Analysis and Design in Stochastic Optimization," *Structural Optimization*, Vol. 2, No. 3, pp. 163-172.

Maglaras, G., Nikolaidis, E., et al., 1997, "Analytical-experimental Comparison of Probabilistic Methods and Fuzzy Set Based Methods for Designing under Uncertainty," *Structural Optimization*, Vol. 13, No. 2/3, pp. 69-81.

Mahadevan, S., Liu, X., 1998, "Probabilistic Optimum Design of Composite Laminates", *Journal of Composite Materials*, Vol. 32, No.1, pp. 68-83.

Marek, P., 1999, "Transitions from Partial Factors Method to Simulation-Based Reliability Assessment in Structural Design", *Probabilistic Engineering Mechanics*, Vol. 14, No.1/2.

Mullen, R. L., and Muhanna, R. L., 1999, "Interval-Based Geometric and Material Uncertainty for Mechanics Problems," *13th ASCE Engineering Mechanics Division Conference*, ASCE, Baltimore, Maryland, 1999.

Myers, R.H., 1992, "Response Surface Alternatives to the Taguchi Robust Parameter Design Approach", *The American Statistician*, Vol. 46, No.2, pp. 131-139.

Nikolaidis, E., Haftka, R., and Rosca, R., 1998, "Comparison of Probabilistic and Possibility-Based Methods for Design Against Catastrophic Failure Under Uncertainty," Aerospace and Ocean Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-203. (can be downloaded from http://www.aoe.vt.edu/~nikolai)

Otto, K.N., Antonsson, E.K., 1993, "The Method of Imprecision Compared to Utility Theory for Design Selection Problems", *The Proceedings of the ASME 1993Design Theory and Methodology Conference-DTM'93*, pp.167-173.

Pantelides, C.P., and Tzan, S.R., 1996, "Convex Model for Seismic Design of Structures--I: Analysis," *Earthquake Engineering and Structural Dynamics*, Vol. 25, September, pp. 927-944.

Papoulis, A., 1965, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York.

Pareto, V., 1906, *Manuale di Economica Politica*, Societa Editrice Libraria, Milano, Italy. Translated into English by A.S. Schwier as Manual of Political Economy, MacMillan, New York, 1971.

Pedrycz, W., Gomide, F., 1998, *An Introduction to Fuzzy Sets: Analysis and Design*, The MIT Press, Cambridge, Massachusetts.

Phadke, M., 1989, *Quality Engineering Using Robust Design*, Prentice Hall**.**

Savage, L. J., 1972, *The Foundations of Statistics*, 2nd rev. ed., Dover, New York.

Sawyer, J.P., Rao, S.S., 1999, "Strength-based Reliability and Fracture Assessment of Fuzzy Mechanical and Structural Systems" *AIAA Journal*, Vol. 37, No. 1, pp. 84-92.

Shackle, G. L. S., 1961, *Decision, Order and Time in Human Affairs*, Cambidge University Press, New York.

Scott, M.J., Antonsson, E.K.,1998, "Preliminary Vehicle Structure Design: An Industrial Application of Imprecision in Engineering Design", *The Proceedings of the ASME 1998Design Engineering Technical Conference-DETC'98.* Sept. 13-16, 1998, Atlanta, GA.

Siddall, J.N., 1983, *Probabilistic Engineering Design: Principles and Applications*, Marcel Dekker Inc., New York.

Sugeno, M., 1977, Fuzzy Measures and Fuzzy Intervals: A Survey," In: Gupta, M. M., Saridis, G. N., and Gaines, B. R., eds., *Fuzzy Automata and Decision Processes*, North-Holland, Amsterdam and New York, pp. 89-102.

Sunar, M., and Hyder, S.J. et al., 1998, "Robust Design of Thermopiezoelectric Actuators," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 3, pp. 521-522.

Sundararajan, C., eds., 1995, *Probabilistic Structural Mechanics Handbook: Theory and Industrial Applications*, Chapman & Hall, New York.

Thurston, D. L., 1991, "A Formal Method for Subjective Design Evaluation with Multiple Attributes," *Research in Engineering Design*, Vol. 3, No. 2, pp. 105-122.

Thurston,D.L., Carnahan,J.V.,1994, "Optimization of Design Utility", *Journal of Mechanic Design*, Vol. 116, September, pp. 801-808.

Torng,T., Lin, H.Z., 1998, "Probabilistic Thin Shell Structural Design", *39th AIAA/ASME/ASCE/ AHS/ASC SDM Conference*, April, 1998, Long Beach, CA.

Vadde, S. Allen, J. K., and Mistree, F., 1994, "Compromise Decision Support Problems for Hierarchical Design Involving Uncertainty," *Computers and Structures*, Vol. 52, No. 4, pp. 645-658.

Walley, P., 1991, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall, New York.

Wood, K.L., Antonsson, E.K., 1990, "Representing Imprecision in Engineering Design - Comparing Fuzzy and Probability Calculus", *Research in Engineering Design*, Vol. 1, No.3/4, pp. 187-203.

Yu, X., Choi, K.K., and Chang, K.H., 1997, "A Mixed Design Approach for Probabilistic Structural Durability", *Structural Optimization*, Vol. 14, No. 2-3, October, pp. 81-90.

Zadeh, L. A., 1965, "Fuzzy Sets", *Information and Control*, Vol. 8, No.3, pp. 338-353.

Zadeh, L. A., 1978, "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems*, Vol. 1, No.1, pp. 3-28.

CHAPTER 2 PROBABILISTIC AND FUZZY SET METHODS

This Chapter compares theoretical foundations of probabilistic and fuzzy set method. First, we compare the axioms of probability and possibility by identifying the differences and similarities in these axioms. Next, we compare how probabilistic and possibility-based methods model uncertainties. We also compare how probabilistic and possibility calculi compute the distributions of a dependent variable as a function of the independent variables. Then, we compare these two methods for assessing the safety of a system and demonstrate the weakness in these two methods. We also compare how each method maximizes safety for a given budget. Finally, we summarize our observations based on these theoretical comparisons.

2.1 Comparing axioms of probability and possibility

People had used the classical definition of probability by Laplace (1812) as the fundamental notion of probability theory before Russian statistician Kolmogorov proposed three axioms as foundations of probability in 1933. According to the Laplace definition, if we partition the outcome space of a random experiment into equally likely elementary events, the probability of an event, *A*, is the ratio of the number of elementary events whose occurrence implies the occurrence of *A* over the total number of elementary events. The Laplace definition of probability satisfies the Kolmogorov axioms that are presented in the fist column of Table 2.1.

Kolmogorov actually extended the measure notion in set theory to study probability. Measure is a set function defined on a σ-*algebra* — a class of subsets *S* of a universal space Ω that is closed with respect to complementation and finite unions. This function satisfies the following conditions (Billingsley, 1986):

(i) $\mu(A) \in [0, \infty]$, for $A \in S$;

(ii) $\mu(\emptyset) = 0$;

(iii) for any disjointed $A_1, A_2, \ldots,$ of sets of *S*, such that:

$$A = \bigcup_{i=1}^{\infty} A_i \in S$$
$$\mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \tag{2.1}$$

Probability measures are special measures where the conditions of their set functions are restricted to the form presented in the first column of Table 2.1.

Table 2.1 Axioms of Probability Measures, Possibility Measures and Fuzzy Measures

Probability measure, $P(\cdot)$	**Possibility measure, $\Pi(\cdot)$**	**Fuzzy measure, $g(\cdot)$**
1) Boundary requirement: $P(\Omega)=1$	1) Boundary requirements: $\Pi(\varnothing)=0$, $\Pi(\Omega)=1$	1) Boundary requirements: $g(\varnothing)=0$, $g(\Omega)=1$
2) Non-negativity: $P(A) \geq 0\ \forall A \in S$	2) Monotonicity: $\forall A, B \in S$, *if* $A \subseteq B$, then $\Pi(A) \leq \Pi(B)$	2) Monotonicity: $\forall A, B \in S$, *if* $A \subseteq B$, then $g(A) \leq g(B)$
3) Additivity: $\forall A_i, i \in I$, A_i are disjoint $P(\bigcup_{i=1}^{I} A_i) = \sum_{i \in I} P(A_i)$	3) Sub-additivity: $\forall A_i, i \in I$ $\Pi(\bigcup_{i=1}^{I} A_i) = \max_{i \in I}(\Pi(A_i))$	3) Continuity: $\forall A_1 \subseteq A_2 \subseteq \ldots, or\ A_1 \supseteq A_2 \supseteq \ldots$ $i \in N$ $\lim_{i \to \infty} g(A_i) = g(\lim_{i \to \infty} A_i)$

Sugeno (1977) introduced fuzzy measures as a generalization of real measures. Fuzzy measure theory studies the uncertainty due to lack of perfect evidence to identify which set a specific element belongs to. Here every set is crisply defined. For example, a doctor diagnoses a patient who might have ordinary flu or pneumonia. From the patient's symptom, he can assign two values that represent the degree to which the evidence supports the fact that the patient has flu or pneumonia, respectively. These values are known as fuzzy measures. Fuzzy measures are functions that satisfy axioms listed in Table 2.1, which state that:

1) The degree to which the evidence supports the fact that any element belongs to the empty set is zero, whereas the degree to which the evidence supports the fact that any element belongs to the universal set is one.

2) The degree to which evidence supports the fact that an element belongs to a set must be greater than or equal to the degree that the evidence supports the fact that the element belongs to any subset of that set.

3) For an infinite sequence of monotonic sets, which converge to a certain set, the respective degrees of evidence that an element belongs to each set construct an infinite sequence, which converges to the degree of evidence that the element belongs to that set the sequence converge to.

Belief measures and plausibility measures, which were defined in Section 1.4, are fuzzy measures in that they satisfy the axioms of fuzzy measures. As we have stated, the possibility measure is a special plausibility measure in evidence theory. Possibility measures satisfy the axioms of fuzzy measures and an additional axiom, which states that the possibility of the union of events is equal to the maximum of the possibilities of the individual events.

Table 2.1 also compares the underlying axioms for probability and possibility measures. Let Ω be the universal set and S a set of crisp subsets of Ω. The boundary axiom for probability and possibility measures indicates that they are mappings from the universal set space onto a unit interval. These measures assume values between zero and one. The boundary axiom states that the probability (or possibility) that an element belongs to the universal set Ω is one. Probability and possibility measures are monotonic. The second axiom of possibility in Table 2.1 dictates that possibility is a monotonic measure. We can show that probability is also a monotonic measure from the additivity axiom of probability measures. Let $A \subseteq B$ and $C = B–A$. Then $A \cup C = B$ and $A \cap C = \emptyset$. Plugging A and C to the equation in the additivity axiom, we have: $P(A \cup C) = P(B) = P(A) + P(C) \geq P(A)$.

The major difference between the axioms of possibility and probability measures is that probability is additive whereas possibility is subadditive (i.e., the possibility of an event, which can be partitioned into smaller events is less than or equal to the sum of the possibilities of the constituent events). On the contrary, the probability of the union of a set of disjoint events is equal to the sum of the probabilities of each event. As a result, if $\{A_1, \ldots, A_n\}$ is a partition of the universal event, Ω, the probabilities of A_i must add up to one, whereas there is no such constraint for the possibilities of A_i. For example, the probabilities of the events "Tomorrow it will rain" and "Tomorrow it will not rain" must sum up to one. On the other hand, if one estimates the possibility that tomorrow it will rain is 0.7, he/she has to assign the possibility of "Tomorrow it will not rain" as 1.0. The reason is that, since the possibility of Ω is equal to the maximum of the possibilities of events A_i, the possibility of at least one of these events should be one. Therefore:

$$\sum_{i=1}^{n} \Pi(A_i) \geq 1 \qquad (2.2)$$

An important difference between the axiomatic foundations of probability and possibility is that we can only assign a probability measure to a *σ-algebra*. On the contrary, we can assign possibilities to any universe, since possibility is both a measure and a function. The class of all subsets of the real line is not a σ-algebra. A probability measure can be assigned to the smallest *σ-algebra* that contains all intervals $(-\infty, x_1]$, where x_1 is a real number (Papoulis, 1965), whereas we can assign a possibility to any class of subsets of the real line.

2.2 Comparing uncertainty models by probabilistic and fuzzy set methods

In the previous section, we introduced axioms of probability measures in the context of set theory. In many problems, the outcomes of a random event can also be identified through the values of a function, which is called a random variable. This function maps the events of the outcome space of random experiments onto the set of real numbers. For example, if the value of X represents the number of vehicles passing an intersection in a day, $X > 1000$ represents the set of events that more than 1000 cars pass an intersection in a day. Instead of using sets to represent random events in a universe, we use values of random variables to represent different events.

Random variables can either be discrete or continuous. The range of a discrete random variable is a set of isolated points, whereas the range of a continuous random variable is a continuum. Probability theory models a discrete random variable in terms of a probability mass function and models a continuous random variable in terms of a probability density function. The probability distribution of a random variable *X* is:

$$F_X(x) = P(X \leq x) \tag{2.3a}$$

For a discrete random variable *X*, the above definition can be written as follows:

$$F_X(x) = P(X \leq x) = \sum_{all\ x_i \leq x} p_X(x_i) \tag{2.3b}$$

where $p_X(x_i)$ is the probability mass function. The probability mass function of a discrete variable, *X*, assigning to each real number, x_i, is the probability of the random variable *X* being equal to x_i. If *X* is continuous, the probability mass function is not defined. Instead, we use the probability density function $f_X(x)$, which is the derivative of the probability distribution function, $F_X(x)$. Probability distribution functions must also satisfy the same axioms satisfied by probability measures.

The counterparts of random variables in fuzzy set theory are fuzzy variables. Fuzzy set theory uses the membership function (possibility distribution function), $\Pi_X(x)$ (possibility of *X* being equal to *x*) for both discrete and continuous variables. Possibility distribution functions must satisfy the same axioms as possibility measures. Following are some differences in the properties of a probability density function and a possibility distribution function of a continuous variable (Fig. 2.1):

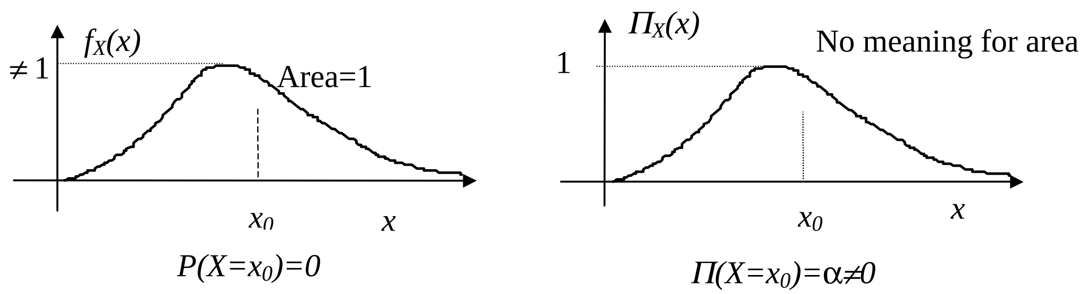

Figure 2.1 Comparison of probability density function and possibility distribution function

1. The area under the probability density function of a variable in an interval is equal to the probability of the variable assuming any value in that interval, whereas the area under the possibility distribution function has no meaning.
2. The total area under any probability density function is one whereas the area below a possibility distribution can be less or greater than one.
3. The probability of a variable, whose probability density function is a continuous function, assuming a specific value is zero, whereas the possibility of the same event can be any value between zero and one.
4. Both the probability density function of a random variable and the possibility distribution of a fuzzy variable must be nonnegative. In the case of the possibility distribution function, we have the additional restriction that its maximum value must be one.
5. The horizontal interval between two points with possibility α on the possibility distribution curve represents an α-cut, which is a crisp set that contains all the variables with possibility greater or equal than α. For example, the interval [a, b] in Fig. 2.2 represents a set for which the possibility of each variable is greater or equal to α. For a probability density function, there is no analogy for this representation.

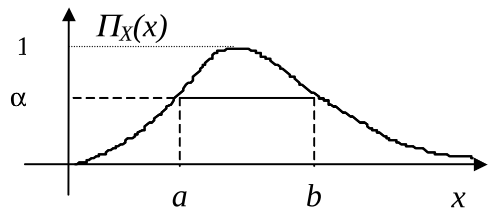

Figure 2.2 α-cut representation in a possibility distribution function

An important difference between probability and possibility is in the ways they model uncertainty when there is limited data. We will show that, in many problems, the same assumptions about the possibility distribution of the uncertain variables always make the model more conservative. On the other hand, an assumption about the probability density of a random variable may make the model more or less conservative depending on the particular problem. Consider a problem where the frequency of excitation applied to a lightly damped mechanical system, ω, is an uncertain variable. The system fails if the frequency of excitation falls in a danger zone shown in Fig. 2.3. We want to construct probabilistic and possibilistic models for ω and to assess the safety of the system. However, we do not have sufficient data for a probability distribution. The only information available is that the frequency falls in some interval. We do not know the type of its distribution and we do not have confidence in the definition of the interval. In this case of total ignorance, both probability and possibility theories use a uniform density or distribution, respectively. The probability method will assign a probability density of $1/\Delta$ where Δ is the length of the interval, whereas the possibility distribution will be one over the interval (Fig. 2.3).

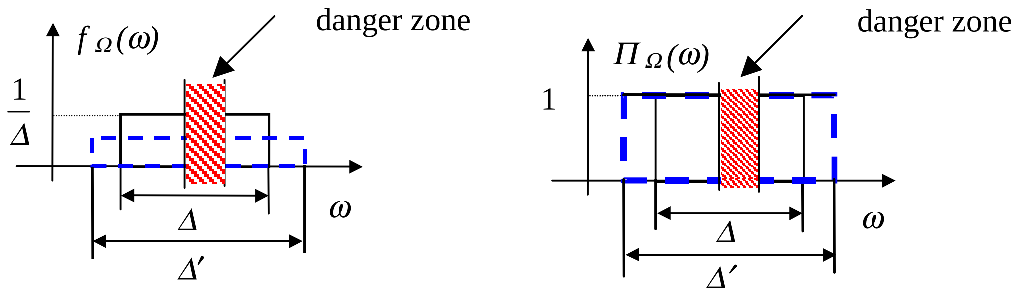

Figure 2.3 Probabilistic and possibilistic models when there is very limited information about a variable, such as the excitation frequency. The possibilistic model is guaranteed to become more conservative as we increase the length of the interval of variation of the excitation frequency but this is not the case for the probabilistic model.

Since in this problem there is uncertainty about the model we use to characterize uncertainty, it is reasonable to use the most conservative model in each case, which is consistent with the available information. A general approach is to enlarge the span of the interval to represent a higher degree of uncertainty about the model and thus make the model conservative. The probabilistic model increases the interval to Δ' and accordingly lowers the probability density to 1/Δ' so that the total area under the probability density function is kept equal to one. On the other hand, we do not have to lower the value of the possibility distribution since there is no restriction about the total area under this distribution. As a result, although we are trying to make the probabilistic model more conservative, in reality, we underestimate the probability of failure, which occurs if the frequency of excitation falls in the shaded failure zone. On the other hand, the two possibility models with different intervals are equally conservative because the possibility distribution remains equal to one as the interval length increases (Fig. 2.3). One may argue that it is easy in this problem to make the probabilistic model more conservative by setting the lower and upper limits of the probability density function of the frequency of excitation equal to the bounds of the failure zone. However, in real-life complex systems involving many failure modes and many uncertain variables we may not know how to make a probabilistic model more conservative. Therefore, possibility-based models can be useful in assessing the safety of such systems.

Another difference is in the notion of independence of events. In probability, two events are independent if the occurrence of one event does not affect the probability of the other event. Mathematically, independence is satisfied if and only if the probability of their intersection is equal to the product of each probability. In possibility theory, *non-interaction* is analogous to the concept of independence in probability theory (Zadeh, 1975). Two events are non-interactive if the possibility of their intersection is equal to the smallest possibility of these events. Non-interaction is a form of non-compensation because an increase in the possibility of one event can not compensate the decrease in the possibility of the other event. On the contrary, an increase of probability can compensate the decrease of the other event if these two events are independent. We will show later that the assumption of non- interaction always makes a possibility-based model more conservative than other models whereas the assumption of independence in probability can make a probabilistic model less or more conservative.

2.3 Comparison of the probability calculus and fuzzy set calculus

The probability density of a function of independent variables is calculated using the following concept. If the function transforms a small element of volume in the space of the independent variables to another element of volume in the space of dependent variables, then the probabilities corresponding to the two elements of volume are equal. Specifically, for a random variable Y, which is a function of random variables X_i, i=1, …, n. Suppose we know that the mapping function between Y and X_i is $Y=g(X_1,\dots,X_n)$. The following equation is used for calculating the probability distribution function of Y, $F_Y(y)$, (Davenport, 1970):

$$F_Y(y) = P(Y \le y) = P(g(X_1,\dots,X_n) \le y) = \iiint_{g(x_1,\dots,x_n)\le y} f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\dots dx_n \quad (2.4)$$

where $f_{X_1,...,X_n}(x_1,...,x_n)$ is the joint probability density of variables X_i, i=1, …, n. The probability density function of Y, $f_Y(y)$, can be found by differentiating $F_Y(y)$ with respect to y.

In the special case where $Y = g(X_1)$, and g is a one to one transformation, the probability density function of y is found using the following equation:

$$f_Y(y) = f_X(h(y))\left|\frac{dh(y)}{dy}\right| \qquad (2.5)$$

where $h(y)$ is the inverse function of $Y = g(X_1)$ and $\frac{dh(y)}{dy}$ is the derivative of $h(y)$ with respect to y.

Consider, for example the function $Y=X^2$ applied to X with a uniform probability density function as shown in Fig. 2.4. According to the above concept, the probability density of dependent variable Y, at any value y_0 in the interval [64,144] is equal to the density of X at x_0, divided by the slope of x^2, at x_0. This slope is $2x_0$, where $y_0= x_0^2$. Note that this means that the value of $Y=X^2$ with the largest probability density is 64, corresponding to the lower end of the range [8,12] in Fig. 2.4.

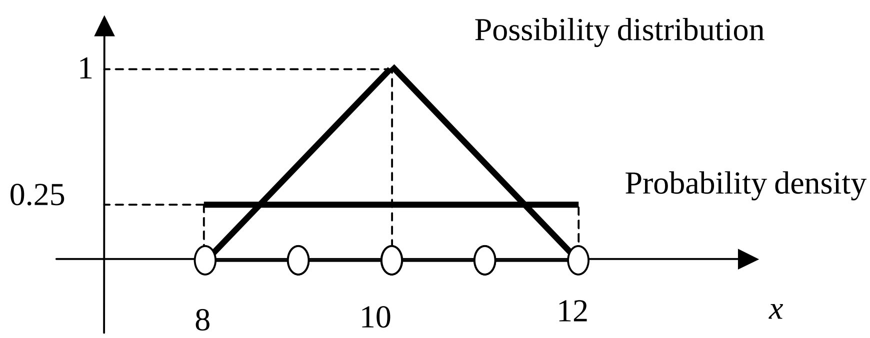

Figure 2.4 Probability density and possibility distribution of *X*, for the statement `*X* is about 10'.

The *extension principle* (Zimmerman, 1996) can be used to derive possibility distribution of a dependent variable. According to this principle, the possibility of $Y = y$ is equal to the maximum of the possibilities of the all combinations of values of the independent variables for which $y = g(x_1,...,x_n)$:

$$\Pi_Y(y) = \max_{y=g(x_1,...,x_n)} \Pi(X_1 = x_1,..., X_n = x_n) \tag{2.6}$$

The possibility of the dependent variable Y becoming equal to y_0, is equal to the maximum of the possibilities of all combinations of the independent variables mapped by the function to y_0. Thus, if $Y=X^2$ is applied to X defined in Fig. 2.4, the possibility of $Y=100$ is one, and this is the only value of Y with this high possibility. *That is, the most possible value of the function corresponds to the most possible value of the argument, while this is rarely the case in probabilities*. This result was presented by Wood et al. (1990), in a comparison of probability and fuzzy set calculi for modeling preferences.

In possibility theory, there is no law corresponding to the *law of large numbers*, which is a fundamental law in probability. This principle states that the probability density of the average of the n independent variables (called sample mean) tends to become less scattered than the density of the original variables and, as the number of variables tends to infinity, the average converges to the mean value of the variables. In possibility theory, the average of n identically distributed, non-interactive fuzzy variables has the same possibility distribution as the original variables, regardless of n.

Because of the above differences, we conclude that one cannot simulate the results of possibility calculus by deriving the parameters of the possibility distributions of the independent variables from probability density function or vice versa. Even if the input variables are consistent, the output possibility and probability may not be comparable.

2.4 Comparison of the ways probability and possibility measure safety

As a result of the additivity axiom of the probability measure, the sum of the probabilities of survival and failure of a system is always one. Therefore, the probability of survival of a system is always less than one. In contrast, the axiom for the possibility of the union of events leads to the conclusion that either the survival or the failure of the system has a possibility of one. Therefore, any practical system has a possibility of survival of one because it should be perfectly possible that it can survive.

Let F denote the failure event of a system. The following approach is followed to find the probability and possibility of failure. We partition F into elementary events whose probabilities or possibilities are easy to calculate. Then, according to the axiom for the probability of the union of disjoint events, we add the probabilities of the elementary events to find the probability of failure. Similarly, according to the axiom for the possibility of the union of events, the possibility of failure is the maximum of the possibilities of all the elementary events, whose union is the failure event, F.

Possibility is more conservative than probability in assessing the risk of failure of systems whose failure requires the simultaneous occurrence of many unfavorable independent events. Consider a parallel system for which the failures of all the components are mutually independent, and system failure is defined as simultaneous failure of all components. The probability of failure of the system decreases with the number of components, whereas the possibility of failure of the same system is the minimum of the possibility of failure for each component. The system failure region is usually small, which tends to make the probability of failure small, whereas the possibility of failure can be still high if all the components have large possibilities of failure. For example, Fig. 2.5(a) illustrates a system of n nominally identical, independent components connected in parallel. The probability of failure of the system $P(F)$ is the n^{th} power of the probability of failure of one component and decreases with n as shown in Fig. 2.5(b). In contrast, the system possibility of failure $\Pi(F)$ is equal to the possibility of failure of a single component, regardless of n, as shown in Fig. 2.5(c).

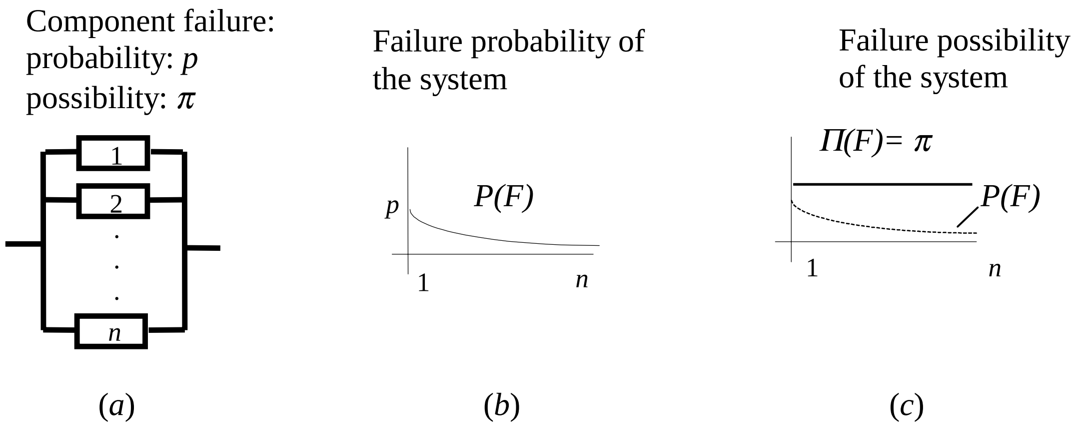

Figure 2.5 A parallel system and its probability and possibility of failure as a function of the number of components

However, possibility can be non-conservative when assessing the risk of failure of a system that has many failure modes. An example is a series system of *n* nominally identical, independent components, such as the system in Fig. 2.6(a). From the axiom for the possibility of the union of events, the possibility of failure of this system is equal to the possibility of failure of a single component, regardless of *n* (see Fig. 2.6(b)). On the other hand, the probability of survival of the system is the n^{th} power of the probability of survival of each component, which is (1–*p*). Therefore, the probability of system failure is $1-(1-p)^n$ and increases with *n*. As a result, we can always find a system that consists of a large number of components for which its possibility of failure is smaller than its probability of failure. For example, if $\pi = 0.8$, $p = 0.2$ and $n > 7$, then $P(F) > \Pi(F)$. This violates the probability-possibility consistency principle, according to which the possibility of an event must always be greater or equal to its probability.

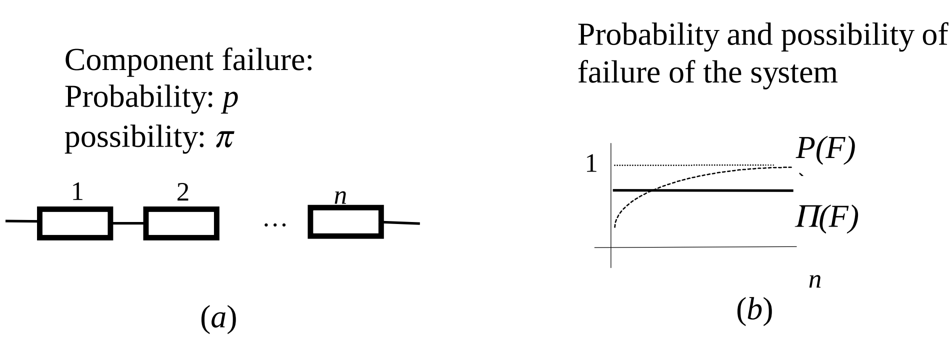

Figure 2.6 A series system and its probability and possibility of failure as a function of the number of components

Similar results are observed when we try to model extreme events using probability and possibility. The probability distribution of the maximum of *n* independent, identically distributed random variables, e.g., the maximum 50-year wind speed at some location, follows an asymptotic distribution derived from the annual maximum wind speed using order statistics (Ochi, 1990). On the other hand, the possibility distribution of the maximum is identical to the possibility distribution of any of the *n* variables. Thus, the possibility distribution of the maximum 50-year wind speed is identical to the possibility distribution of the maximum annual wind speed. This means that if *n* is large, a possibility-based approach will underestimate the risk of failure of a design (say a building subjected to wind loads over a 50-year period). On the other hand, wind loads over a 50-year period, which are derived from year to year distributions by order statistics, are severer than annual data and make probabilistic design more conservative. Fig.2.7 shows that possibility theory is not always more conservative than probability.

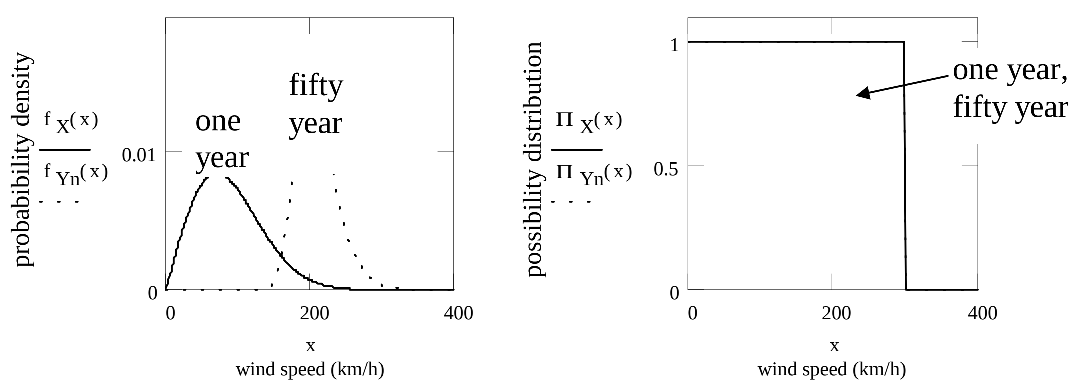

Figure 2.7 Probability density and possibility distributions of one and 50 year wind speed. Possibility distributions of one and 50 year speed are the

The prediction of the possibility of failure of a system with many failure modes and the prediction of the possibility of extreme events appear to be the weak points of possibility theory since it yields counter-intuitive results and contradicts with the probability-possibility consistency principle. In these cases, a low calculated system possibility of failure by possibility-based methods does not guarantee a low actual risk of failure for such a system. *The axiom on the union of events is responsible for this counterintuitive result of possibility-based methods*.

One may try to overcome this weakness by defining the joint possibility distribution of the components (Klir and Yuan, 1995), which is the possibility distribution defined on the Cartesian product space of all the components, in a way that the probability-possibility consistency principle is satisfied for any possible event. In this case, one will need to model failures of components using the following properties, which seem counterintuitive:

a) The possibility of failure of a component would depend on the number of components. This is surprising because the possibility of failure of a component should depend only on its properties and the operating environment.
b) To avoid the dependence of the possibility of failure on the number of components listed in the previous conclusion, one can try to define the possibilities of failure of the components so that the consistency principle is guaranteed to be satisfied for any number of components. That is, for any n, the possibility of failure should be greater than or equal to the probability of failure. But this will yield the possibility of failure to be one. This is because, in general, one can find that, in a problem involving n uncertain variables $X_1, \ldots, X_n$, the only joint possibility distribution guaranteed to satisfy the consistency principle should be equal to one for all combinations of values of the uncertain variables that have non-zero probability density. Otherwise, one can always find a case where the probability of the event: $x_1 \leq X_1 \cup x_2 \leq X_2 \ldots \cup x_n \leq X_n$ exceeds its possibility simply by increasing the number of variables. Of course, such a possibility distribution would be of little use because it would lead to overly conservative designs.

In conclusion, possibility theory is of little use in design of systems with a large number of independent failure modes.

On the other hand, the weak link of probability theory appears to be the way it handles insufficient data. Specifically, *if little data is available, then a probabilistic designer does not know what assumptions will increase the conservatism of his/her model*. The following example demonstrates this point.

In Fig. 2.4, we describe the uncertainty in the statement "*X* is about 10" with a uniform distribution over the interval [8, 12]. If we are less confident about the range of variable *X*, we can enlarge the interval, for example, to [7, 13] to reflect our increased uncertainty in the probabilistic model. This method of increasing the length of an interval in which an uncertain variable can vary maximizes Shannon Entropy (Shannon and Weaver, 1964). Entropy, which forms the basis of information theory, measures the uncertainty associated with predicting the results of a random experiment. We have to lower the probability density from 0.25 to 0.16667 so that its integral over the interval will remain one. This means that if there is a danger zone for failure localized to the interval [8.5, 9], the probability of failure is reduced from 12.5% to 8.3%. Thus, our increased uncertainty about the range of *X* reduces the probability of failure, which defies common sense. On the other hand, it is easy to increase the degree of conservatism of a possibility-based model in cases where little data is available. Here, we compare a triangular possibility distribution with a uniform probability density function because the former is the least conservative distribution, which is consistent with a uniform probability density, as we will explain in Chapter 3. As shown in Fig. 2.8, if we use triangular membership functions to describe the uncertainty in *X*, the possibility of failure assumes the maximum value at $X = 9$ for the failure zone [8.5, 9]. The possibility of failure increases from 50% to 66.7% as we extend the interval from [8, 12] to [7, 13] to account for lack of data.

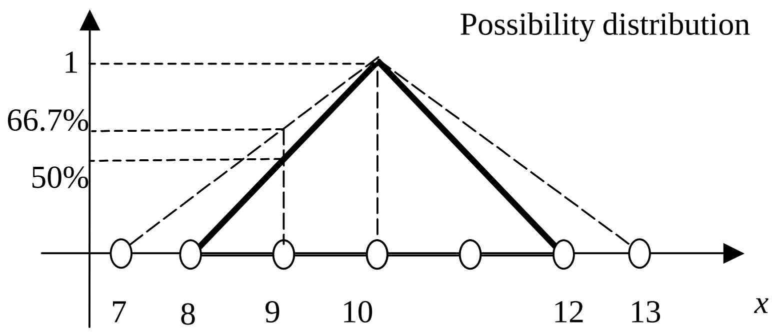

Figure 2.8 Possibility distribution of *X*, for the statement `*X* is about 10' when extending the interval from [8, 12] to [7, 13]

In highly redundant systems, many rare events have to happen simultaneously for failure to occur. For example, an offshore platform may only fail if several of its structural members fail and both the wind and wave loads on the structure are unusually high. In these cases, probabilistic methods tend to be less conservative in estimating the risk of failure than possibility-based methods because the size of failure zone is very small as explained in Section 2.2. If uncertainties are modeled accurately, then probabilistic methods are better than possibility-based methods because they yield more efficient designs. Consider however a scenario where these rare events that may lead to failure are highly correlated, and the correlation is not known. Then, if one ignores correlation and uses probabilistic methods one may come up with an unsafe design, whereas a possibility-based approach can yield a safer design. Indeed, we will show that assuming that the uncertainties are non-interactive when using a possibility-based approach always yields more conservative results.

Lemma 2.1 (Extracted from Nikolaidis et al., 1997):

The most conservative joint possibility distribution that is consistent with the marginal possibility distributions of the individual variables is the one corresponding to the assumption that the variables are non-interactive.

Proof: Let $\Pi'_{X_1,X_2}(x_1,x_2)$ be a joint possibility distribution of X_1 and X_2 that is consistent with the marginal possibility distributions of these variables. We will prove that $\Pi_{X_1X_2}(x_1,x_2)$ is greater or equal to $\Pi'_{X_1X_2}(x_1,x_2)$.

Let x_{10}, y_{20} be any pair of values of variables X_1 and X_2. Then,

$$\Pi_{X_1}(x_{10}) = \max_{x_2}(\Pi'_{X_1,X_2}(x_{10},x_2)) \tag{2.7a}$$

$$\text{and } \Pi_{X_2}(x_{20}) = \max_{x_1}(\Pi'_{X_1,X_2}(x_1,x_{20})) \tag{2.7b}$$

In Equations (2.7a) and (2.7b), the maximum is taken over all values of X_1 and X_2, respectively.

From equation (2.7a) we conclude that $\Pi_{X_1}(x_{10}) \geq \Pi'_{X_1,X_2}(x_{10},x_2)$, and from equation (2.7b) we conclude that $\Pi_{X_2}(x_{20}) \geq \Pi'_{X_1,X_2}(x_1,x_{20})$. Therefore, $\min(\Pi_{X_1}(x_{10}),\Pi_{X_2}(x_{20})) \geq \Pi'_{X_1,X_2}(x_{10},x_{20})$.

But $min(\Pi_{X_1}(x_{10}),\Pi_{X_2}(x_{20})) = \Pi_{X_1,X_2}(x_{10},x_{20})$.

Therefore, $\Pi_{X_1,X_2}(x_{10},x_{20}) \geq \Pi'_{X_1,X_2}(x_{10},x_{20})$.

Since the pair x_{10}, x_{20} is arbitrary, we can conclude that, out of all joint possibility distributions consistent with the marginal possibility distributions, the possibility distribution corresponding to the non interaction assumption is the most conservative, that is the one that yields the largest probability of failure of an event. Q.E.D.

It is also important to note that probabilistic methods may fail to predict the trends in the probability of failure because of design modifications. Suppose the true probability distribution is uniform over the range [8, 12]. Suppose that we could modify a design to move the failure zone in the example of the previous paragraph from the interval [8.5, 9] to [7.5, 8]. If we model X as a uniformly distributed random variable from [7, 13], we will conclude that this modification is useless because the probabilistic approach shows that the probability of failure remains 8.3%. In reality this modification is effective because it reduces the true failure probability from 12.5% to 0 because the modified interval [7.5, 8] is out of the failure zone. Note that a possibility model correctly predicts that this modification will increase safety.

In this section, we presented differences in probability and possibility measures when they are applied to estimate the safety of systems. We also explained why it is easier to determine the most conservative possibility model that is consistent with the data than to determine the most conservative probabilistic model consistent with the data.

2.5 Comparison of the ways probability and possibility maximize safety for a given budget

Consider a design problem in which two designers try to design a system to maximize safety using probabilistic and possibility based methods, respectively. For simplicity, consider that the system has only two failure modes. Consider that the events of failure under the first mode and the second mode are statistically independent. We will investigate how probabilistic and possibility-based methods allocate resources to maximize safety.

The probabilistic designer formulates the design problem as follows:

Find the optimum values of the design variables, $d_1,...,d_n$,
to minimize PF (2.8)
such that $g_j(d_1,...,d_n) \geq 0$, $j = 1,\ldots, m$.

In the above formulation, PF is the probability of failure of the system. The constraints prevent a designer from exceeding a budget and/or a maximum allowable weight.

The possibilistic designer uses the same formulation as the probabilistic designer except that his/her objective is to minimize the possibility of system failure, ΠF, instead of the probability of failure, PF.

The necessary conditions for the optimum of (2.8) are:

$$\frac{\partial PF}{\partial d_i} + \sum_{j=1,\dots m} \lambda_j g_j = 0 \quad \textit{for } i=1,\dots,n \tag{2.9}$$

Consider the case when λ_i=0, i.e. when none of the constraints is active at the optimum. Then, at the optimum, the derivatives of the system failure probability with respect to the design variables should be equal to zero. Since there are two failure modes:

$$\begin{aligned} \frac{\partial PF}{\partial d_i} &= \frac{\partial P(F_1 \cup F_2)}{\partial d_i} \\ &= \frac{\partial PF_1}{\partial d_i} + \frac{\partial PF_2}{\partial d_i} - \frac{\partial (PF_1 \cdot PF_2)}{\partial d_i} \\ &= 0 \\ \text{for } i &= 1,\dots,n \end{aligned} \tag{2.10}$$

If the probabilities of the two failure modes are small (e.g., 10^{-3}), then the third term is much smaller than the individual failure probabilities and can be omitted. Then the optimality condition (2.10) reduces to the following equation:

$$\left| \frac{\partial PF_1}{\partial d_i} \right| \approx \left| \frac{\partial PF_2}{\partial d_i} \right| \quad \textit{for } i=1,\dots,n \tag{2.11}$$

According to the above optimality condition, probabilistic optimization will *try to equate the absolute values of the derivatives of the probabilities of failure under the two failure modes*.

The possibility of system failure, consisting of two failure modes, is calculated by:

$$\Pi F = \Pi(F_1 \cup F_2) = \max(\Pi F_1, \Pi F_2) \tag{2.12}$$

The optimum design will minimize the possibility of system failure:

$$\min(\Pi F) = \min[\max(\Pi F_1, \Pi F_2)] \tag{2.13}$$

The optimum occurs when $\Pi F_1 = \Pi F_2$. Otherwise, any offset from this condition will either increase ΠF_1 or ΠF_2 or both ΠF_1 and ΠF_2. Therefore, for the possibilistic optimum design, if all constraints are inactive at the optimum, *the optimizer will try to equalize the possibilities of failure under the two failure modes*:

$$\Pi F_1 = \Pi F_2 \tag{2.14}$$

The different optimization mechanisms of the probabilistic and possibility-based designs can result in dramatically different optimum designs, as we will see in Chapter 3. Moreover, generally, the optimum of the possibilistic approach is less sensitive than the optimum of the probabilistic design. The reason is that, generally, the sensitivity derivatives of a quantity are more sensitive to errors than the quantity itself. For example, in strength of materials, stresses and strains are more sensitive to discretization errors than displacements. Another reason that the possibility-based optimum is less sensitive than the probabilistic optimum is that the possibility of an event is bounded between zero and one whereas the sensitivity derivative of the probability is unbounded.

As we will see in chapters four and five, if there is little data about the uncertain variables one has to make strong assumptions to construct models of the uncertain variables. The probabilistic optimum can be very sensitive to these assumptions. This indicates that a designer ought to invest more resources to collect more data so that he/she can build more accurate models of the random variables. If it is impractical to do so, then the probabilistic approach is of little use to the designer. In many real life design problems, probabilistic designers do not perform a parametric study to assess the sensitivity of the optimum designs to errors in modeling uncertainties. One reason is the high computational cost. In this case, the optimum probabilistic design can be too expensive or even unsafe, that is its true probability of failure can be unacceptably high.

2.6 Observations based on the comparison of theoretical foundations

The fundamental axiomatic difference between possibility and probability is that the probability of the union of a set of disjoint events is equal to the sum of the probabilities of each event. On the other hand, the possibility of the union of a finite number of events is equal to the maximum of the possibilities of these events. This leads to the following observations:

1. The possibilities of an event and its complement may add up to a number greater or equal to one. Moreover, either the possibility of that event or the possibility of its complement should be one. In contrast, the probabilities of an event and its complement must add up to one.
2. The possibility of failure of a series system consisting of identical, independent components is equal to the possibility of failure of one component. On the contrary, the probability of failure of the system increases as the number of components in series increases. As a result, a possibility-based method is likely to underestimate the chance of failure of a series system with large numbers of components. This property of possibility violates the possibility-probability consistency principle.
3. The possibility of failure of a system, consisting of identical, independent components connected in parallel, is equal to the possibility of a single component. The probability of failure of the system decreases as the number of components in the system increases. The failure zone for such a system is usually small compared to the range of the uncertain variables. The possibility-based method appears too conservative to predict the chance of the system failure. Compared to the possibility-based method, the probabilistic method may estimate more accurately the chance of failure of this system.
4. If little data is available to a probabilistic designer, he/she does not know how to increase the conservatism of the probabilistic model. In some problems, his/her rational effort to increase the uncertainty in a model leads to underestimation of the true probability of failure. In this dissertation, the term "true probability of an event", refers to the limit of the relative frequency of the event obtained from a number of independent replications of the same random experiment as the number of replications tends to infinity. On the contrary, a

designer who uses possibility can always increase the conservatism of his/her model when little data is available. This can be done by increasing the range in which the variables are assumed to vary and by assuming that the variables are non-interactive.

5. For a design problem to maximize the safety of a series system where two failure modes are present, the probabilistic optimization approach will try to equate the absolute values of the derivatives of the probabilities of failure under the two failure modes. On the contrary, the possibility-based optimizer will try to equate the possibilities of failure under the two failure modes. The different optimization mechanism in the probabilistic and possibility-based designs can result in dramatically different optimum designs.

From the above observations, we conclude that one cannot simulate the results of possibility calculus using probability calculus by scaling of the parameters of the probability distributions of the random variables.

In Chapter 3, we present an approach, which uses design, to compare probabilistic and possibility-based methods.

2.7 References

Billingsley, P., 1986, *Probability and Measure*, John Wiley & Sons, New York.

Klir, G., and Yuan, G., 1995, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey.

Laplace, P.S., 1812, *Theorie Analytique des Probabilities*, Courcier, Paris.

Nikolaidis, E., Haftka, R., and Rosca, R., 1998, "Comparison of Probabilistic and Possibility-Based Methods for Design Against Catastrophic Failure Under Uncertainty," Aerospace and Ocean Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-203. (can be downloaded from http://www.aoe.vt.edu/~nikolai)

Ochi, M. K., 1990, *Applied probability and stochastic processes in engineering and physical sciences*, Willey, New York

Papoulis, A., 1965, *Probability, Random Variables and Stochastic Processes,* McGraw-Hill, New York.

Shannon, C. E., and Weaver, W., 1964, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Illinois.

Sugeno, M., 1977, Fuzzy Measures and Fuzzy Intervals: A Survey," In: Gupta, M. M., Saridis, G. N., and Gaines, B. R., eds., *Fuzzy Automata and Decision Processes*, North-Holland, Amsterdam and New York, pp. 89-102.

Wood, K.L., Antonsson, E.K., 1990, "Representing Imprecision in Engineering Design - Comparing Fuzzy and Probability Calculus", *Research in Engineering Design*, Vol. 1, No.3/4, pp. 187-203.

Zadeh, L. A., 1975, "The Concept of a Linguistic Variable and Its application to Approximate reasoning I ," *Information Sciences*, Vol. 8, pp. 199-251.

Zimmermann, H. J., 1996, *Fuzzy Set Theory*, Kluwer Academic Publishers, Norwell, Massachusetts.

CHAPTER 3 TWO ANALYTICAL STUDIES COMPARING THE EFFICACIES OF THE METHODS IN MAXIMIZING SAFETY FOR A GIVEN BUDGET

In the previous Chapter, we compared theoretical foundations of probabilistic and fuzzy set design methods and concluded that we should use design rather than analysis to compare these two methods. In this Chapter, we first explain a general approach to compare probabilistic and fuzzy set methods using designs. Then, we describe a tuned-damper system and formulate probabilistic and fuzzy-set design optimization problems to minimize the risk of failure of the system for a given budget. Next, we present in detail the probabilistic approach to find an optimum as well as its fuzzy set counterpart approach. We also explain how to ensure these two design methods to use the same amount of information by building a probability distribution of the uncertain variable based on the available data and transforming the probability distribution into a possibility distribution that is consistent with the probability distribution.

3.1 An approach to compare probabilistic and fuzzy set methods

Figure 3.1 explains our approach of comparing methods for design in the presence of uncertainty. The key idea is to compare the safety of competing designs obtained by fuzzy set and probabilistic optimizations using the same resources and same amount of data about uncertainties.

To simulate real life design, where we rarely have enough data about uncertainties, we design using only a portion of the available information. Both probabilistic and fuzzy set based optimizations maximize safety, but they use different metrics of safety. Probabilistic design minimizes the failure probability, whereas fuzzy set design minimizes the failure possibility. Both use the same design variables. Because of the incomplete information, both techniques must work with inaccurate models of the uncertainties.

Once we have obtained two competing designs we compare them using complete information about uncertainties. Complete information means all the data needed to construct the true probability distributions of uncertain variables, which is infinite in the number of values. As noted in Chapter 1, this study considers only random uncertainties. Therefore, the designs are compared on the basis of their relative frequency of failure or probability of failure. This means that with complete information available, the probabilistic design is safer.

However, if little information is available, probabilistic design optimizes a design using estimates of the probability of failure, which can be significantly different from the true probability of failure. As explained in Chapter 2, designers who use probabilistic methods do not know how to make models more conservative to protect themselves from building unsafe designs because of lack of information. As a result, they may end up with designs that are less safe than their fuzzy set counterparts.

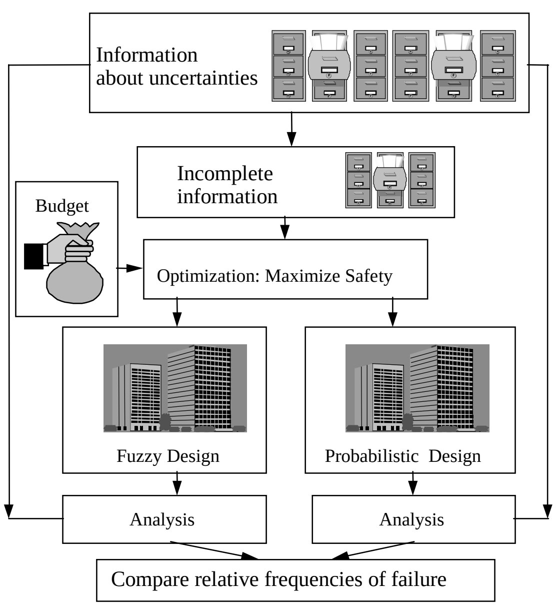

Figure 3.1. Analytical Comparison of Method

To demonstrate our approach of comparison, we present a design problem that involves the design of a damped single-degree of freedom system with a dynamic vibration absorber. The objective is to minimize the risk of failure due to cost overrun or performance shortfall. There are uncertainties in the system physical properties as well as in the construction budget. There are two narrow failure zones in the vicinity of the nominal values of the uncertain variables. In these types of problems, users of probabilistic design methods may seriously underestimate the risk of failure, if they assume large tolerances in the uncertain variables to account for the lack of data. These problems demonstrate the advantages of the fuzzy set based designs when there is limited information.

3.2 Design of a tuned damper system with parameter uncertainties

3.2.1 System description

Figure 3.2 illustrates a tuned damper system consisting of a single-degree of freedom system, called the original system, and a dynamic vibration absorber, which is used to reduce the vibration of the original system due to a harmonic excitation force. The amplitude of the displacement of the tuned damper system is normalized by the amplitude of the quasi-static response, and is denoted by *y*. The normalized amplitude of the vibration depends on the following system parameters: the mass ratio of the absorber to the original system (R); the damping ratio of the original system (ζ); the ratio of the natural frequency of the original system to the excitation frequency (β_1) and the ratio of the natural frequency of the absorber to the excitation frequency (β_2).

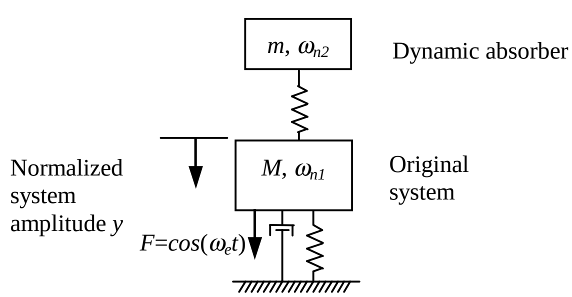

Figure 3.2 Tuned damper system

The normalized amplitude of the tuned damper system is calculated using the following equation:

$$y(\beta_1,\beta_2,R,\zeta)=\frac{\left|1-(\frac{1}{\beta_2})^2\right|}{\sqrt{\left[1-R(\frac{1}{\beta_1})^2-(\frac{1}{\beta_1})^2-(\frac{1}{\beta_2})^2+\frac{1}{\beta_1{}^2\beta_2{}^2}\right]^2+4\zeta^2\left[(\frac{1}{\beta_1})-\frac{1}{\beta_1\beta_2{}^2}\right]^2}} \tag{3.1}$$

This tuned damper system is ideal to demonstrate advantages of possibility-based methods and the pitfalls of probabilistic methods when little information is available for modeling uncertainties. The reason is that the probabilistic designer does not know how to make his/her model of the uncertainties more conservative to account for the lack of information. Figure 3.3 and 3.4 show the vibration amplitude of the tuned damper system as a function of β_1 and β_2. In Fig. 3.4, there is only one random variable β since β_1 and β_2 are assumed equal in the figure. Failure due to excessive vibration occurs in two narrow zones near the nominal values of β_1 and β_2.

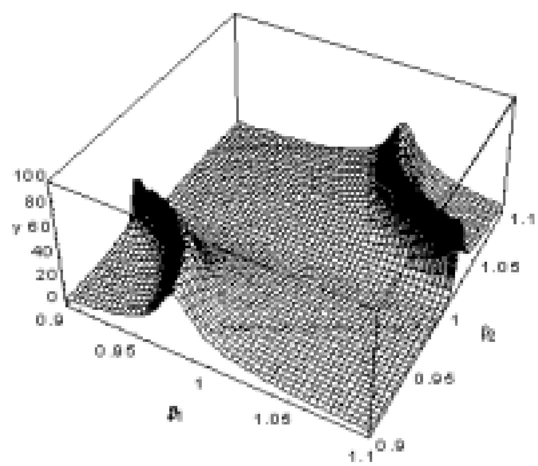

Figure 3.3 The normalized amplitude of the tuned damper system as a function of β_1 and β_2. ζ=1%, R=1%

In design problems where little information is available, designers who use probabilistic methods tend to assume a large tolerance on the uncertain variables to be conservative. Otherwise, they may miss the peak values of the amplitude (Fig. 3.4).

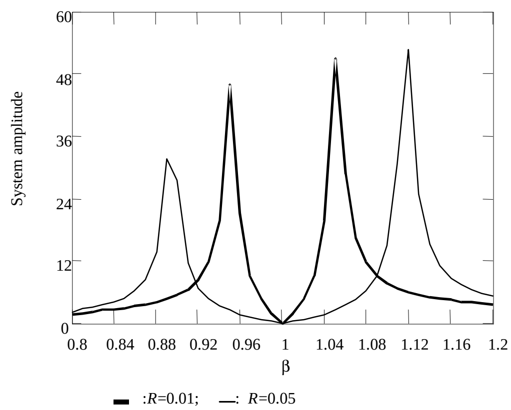

Figure 3.4 The normalized amplitude of the tuned damper system when $\beta_1 = \beta_2$ and ζ =1%

However, if a system has narrow failure zones, like the system in Fig. 3.4, increasing tolerances may produce a less conservative probabilistic model and lead to large errors in predicting the effect of design modifications on reliability. We have discussed this weakness of the probabilistic method in detail in Section 2.3. Figure 3.5 demonstrates how the standard deviation of normalized frequencies β_1 and β_2 affects the estimated probability of failure due to excessive vibration as a function of mass ratio *R*. The failure due to excessive vibration is defined to occur when the normalized amplitude exceeds 20. It is observed that even a small change in the calculated standard deviation of β_1 and β_2 dramatically changes the sensitivity of the probability of failure with respect to *R*.

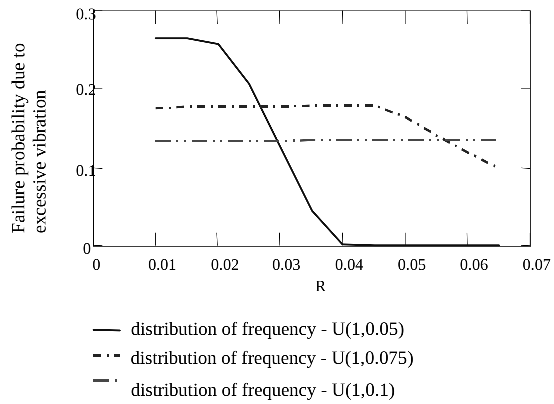

Figure 3.5 Effect of standard deviations of β_1 and β_2 on the probability of failure (y > 20) due to excessive vibration, β_1 and β_2 are equal

Specifically, if the normalized frequency is uniformly distributed about one with a standard deviation of 5%, the probability of failure due to excessive vibration decreases dramatically as the mass ratio, *R*, increases. However, if the standard deviation is 10%, increasing the mass ratio does not reduce the failure probability. When the standard deviation is small, increasing the mass pushes the failure zones outside the range of variation of the normalized frequency, which reduces the probability of failure (Fig. 3.4). When the standard deviation is larger (say 10%), the failure zones of all systems with *R* between 1% and 5% fall within the range of variation of the normalized frequency β so all these systems have nearly the same failure probability. Probabilistic methods can lead to poor designs in this type of problems if there is limited data about uncertainties because they cannot estimate the sensitivities of the probability of failure with respect to the design variables.

3.2.2 Design problem formulation

3.2.2.1 Failure modes

The following scenario is considered for the design problem: a tuned damper system is designed to have low vibration, to be insensitive to variations in the normalized natural frequencies β_1 and β_2, and to simultaneously satisfy the construction budget requirement of a client. The budget of the client has not been decided — the designer just knows the upper and lower limit in which it varies. Also, the cost is assumed proportional to the mass of the dynamic vibration absorber. The failure of this system consists of failure due to excessive vibration or construction cost overrun.

3.2.2.2 Uncertainties in the design problem

In Chapter 3 and 4, we assume that the random variability in β_1 and β_2 is the only source of uncertainty in the system parameters. Two cases are studied:

a) β_1 and β_2 are statistically independent. This is possible when the natural frequencies of the original system and the absorber are independent random variables, and the excitation frequency is deterministic.

b) β_1 and β_2 are equal (The two variables are perfectly correlated). This can happen if both natural frequencies are deterministic and the excitation frequency is uncertain.

To analytically simulate the real life scenario where a designer builds the probabilistic model with incomplete information about uncertainty, we use a random generator to produce sample values of β_1 and β_2. With these data as samples we calculate the statistics of β_1 and β_2.

The sample size is a main factor to influence the accuracy of probabilistic model. In this dissertation, we evaluate the effectiveness of the probabilistic and possibility-based methods as a function of the sample size.

Besides the uncertainties in the system properties, there is uncertainty in the budget, which is the maximum amount of money a client is willing to pay for a tuned damper system. In this dissertation the probability distribution of the budget is assumed known.

3.2.2.3 Formulation of design optimization problem

In this design problem, both failure modes involve uncertainties. The design variable is the mass ratio R. A light absorber is cheaper but not effective in reducing the vibration. On the other hand, a heavy absorber successfully dampens the vibration but is more likely to exceed the budget.

The objective of the design problem is to minimize the risk of excessive vibration or construction cost overrun by adjusting the mass ratio *R*. This ratio can vary between bounds R_l and R_u which are known. An optimization design problem is thus formulated as follows:

Find** R **to

Minimize: Risk of failure of the system due to excessive vibration or construction cost overrun

Subject to: $g_i(R, \zeta, \beta_1, \beta_2) \leq 0 \quad (i = 1, \ldots, m)$

In this problem, constraints are imposed on the allowable range of the mass ratio:

$$\begin{aligned} g_1(R) &= R_l - R \leq 0 \\ g_2(R) &= R - R_u \leq 0 \end{aligned} \tag{3.2}$$

3.3 Design using probabilistic method

3.3.1 The probabilistic design problem formulation

The probabilistic method minimizes the probability of failure in this problem. The design problem is formulated as:

Find** R **to

Minimize: $P(FS) = P(A \cup B)$ (3.3)

Subject to: $g_i(R) \leq 0 \quad (i = 1, 2)$

where $g_1(R)$ and $g_2(R)$ were defined in inequality(3.2).

P(A) stands for the probability that the construction cost of the system exceeds the budget; *P(B)* is the probability that vibration above a required safe level occurs, and *P(FS)* is the probability of system failure, which is the union of the above two failure modes. Section 3.3.3 presents the equations for determining if failure due to cost overrun or failure due to excessive vibration occurs, and the equations for computing the probabilities of failure under these two modes.

3.3.2 Creating conservative probability distributions from sample data

The probabilistic design models uncertainties with probability distributions. In this dissertation, the main objective is to study the effectiveness of probabilistic and possibility-based methods in the presence of incomplete information about the frequency ratios β_1 and β_2. In real life, the probabilistic designer first determines the distribution type available information or based on experience, and then he/she uses a statistical inference method to estimate the distribution parameters using the available sample values of uncertain variables. In this dissertation, two statistical inference methods are used for estimating the probability distributions of β_1 and β_2, the *standard statistical method* and *Bayesian method*. These methods are described below.

3.3.2.1 Standard statistical method

The standard statistical method uses the statistics of sample data to infer the distribution parameters of a random variable. For example, statistics such as mean and variance of a sample of size *N*, are used to estimate the mean and variance of the population, which are the true mean and standard deviation of the random variable respectively. As *N* tends to infinity, these estimates converge to the population parameters. In the tuned damper design problem, the population parameters of uncertain variables β_1 and β_2 are calculated using these estimates.

In this design problem, we assume that the population means of β_1 and β_2 are known to be one. This means that, on average, the dynamic vibration absorbers are perfectly tuned. The standard deviations of β_1 and β_2 are estimated from samples. Consider a random variable, *X*, and a sample $x_1, \ldots, x_N$ of size N. The *mean square error* of the sample, δ^2, is an unbiased estimate of the population variance σ^2 (Deming, 1966):

$$\delta^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{3.4}$$

The standard error of δ^2 can be derived using a standard inference method:

$$\begin{aligned} {\sigma_{\delta^2}}^2 &= \frac{1}{N}(\mu_4 - \sigma^4) \\ where: \mu_4 &= \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^4 \end{aligned} \tag{3.5}$$

where μ is the population mean of the variable X;

and σ_{δ^2} is the standard error of δ^2.

When a small sample (sample size less than 30) is used to estimate the parameters of the probability distribution of a random variable there is considerable statistical error in the estimates. It is common practice to assume a large tolerance for this random variable to account for the statistical error. Since the true values of the distribution parameters of the random variable are unknown (e.g., the mean, and the variance) they are treated as random variables. To account for the error in the estimate of the variance we scale the standard error in the variance by an inflation factor greater than one. This increases the range in which the variable can vary. In the tuned damper design problem, the variances of β_1 and β_2 are estimated from Equation (3.4) and the standard errors in these variances are calculated from Equation (3.5). To account for the statistical error, we increase the estimates of the variances obtained from Equation (3.4) by two times the square of the standard error:

$$\sigma^2 = \delta^2 + 2\sigma^2_{\delta^2} \tag{3.6}$$

Note that as the sample size used to estimate the variance of the normalized frequency increases the error in the estimate of the variance tends to zero, and so does the increase in the estimate of the variance on the right hand side of Equation (3.6).

3.3.2.2 Bayesian method

In Bayesian statistics, the unknown parameters of the probability distributions of the uncertain variables (e.g., the mean and the variance) are considered as random variables and have their own probability distributions.

Bayesian inference combines the objective information from numerical data and the prior information about a distribution parameter (which is usually based on judgment) into a posterior distribution of that parameter through a likelihood function.

A prior probability distribution for each distribution parameter is selected based on the available prior information, which can be either subjective or objective. If there is little knowledge about the system, the prior that has a minimal effect on the posterior distribution should be chosen. This kind of prior is called a *noninformative prior*. One approach to create a noninformative prior, which is called the *invariance principle*, is based on the idea that two variables that are related by a one-to-one transformation should have the same noninformative prior. For example, if a probability density with parameter σ has the form of $\sigma^{-1}f(x/\sigma)$, σ is called a *scale parameter*. Scale parameters control the shapes of probability density functions. The standard deviations in both normal and uniform distributions are examples of scale parameters. The scale parameters of two variables X and Y, related by a transformation $X = c\ Y$, should have the same noninformative prior according to the invariance principle. The noninformative prior for a scale parameter is $1/\sigma$. (Berger, 1985).

Assume that X is a random variable with uniform distribution $U(\mu,\sigma^2)$. The population mean, μ, is known to be one. The standard deviation, σ, is uncertain and its posterior distribution is estimated using information from the sample and an assumed prior. σ is a scale parameter and we adopt the noninformative prior, which is $1/\sigma$. The posterior probability density function and probability distribution of σ are derived as follows:

Using Bayes theorem, the posterior probability density function of σ is:

$$p(\sigma|X) = \begin{cases} 0 & for\ \sigma < \sigma_1 \\ \dfrac{N\sigma_1^N}{\sigma^{N+1}} & for\ \sigma \geq \sigma_1 \end{cases} \tag{3.7}$$

where: N is the sample size, X is the sample set. σ_1 is the lower boundary of σ, which is determined using the following equation:

$$1-\sqrt{3}\sigma_1 \leq x_1 \leq x_N \leq 1+\sqrt{3}\sigma_1 \tag{3.8}$$

where x_1 is the smallest value, x_N is the biggest value in the sample.

Thus, σ_1 is the bigger of the following quantities:

$$\frac{1-x_1}{\sqrt{3}}\ and\ \frac{x_N-1}{\sqrt{3}} \tag{3.9}$$

The posterior probability distribution of σ is found by integrating the probability density function of σ in Equation (3.7):

$$P(\sigma|X) = \begin{cases} 0 & for\,\sigma < \sigma_1 \\ 1-\dfrac{\sigma_1^N}{\sigma^N} & for\,\sigma \geq \sigma_1 \end{cases} \tag{3.10}$$

3.3.3 Calculating the probability of failure

The probability of failure due to excessive vibration over the maximum allowable limit is evaluated by *Monte Carlo simulation*. 100,000 repetitions are performed. If the standard statistical method is adopted to infer the population parameter, samples for β_1 and β_2 are generated from the estimated probability distributions of β_1 and β_2.

When a Bayesian method is used to infer a population parameter, σ, the posterior distribution of σ is obtained from a set of sample data. In that case, the probability of failure, $P(B)$, is the expectation of the conditional probability of failure given the value of σ, over all possible values of σ (Der Kiureghian, 1990):

$$\begin{aligned} P(B) &= E[P_f(\sigma)] \\ &= \int_{\sigma} P_f(\sigma)\, p(\sigma)\, d\sigma \\ &= \iint_{Failure\ Region} p(\boldsymbol{x},\sigma)\, p(\sigma)\, d\boldsymbol{x}\, d\sigma \end{aligned} \tag{3.11}$$

where $P_f(\sigma)$ is the conditional probability of failure given the value of σ, $p(\sigma)$ is the posterior density of σ, and $p(\boldsymbol{x}, \sigma)$ is the posterior joint density of the vector of random variables, $\boldsymbol{x}$, and σ.

In the problem of finding the probability of failure of the system with the dynamic vibration absorber, $\boldsymbol{x}=\{\beta_1, \beta_2\}^T$. Using Equation (3.11) the probability of failure could be evaluated from the joint distribution of $\boldsymbol{x}$ and σ by Monte Carlo simulation.

In this dissertation, we assume that the designer knows the true distribution of the budget. The construction cost of the system consists of a constant component and a component that is a linear function of the tuned damper mass, R. The equation for the construction cost, ic, is as follows:

$$ic = 20 + 500R \tag{3.12}$$

We also assume that the budget is a random variable uniformly distributed between [20, 200]; therefore, the probability of failure due to cost overrun can be evaluated as:

$$P(A)=\frac{(ic-20)}{180} \tag{3.13}$$

The probability of failure due to excessive vibration or cost overrun is:

$$P(FS) = P(A\cup B)= P(A)+P(B)-P(A)P(B) \tag{3.14}$$

3.3.4 Finding the optimal design

In this dissertation, the mass ratio can assume 11 discrete values equally distributed from 1% to 5.5%. For each mass ratio, the probability of failure due to excessive vibration or construction cost overrun can be determined as described in the previous subsection. Among all the designs with different mass ratios, the one with the smallest estimated failure probability is the optimum.

3.4 Design using the fuzzy set method

3.4.1 The fuzzy set design problem formulation

The possibility-based method minimizes the possibility of failure of the system. The formulation of the possibility-based optimization is:

Find** R **to

Minimize: Max [$\Pi(A),\Pi(B)$]

Subject to: $g_i(R) \le 0$ ($i = 1, 2$)

where $g_1(R)$ and $g_2(R)$ were defined in the inequality (3.2).

Π (A) is the possibility that the construction cost of the system exceeds the allowable budget; *Π(B)* is the possibility that vibration exceeds the safety level. Note that both probabilistic and possibility-based designers use the same amount of information and have to satisfy the same constraints. However, they use different metrics to define the chance of failure of a design.

3.4.2 Creating a possibility distribution that is consistent with the probability distribution

The probabilistic design method and the possibility-based method differ in the way they model uncertainties and evaluate the risk of system failure. It is important to ensure that both models are constructed using the same data. For this purpose, first we construct a probabilistic model of an uncertain variable, as described in the previous section, and then we transform this model into a possibility distribution.

The least conservative principle is used to construct a possibility distribution that is consistent with the given probability distribution (Nikolaidis et. Al., 1997). *This principle is based on the concept that, among all transformations that yield possibility distributions consistent with a given probability distribution, the one that results in the minimum loss of information is the best*. A possibility distribution is said to be consistent with a probability distribution if the possibility of any event is greater than or equal to its probability. In the following we will derive a necessary condition that a possibility distribution must satisfy so that it is consistent with a given probability distribution.

Figure 3.6 shows the probability density function *f(x)* for variable *X*. Assume that this distribution is unimodal. The corresponding probability distribution of *X* is *F(x)*. Suppose *Π(x)* is a possibility distribution consistent with the probability distribution of *X*.

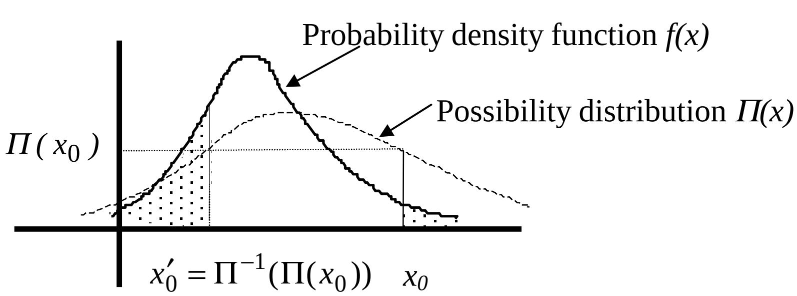

Figure 3.6 A probability density function and its consistent possibility distribution function

Consider a value x_0 to the right of the apex of the possibility distribution function. All events that have possibility of occurrence, $\Pi(x_0)$, have the form $(x \le x_l \cup x \ge x_0)$ or $(x \le x_0' \cup x \ge x_u)$. Let x_0' is the value of X on the increasing side of the possibility distribution with the same possibility as x_0, x_l is any value of X smaller than x_0' and x_u is any value of X larger than x_0. Of all events that have possibility $\Pi(x_0)$, the one that has the highest probability is the event $A = (X \le x_0' \cup X \ge x_0)$. The probability of this event is the shaded area in Fig. 3.6. The reason is that the probabilities of events $X \le x_l$ and $X \ge x_u$, becomes maximal when $x_l = x_0'$ and $x_u = x_0$.

The probability of the event E is represented by:

$$\begin{aligned} P(E) &= P(X \le x_0') + P(X \ge x_0) \\ &= F(x_0') + 1 - F(x_0) \end{aligned} \tag{3.15}$$

Therefore, if the possibility distribution of X is consistent with the probability distribution of the same variable, then the following inequality holds:

$$\begin{aligned} &\Pi(A) \ge P(A) \Leftrightarrow \\ &\Pi(x_0) \ge F(x_0') + 1 - F(x_0) \end{aligned} \tag{3.16}$$

If x_0 locates to the left of the apex, a similar inequality can be obtained:

$$\Pi(x_0) \ge F(x_0) + 1 - F(x_0') \tag{3.17}$$

We have proven that, inequalities (3.16) and (3.17) are necessary conditions so that the possibility distribution $\Pi(x)$ is consistent with the probability distribution $F(x)$.

We will prove that inequalities (3.16) and (3.17) are also sufficient conditions. Suppose that the possibility distribution and the probability distribution of a variable, *X*, satisfy these inequalities for any value that variable *X* can assume. Consider an arbitrary event, *E'*, with possibility Π. Because Inequlities (3.16) and (3.17) are satisfied, the probability of any event that has possibility Π, is equal to or smaller than Π. Therefore, the probability of *E'* is equal to or smaller than the possibility of *E'*, which means that the possibility distribution of *X* is consistent with the probability distribution of the same variable.

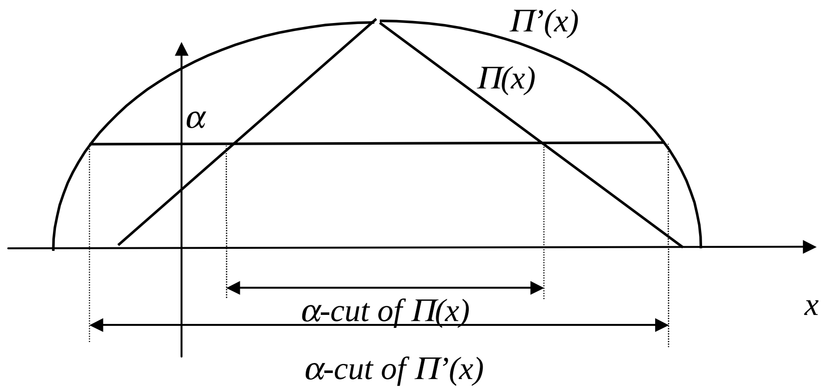

Figure 3.7 Demonstration of the proposition "possibility distribution *Π'(x)* is more conservative than *Π(x)*"

For a given probability density function, there are infinite possibility distributions satisfying the consistency principle expressed by inequalities (3.16) and (3.17). We will impose an additional requirement that the possibility distribution be the least conservative among all possibility distributions consistent with a given probability distribution. A possibility distribution function, *Π(x)*, is defined to be less conservative than another possibility distribution function *Π'(x)* if for any *x*, $\Pi(x) \le \Pi'(x)$. For example, if we are not sure about tomorrow's weather, we may want to say that the possibility that it will rain tomorrow is one instead of 0.3.

All α-cuts of $\Pi(x)$ are subsets of the corresponding α cuts of $\Pi'(x)$. Fig. 3.7 demonstrates two possibility distributions with different degrees of conservatism.

Unfortunately, even when inequality (3.16) becomes equality, it does not completely define the least conservative possibility distribution for a given probability distribution. The reason is that for a given value of x_0, there are infinite choices for x'_0. However, if the probability density is symmetric, by introducing symmetry as an additional requirement for the possibility distribution we can resolve this difficulty (Fig.3.8). In the following, we consider only symmetric probability density functions. To find a unique solution to the above equations, we impose the restriction that the possibility distribution is symmetric and its peak has the same value as the mode x_{mod} of the probability density function. The mode in a probability density function is the value with the biggest probability density. It is reasonable to assign the highest possibility to this value.

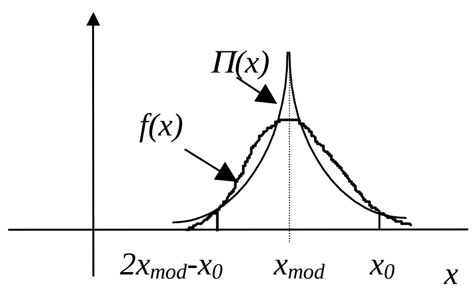

Figure 3.8 A probability density function and its least conservative consistent possibility distribution, which is symmetric and the symmetric axis coincides x_{mod}

With above additional restrictions, inequalities (3.16) and (3.17) become:

$$\Pi(x) = \begin{cases} F(x) + 1 - F(2x_{\text{mod}} - x) & x < x_{\text{mod}} \\ F(2x_{\text{mod}} - x) + 1 - F(x) & x \geq x_{\text{mod}} \end{cases} \tag{3.18}$$

If the random variable x has a uniform distribution over interval $[a, b]$, the probability distribution of this random variable is:

$$F(x) = \frac{x-a}{b-a} \text{ for } a \le x \le b. \tag{3.19}$$

In a uniform probability distribution, there is no mode over the interval. All x values are equally probable. We assume the symmetric apex of the least conservative possibility distribution locates at the middle point of the interval [a, b]. Applying Equation (3.18) with $x_{mod} = (a+b)/2$:

$$\Pi(x) = \begin{cases} 2\dfrac{x-a}{b-a} & a \le x < \dfrac{a+b}{2} \\ 2(1-\dfrac{x-a}{b-a}) & \dfrac{a+b}{2} \le x \le b \end{cases} \tag{3.20}$$

The least conservative possibility distribution consistent with a uniform density function is a symmetric triangular distribution with apex at $(a+b)/2$. Its support is the interval [a, b] (Fig. 3.9):

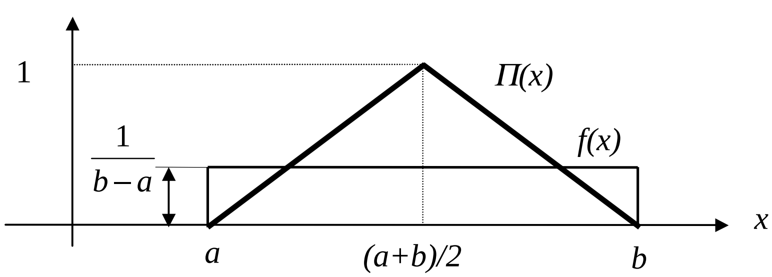

Figure 3.9 Least conservative possibility distribution that is consistent with a uniform probability density function

When using the Bayesian approach to model the uncertainty in the normalized frequencies, the standard deviations of these frequencies are treated as random variables. We could use the mean value of the standard deviation of a frequency, σ, as an estimate of σ. However, since little information is available about the frequencies of the original system and the absorber, we want to increase the assumed tolerance in the uncertain frequency. To do so, we calculate the standard deviation of σ and increase this mean value of σ by two standard deviations

$$\sigma = E(\sigma) + 2\sigma_{\sigma} \tag{3.21}$$

where $E(\sigma)$ is the posterior mean of σ, and σ_σ is the standard deviation of σ. The latter statistic is calculated using the posterior distribution $P(\sigma/X)$ obtained from Equation (3.10).

After estimating the standard deviation of the normalized frequency, we transform the probability distribution of the normalized frequency into a possibility distribution using the least conservative distribution principle.

3.4.3 Calculating the possibility of failure

The possibility of failure due to excessive vibration is evaluated using the *vertex method* (Dong and Shah, 1987). This method allows us to calculate the possibility distribution for the system amplitude that is a function of R, ζ, β_1 and β_2, from the possibility distributions of β_1 and β_2. The possibility of failure due to excessive vibration, $\Pi(B)$, is the maximum value of the possibility distribution of the vibration amplitude over all values of the amplitude that are greater or equal to the maximum allowable amplitude. As we know, the budget is uniformly distributed between [20, 200]. Its consistent possibility distribution $\Pi(A)$ is a triangular function over the same interval that has the maximum at the middle of the interval.

The possibility of failure of events A and B (failure due to budget overrun or excessive vibration, respectively) is:

$$\Pi(A \cup B) = Max(\Pi(A), \Pi(B)) \tag{3.22}$$

3.4.4 Finding the optimal design

To determine the optimum possibilistic design we compute the possibilities of failure of designs with mass ratios, *R*, from 1% to 5.5%. For a given mass ratio, the possibility of failure due to excessive vibration or construction cost overrun can be determined using Equation (3.22). The optimum mass ratio is selected as the value that corresponds to the smallest possibility of system failure.

3.5 References

Berger, J. O., 1985, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.

Der Kiureghian, A., 1990, "Bayesian Analysis of Model Uncertainty in Structural Reliability", 3^{rd} IFIP, March.

Deming, W. E., 1966, *Some Theory of Sampling*, Dover, New York.

Dong, W. and Shah, H. C., 1987, "Vertex Method for Computing Functions of Fuzzy Variables", *Fuzzy Sets and Systems*, 24, pp. 65-78

Nikolaidis, E., Haftka, R., and Rosca, R., 1998, "Comparison of Probabilistic and Possibility-Based Methods for Design Against Catastrophic Failure Under Uncertainty," Aerospace and Ocean Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-203. (can be downloaded from http://www.aoe.vt.edu/~nikolai)

CHAPTER 4 COMPARING THE RESULTING OPTIMAL DESIGNS

In Chapter 3, we developed analytical models of probabilistic and fuzzy set methods for the tuned-damper design problem. This Chapter we performs numerical analysis to compare these two methods. First, we explain the criteria to justify which method is better. Next, we list factors that we consider in comparisons. Finally, we present comparison results with bar charts and summarize observations from these results.

4.1 Calculating the true probability of failure for both approaches

This design problem considers only crisply defined failure. Therefore, the designs are compared on the basis of their relative frequency of failure (in an experimental comparison) or the probability of failure (in a comparison based on computer simulation). This means that with complete information available, the probabilistic design is the safest one.

However, if little information is available, probabilistic methods optimize a design using estimates of the probability of failure, which can be significantly different than the true probability of failure. Moreover, probabilistic design may fail to predict accurately the sensitivity derivatives of the probability of failure with respect to the design variables as we show in Section 2.4. In these cases, it can yield less safe designs than possibility-based methods.

For each design, the probability of failure due to excessive vibration is calculated from Monte Carlo simulations using sample values of uncertain variables β_1 and β_2. The probability of failure due to construction cost overrun is determined analytically as a function of the mass ratio. The true optimum mass ratio is the one with the smallest true probability of failure.

For comparison, the possibility distributions of β_1 and β_2 are derived according to the least conservative principle from the true probability distributions of β_1 and β_2. The optimum mass ratio R corresponding to the probabilistic method is thus calculated.

Table 4.1a shows true probability distributions for β_1 and β_2, as well as the probability distribution for budget.

Table.4.1a The true distribution parameters of the budget, β_1 and β_2

Distribution parameters	β_1 and β_2	Budget
Mean	1.0	100
Standard deviation	0.05	$100\sqrt{3}/3$

Table 4.1b presents the optimum probabilistic and possibility-based designs as described in the previous two paragraphs. The true probability distributions of the frequency ratios β_1 and β_2 are uniform with mean value of one and standard deviation of 0.05. From Table 4.1b we observe that the optimum mass ratios obtained using Equation 3.3 are sensitive to the correlation and the type of distribution of β_1 and β_2, while the optimum mass ratios obtained using the possibility-based method are the same for different degrees of correlation and the type of distribution of β_1 and β_2.

Table.4.1b The true optimum *R* and true probability of failure, *P(FS)*, when the type of distributions and the standard deviations of β_1 and β_2 are known

Degree of correlation and type of distribution of β_1 and β_2	**Probabilistic method**		**Possibility-based method**	
	True *P(FS)* at optimum	Optimum *R*	True *P(FS)* at optimum	Optimum *R*
β_1 and β_2 are equal (True distribution is uniform)	11.3%	4%	19.7%	3%
β_1 and β_2 are independent (True distribution is uniform)	9.85%	3%	9.85%	3%

Figure 4.1 plots the probability of the system failure vs. mass ratio. β_1 and β_2 are equal and follow a uniform distribution. We observe that when mass ratio is small, the system has a high probability of failure due to the inefficacy of a small mass to absorb vibrations. Though a

damper with big mass is efficient to reduce vibrations, it also tends to increase the cost of construction. The optimum tuned-damper has a mass ratio 4%, which minimizes the probability contributed by the two failure modes in the problem.

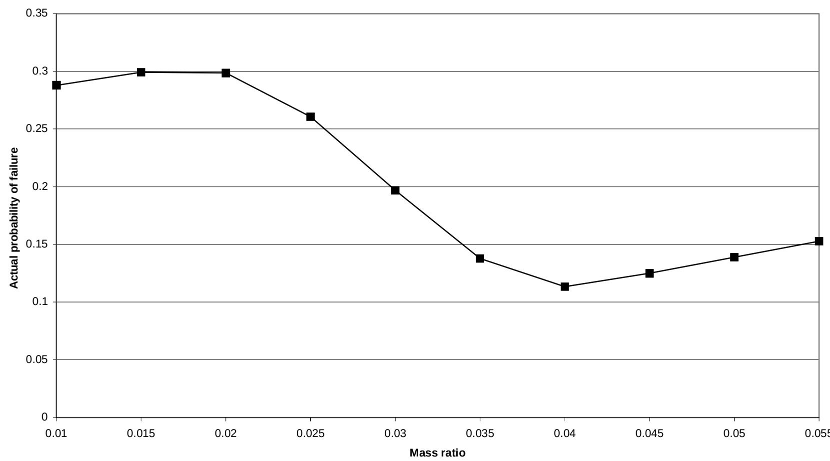

Figure 4.1 Probability of failure of the system vs. mass ratio, $\beta_1 = \beta_2$

4.2 Factors considered in the comparison

In this dissertation, the following four factors are investigated for comparing the efficacy of probabilistic and possibility-based methods: a) the size of the samples of values of β_1 and β_2 used to estimate the probability distributions of these variables; b) the correlation between β_1 and β_2; c) the method used to infer the statistics of β_1 and β_2; and d) the error in the type of probability distribution of β_1 and β_2. In this dissertation, we do not study the effect of errors in the correlation between variables.

Sample size reflects the amount of information one has to estimate the probability distributions of uncertain variables. The larger the sample size, the more accurate the probability distributions estimated from the sample are. Seven sample sizes of 3, 5, 10, 20, 100, 1000 and 3000 are considered in this dissertation. For the first four sample sizes, it is difficult to identify the true probability distribution using hypothesis tests, such as the χ^2 test. That is, probability distributions that are significantly different from the true probability distribution pass the χ^2 test.

The last three sample sizes represent situations when information about uncertainty is sufficient to accurately identify the true type of probabilistic distribution with χ^2 test.

In this dissertation, random variables β_1 and β_2 are either statistically independent or equal. The statistics of β_1 and β_2 are inferred by either the standard statistical method or by the Bayesian method. Two cases are studied to investigate the effect of the error in the type of the probability distribution: the true probability distribution of the uncertain variables is uniform while a normal distribution is assumed and vice-versa.

4.3 Results of the analytical comparison

The effectiveness of the two methods is assessed by considering:

1. Which method produces safer designs, on average?
2. Which method produces designs whose optimum masses or failure probabilities are less sensitive to sample-to-sample variation?

The mean values of the normalized frequencies β_1 and β_2 both are equal to one. The true standard deviations are 5%. Failure due to excessive vibration is assumed to occur when the normalized amplitude, y, exceeds 20. Note that without the dynamic vibration absorber, the normalized amplitude at resonance is 50. 100,000 replications are used in Monte Carlo simulation to calculate the true probability of failure due to excessive vibration.

Figures 4.2 and 4.3 compare the results of the two methods, when the standard deviations of the probability distributions of β_1 and β_2 are unknown and estimated from 3 and 3000 sample points, respectively using the probabilistic/standard statistical method.

Figure 4.2 presents 10 pairs of designs obtained using 10 sets of sample values. Each sample has three points, generated from a uniform distribution. The true standard deviations β_1 and β_2 are both 5%. This figure corresponds to a case where: a) β_1 and β_2 are equal, b) the designer knows the true type of correlation and distribution of these random variables, and c) he/she uses the standard statistical method to infer the standard deviations of β_1 and β_2. Table 4.2 shows the standard deviations estimated using Equation 3.6. The same table shows the optimum mass ratios of the probabilistic and possibility-based designs.

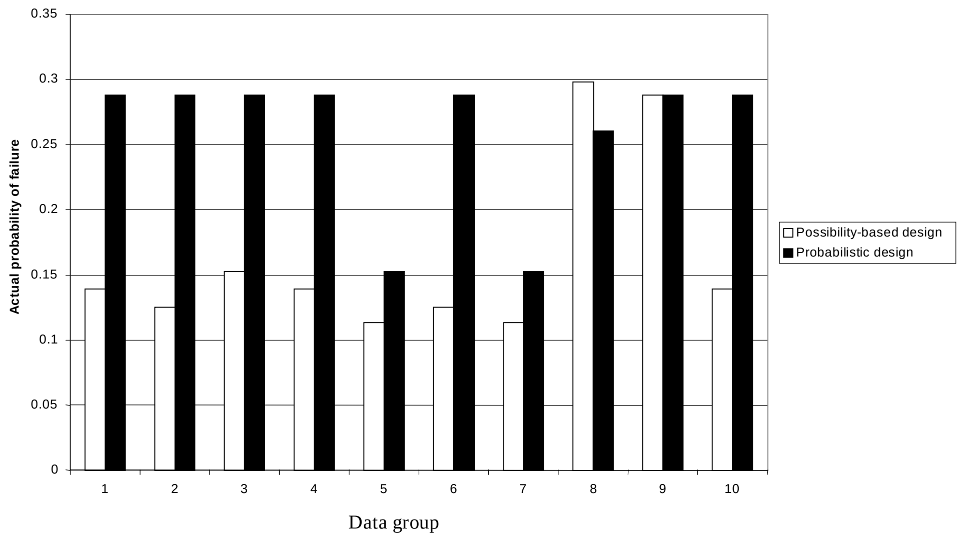

Figure 4.2 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. Three sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.2 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (3 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	1%	1%	1%	1%	5.5%	1%	5.5%	2.5%	1%	1%
Optimum *R*: Possibility-based method	5%	4.5%	5.5%	5%	4%	4.5%	4%	2%	1%	5%
σ (from Equation 3.6)	7.54%	6.73%	8.21%	7.58%	6.29%	7.09%	6.17%	3.65%	1.96%	7.70%

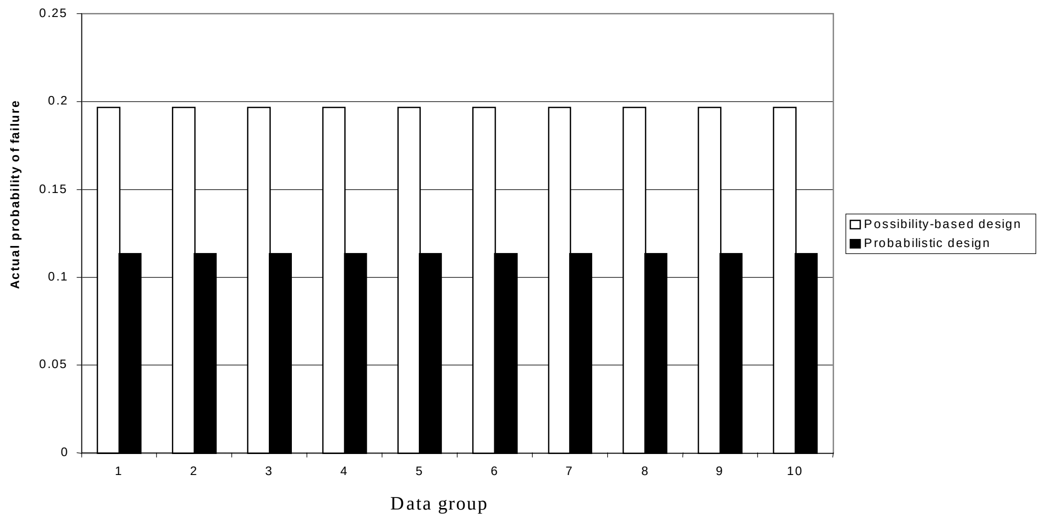

Figure 4.3 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. 3000 sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.3 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (3000 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	4%	4%	4%	4%	4%	4%	4%	4%	4%	4%
Optimum *R*: Possibility-based method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
σ (from Equation 3.6)	5.03%	5.08%	5.00%	5.03%	5.09%	5.09%	5.09%	5.09%	5.06%	5.11%

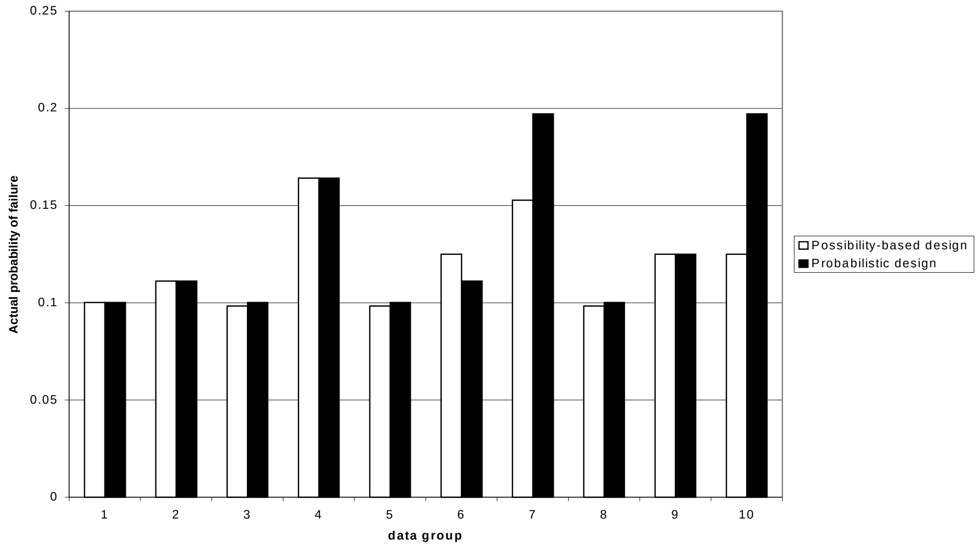

Figure 4.4 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are independent. Three sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.4 Probabilistic vs. possibility-based optima when β_1 and β_2 are independent (3 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	3.5%	4%	3.5%	1.5%	3.5%	4%	1%	3.5%	4.5%	1%
Optimum *R*: Possibility-based method	3.5%	4%	3%	1.5%	3%	4.5%	5.5%	3%	4.5%	4.5%
σ_1 (from Equation 3.6)	7.61%	8.13%	6.42%	5.98%	5.90%	7.59%	8.19%	6.92%	5.23%	8.78%
σ_2 (from Equation 3.6)	4.58%	4.95%	4.31%	1.92%	4.74%	6.93%	8.49%	4.00%	8.13%	6.38%

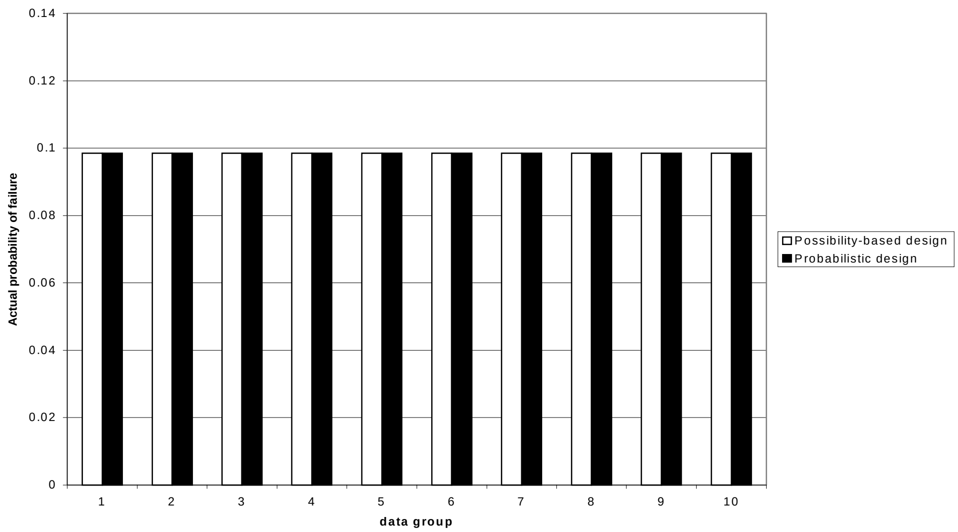

Figure 4.5 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are independent. 3000 sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.5 Probabilistic vs. possibility-based optima when β_1 and β_2 are independent (3000 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
Optimum *R*: Possibility-based method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
σ_1 (from Equation 3.6)	5.04%	5.05%	5.11%	5.12%	5.07%	5.03%	5.09%	5.07%	5.10%	5.01%
σ_2 (from Equation 3.6)	5.02%	5.07%	5.05%	5.05%	5.05%	5.12%	5.07%	5.09%	5.01%	5.10%

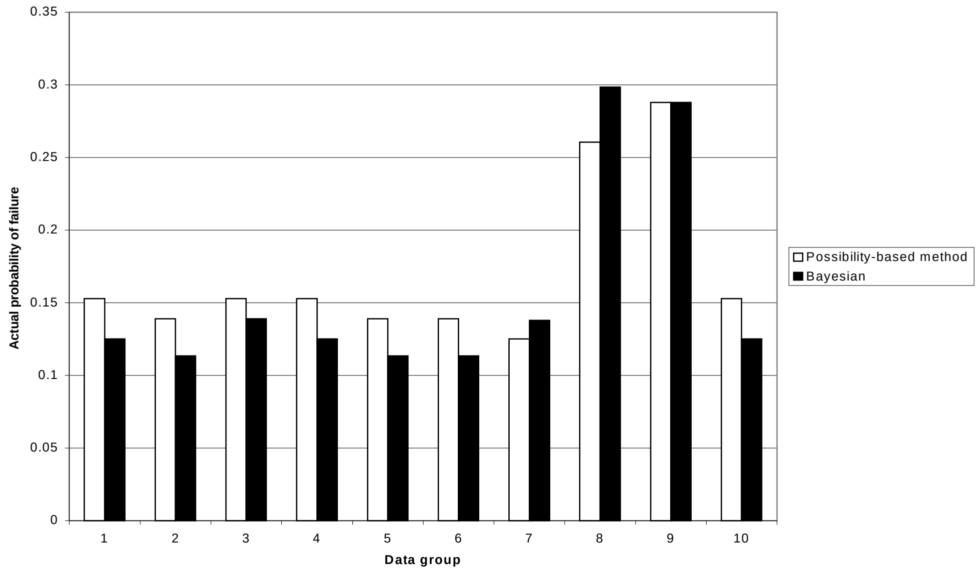

Figure 4.6 Bayesian probabilistic vs. possibility-based method, β_1 and β_2 are equal. Three sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.6 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (3 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	4.5%	4%	5%	4.5%	4%	4%	3.5%	2%	1%	4.5%
Optimum *R*: Possibility-based method	5.5%	5%	5.5%	5.5%	5%	5%	4.5%	2.5%	1%	5.5%
σ	4.58%	3.81%	5%	4.54%	3.85%	4.01%	3.53%	2.12%	1.1%	4.73%

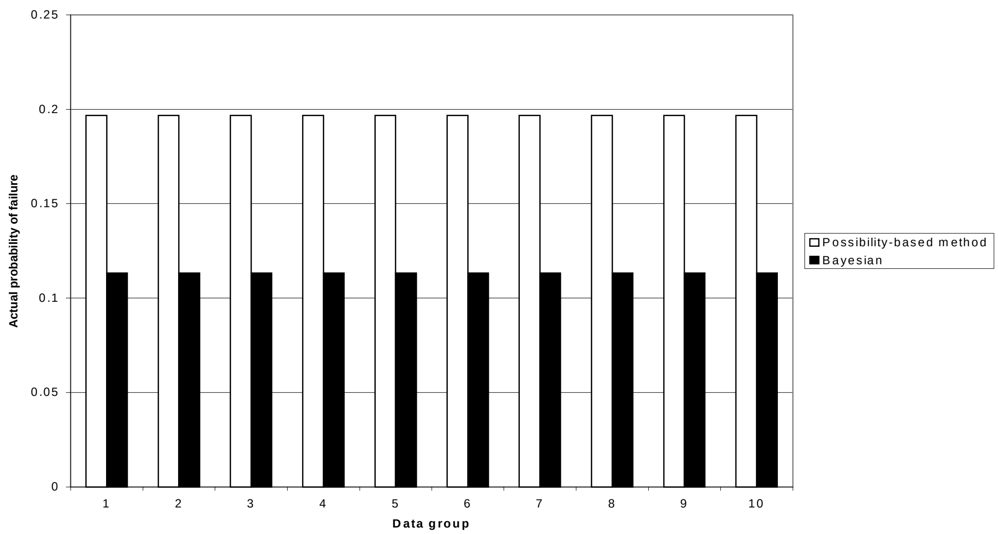

Figure 4.7 Bayesian probabilistic vs. possibility-based method, β_1 and β_2 are equal. 100 sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.7 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (100 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	4%	4%	4%	4%	4%	4%	4%	4%	4%	4%
Optimum *R*: Possibility-based method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
σ	5%	4.92%	4.98%	4.91%	4.99%	5%	4.97%	4.99%	5%	4.9%

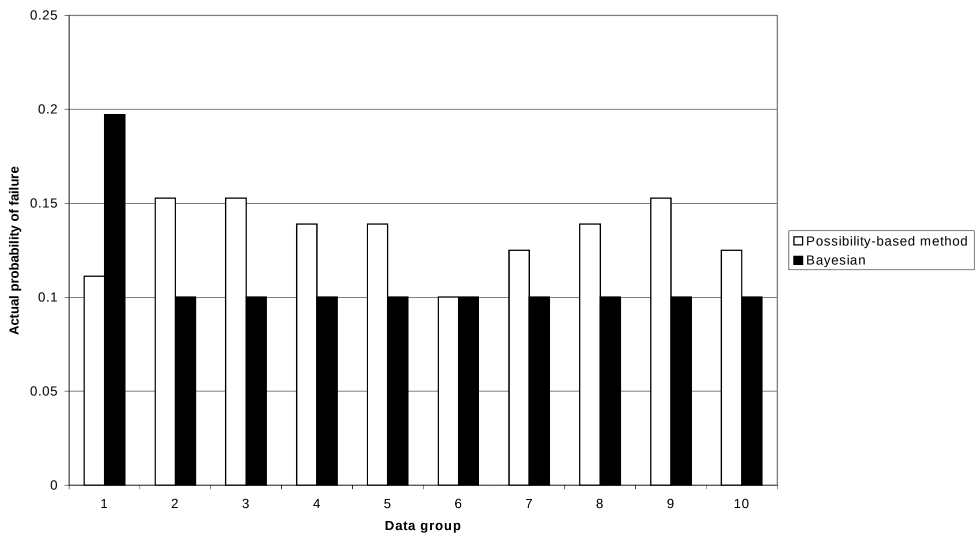

Figure 4.8 Bayesian probabilistic vs. possibility-based method, β_1 and β_2 are independent. Three sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.8 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (3 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	1%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%
Optimum *R*: Possibility-based method	4%	5.5%	5.5%	5%	5%	3.5%	4.5%	5%	5.5%	4.5%
σ_1	4.6%	4.8%	4.9%	4.7%	3.5%	2.5%	3.5%	4.0%	4.1%	2.9%
σ_2	2.6%	3.9%	3.8%	3.8%	4.4%	2.9%	3.5%	3.8%	4.2%	4.1%

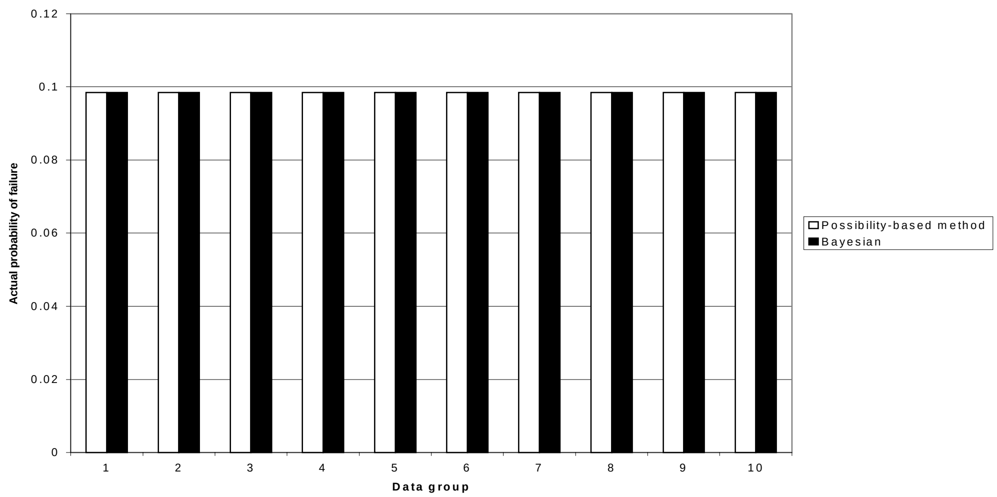

Figure 4.9 Bayesian probabilistic vs. possibility-based method, β_1 and β_2 are independent. 100 sample values of β_1 and β_2 are used to estimate their standard deviations

Table 4.9 Probabilistic vs. possibility-based optima when β_1 and β_2 are equal (100 sample points)

	Design 1	Design 2	Design 3	Design 4	Design 5	Design 6	Design 7	Design 8	Design 9	Design 10
Optimum *R*: Probabilistic method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
Optimum *R*: Possibility-based method	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
σ_1	4.95%	4.98%	4.95%	4.98%	5%	4.92%	4.99%	5%	4.97%	4.89%
σ_2	5%	4.97%	4.99%	4.97%	4.94%	4.96%	4.97%	4.93%	4.99%	4.91%

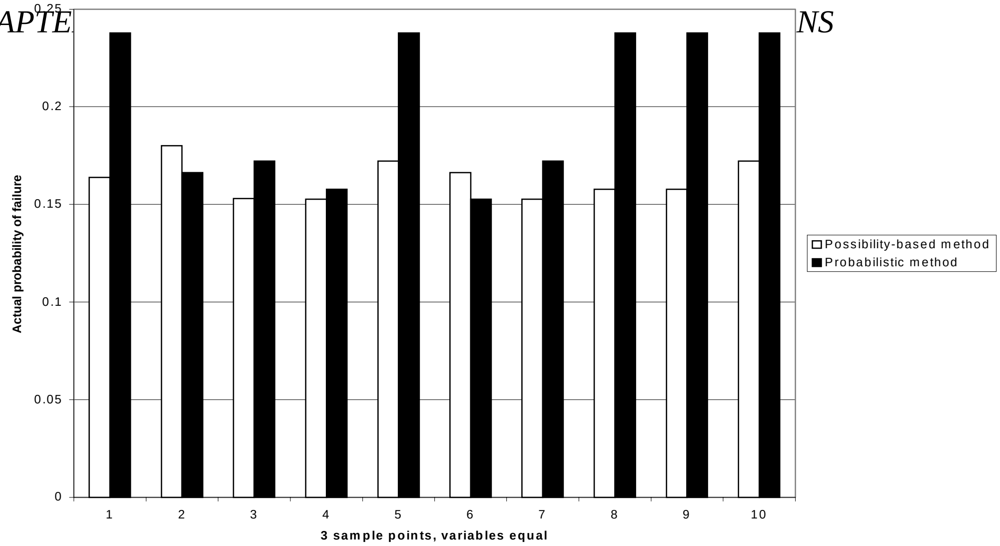

Figure 4.10 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. True distribution is normal, assumed is uniform. Three sample values of β_1 and β_2 are used to estimate their standard deviations

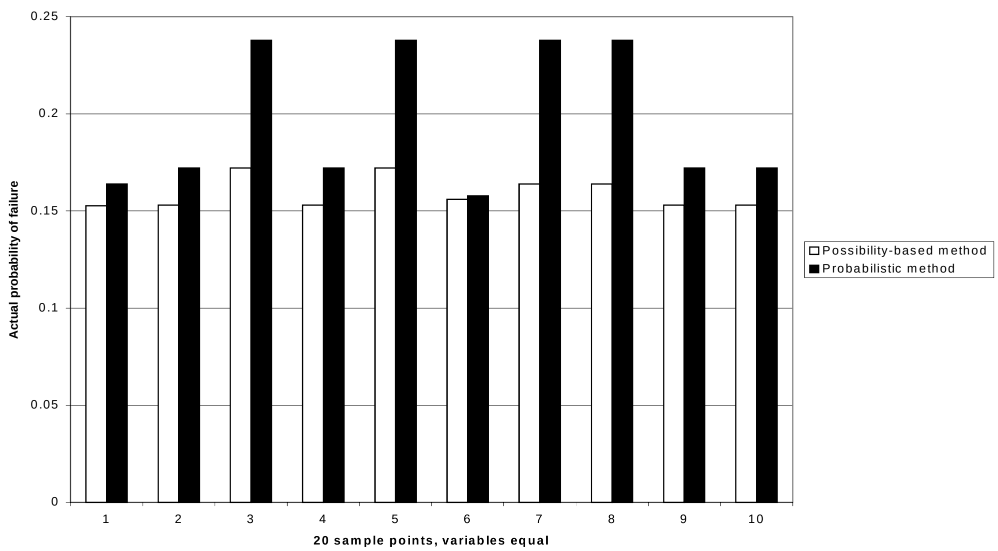

Figure 4.11 Standard statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. True distribution is normal, assumed is uniform. 20 sample values of β_1 and β_2 are used to estimate their standard deviations

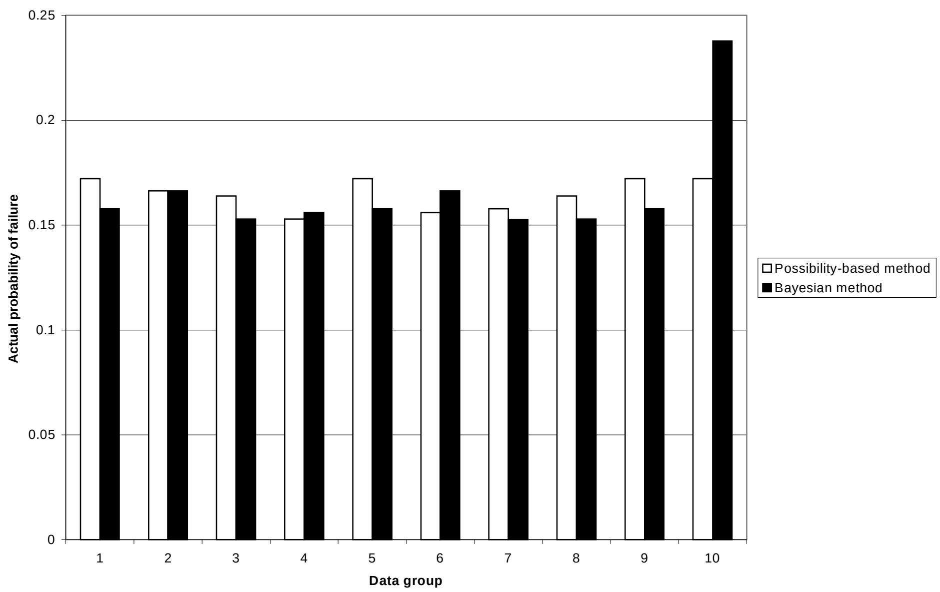

Figure 4.12 Bayesian statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. True distribution is normal, assumed is uniform. Three sample values of β_1 and β_2 are used to estimate their standard deviations

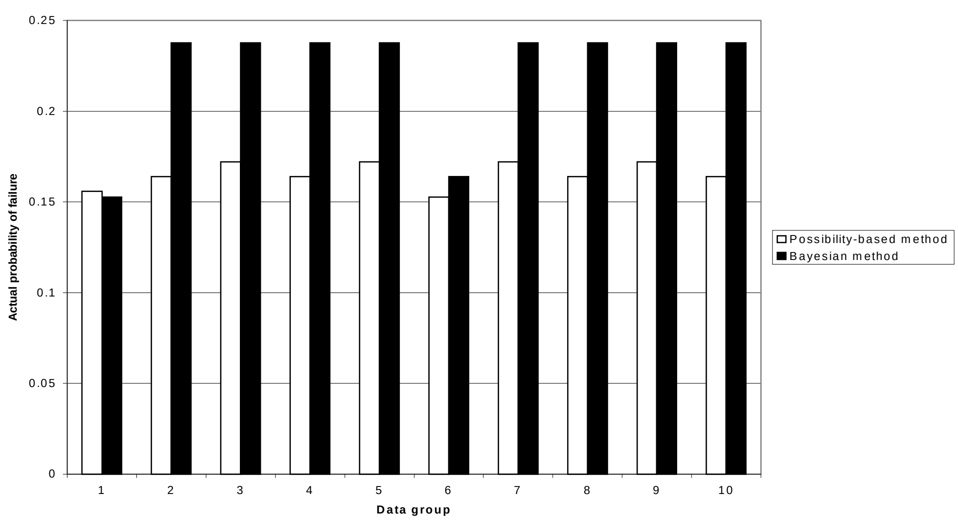

Figure 4.13 Bayesian statistical probabilistic vs. possibility-based method, β_1 and β_2 are equal. True distribution is uniform, assumed is normal. 20 sample values of β_1 and β_2 are used to estimate their standard deviations

Eight out of 10 probabilistic designs have higher probabilities of failure than their possibility-based counterparts. Among these probabilistic designs, seven designs have optimum mass ratio of 1%, which means that the lightest possible absorber is used. In reality, this light absorber is cheap but unsafe in terms of excessive vibration. This absorber has a higher probability of failure of 28.79%. The possibility-based method, on the other hand, is likely to choose a heavier absorber, and on average, yields safer designs than the probabilistic method.

Figure 4.3 and Table 4.3 compare probabilistic and possibility-based designs for a similar problem, except that the sample size is 3000 instead of three. The estimated standard deviations of β_1 and β_2 are very close to their true values. All probabilistic designs have optimum R of 4%. On the other hand, all the possibility-based designs have R equal to 3%. The true probability of failure of these possibility-based designs 19.68% is considerably larger than the failure probabilities of their probabilistic counterparts 11.33%.

Figure 4.4 and Table 4.4 compare probabilistic and possibility-based designs when random variables β_1 and β_2 are independent. Three sample values are used in each design to create probabilistic models. The standard statistical method is used to infer the standard deviations of β_1 and β_2. Two out of 10 probabilistic designs produce a mass ratio of 1%, which has significant higher true probabilities of failure than their possibility-based counterparts. Three other probabilistic designs slightly exceed the possibility-based designs in the true probability of failure. Four designs have the same probabilistic and possibility-based results. The sixth data group has an inferior possibility-based design. Possibility-based designs are better than probabilistic designs on average even if it is not so evident as the case when two uncertain variables are equal.

Figure 4.5 and Table 4.5 compare probabilistic and possibility-based designs when sample size is 3000 and two variables are independent. The standard statistical method is used to estimate the standard deviations of β_1 and β_2 which are very close to their true values in Table 4.5. All probabilistic and possibility-based designs have the same optimum R of 3% with a true probability of failure as 9.85%.

Figure 4.6 and Table 4.6 compare probabilistic and possibility-based designs when Bayesian statistics is used to estimate the distribution parameters for β_1 and β_2 when they are equal. Three sample values are used in each design to create probabilistic models. Seven out of 10 possibility-based designs produce higher true probabilities of failure than their probabilistic counterparts. Two possibility-based designs are safer than the probabilistic designs in the true probability of failure. One design has the same probabilistic and possibility-based results. Probabilistic designs are better than possibility-based designs on average if Bayesian method is used to estimate the standard deviations of uncertain variables. From Table 4.6, we observe that the standard deviations estimated by Bayesian method are close to the true value even if only three sample values are available.

Figure 4.7 and Table 4.7 compare probabilistic and possibility-based designs when sample size is 100 and two variables are equal. Bayesian method is used to estimate the standard deviations of β_1 and β_2. All probabilistic designs have the same optimum R of 4% compared with all same optimum designs at sample size 3000 with the standard statistical method. Bayesian method is more efficient than the standard statistical method to create a probabilistic model using few sample data for the prior distribution we chose.

Figure 4.8 and Table 4.8 compare probabilistic and possibility-based designs when Bayesian statistics is used to estimate the distribution parameters for β_1 and β_2 when they are independent. Three sample values are used in each design to create probabilistic models. Eight possibility-based designs produce higher true probabilities of failure than their probabilistic counterparts. One possibility-based design is safer than the possibility-based designs in the true probability of failure. One design has the same probabilistic and possibility-based results. Probabilistic designs are better than possibility-based designs on average if Bayesian method is used to estimate the standard deviations of uncertain variables. From Table 4.8, we notice that Bayesian method does not inflate the standard deviation of the uncertain variable. This is the reason that why most Bayesian probabilistic designs do not underestimate the probability of failure and therefore are conservative.

Figure 4.9 and Table 4.9 compare probabilistic and possibility-based designs when sample size is 100 and two variables are independent. Bayesian method is used to estimate the standard deviations of β_1 and β_2. All probabilistic and possibility-based designs have the same optimum R of 3%.

Figure 4.10 compares probabilistic and possibility-based designs when true distribution is normal and the designer assumes a wrong distribution type, which is uniform. β_1 and β_2 are equal. Three sample points are used to create probabilistic model. Here the standard statistical method is used to estimate the standard deviation of distributions. Eight probabilistic designs have higher true probability of failure, compared to one possibility-based design with higher true probability of failure. The other group has the same optimum for both methods. Possibility-based methods are better in this case.

Figure 4.11 compares probabilistic and possibility-based designs when true distribution is normal and the designer assumes a wrong distribution type, which is uniform. β_1 and β_2 are equal. 20 sample points are used to create probabilistic model. In this case, a χ^2 test can not distinguish two distributions. The standard statistical method is used to estimate the standard deviation of distributions. All probabilistic designs have higher true probability of failure. The accumulation of data for a wrong probabilistic model will worsen probabilistic designs.

Figure 4.12 compares probabilistic and possibility-based designs when true distribution is normal and the designer assumes a uniform distribution. β_1 and β_2 are equal. Three sample points are used to create probabilistic model. The Bayesian statistical method is used to estimate the standard deviation of distributions. Three probabilistic designs have higher true probability of failure, compared to six possibility-based designs with higher true probability of failure. The other group has the same optimum for both methods. Probabilistic methods are better in this case.

Figure 4.13 compares probabilistic and possibility-based designs when true distribution is normal and the designer assumes a uniform distribution. β_1 and β_2 are equal. 20 sample points are used to create probabilistic model. The Bayesian statistical method is used to estimate the standard deviation of distributions. Nine out of 10 probabilistic designs have higher true probability of failure, whereas the other one has lower true probability of failure.

Figure 4.14 shows a coordinate system P1 - P3 whose axes represent the factors considered when assessing the effectiveness of the two methods. P1 represents the correlation of the frequencies, P2 the type of probabilistic analysis and P3 the type of probability distribution of the frequency. A cube, whose vertices correspond to the combination of factors considered in each comparison, is used to summarize the results (Figure 4.14). The vertex at the origin represents the case where: a) the frequencies are independent, b) we use a probabilistic/standard statistical method to estimate the standard deviations of β_1 and β_2, c) use the true type of probability distribution of the frequencies. For each vertex of the cube in Fig.4.14, we compare probabilistic and possibility-based methods using samples of seven different sizes ranging from 3 to 3000.

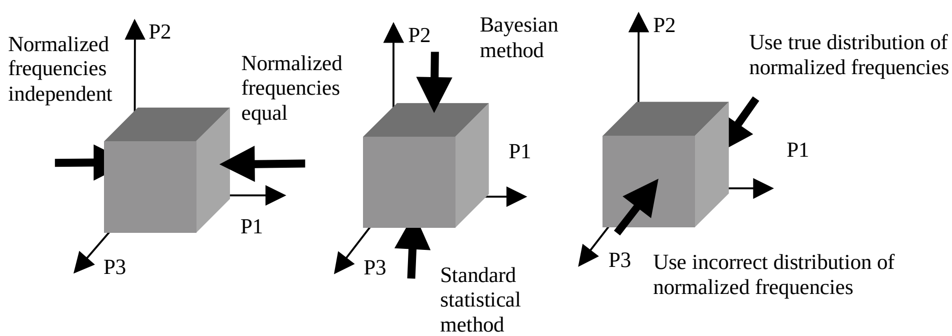

Figure 4.14 Summary of factors considered

According to the above paragraph, a total of 7×2×2×2 = 56 cases are studied. In each case, 10 sets of sample values of frequencies are generated, and 10 pairs of alternative probabilistic and possibility-based designs are compared — one pair for each sample. Thus, 560 comparisons are performed.

Figures 4.15 and 4.16 compare the methods in terms of a) safety of the optimum designs and b) the variability in the failure probabilities. Black, white and gray bullets are used to show which method was found better in each case. A black bullet means that the probabilistic method is better, a white bullet means that the possibility-based method is better, and a gray bullet means that the results did not make clear which method is better.

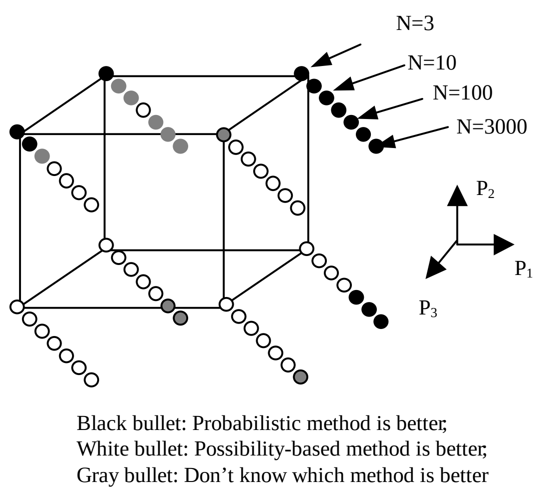

Figure 4.15 Effect of sample size on effectiveness of method in terms of true failure probabilities at optimum designs

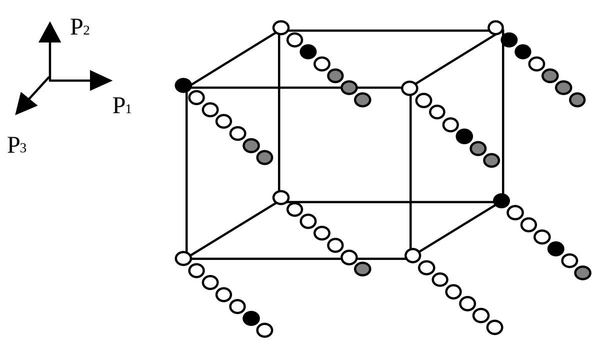

Figure 4.16 Effect of sample size on method in terms of consistency of the true failure effectiveness of probabilities of optimum designs

4.4 Observations on the relative merits of each design method

The following trends are observed from the results in the previous section:

- When the probabilistic/standard statistical method is used (*P2*=0), the possibility-based method is better for small sample sizes ($N \leq 20$) regardless of type of correlation and probability distribution used (true or wrong).
- When we know the true distribution (*P3*=0) and use large sample size ($N \geq 100$), the probabilistic/ standard statistical method yields safer designs.
- In cases where the true type of the probability distribution is not known (*P3*=1), a possibility-based method yields safer designs. This observation applies to cases with small sample sizes.
- When the sample size is large enough ($N \approx 40$), a hypothesis test can be used to identify if the sample data fit to the assumed distribution.
- In cases where the true type of the probability distribution is known (*P3*=0), the Bayesian probabilistic method (*P2*=1) is better than the possibility-based method, even for small samples.
- In cases where the wrong type of distribution is used (*P3*=1), possibility-based method yields safer designs than the Bayesian method (*P2*=1). This is more significant when two variables are equal (*P1*=1).
- Overall, the true failure probabilities and the optimal mass ratios of the possibility-based designs are less sensitive to sample-to-sample variation than those of the probabilistic designs are. There are 36 white bullets, 8 black bullets and 12 gray bullets in Figure 4.15.

In general, probabilistic design is better when sufficient information is available about uncertainties because, in contrast to possibility-based design, it accounts for the sensitivity of the failure probability to the mass ratio when seeking the optimum design. This allows the probabilistic design to trade effectively requirements for low cost and high performance(low vibration).

However, the probability of failure and the sensitivity of the failure probability with respect to design variables can be completely wrong if little information is available (Figure 3.5). For small samples and/or when the wrong type of probability distribution is used, the probabilistic design overestimates the variability in the frequencies (Table 4.2), and grossly underestimates the probability of failure for low mass ratios (Figure 3.5). This also results in severe errors in the sensitivity of the probability of failure with respect to the mass ratio. (Figure 3.5) The probabilistic design opts for the lowest possible mass to minimize the probability of budget overrun, because it neglects the effect of the mass on the probability of failure (see the curve corresponding to σ =10% in Figure 3.5). This happened in most of the cases where the sample size was three or five and when the wrong type of probability distribution was assumed. For example, six out of the 10 optimum probabilistic designs in Figure 4.2 have a mass ratio of 1% instead of 4%, which is the true optimum. These designs have high system failure probability (about 28%).

The Bayesian method is very effective when the true type of probability distribution of the frequency is known, because it uses this information when estimating the standard deviation of the frequency. However, when the type of distribution is not known, the use of information based on the wrong type of distribution becomes a disadvantage and the Bayesian approach tends to yield inferior designs to the possibility-based designs (Figure 4.12~4.13).

CHAPTER 5 ANALYTICAL-EXPRIMENTAL COMPARISON OF PROBABILISTIC AND POSSIBILITY-BASED METHODS USING A PROBLEM INVOLVING DESIGN AND CONSTRUCTION OF DOMINO STACKS

In chapter two, we examined the theoretical foundations of probability and possibility. We showed that possibility-based methods can be better than probabilistic methods in design of systems that fail only when many things go wrong simultaneously, when there is little data about uncertainty. In chapters three and four, we demonstrated this on a numerical example. Here we use a problem involving design and construction of stacks of dominoes to demonstrate the same assertion experimentally. In section 5.1 we describe this design problem. Sections 5.2 and 5.4 describe a probabilistic approach and a possibility-based approach to the design problem, respectively. Section 5.5 explains how to compare these two methods analytically and experimentally. Section 5.6 presents the comparison results and draws conclusions about the effectiveness of the probabilistic and possibility-based methods.

5.1 Design problem formulation

5.1.1 Design problem description

The objectives of this chapter are to present an efficient experimental approach for comparing probabilistic and possibility-based methods for design under uncertainty and apply this approach to demonstrate that possibility-based methods can be better than probability-based methods when there is little data about uncertainty. Experimental comparison of methods for design under uncertainty requires a large number of experiments so that statistically significant results can be obtained. Very few experimental comparisons of methods for design under uncertainty have been performed because the cost for performing such experiments is too high. We need a type of problem for which there exists an experimental procedure for comparing methods for design under uncertainty in the problem that is cheap and easy to repeat many times. The design problems should involve modeling errors, human errors and randomness. A design problem, which involves building stacks of dominoes (Rosca and Haftka, 2000), is suitable for comparing methods for design under uncertainty because it has the following features: There are minimal requirements for a builder's skill and building facilities. Experiments can be performed rapidly (in few minutes) and failure (collapse of domino towers) is not destructive. Therefore, the cost of experiments is low. The experiments involve the most important types of uncertainties encountered in real-life design problems, such as randomness because of geometrical imperfections in domino blocks, human errors in the building process and modeling errors in developing analytical models for predicting the collapse of a domino tower. We will implement this idea to compare probabilistic and possibility-based methods.

Consider a design situation where two designers are asked to build a domino tower (called tower A) close to another tower (called tower B) using probabilistic and possibility-based methods, respectively (Figure 5.1). Each designer tries to build a tall and stable tower. However, the designer does not know exactly the minimum acceptable height for a tower and of course he/she does not know if the tower she/he will try to build will collapse. Moreover, the designer has to design tower A before he/she is told where to build it. Therefore, he/she does not know the exact location where he/she has to build the tower.

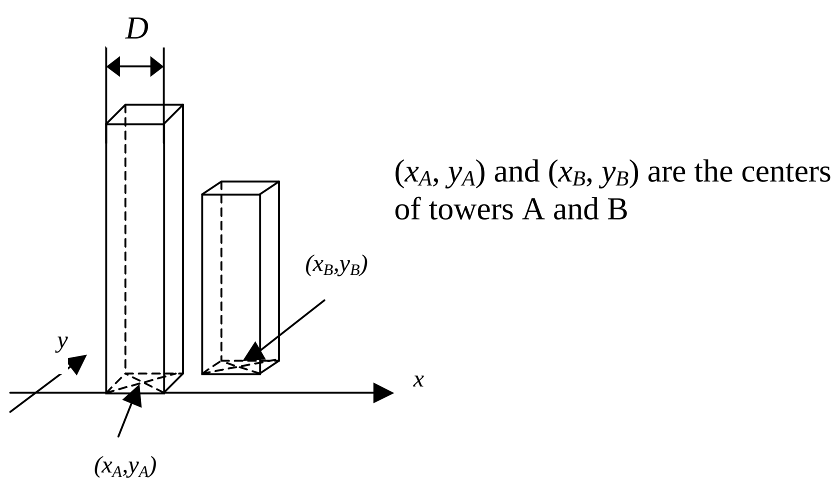

Figure 5.1 Building Tower A close to Tower B

Suppose that the coordinates of the centers of these two towers are (x_A, y_A) and (x_B, y_B), respectively. y_A and y_B are almost equal and known to the designers. When towers *A* and *B* overlap in the *x* direction, the distance between x_A and x_B, which is denoted as d_0, is smaller than the length of a domino block *D*. The designers also know that x_A and x_B are uniformly distributed, however, they have to estimate the distributions of these variables from a few measurements. In this problem, when x_A and x_B are equal to their mean values respectively, the towers do not overlap.

We can find an analogous scenario in real life. One company is going to construct a building near another building in an earthquake-prone area. One of the buildings is short and safe to earthquake. The company buys insurance against earthquake for the other building whose height is the design variable. The company will determine the height of this building based on two concerns: 1. If the building is tall, it is likely to fall in the event of an earthquake and might damage the other building; 2. If the building is short, it might fail to meet future demand for space, and the company will have to build another building. Each event will cause the company a significant monetary loss. The company will suffer the highest loss if one building collapses in the earthquake and destroys the neighboring building. The next highest loss will occur if the building is short because the company must construct another building to meet space demand. In the event that the insured building collapses but does not damage the uninsured building, the insurance company will pay most of the loss. Therefore, this is the event with the smallest loss. The exact locations of the two buildings are not known because they depend on zoning decisions, and there is very limited information about this uncertainty. Since this design involves uncertainties, the company gives this project to two designers, who use probabilistic and possibility-based designs respectively, to find out the optimum building height that minimizes the expected monetary loss.

We use the domino design problem to simulate this real life situation. In this domino design problem, the designer selects the height of a tower A built next to tower B. Either a short tower or a tower that has collapsed is a failure. Each failure mode entails a monetary loss. If a designer builds a tall tower it is likely to collapse. In this case, it may knock down tower B if the latter is close to tower A. On the other hand, if the designer sets up a modest target for the tower height he may end up with a stable but short tower. Monetary loss for this failure mode is medium. If the tower collapses but misses B, the loss will be the lowest. Each designer has to minimize the risk of monetary loss caused by building a tower that is too short or a tower that collapses.

The two designers obtain the same information for their designs. They know the probability distribution of the height at which tower A collapses. They know the height of tower B. The designers also obtain a few measurements of the locations of a tower A and B to model the uncertainty in their locations. These measurements are extracted randomly from the probability distributions of the locations of two towers. Though the designers do not know the exact minimum acceptable height, they know the range within which the minimum acceptable height varies.

One designer uses probability while the other uses possibility to model the uncertainties. Designs obtained by these two methods are evaluated based on the incurred losses and the probability distributions at this height. The best design has the lowest loss.

5.1.2 Uncertainties in the domino design problem

In this problem there are irreducible uncertainties in the locations of the towers, the minimum acceptable height of tower A and the collapse height of tower A. First, the exact locations of towers *A* and *B* are unknown. This uncertainty contains irreducible and reducible parts. The available information about locations is a sample of measurements, which the designers will use to construct probabilistic or possibilistic models of the uncertainty in the locations of the towers, respectively. If we had infinite measurements we could construct accurate probability distributions for these locations and eliminate the reducible part of the uncertainty. However, even though there are sufficient measurements to build the true probability distributions of locations of the towers, the designer still does not know the exact location when designing a new tower. Second, the two designers do not know the minimum acceptable height. They know that the minimum acceptable height is a random variable that is uniformly distributed between known lower and upper limits. These uncertainties are irreducible. Even though there is sufficient data to construct the true probability distribution of the minimum acceptable height, the value of this variable for a particular design is still unknown. Finally, there is uncertainty in the collapse height of tower A. This uncertainty consists of a reducible part and an irreducible part. Even if we know the true probability distribution of the collapse height we can not predict the collapse height each time we build a new tower.

There is also reducible uncertainty due to lack of knowledge, which can be eliminated by collecting data or refining our predictive model. As we mentioned before, there is reducible uncertainty in the locations of tower A and B. Moreover, the true probability distribution of the collapse height of a tower is estimated using data from one hundred experiments. Designers assume that the distribution is normal. This uncertainty is reducible because the designers can estimate the probability distribution more accurately by collecting more data about the collapse height. There is reducible uncertainty because of assumptions the designers make to predict whether collapse of tower A will entails collapse of tower B.

In Fig.5.1, the longer side of a domino block is along the *x* direction. When tower A collapses, it falls in the *y* direction. The designers assume that the distance between tower A and *B* in the *y* direction is so small that, if the two towers overlap in the *x* direction, the collapse of tower A always entails collapse of tower B. As we will see in the later sections, this is a strong assumption.

This problem is designed to demonstrate the limitations of probabilistic methods. If the probabilistic designer has complete information about the probability distribution of locations, he/she will find an optimum height, which minimizes the expected loss of all modes of failure. When there is very few data about the locations of towers A and B (e.g., three data points), it is common practice to assume large standard deviations for these locations to account for the large reducible uncertainty. However, in this problem, this practice may lead to underestimation of the probability that both towers will collapse. If the standard deviations of distributions of x_A and x_B are small, the probability of the failure mode that has the highest penalty (collapse of both towers) is small. As the probabilistic designer increases the standard deviation to be conservative, the probability increases because x_A and x_B are more likely to overlap with each other. However, when the standard deviations exceed some value, this probability will start decreasing as the standard deviations of x_A and x_B keep increasing. Thus, the probabilistic designer may underestimate the probability of the most severe failure mode and build designs that are found unsafe when perform experiments.

5.1.3 Analytical model of the design problem

In this design problem, failure (denoted as F) occurs if tower A fails to meet the height requirement or tower A collapses. Let S be the event that the design height is shorter than the minimum acceptable limit. Let E_C be the event that tower A collapse and E^c_C be its complementary event that tower A is stable. Failure can be interpreted as the union of *A* being short while being stable and *A* collapsing.

$$F = (S \cap E^c_C) \cup E_C \quad (5.1)$$

We further divide the event E_C into two sub-failure-modes: *A* collapses while knocking down B and A collapses while missing B. Thus, there are three failure modes at the needed height: (1) E_1 — A is stable but too short; (2) E_2 — A collapses but misses tower B; (3) E_3 — A collapses and knocks B down. As we mentioned earlier, we assume that if tower A collapses it will knock down tower B if and only if towers A and B overlap. Let E_O be the latter event and E^c_O be its complementary event, then:

$$E_C = E_C \cap (E^c_O \cup E_O) = (E_C \cap E^c_O) \cup (E_C \cap E_O) \quad (5.2)$$

Therefore, Equation (5.1) is extended to:

$$F = (S \cap E^c_C) \cup (E_C \cap E^c_O) \cup (E_C \cap E_O) \quad (5.3)$$

For each failure mode, there is a corresponding monetary loss. A loss function measuring the degree to which a designer considers each penalty is associated to each loss considerable is used. The objective of the probabilistic method is to minimize expected value of the loss function. The possibility-based method is to minimize the possibility of considerable loss due to each failure mode. We need to develop a membership for each loss measuring the degree to which a given amount belongs to the set "considerable loss". In order to do this, we first transform monetary losses to normalized loss factors. Let ML1, ML2, ML3 be the monetary losses for the three failure modes, ML3 > ML1 > ML2 as we have stated before. The corresponding loss factors will be: γ_1 = ML1/ML3, γ_2 = ML2/ML3, γ_3 = 1.0. To ensure that the designer who uses probability and the designer who uses possibility have the same preferences, we assume that the membership of each monetary loss is equal to the corresponding loss factor. Details are given in Section 5.4.

5.2 Probabilistic and possibilistic models of the uncertainties

In this section we will explain how we collect data and construct probabilistic and possibility-based models for the uncertain variables.

5.2.1 Probabilistic models

In section 5.1, we explained that there are four uncertain variables: the height at which a tower will collapse, the x-coordinates of towers A and B (Fig. 5.1) and the minimum acceptable height. Data for constructing the probability distribution of the collapse height of tower A is obtained from domino stacking experiments, while data for the locations of the towers are generated from the true probability distribution of the locations of the towers using a random number generator. The true probability distribution of the minimum allowable height is known.

First, we explain how to collect data for the collapse height. We mix domino blocks thoroughly before we randomly pick a block and place it dotted face up. Then, we randomly pick a block again and stack it on top of the previous domino block, with the same face orientation. We repeat this procedure until the tower collapses. By repeating the stacking for a number of times, we can plot a histogram for the number of stacks when a tower topples. Here we assume that the number of blocks represents the height of a tower since the height of each block is approximately the same. We find that both the normal and gamma probability distributions fitted well to the histogram of the collapse height. In this dissertation, we use the normal distribution for convenience.

The x-coordinates of towers *A* and *B* are uniformly distributed. However, we do not know the mean values and standard deviations of these locations. We estimate these parameters from sample values obtained from the true probability distributions of these variables using a random number generator by treating these values as measurements. Standard statistics is used to estimate distribution parameters from the sample values. To account for the lack of data, the standard deviation of the mean and the standard deviation of the standard deviation are also calculated. The range of the x-coordinates of the towers is inflated to account for the statistical errors in the mean and standard deviations. The procedure for estimating the distribution parameters is described in detail in Section 3.3.2.

The minimum allowable height follows a uniform distribution with known distribution parameters.

5.2.2 Possibilistic models of the uncertain variables

In order to ensure that the probabilistic and possibility-based designs are obtained using the same amount of information, all possibility distributions for the uncertain variables are developed using the least conservative principle detailed in Section 3.4.2.

5.3 Probabilistic analysis

In Equation (5.3), the three failure modes are mutually exclusive events. Therefore, the probability of failure, P(F), is expressed as:

$$P(F) = P(S \cap E^c_C) + P(E_C \cap E^c_O) + P(E_C \cap E_O) \tag{5.4}$$

Equation (5.4) implies that:

$$P(F) = P(S|E^c_C)\, P(E^c_C) + P(E_C|E^c_O)\, P(E^c_O) + P(E_C|E_O)\, P(E_O) \tag{5.5}$$

Since E_C and E^c_O, E_C and E_O are statistically independent pairs of events, respectively, $P(E_C|E^c_O) = P(E_C)$ and $P(E_C|E_O) = P(E_C)$.

$P(E_C)$, which is the probability that a tower will collapse, is denoted as p_1. This probability is equal to the probability that the collapse height is less than or equal to the target height. The probability density function of the collapse height can be approximated from the histogram of measured heights of dominoes when they collapse. Figure 5.2 shows a histogram of collapse height of 100 towers. A normal distribution, whose mean is equal to the sample mean and the standard deviation is equal to the sample standard deviation, is used to model the collapse height. The sample mean and the sample standard deviation are 31.5 and 3.6, respectively. This distribution is denoted as N(31.5, 3.6).

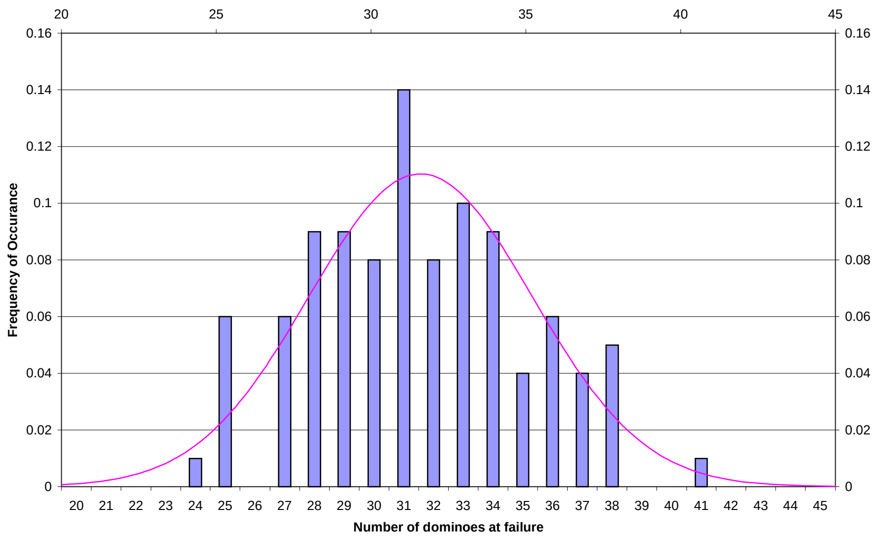

Figure 5.2 Histogram and fitted normal distribution of collapse height (from the 1st 100 repetitions of building)

$P(S|E^c_C)$ is the probability that, given that we have been able to build a stable tower with height equal to the target height, $H=h$, this tower fails to meet the requirement about the minimum acceptable height, h_0. Since the tower is stable, its height is equal to the target height. Therefore, $P(S|E^c_C)$ is equal to the probability that the target height, $H = h$, is less than the minimum acceptable height h_0 and is denoted by p_2. We assume that h_0 varies uniformly within lower and upper limits h_1 and h_2 (Fig. 5.3). Then, p_2 is:

$$p_2 = \begin{cases} 1 & h \le h_1 \\ \dfrac{h_2 - h}{h_2 - h_1} & h_1 < h \le h_2 \\ 0 & h > h_2 \end{cases} \tag{5.6}$$

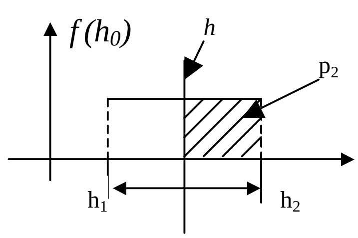

Figure 5.3 Calculation of *p₂*

$P(E_O)$, which is the probability that the two towers overlap in the *x* direction, is denoted by p_3. It is calculated by Monte Carlo simulation or methods explained in Appendix A.1.

Since each failure mode is associated with a different loss factor, the expected loss due to failure is expressed as:

$$PLS(h) = p_2 (1 - p_1)\, \gamma_1 + p_1 (1 - p_3)\, \gamma_2 + p_1\, p_3\, \gamma_3 \tag{5.7}$$

The objective of the probabilistic designer is to minimize the expected loss in Equation (5.7).

5.4 Possibility-based analysis

In possibility-based design, we characterize the uncertainty in the outcome of the process of building tower A with a given target height using possibility. We use a membership function for each outcome of the building process to express the degree to which the loss is judged to be considerable. The membership of loss for each failure mode represents the severity of that mode. E_3 , is the most disastrous outcome and it is assigned a membership $\mu_3 = 1$. As explained in Section 5.1, failure to meet the requirement for height is less severe and has a membership μ_1. E_2 causes the minimal loss because the insurance will cover most of the damage. The membership of loss for E_2 is μ_2.

In the probabilistic formulation, there are loss factors γ_1, γ_2 and γ_3 for each failure outcome and the objective is to minimize the average loss. To make the possibility-based problem equivalent to the probabilistic counterpart, we assume that $\mu_1 = \gamma_1$, $\mu_2 = \gamma_2$ and $\mu_3 = \gamma_3$.

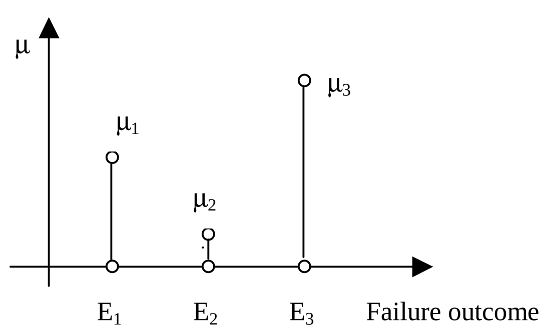

Figure 5.4 Membership function of loss for the three failure modes

Let π_1 be the possibility that a tower with height equal to the target height, *h*, will collapse, and π_1^C be the possibility that the tower will be stable. In this problem, the collapse height H_c satisfies a normal distribution $N(\mu_{Hc}, \sigma_{Hc})$. Figure 5.5 plots a normal probability density function and its corresponding least conservative possibility distribution. Applying the least conservative principle, the possibility of the collapse height $Hc = h$ is expressed as:

$$\Pi_1(h) = 1 - p_1(h) + p_1(h') \tag{5.8}$$

where *h*' is the point with the same possibility value with *h* at the other side of the apex. Therefore,

$$\Pi_1(h) = 1 + p_1(2\mu_{Hc} - h) - p_1(h) = 2(1 - \Phi(\frac{h - \mu_{Hc}}{\sigma_{Hc}})) \tag{5.9}$$

$\Pi_1^C(h)$ has the same support and mode value as $\Pi_1(h)$.

The possibility that a tower with height equal to the target height will collapse is equal to the possibility that the collapse height *Hc* will be less or equal to the target height *h*:

$$\pi_1 = \begin{cases} 2\Phi(\frac{h - \mu_{H_C}}{\sigma_{Hc}}) & \text{for } 0 \le h < \mu_{Hc} \\ 1 & \text{for } \mu_{Hc} \le h \le \infty \end{cases} \tag{5.10a}$$

The possibility of the tower not collapsing is:

$$\pi_1^C = \begin{cases} 2(1 - \Phi(\frac{h - \mu_{H_C}}{\sigma_{Hc}})) & \text{for } \mu_{Hc} \le h \le \infty \\ 1 & \text{for } 0 \le h < \mu_{Hc} \end{cases} \tag{5.10b}$$

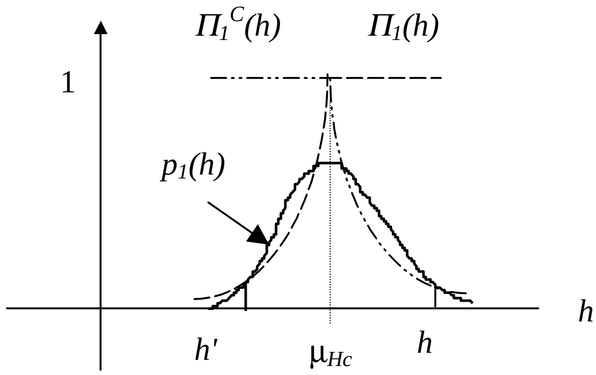

Figure 5.5 Least conservative possibility distributions that are consistent with a normal probability density function *$p_1(h)$* Dashed line: *$\Pi_1(h)$*; Double-Dotted line: *$\Pi_1^C(h)$*

Let π_2 represent the possibility of the design height being too short. π_2^C represents the possibility of the design height being acceptable. The possibilistic designer transforms the probability distribution of the minimum acceptable height (Equation 5.6) into a possibility distribution using the least conservative principle. Therefore, the possibility distribution of the minimum acceptable height is:

$$\Pi_2(h)=p_2(h)+1-p_2(h') \tag{5.11}$$

where h' is the point with the same possibility value with h at the other side of the apex of possibility distribution. Here,

$$\Pi_2(h)=2(1-\frac{h-h_1}{h_2-h_1}), \text{ for } \frac{h_1+h_2}{2} \le h \le h_2 \tag{5.12}$$

This is a symmetric triangular possibility distribution with the symmetric axis at $(h_1+h_2)/2$ over the interval $[h_1, h_2]$. $\Pi_2^C(h)$ has the same support and mode value as $\Pi_2(h)$. Figure 5.6 plots the probability density function and its corresponding least conservative possibility distribution. Therefore:

$$\pi_2 = \begin{cases} 1 & h_1 \le h < \frac{h_1 + h_2}{2} \\ 2(1 - \frac{h - h_1}{h_2 - h_1}) & \frac{h_1 + h_2}{2} \le h \le h_2 \end{cases} \quad (5.13a)$$

$$\pi_2{}^C = \begin{cases} 1 & \frac{h_1 + h_2}{2} \le h \le h_2 \\ 2\frac{h - h_1}{h_2 - h_1} & h_1 \le h < \frac{h_1 + h_2}{2} \end{cases} \quad (5.13b)$$

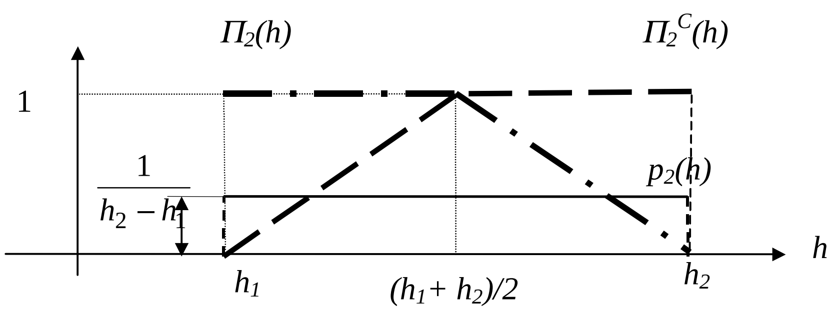

Figure 5.6 Least conservative possibility distributions that are consistent with a uniform probability density function $p_2(h)$
Dotted line: $\Pi_2(h)$; Dashed-dotted line: $\Pi_2{}^C(h)$

π_3 is the possibility that the towers overlap, that is the possibility that the distance of the centers of two towers is smaller than *D*. $\pi_3{}^C$ is the possibility that the distance of the centers of two towers is larger or equal to *D*. Figure 5.7 demonstrates the relative locations of towers *A* and *B*. As we mentioned, the centers of the two towers, x_A and x_B, are estimated from measurements and are assumed to have a uniform probability distribution between [x_0', x_1'] and [x_0 , x_1], respectively. The corresponding possibility distributions, which satisfy the least conservative principle, have a triangular shape with an apex of one at the middle of the intervals in which x_A and x_B vary. Let x_m' and x_m be the apexes for the locations of towers *A* and *B*, respectively.

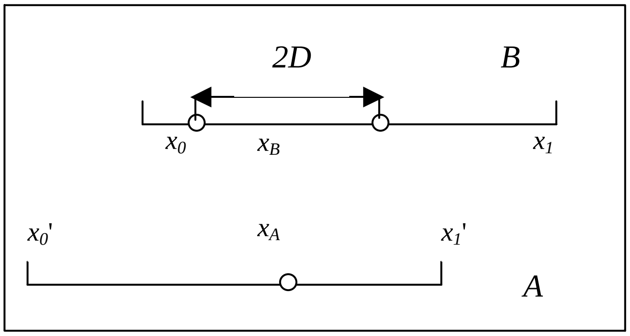

Figure 5.7 x-coordinates of Tower A and Tower B

Using possibility calculus we find that the possibility that the towers overlap is:

$$\pi_3 = \sup_{x_0' \le x_A \le x_1'} \{\min[\pi(x_A), \sup_{x_A - D \le x_B \le x_A + D} \pi(x_B)]\} \tag{5.14}$$

The possibility that the towers do not overlap is:

$$\pi_3^C = \sup_{x_0' \le x_A \le x_1'} \{\min[\pi(x_A), \sup_{x_B < x_A - D \text{ or } x_A + D < x_B} \pi(x_B)]\} \tag{5.15}$$

The detailed calculations are attached in Appendix A.2.

The possibility of occurrence of failure modes E_1, E_2, E_3 are:

$$\pi_{E1} = \min(\pi_1, \pi_2) \tag{5.16a}$$

$$\pi_{E2} = \min(\pi_1, \pi_3^C) \tag{5.16b}$$

$$\pi_{E3} = \min(\pi_1, \pi_3) \tag{5.16c}$$

According to the principle of fuzzy set events, the possibility of a considerable loss at a given height $H = h$ is calculated as:

$$\Pi(h) = \max[\min(\mu_i, \pi_{Ei})]_{i=1,\ldots,3} \tag{5.17}$$

The objective of the possibility-based approach is to find an optimum height, which minimizes the possibility of considerable loss, $\Pi(h)$.

5.5. Scope and method for comparison

First, we describe the procedure for analytical comparison of the methods in subsection 5.5.1. Subsection 5.5.2 explains the procedure for experimental comparison of the methods. The goal of the comparison is to demonstrate both analytically and experimentally that, in problems involving systems with narrow failure zones, probabilistic methods may yield poor designs when little data is available about uncertainties. For this purpose, we select the parameters of the design problem so as to maximize the sensitivity of the probabilistic optimum design to statistical errors. Subsection 5.5.3 explains how the parameters of the design problem are selected.

5.5.1 Procedure for analytical comparison

In the analytical comparison, the optimum probabilistic and possibility-based designs are calculated using the same data about uncertainties. To be able to compare designs we need a measure of goodness of a design. To select such a measure we assume that in a problem involving only randomness (irreducible uncertainty), the best design is the one with the largest expected utility (or equivalently smallest expected loss). Therefore, to test which design is better the true average loss of each optimum is calculated using the true probability distribution of the uncertain variables. The design that has the lowest expected loss is the best. In the analytical model, two assumptions are essential: (1) The probability distribution of the collapse height of tower A is normal; and (2) Tower A will knock down tower B if and only if the x-coordinates of the two towers overlap.

As we mentioned in Section 5.3, the designers know that the probability distributions of the x-coordinates of the towers are uniformly distributed and they estimate the mean and the standard deviations from sample values. They obtain conservative estimates of these parameters using the standard statistical method described in Sec.3.3.2. The possibilistic designer transforms the probability distributions of the uncertain variables into possibility distributions using the least conservative principle in Sec.3.4.2. Then, the expected losses of the probabilistic and the possibility-based designs are computed and compared based on the true probability distributions of the x-coordinates of the towers. The expected loss of a design is computed using Monte-Carlo simulation with 10,000 replications.

5.5.2 Procedure for experimental comparison

In the analytical comparison, we compare the probabilistic and possibilistic designs after removing a portion of the reducible uncertainty involved in this problem (the statistical errors in the mean and standard deviation of the x-coordinates of the two towers). However, we cannot remove all reducible uncertainty because the analytical models involve many assumptions. The objective of the experimental comparison is to compare the probabilistic and possibilistic designs after removing almost all the reducible uncertainties. However, because a finite number of pairs of probabilistic and possibilistic designs are compared, still, some statistical uncertainties in the estimates of the expected losses of the two designs remain, which cannot be removed unless a very large number of experiments are performed.

The following hybrid experimental-analytical procedure allows us to compare designs much more efficiently than a purely experimental comparison, which would compare a very large number of pairs of probabilistic and possibility-based designs. In each experiment, we build two towers, called towers A and B. We keep stacking blocks for tower A until it collapses. Tower B has a fixed short height so that it will not topple by itself. In each experiment, we build these towers on the locations obtained from the true probability distributions of the x-coordinates of the two towers using a random number generator. We record the height at which tower A collapses and record if it destroys B when it falls. After repeating this procedure for one hundred times, we compute the frequencies of each failure mode and the true expected loss for every target height. This allows us to construct a look-up table for the true expected loss of any design whose target height is given. Using this table, we can compare the probabilistic and possibility-based designs determined in the analytical model in terms of the their true expected losses.

5.5.3 Selection of the parameters of the design problem

As mentioned in Section 5.5.1, we need to select the values of the problem parameters to maximize the sensitivity of the probabilistic optimum design to statistical errors. This process of finding a combination of parameters that maximizes the sensitivity of a method to modeling or statistical errors is called *antioptimization*. The only fixed design parameter is the length of a domino block, *D*, which is 2" (51 mm). Besides loss factors γ_1~ γ_3 and the corresponding memberships of the penalties for the three failure modes, μ_1~μ_3, we have to select the values of the following parameters:

- *NH1*, *NH2*: the lower and upper limits of the minimum acceptable height;
- *xAB*: the difference in the mean values of the x-coordinates of the two towers;
- σ_x/D: the normalized standard deviations of the x-coordinates of the two towers. For simplicity, we assume that two towers have the same standard deviation.

These last two parameters control the probability distribution of the x-coordinates of the two towers, from which we compute the probabilities of failure modes E_2 and E_3.

First we select the loss factors. The 3rd failure mode, where both towers A and B collapse, is the most severe. We will normalize its loss factor so that γ_3 =1.0. There is a narrow failure zone for the 3rd failure mode (tower A collapses and knocks down tower B), which makes the problem difficult to solve using probabilistic methods. To increase the importance of this failure mode relative to the second mode, we will assign a small loss factor to this mode. This corresponds to the real life scenario where the insurance company covers most of this failure damage. We assume that γ_2 =0.1. The loss factor for the 1st failure mode will be determined in order to maximize the difference between probabilistic and possibility-based designs.

We found that an average builder can build a stable tower with at least 20 blocks. Therefore, we assume that *NH1*=20. We also assume that the target height should be at least 20.

Next, we select the remaining design parameters in the problem. Since the possibility-based method is insensitive to sample to sample variations (a conclusion derived from the tuned-damper design problem), we will select the combination of values of the design parameters that maximizes the sample to sample variations of the optimum probabilistic design. There is sample to sample variation because different sample data will give a different estimate of the probability distribution of the x-coordinates of the towers compared to the true distribution. The larger the difference between the estimated distribution and the true distribution is, the larger the variation in the optimum height will be.

Our strategy is to select values for the problem parameters that produce a significant variation between the optimum heights of the towers obtained using the estimated and the true distributions of the x-coordinate of Tower A.

Next, we will explain how to measure the error in the estimated standard deviation of the x-coordinate of a tower from sample data with a given size *N* and the true standard deviation. Assume that we know the true type of the probability distribution for the tower locations. We generate *N* sample data from the true distribution and calculate the estimated standard deviation. Let us denote the true standard deviation as σ_t and the estimated as σ_e. After we repeat this procedure 10 times, we calculate the standard error in the standard deviation, which is defined as:

$$\delta = \frac{|\sigma_e - \sigma_t|}{\sigma_t} \tag{5.18}$$

Figure 5.8 illustrates the sensitivity in the estimated standard deviations when the sample size varies. To construct this figure, we generated ten samples of the same size of the x-coordinate of Tower A and estimated the standard deviation from each sample. Then we computed the standard error of each sample. Figure 5.8 shows the maximum and the mean value of the standard error as a function of the sample size. They are denoted as δ_{max} and δ_{mean} respectively in Fig. 5.8. We vary the sample size *N* from three to 50. These two curves represent roughly the variations a group of sample data can produce to the standard deviation of a distribution. The smaller the sample size is the larger the standard error and its variation will be. We use small sample sizes to study variations. Any value between the two curves in Fig.5.8 is a valid estimate of the standard deviation. The figure indicates that for a sample size of three the estimated standard deviation can be as high as $2.5\sigma_t$ if the designer uses an inflation factor of two. We will find out the parameter combination which maximizes the difference between the optimum heights of the towers when these optimum heights are calculated using the true value of the standard deviation, σ_t, and $\sigma_e=2.5\sigma_t$.

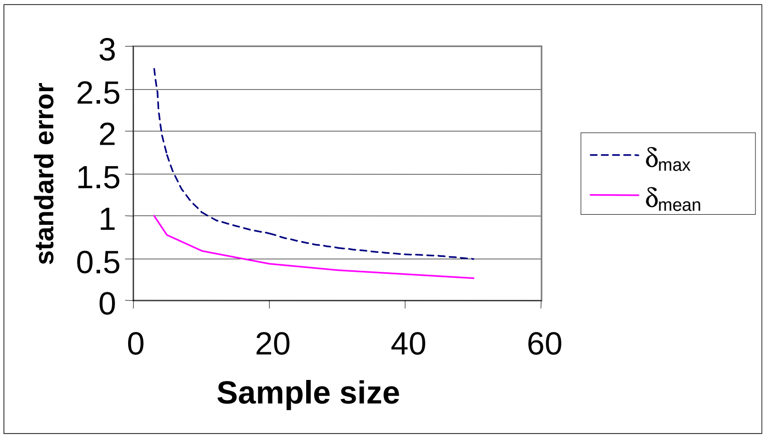

Figure 5.8 Standard error of the estimated standard deviation as a function of sample size

The remaining parameters to be determined are NH2, γ_1 , xAB/D and σ_x/D. Here we use the same distribution as Fig 5.2 for the collapse heights, which is based on the experiments performed with the 1st 100 repetitions of building. It is a normal distribution N(31.5, 3.6). Figure 5.9 shows the optimum height variation, which is the largest sample-to-sample variation in optimum heights of 10 designs, as a function of NH2 and γ_1, when NH2 varies from 30 to 60 and γ_1 varies from 0.1 to 0.5. We notice that the larger the NH2 and γ_1, the bigger the variation in the optimum height is. γ_1 is the loss factor for the failure mode "the tower is stable but too short". In real life, it should be a relative small portion of γ_3, which is the loss due to collapse of both buildings. Therefore, we select NH2=60 and γ_1 = 0.3 for the domino problem. Next, we investigate the impact of the other two location parameters xAB/D and σ_x/D on the optimum height variation. Table 5.1 lists the maximum optimum height variation and the corresponding σ_x/D for a given xAB/D. From this table we observe that the smaller the distance xAB/D, the larger the sensitivity of the optimum heights to errors. Therefore, we should select a small distance and a large standard deviation for the distance. However, for convenience we want to perform the experiments in a small area. Balancing these two concerns, we select location parameters as xAB/D = 1.0 and σ_x/D = 1.25. Other parameters are: NH1=20, NH2=60 and γ_1 = 0.3. In later sections, we will find that this combination creates significant differences in probabilistic designs due to sample to sample to sample variations.

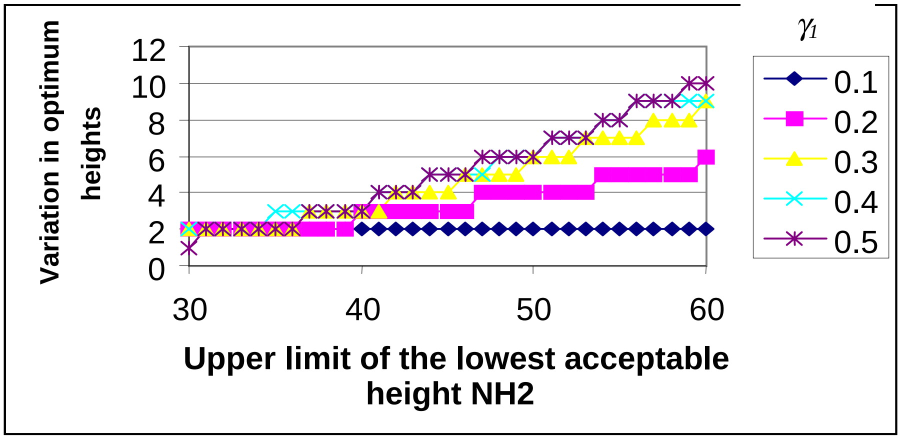

Figure 5.9 Optimum height variations vs. NH2 and γ_1

Table 5.1 Maximum variations in optimum height for different location parameters

xAB/D	0	0.25	0.5	0.75	1	1.25	1.5	1.75
σ_x/D	2.25	2.25	2.5	2.5	2.75	2	2.25	2.5
Max variation	9	9	9	9	9	8	8	8

5.6 Results

In this section, we perform analytical and experimental comparisons according to the methods presented in the previous section. After presenting comparison results, we summarize the observations and draw conclusions.

5.6.1 Analytical comparison

The parameters of the design problem in the analytical comparisons depend on the distribution parameters of the collapse height, which are derived from the histogram of the collapse height. Therefore, first have to perform experiments on one tower to collect sample values of the collapse height, and then we select the problem parameters and perform experiments involving two towers. When we perform the experiments with the two towers we collect data about the collapse heights of the towers and estimate the probability distribution of the tower at which a tower topples again. In the project presented in this dissertation, an undergraduate student, Jonathan Abbott, collected the 1st 100 data about the collapse height of a single tower first. Then, the author estimated the probability distribution of the collapse height and performed antioptimization to determine the problem parameters. She also conducted the 2nd 100 repetitions of experiments with the two towers. Here, the design problem parameters were determined from the normal distribution of collapse height, N(31.5, 3.6), as mentioned in the previous section. The only exception is the parameter σ_x/D. In the experiments, we assumed that the probability distributions of the locations of the two towers to be uniform in the ranges [-2", 6"] and [0, 8"], respectively. Thus, $\sigma_x/D = 8''/(2\sqrt{3}D)$=1.155 instead of 1.25. Figure 5.10 is the histogram and the fitted normal distribution of 100 experiments on two towers conducted by the author. The fitted normal distribution is N(34.47, 4.18). The mean and standard deviation is calculated from the sample mean and sample variance respectively. We performed a χ^2 test with 10 intervals with a significance level of 0.05. The degrees of freedom of χ^2 is 10–2–1=7. We found that the chi-square statistic was: χ^2 = 4.6. This value is smaller than the test statistics $\chi^2_{0.05}(7)$ = 14.067. As a result, we can not reject the null hypothesis that these 100 data follow a normal distribution.

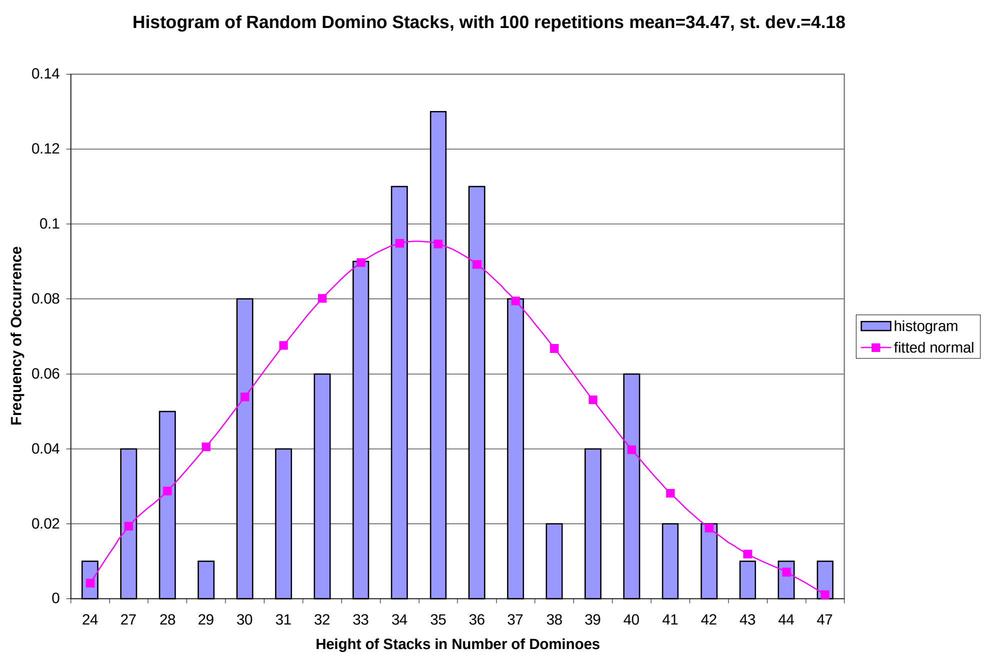

Figure 5.10 Histogram of collapse height and the fitted normal distribution (from the 2nd 100 experiments)

First we obtained six sets of 20 optimum designs using probabilistic optimization. For each set we used a different sample size of data to estimate the mean and standard deviation of the x-coordinates of the towers, ranging from 3 to 1000. Table 5.2 shows the analytical optimum heights of the six sets of 20 probabilistic designs. Using the true probability distributions of the x-coordinates of the towers, instead of those estimated from data, we found that the optimum height from the probabilistic design is 28.

Table 5.2 Probabilistic optimum heights for estimated distributions of the x-coordinates of the towers using 3~1000 sample values

Sample size	**Design No.**																			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
3	32	31	27	35	32	31	31	29	35	33	34	29	31	32	35	32	33	31	33	34
5	30	31	29	30	32	30	30	31	32	33	29	32	32	31	32	31	30	31	32	31
10	30	30	30	30	31	31	30	30	30	30	31	30	30	30	30	30	31	31	30	29
20	29	30	30	30	29	30	29	30	30	29	30	30	30	29	30	29	29	29	29	30
100	29	29	29	29	28	29	29	28	28	28	28	29	29	29	28	29	29	29	28	29
1000	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28

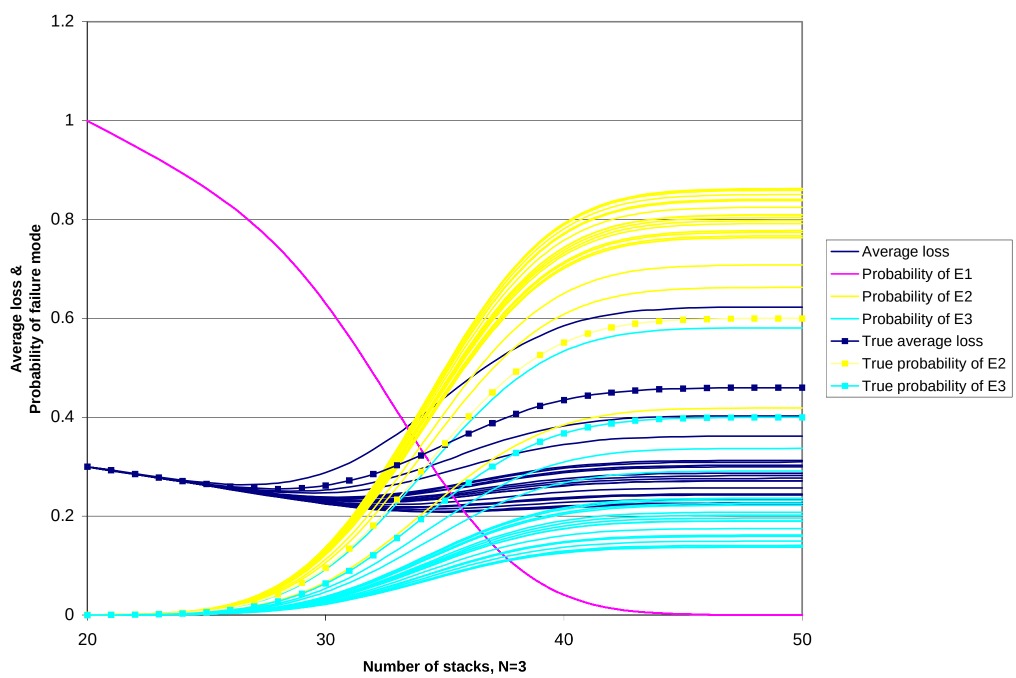

Figure 5.11 Composition of probability of loss of each failure mode, N=3

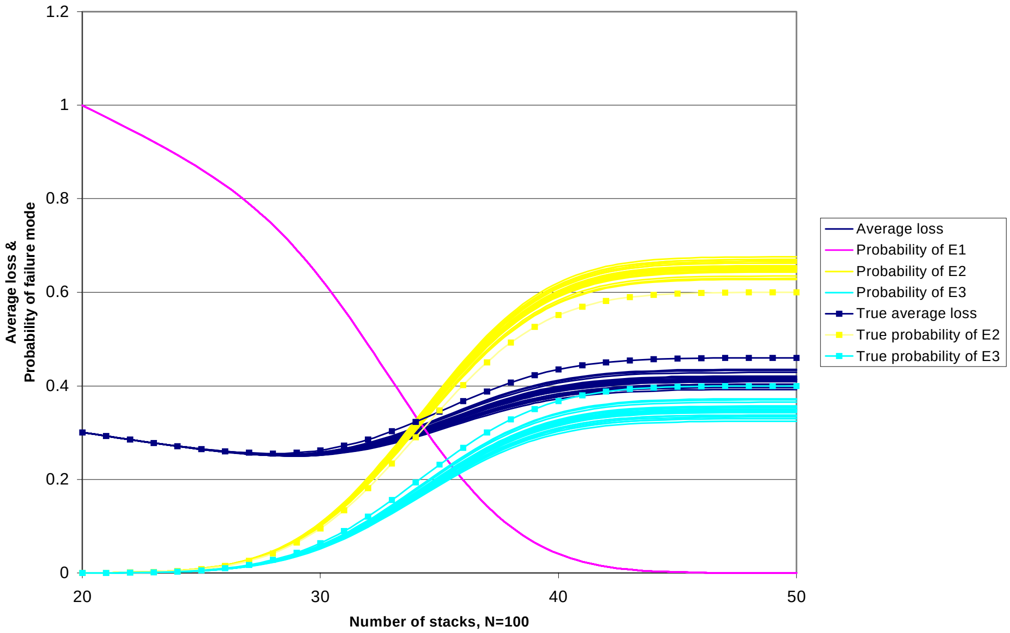

Figure 5.12 Composition of probability of loss of each failure mode, N=100

Figures 5.11 and 5.12 show the expected losses, and the probabilities of failure for each failure mode as a function of the target height for 20 designs in Table 5.2 when the sample size is three and 100, respectively. When the sample size is three, the optimum heights of all designs except No.3 are higher than the true optimum height. This is because when designing these towers we overestimate the variability in their locations. This happens because we apply an inflation factor to the sample standard deviation of the x-coordinates of the towers to account for the statistical error. This strategy underestimates the probability of the 3rd failure mode, because when the variability in the locations of towers A and B is assumed large, the towers are unlikely to overlap. Therefore, when designing these towers the optimizer incorrectly increases the height to reduce the chance that a tower is too short.

When 1000 measurements are available, the statistical error in the mean and the variance is negligible so the estimated distribution of the location of a tower approaches the true distribution even when using inflation factors. Therefore, the estimated probability of the 3rd failure mode and the expected loss are close to their respective true values. In this case, all 20 probabilistic designs are identical with the true optimum designs.

Figure 5.13 presents the possibility of considerable loss of 20 designs, and a breakdown of this possibility according to equation 5.17, when we use sample sizes of three for the locations of the towers. The possibility of considerable loss at each target height is the maximum of the possibilities of considerable loss for each of the three failure modes. The possibility of considerable loss due to failure to meet the minimum height requirement is the minimum of the possibility of the tower being too short and the membership function of the penalty to the set "considerable loss" (equation 5.17). This failure mode dominates for heights ranging from 20 to 30 and is constant. The possibility of loss due to collapse of tower A only is the minimum of the possibility of this event and the membership of the penalty to the set "considerable loss". Since the latter is minimal, this failure mode is less significant than the other two modes. The possibility of loss due to both tower A and B falling is the minimum of the possibility of this event and the loss membership of both tower A and B falling. The failure mode "both towers collapse" dominates for tower height greater than 31. The possibility of considerable loss for this mode is equal to the possibility of both towers collapsing because the membership of the penalty is always one. The possibility of loss due to this failure mode is also insensitive to the height when the height exceeds 35, because the possibility of tower A knocking down tower B is determined by the possibility of the two towers overlapping. It is observed that all towers with heights between 20 and 30 have the same minimum possibility of considerable loss.

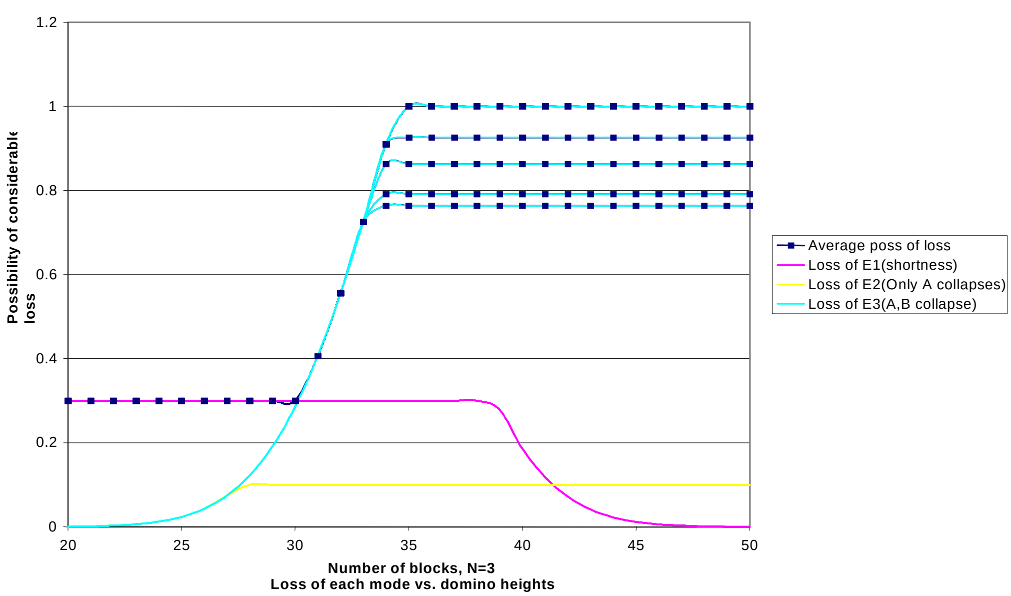

Figure 5.13 Breakdown of the possibility of considerable loss, N=3

The possibility-based counterparts of the probabilistic designs have optimum heights in a range of 20~30 regardless of the sample size. This means that even if each design is developed from different sample points, the possibility-based method will always provide a range instead of a specific number for the choice of the optimum height. Therefore, the possibility-based designer has to use other criteria to determine the optimum. The designer has two reasonable options after studying the composition of failure modes within the range from 20 to 30 as shown in Fig.5.13. Since the failure mode due to shortness dominates in this range, the possibility of loss is the minimum of the possibility of a tower failing to meet the minimum height requirement and the membership function of the penalty to the set "considerable loss". When the height increases, the possibility of a tower being too short decreases. Based on this concern, a designer will choose 30 as the optimum. However, a designer might also choose 20 as the optimum. His/her concern is that if all heights have the same possibility of loss for the dominant failure mode, then he/she has to find an optimum to minimize possibility of loss for the other two modes. Increasing the height will increase these possibilities of loss. This designer will choose the smallest value that is 20. We will compare probabilistic designs with these two opposite possibility-based options, respectively.

Each pair of bar charts in Figures 5.14 ~ 5.16 compares the true average loss of a probabilistic and possibility-based designs obtained from the same sample data. The possibility-based designs have a height of 30 blocks. The true average loss of a design height is calculated from Equation 5.7 using the true probability distribution of the locations of the towers. Each figure contains 20 pairs of designs. Each pair is obtained using the same sample of data. Thus, we can investigate the sample to sample variations of the probabilistic and possibilistic optima and compare which method yields designs with lowest expected loss on average.

In Fig. 5.14~17, the probabilistic designs have higher true average loss than their possibility-based counterparts with a sample size of three. When the sample size increases to five, 13 probabilistic designs are worse than their possibilistic counterparts in terms of true expected loss. When there are 10 sample points, five probabilistic designs are worse than their possibilistic counterparts, and 14 are the same. However, this result is not shown. When the sample size increases to 20 as shown in Fig.5.15, on average, probabilistic designs are better than the corresponding possibility-based designs. Specifically, 11 probabilistic designs are the same as the corresponding possibility-based designs, while nine probabilistic designs are better than the corresponding possibilistic designs. When the sample size is 100, all probabilistic designs are better than the corresponding possibility-based designs. The heights of the probabilistic designs are close to the true optimum height of 28 (Table 5.2). In Fig.5.16, when 1000 sample points are available, all the probabilistic designs have the same optimum heights as the true optimum height. On the other hand, the possibility-based designs are the same as those obtained using three sample values for the locations of the towers.

Figures 5.17 ~ 5.19 compare the true average losses of probabilistic and possibility-based designs when possibility-based designs have 20 as their optimum. In Fig. 5.17, eight probabilistic designs have higher true average loss than their possibility-based counterparts with a sample size of three. Therefore, the probabilistic designs are better than the possibility-based designs on average. When sample size increases to five, 19 probabilistic designs are better in terms of true average loss than the corresponding possibilistic designs in Fig.5.18. In Fig.5.19, when 1000 sample points are available, all the probabilistic designs have the same optimum heights as the true optimum height. Probabilistic designs are better in terms of analytical true average loss.

The analytical comparisons reveal that probabilistic designs can generate very poor designs when few sample values are available (e.g. three). The worst designs with the highest true average losses had large heights (34 or 35) and were obtained using the probabilistic method. However, the verdict as to what method is better on average when the sample size is small, depends on what optimum the possibility-based method chooses. When sample size is greater than 20, probabilistic designs are better in terms of true average loss.

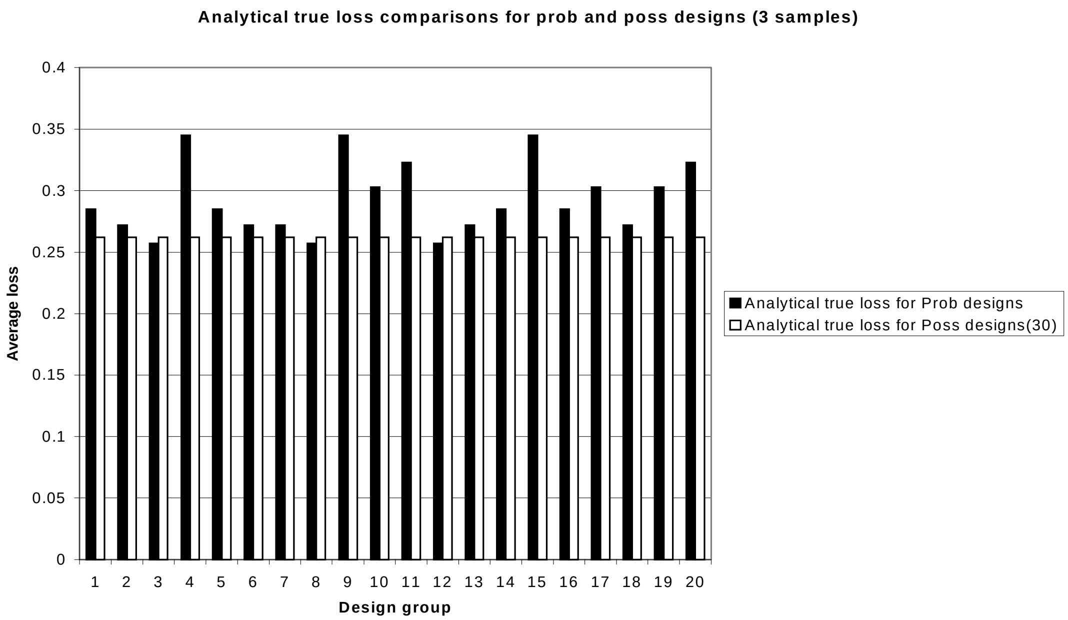

Figure 5.14 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum (sample size=3)

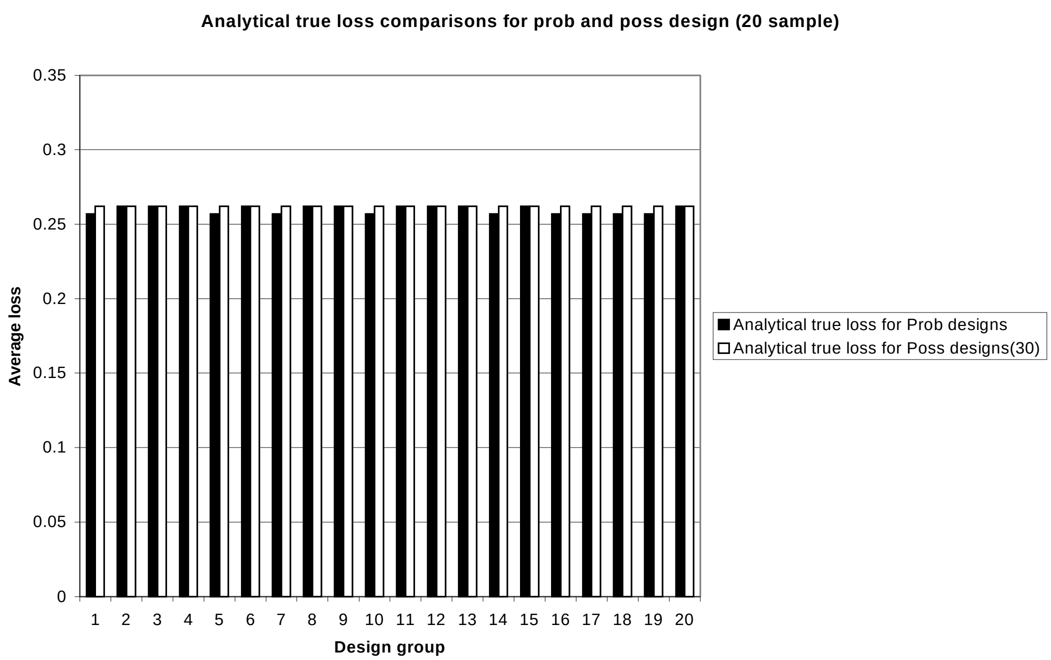

Figure 5.15 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum (sample size =20)

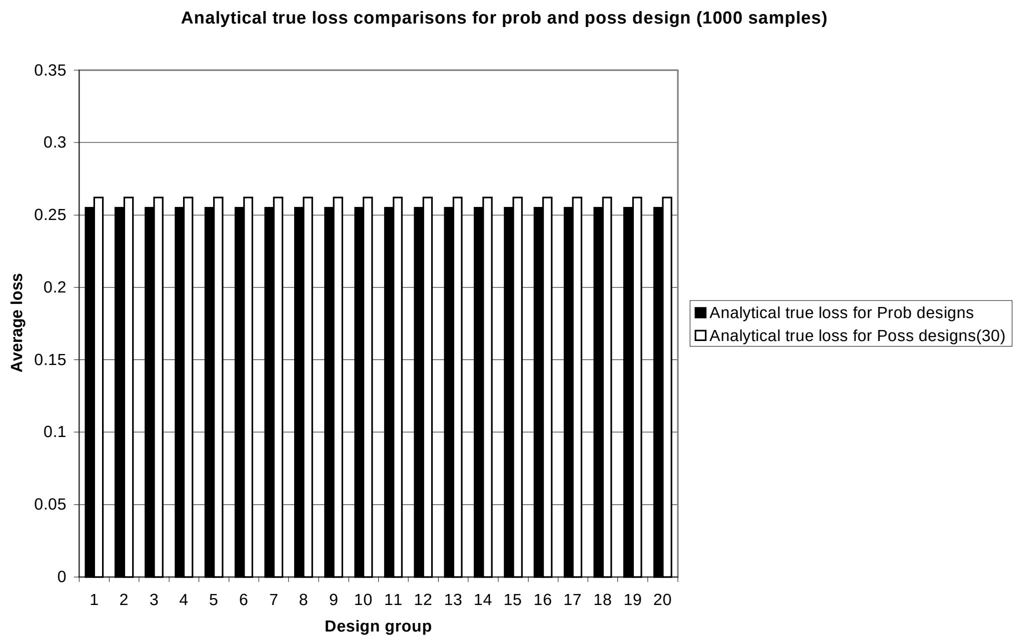

Figure 5.16 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum (sample size = 1000)

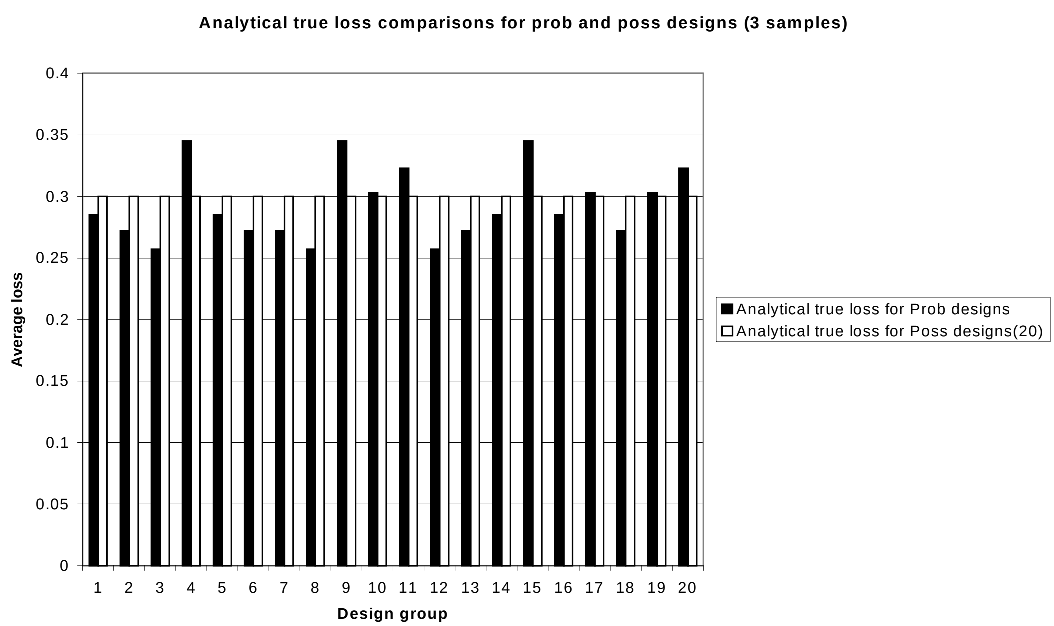

Figure 5.17 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 3)

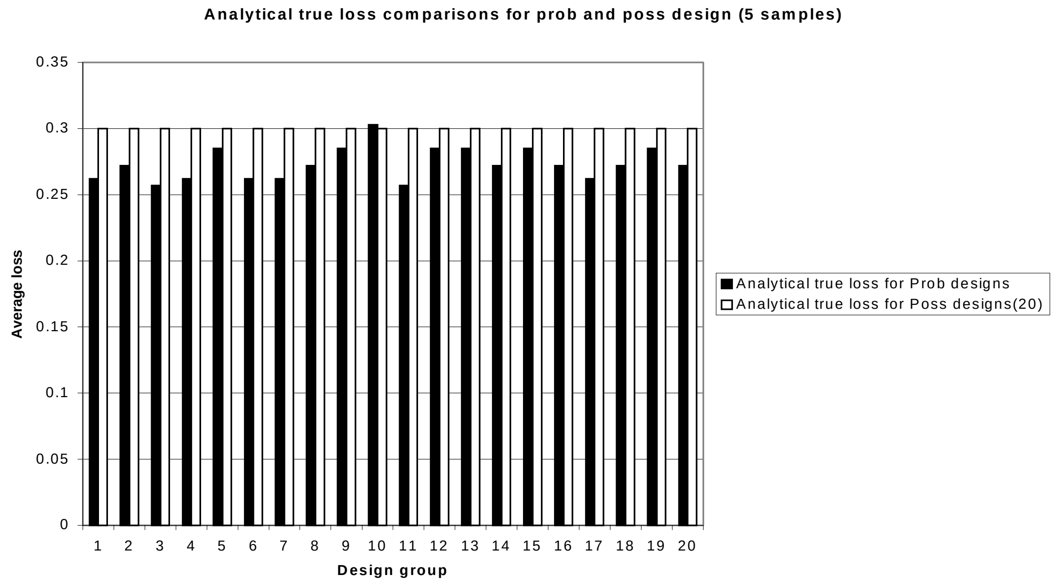

Figure 5.18 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 5)

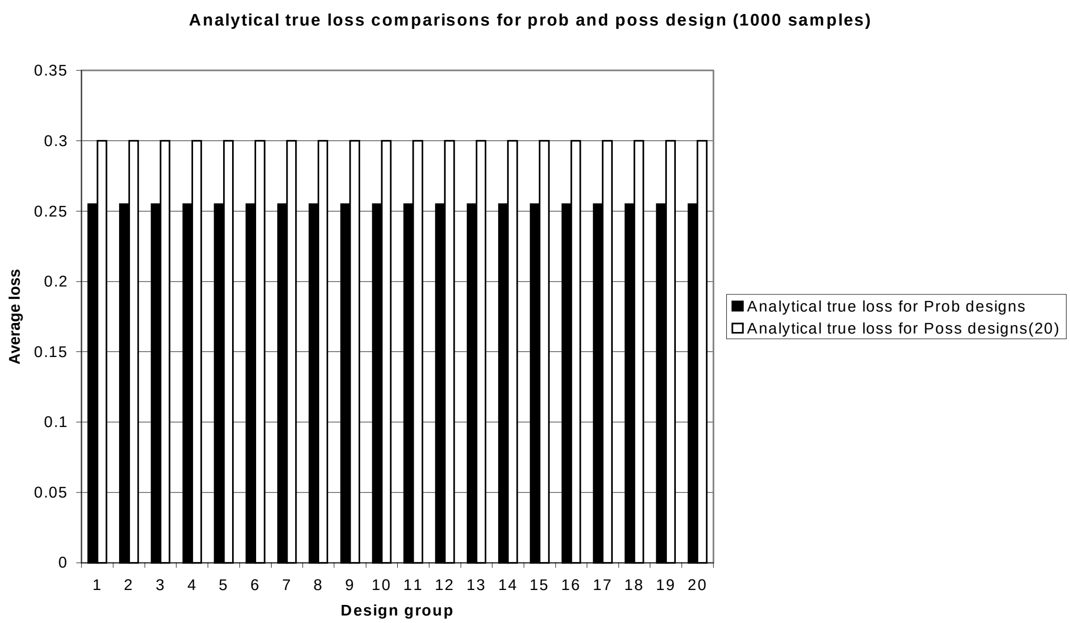

Figure 5.19 Comparison of true expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 1000)

5.6.2 Experimental comparison

As we have explained in Sec.5.5, the experimental comparison removes the major portion of the reducible uncertainties in the analytical model of this design problem. In the analytical model, the probability distribution function of the collapse height is assumed normal. Also, the probabilities of occurrence of the 2^{nd} and 3^{rd} failure modes are computed based on the assumption that the collapse of tower A will definitely hit B if and only if they overlap in x-direction. In experimental comparisons, these probabilities are estimated directly from the frequency of occurrence of these modes in the experiments.

Figure 5.20 plots the true expected loss from the analysis and the experiments. In experiments, we built 100 towers. From this figure, according to the experimental results, the optimum height is 32, whereas, according to the analytical results, the optimum height is 28. There is a significant drop in the expected loss in the neighborhood of 32 blocks. This could be due to the discrepancy between the relative frequencies of collapse of a tower with a height of 32 blocks obtained from the histogram of the collapse height and from the idealized normal distribution (Fig.5.10). Also, because only 100 experiments were performed we did not get any observations of the event "Tower A collapses and knocks down tower B" for heights greater than 37 blocks. Therefore, the experimental results for heights greater than 37 yield a lower average loss compared with analytical results.

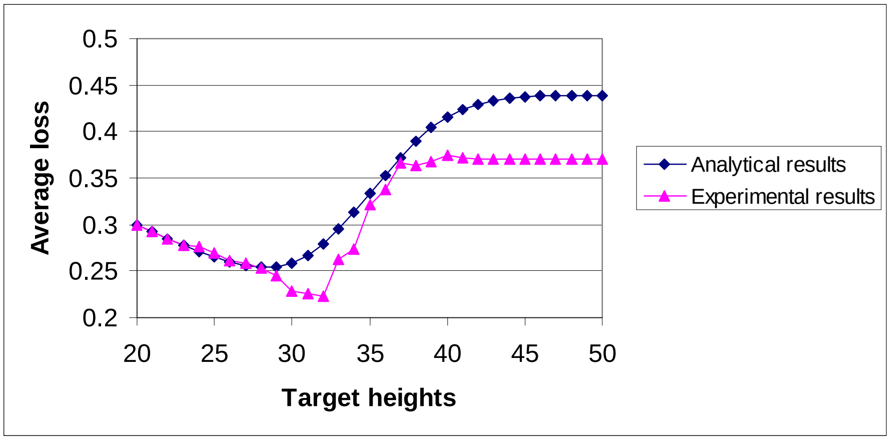

Figure 5.20 The true average losses of target heights by analytical and experimental results

Each pair of bar charts in Figures 5.21 ~ 5.23 compares the true average loss of a probabilistic and possibility-based designs obtained from the same sample data. The possibility-based designs have a height of 30 blocks. The true average loss of a design height is obtained from the experimental curve in Figure 5.20. In this way, we avoid the modeling errors in the calculations of analytical true losses. Similar to analytical comparisons in Figures 5.14~5.16, we investigate the sample to sample variations of the probabilistic and possibilistic optima and compare which method yields designs with lowest expected loss on average.

In Fig.5.21, there are 11 probabilistic designs whose true average losses exceed those of the corresponding possibility-based designs for a sample size of three. Probabilistic designs No. 4, 9 and 15 have significantly higher expected losses than the other designs in Fig. 5.21. When the sample size increases to five, the probabilistic designs with a higher true average loss than the corresponding possibility-based designs decline to three. Five probabilistic designs have the same optimum height with their possibility-based counterparts, whereas the average losses of the remaining twelve probabilistic designs are slightly lower than those of the corresponding possibilistic designs. When the sample size increases further to 10, in Fig. 5.22, most probabilistic designs are close to the possibility-based designs. When the sample size exceeds 20, probabilistic designs become worse again in terms of their expected loss than the corresponding possibilistic designs. The reason is that the optimum height of the possibilistic designs (30 blocks) is closer to the experimental optimum of 32 blocks compared to the probabilistic optimum height (28 blocks).

Figures 5.24 ~ 5.27 compare probabilistic and possibility-based designs using experimental true losses when possibility-based designer chooses the height of 20 blocks. In Fig.5.24, there are three probabilistic designs whose true average losses exceed those of possibility-based designs for a sample size of three. In Fig.5.25, when sample size is five, all probabilistic designs have lower expected losses than the corresponding possibility-based designs. When the sample size increases to 1000 in Fig.5.26, all probabilistic designs are still better than the corresponding possibility-based designs in terms of experimental true loss. In conclusion, when the possibility-based optimum is 20, the probabilistic designs are better in terms of experimental true losses regardless of the sample size.

The results of the experimental comparisons are inconclusive. The conclusions depend on what optimum the possibility-based designer chooses.

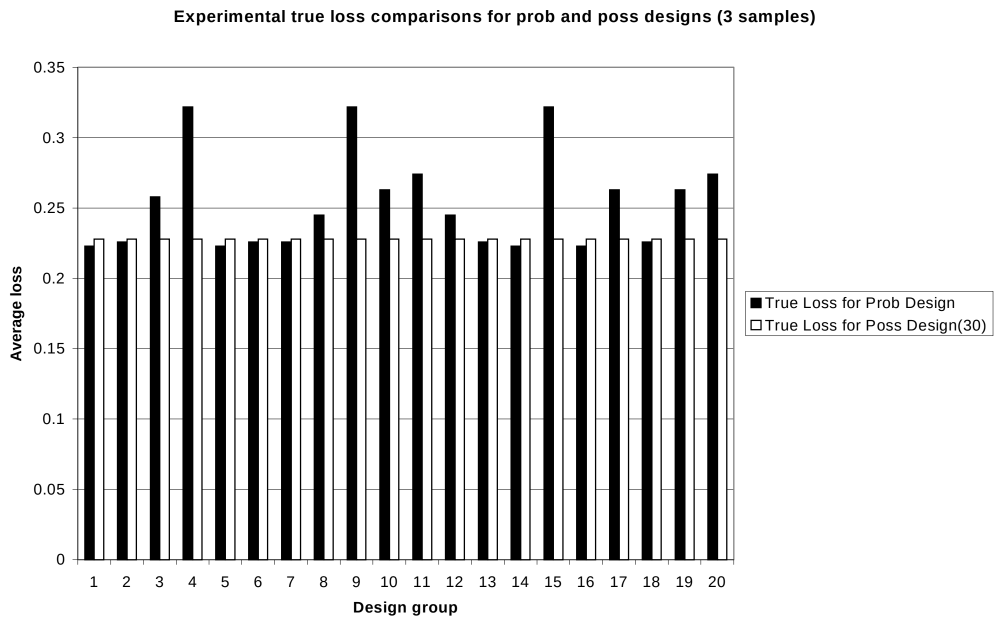

Figure 5.21 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum (sample size=3)

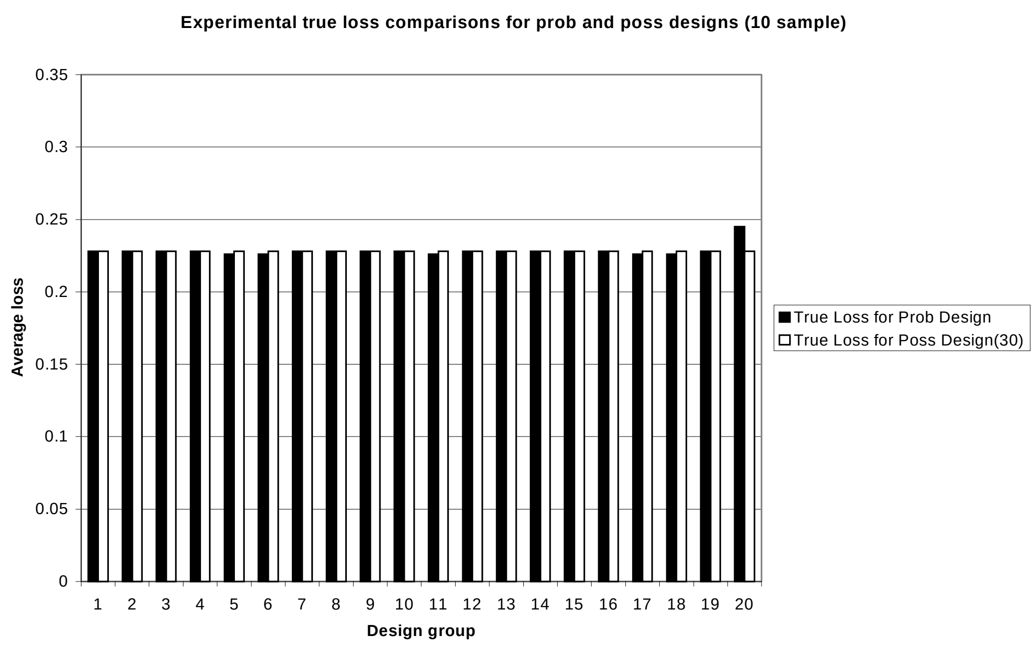

Figure 5.22 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum (sample size = 10)

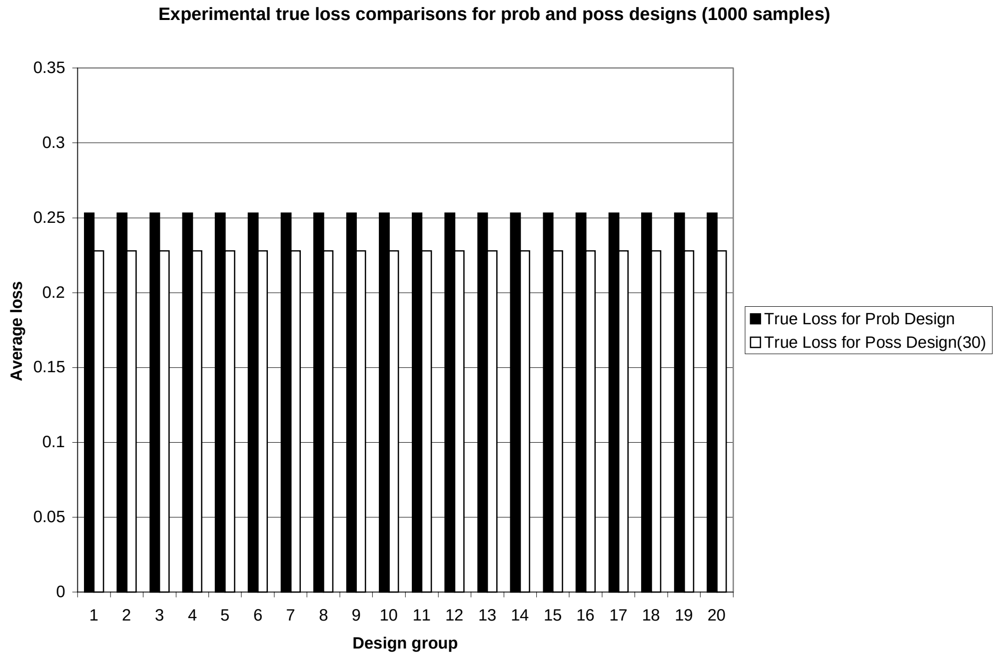

Figure 5.23 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 30 as optimum(sample size = 1000)

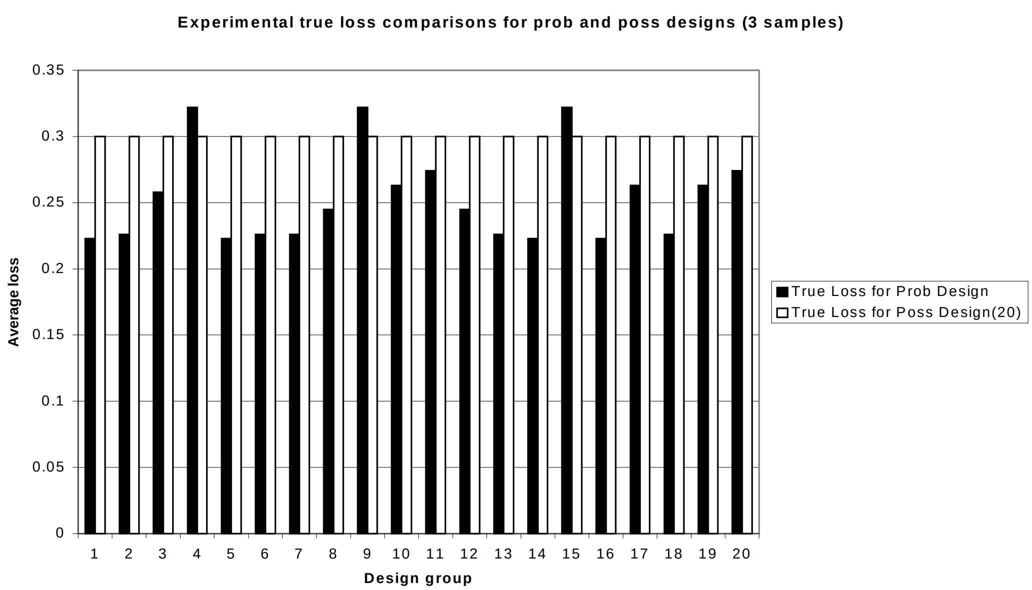

Figure 5.24 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 3)

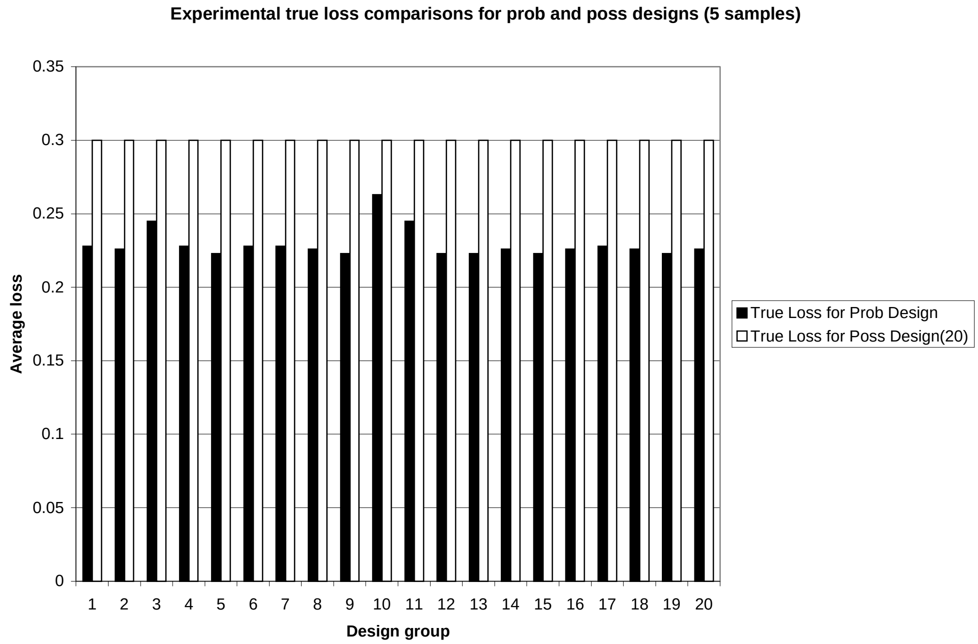

Figure 5.25 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 5)

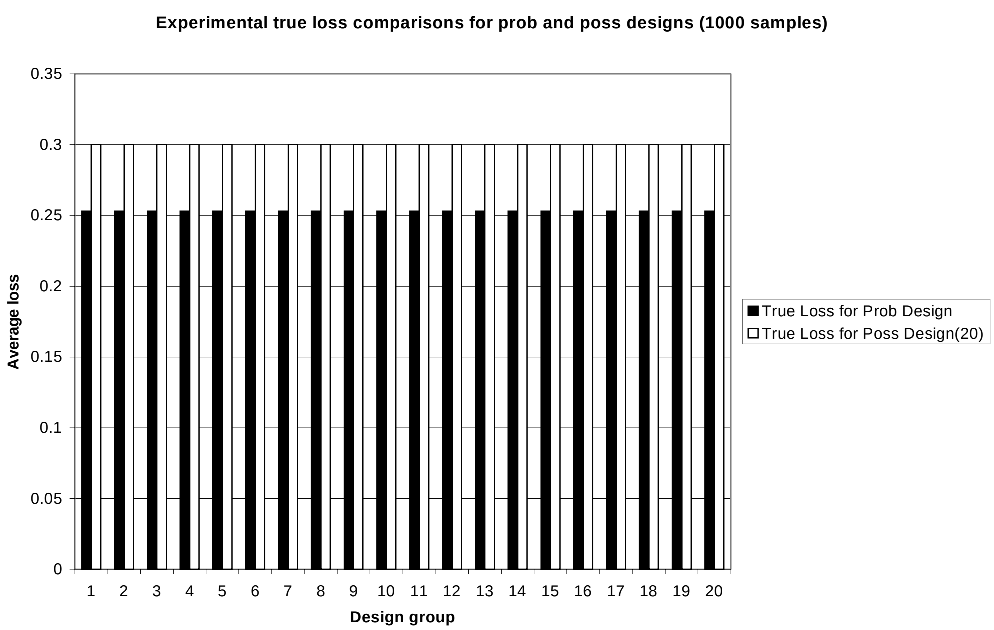

Figure 5.26 Comparison of experimentally estimated expected losses for 20 sets of probabilistic and possibility-based designs with 20 as optimum (sample size = 1000)

5.6.3 Discusssion of results

We compare probabilistic and possibility-based designs using an analytical approach and an experimental approach. The following conclusions can be drawn from comparison results:

- When the sample size is small (e.g. three or five), both analytical and experimental comparisons demonstrate that probabilistic methods can yield designs with considerably higher true expected loss than their possibilistic counterparts.
- Possibility-based methods yield a multiple optimum designs and they can not discriminate between these designs. As a result, other criteria have to be added to specify the optimum design. Two criteria are applied in this study to obtain two different possibility-based optimum designs, with heights of 20 and 30 blocks. Based on these two criteria, we have different conclusions on which methods yield safer designs in terms of analytical and experimental true losses. Table 5.3 shows the comparison results. The conclusion on this issue is conditional.

Table 5.3 Winners of experimental and analytical comparisons in terms of true average losses

	Analytical comparisons		**Experimental comparisons**	
	Small sample size	**Large sample size**	**Small sample size**	**Large sample size**
Possibilistic optimum is 30	Possibility-based method*	Probabilistic method	Possibility-based method*	Possibility-based method
Possibilistic optimum is 20	Probabilistic method*	Probabilistic method	Probabilistic method*	Probabilistic method

* Some probabilistic designs had considerably higher expected loss than the other probabilistic and possibilistic designs

- When the sample size increases, analytical comparisons demonstrate that probabilistic methods improve designs very fast by adjusting the probabilistic models according to the accumulation of data. However, possibility-based methods are insensitive to the change in sample size.

- When very little information is available for uncertainty models, a designer should avoid using these two methods alone. A probabilistic designer can resort to possibility-based method to screen out the poor designs if he/she has many probabilistic candidates. If a probabilistic design has a relative higher possibility of failure, this design should be rejected. If the possibility-based method can not discriminate these probabilistic designs, the designer should try to collect more information to develop a reliable model rather than go ahead and make risky decisions.

5.6 References

Rosca, R., Haftka, R.T., et al, "Block Toppling Model for Testing Procedures for Design Against Uncertainty", *41th AIAA/ASME/ASCE/ AHS/ASC SDM Conference*, April 2000, Atlanta, GA.

CHAPTER 6 CONCLUSIONS

The goal of this dissertation was to understand the advantages and limitations of probabilistic and possibility-based methods for design under uncertainty. We considered problems involving both random uncertainty and reducible uncertainty, but we modeled only random uncertainty using probabilistic and possibility-based methods. In Chapters 2 to 4 we only considered problems where failure is crisply defined, that is the boundary between survival and failure is sharp. In Chapter 5 we considered a design problem in which failure is a matter of degree. We focused on how the lack of information about the uncertainties affects each method. In Chapter 2, we examined the theoretical foundations of probability and possibility-based methods. In Chapters 3-5 we compared them in design for maximum safety.

6.1 Differences in the theoretical foundations of the two design methods

In Chapter 2, we examined the theoretical foundations of probabilistic and possibility-based methods in modeling uncertainties. We reached the following conclusions:

A major difference between probability and possibility is in the axioms for the union of disjoint events: the probability of the union is the sum of the probabilities of these events, whereas the possibility of the union is equal to the largest possibility of these events. Consequently, the probabilities of all events, which partition the universal event, must add up to one. Possibility theory does not impose such a constraint on the possibilities of these events. This makes possibility models more flexible than their probabilistic counterparts. For example, if we have no information about the weather tomorrow we can estimate that both the possibilities of raining and not raining tomorrow are one, to account for the lack of information. On the other hand, if we estimate that the probability of raining is 0.5 then we have to assign a probability of 0.5 to the event "it will not rain tomorrow". If we have enough information about uncertainties and accurate predictive models probability is advantageous, whereas possibility is useful if we have little information.

Possibility and probability calculi are fundamentally different. We cannot simulate the results of possibility calculus using probability calculus by properly selecting the parameters of the probabilistic models.

Possibility can be less conservative than probability in risk assessment of systems with many failure modes. Possibility-based methods tend to underestimate the risk of failure of such systems, especially if the number of modes is large.

Possibility tends to yield more conservative estimates of the risk of failure for systems for which many unfavorable events have to occur simultaneously in order to produce failure. An example is a parallel system.

In many reliability assessment problems, one can easily determine the most conservative possibilistic model that is consistent with the available information. On the other hand, it is difficult to choose the most conservative probabilistic model if little information is available.

If there is little information about uncertainty, probabilistic methods may fail to predict the effect of design modifications on the system probability of failure.

The difference between probabilistic and possibilistic models can lead to a diametrically opposed balancing of risks associated with cost over-runs and risks associated with performance shortfall. The reason is that probabilistic optimization tends to make the derivatives of the probabilities of failure of the failure modes with respect to the design variables equal, whereas possibility-based optimization tends to make the possibilities of the modes equal, rather than their derivatives.

Possibility-based design leads to optimum designs that tend to be less sensitive to errors in the models of uncertainties and in the predictive models than their probabilistic counterparts. The reason is that, as mentioned in the previous paragraph, probabilistic design uses sensitivity derivatives, whereas possibility uses the possibilities of the failure modes, and because the derivatives of a quantity tend to be more sensitive than the value of this quantity to errors.

The difference in the axioms about the union of events is the principal reason for the differences mentioned in the previous paragraphs.

If we have enough information about uncertainties and accurate predictive models, then probability is advantageous. On the other hand, *when making design decisions under limited* *information or using crude predictive models it is important to consider both the probability and possibility of failure of a system.*

6.2 How the two design methods compare in design problems

In Chapters 3 through 5, we demonstrated the advantages and the limitations of probabilistic and possibility-based methods using a problem involving design of a tuned vibration absorber and a problem involving design and construction of domino stacks. The results showed that when limited information is available for the uncertainty and there are narrow failure zones close to the mean values of uncertain variables, the probabilistic method might produce unsafe designs. In these cases, a possibility-based method is useful. A principal reason is that it is easy to determine what assumptions about the distribution parameters and the correlation of the uncertain variables make a possibility-based model more conservative whereas this is not the case with the probabilistic models.

6.3 Guidelines for the use of probabilistic and possibility-based methods

In sections 4.4 and 5.6, we recommend that when limited numerical information is available for the uncertainties in a design problem, one should choose his/her design method based on the following considerations:

1) How much data about uncertainties is available?
2) Is the true type of probability distribution of the random variables known?

When the true type of the probability distribution is known, a Bayesian probabilistic method is better than the possibility-based method. However, when the true type of the probability distribution is unknown and little data is available, we should use both probabilistic methods and possibility-based methods to assess the safety of a design. When there is a large amount of data, one should determine the true distribution type first and then use a probabilistic method. Table 6.1 provides tentative guidelines to select a method based on the amount of information available for a given design problem. These guidelines are based on the results of the tuned damper and domino stack design problems.

Table 6.1 Guidelines for selection of methods for a given design problem

Sufficiently large sample size / **True type of distribution known**	**Yes**	**No**
Yes	Probabilistic/standard statistical method	Bayesian method
No	Try to identify the correct distribution type and use probabilistic methods if *N* > 40.	• Consider both probabilistic and possibility-based designs. Calculate both the possibility of failure and the probability of failure of these designs. If a design scores well with both measures, it should be selected. (*N* < 40) • Try to identify the correct distribution type and use probabilistic methods if *N* > 40.

When very little statistical information is available for uncertainty models and the true type of probability distribution is not known, a designer should avoid using probabilistic methods alone. He/she should use possibility-based methods to screen out those probabilistic designs that might be unsafe. If a probabilistic design has a relatively high possibility of failure, this design should be rejected even if it has a low failure probability. If the possibility-based method can not distinguish probabilistic designs, the designer should collect more information to construct a reliable uncertainty model before performing designs.

6.4 Future research

In this dissertation, we compared probabilistic and possibility-based design methods in the presence of uncertainties. It is useful to compare other methods, such as methods using convex models, imprecise probability and evidence theory with probabilistic and possibilistic methods.

In this study we only dealt with modeling random uncertainties. We have not investigated methods of modeling reducible uncertainties, such as errors in the deterministic models for predicting the performance of a system and errors in uncertainty models. Theories of uncertainty, such as imprecise probability and evidence theory could be better than probability theory for these uncertainties. To the best of the author's knowledge, nobody has investigated applications of these theories to engineering design.

It is important to develop and study hybrid methods for design under uncertainty. For example one could model random uncertainties using probability and reducible uncertainty using possibility. This approach will provide a measure of the significance of the reducible uncertainty. It will yield a possibility distribution of the probability of failure, instead of a single number for the probability of failure. The larger the range of variation of the probability of failure the larger the reducible uncertainty is. It is also important to compare this hybrid probabilistic-possibilistic approach with a probabilistic approach that tries to account for reducible uncertainty by modeling the parameters of the distributions of the random variables as random variables and computes a probability distribution of the probability of failure.

APPENDIX A CALCULATING PROBABILITY AND POSSIBILITY OF TWO TOWER OVERLAPPING

A.1 Deriving the probability of two towers within and out of a certain distance

A.1.1 Calculating the probability of two towers overlap — $p(|x_B - x_A| < D)$

As shown in Fig.5.1, x_A and x_B are x-coordinates of centers of tower A and B. We split $|x_B - x_A| < D$ into two cases.

I. For $0 \leq x_B - x_A < D$:

$$If\ x_0 - D < x_0'\ x_1 - D < x_1'$$

$$\begin{aligned} p(x_B - x_A < D) &= \int_{x_0'}^{x_0} \frac{1}{(x_1' - x_0')} dx_A \int_{x_0}^{x_A + D} \frac{1}{(x_1 - x_0)} dx_B \\ &+ \int_{x_0}^{x_1 - D} \frac{1}{(x_1' - x_0')} dx_A \int_{x_A}^{x_A + D} \frac{1}{(x_1 - x_0)} dx_B \\ &+ \int_{x_1 - D}^{x_1'} \frac{1}{(x_1' - x_0')} dx_A \int_{x_A}^{x_1} \frac{1}{(x_1 - x_0)} dx_B \\ &= \frac{(x_0 - x_0')(2D + x_0' - x_0)}{2(x_1 - x_0)(x_1' - x_0')} + \frac{D(x_1 - x_0 - D)}{(x_1 - x_0)(x_1' - x_0')} + \frac{(x_1 - x_1' + D)(x_1' - x_1 + D)}{2(x_1 - x_0)(x_1' - x_0')} \end{aligned}$$

$$If\ x_0 - D < x_0'\ x_1 - D \geq x_1'$$

$$\begin{aligned} p(x_B - x_A < D) &= \int_{x_0'}^{x_0} \frac{1}{(x_1' - x_0')} dx_A \int_{x_0}^{x_A + D} \frac{1}{(x_1 - x_0)} dx_B \\ &+ \int_{x_0}^{x_1'} \frac{1}{(x_1' - x_0')} dx_A \int_{x_A}^{x_A + D} \frac{1}{(x_1 - x_0)} dx_B \\ &= \frac{(x_0 - x_0')(2D + x_0' - x_0)}{2(x_1 - x_0)(x_1' - x_0')} + \frac{D(x_1' - x_0)}{(x_1 - x_0)(x_1' - x_0')} \end{aligned}$$

$$If\ x_0 - D \geq x_0'\ x_1 - D \geq x_1'$$

$$p(x_B - x_A < D) = \int_{x_0-D}^{x_0} \frac{1}{(x_1'-x_0')} dx_A \int_{x_0}^{x_A+D} \frac{1}{(x_1-x_0)} dx_B + \int_{x_0}^{x_1'} \frac{1}{(x_1'-x_0')} dx_A \int_{x_A}^{x_A+D} \frac{1}{(x_1-x_0)} dx_B = \frac{D^2}{2(x_1-x_0)(x_1'-x_0')} + \frac{D(x_1'-x_0)}{(x_1-x_0)(x_1'-x_0')}$$

$$If\ x_0 - D \geq x_0'\ x_1 - D < x_1'$$

$$p(x_B - x_A < D) = \int_{x_0-D}^{x_0} \frac{1}{(x_1'-x_0')} dx_A \int_{x_0}^{x_A+D} \frac{1}{(x_1-x_0)} dx_B + \int_{x_0}^{x_1-D} \frac{1}{(x_1'-x_0')} dx_A \int_{x_A}^{x_A+D} \frac{1}{(x_1-x_0)} dx_B + \int_{x_1-D}^{x_1'} \frac{1}{(x_1'-x_0')} dx_A \int_{x_A}^{x_1} \frac{1}{(x_1-x_0)} dx_B$$

$$= \frac{D^2}{2(x_1-x_0)(x_1'-x_0')} + \frac{D(x_1-x_0-D)}{(x_1-x_0)(x_1'-x_0')} + \frac{(x_1-x_1'+D)(x_1'-x_1+D)}{2(x_1-x_0)(x_1'-x_0')} = \frac{2D(x_1-x_0)-(x_1-x_1')^2}{2(x_1-x_0)(x_1'-x_0')} \quad (A.1)$$

II. For $0 > x_B - x_A > -D$:

$$If\ x_1' - D > x_0$$

$$p(x_B - x_A > -D) = \int_{x_0}^{x_0+D} \frac{1}{(x_1'-x_0')} dx_A \int_{x_0}^{x_A} \frac{1}{(x_1-x_0)} dx_B + \int_{x_0+D}^{x_1'} \frac{1}{(x_1'-x_0')} dx_A \int_{x_A-D}^{x_A} \frac{1}{(x_1-x_0)} dx_B = \frac{D(x_1'-x_0-D)}{(x_1-x_0)(x_1'-x_0')} + \frac{D^2}{2(x_1-x_0)(x_1'-x_0')} \quad (A.2)$$

$$If\ x_1' - D < x_0$$

$$p(x_B - x_A > -D) = \int_{x_0}^{x_1'} \frac{1}{(x_1'-x_0')} dx_A \int_{x_0}^{x_A} \frac{1}{(x_1-x_0)} dx_B = \frac{(x_1'-x_0)^2}{2(x_1-x_0)(x_1'-x_0')}$$

A.1.2 Calculating the probability of two towers overlap in another way — $p(|x_B - x_A| < D)$

No matter how the relative location of these two towers change, we only need to calculate the overlapping portion in the overall distributions. In Fig.A.1, x_0 and x_1 are the left and right range in which the center of tower B, x_B, may vary. Similarly, x_0' and $x_1'$are the left and right range in which the center of tower A, x_A, may vary.

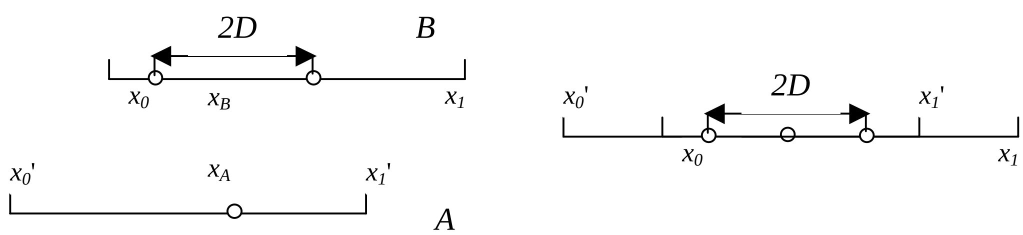

Figure A.1 Alternative calculation of $p(|x_B - x_A| < D)$

If $x_0 - x_0' > D$ and $x_1 - x_1' > D$

$$p(d_0 < D) = \int_{x_0}^{x_1'} \frac{1}{(x_1' - x_0)} dx_A \int_{x_A - D}^{x_A + D} \frac{1}{(x_1 - x_0')} dx_B = \frac{2D}{(x_1 - x_0')} \tag{A.3}$$

If $x_0 - x_0' \leq D$, we have to subtract additional part from Equation (A.3), which is:

$$p_{add} = \int_{x_0}^{x_0' + D} \frac{1}{(x_1' - x_0)} dx_A \int_{x_A - D}^{x_0'} \frac{1}{(x_1 - x_0')} dx_B = \frac{(D + x_0' - x_0)^2}{2(x_1 - x_0')(x_1' - x_0)} \tag{A.4}$$

If $x_1' - x_0' \leq D$, we have to subtract additional part from Equation (A.3), which is:

$$p_{add} = \int_{x_0}^{x_1'} \frac{1}{(x_1' - x_0)} dx_A \int_{x_A - D}^{x_0'} \frac{1}{(x_1 - x_0')} dx_B = \frac{(2D + 2x_0' - x_1' - x_0)}{2(x_1 - x_0')} \tag{A.5}$$

If $x_1 - x_1' \leq D$, we have to subtract additional part from Equation (A.3), which is:

$$p_{add} = \int_{x_1-D}^{x_1'} \frac{1}{(x_1'-x_0)} dx_A \int_{x_1}^{x_A+D} \frac{1}{(x_1-x_0')} dx_B = \frac{(D+x_1'-x_1)^2}{2(x_1-x_0')(x_1'-x_0)} \tag{A.6}$$

If $x_1 - x_0 \le D$, we have to subtract additional part from Equation (A.3), which is:

$$p_{add} = \int_{x_0}^{x_1'} \frac{1}{(x_1'-x_0)} dx_A \int_{x_1}^{x+D} \frac{1}{(x_1-x_0')} dx_B = \frac{(2D-2x_1+x_1'+x_0)}{2(x_1-x_0')} \tag{A.7}$$

A.2 Deriving the possibility of two towers within and out of a certain distance

I. When $x_0 - x_0' < D$ and $x_m - x_m' > D$, as shown in Fig. A.2.

In this case, the maximum value for calculating $_{|d0|\geq D}$ occurs when A locates at the middle point where $\pi(x_A)$=1. From Fig.A.2, we observe that at this point, out of the range $[x_A - D, x_A + D]$, the maximum value for $\pi(x_B)$ is equal to one. Therefore, $\pi_{|d0|\geq D}$ =1.

The maximum value for calculating $\pi_{|d0|<D}$ occurs when the horizontal interval between the right side of triangle A and the left side of triangle B is equal to D. Suppose the right end of that interval is x_{Am}, and the right end of the interval is x_{Bm}. When x_{Am} is smaller than x_A:

$$\min[\pi(x_A), \sup_{x_A-D\leq x_B\leq x_A+D} \pi(x_B)] = \pi(x_A) < \pi(x_{Am}) \tag{A.8}$$

When x_{Am} is bigger than x_A, then:

$$\min[\pi(x_A), \sup_{x_A-D\leq x_B\leq x_A+D} \pi(x_B)] = \pi(x_B) < \pi(x_{Bm}) = \pi(x_{Am}) \tag{A.9}$$

$$\frac{h-h_0}{h_0}=\frac{D}{x_1'-x_0}\Rightarrow \pi_{d_0<D}=h=h_0(1+\frac{D}{x_1'-x_0})$$
$$\frac{h_0}{1}=\frac{l_1}{(x_1'-x_0')/2}\quad \frac{h_0}{1}=\frac{l_2}{(x_1-x_0)/2}$$
$$l_1+l_2=x_1'-x_0 \qquad \text{(A.10)}$$
$$h_0=\frac{2(x_1'-x_0)}{(x_1'-x_0')+(x_1-x_0)}$$
$$\pi_{d_0<D}=\frac{2(x_1'-x_0+D)}{(x_1'-x_0')+(x_1-x_0)}$$

In the equations, *h* represents the height of the horizontal interval with length of *D*. h_0 represents the height of the intersection of right side of *A* with left side of *B*.

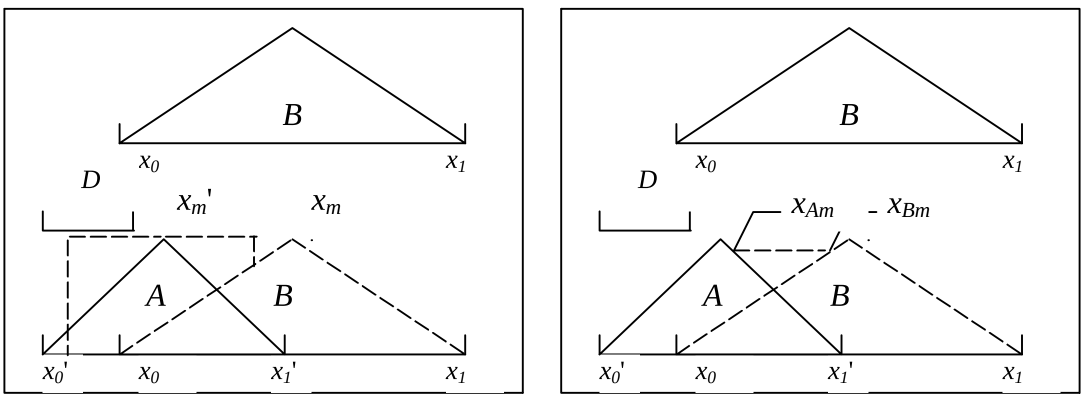

Figure A.2 Illustrations for calculating π| $x_B - x_A$ |$_{<D}$ and π| $x_B - x_A$ |$_{\geq D}$ for case I

II. When $x_0 - x_0' < D$ and $x_m - x_m' \leq D$:

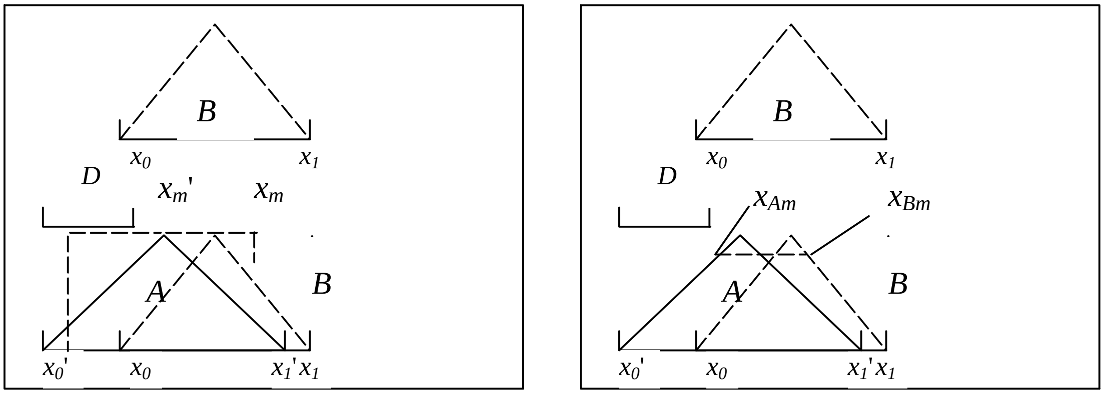

Figure A.3 Illustrations for calculating $\pi|\ x_B - x_A\ |_{<D}$ and $\pi|\ x_B - x_A\ |_{\geq D}$ for case II

In this case, the maximum value for calculating $\pi|\ x_B - x_A\ |_{<D}$ occurs when A locates at the middle point where $\pi(x_A)$=1. From Fig. A.3, we observe that at this point, within the range $[x_A - D, x_A + D]$, the maximum value for $\pi(x_B)$ is equal to one. Therefore, $\pi|\ x_B - x_A\ |_{<D}$ =1.

The maximum value for calculating $\pi|\ x_B - x_A\ |_{\geq D}$ occurs when the horizontal interval between the left side of triangle A and the right side of triangle B is equal to D. Suppose the right end of that interval is x_{Am}, and the right end of the interval is x_{Bm}.

$$\frac{h-h_0}{h_0} = \frac{l_1}{x_1'-x_0} \quad \frac{1-h}{1} = \frac{l_2}{(x_1'-x_0')} \quad \frac{1-h}{1} = \frac{l_3}{(x_1-x_0)}$$
$$l_1 + l_2 + l_3 = D$$
$$h_0 = \frac{2(x_1'-x_0)}{(x_1'-x_0') + (x_1 - x_0)} \tag{A.11}$$
$$\pi_{d_0 \geq D} = h = \frac{2(x_1 - x_0'-D)}{(x_1'-x_0') + (x_1 - x_0)}$$

III. When $x_0 - x_0' \geq D$ and $x_m - x_m' > D$:

In this case, the maximum value for calculating $\pi|\ x_B - x_A\ |_{\geq D}$ will occur when A locates at the middle point where $\pi(x_A)$=1, from a similar approach to Case I. Therefore, $\pi|\ x_B - x_A\ |_{\geq D}$ =1. From Fig.A.4, we observe that the maximum value for calculating $\pi|\ x_B - x_A\ |_{<D}$ will occur when the horizontal interval between the right side of triangle A and the left side of triangle B is equal

to D. Let the right end of that interval is x_{Am}, and the right end of the interval is x_{Bm}. Similar to Case I, $\pi|\ x_B - x_A\ |_{<D}$ is calculated as:

$$h_0 = \frac{2(x_1' - x_0)}{(x_1' - x_0') + (x_1 - x_0)} \quad \pi_{d_0 < D} = \frac{2(x_1' - x_0 + D)}{(x_1' - x_0') + (x_1 - x_0)} \tag{A.12}$$

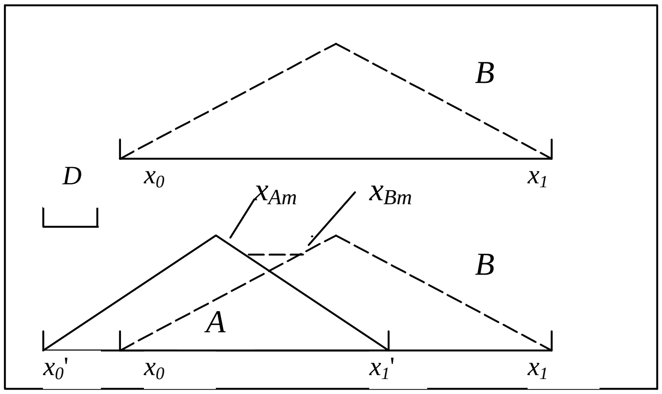

Figure A.4 Illustrations for calculating $\pi|\ x_B - x_A\ |_{<D}$ and $\pi|\ x_B - x_A\ |_{\geq D}$ for case III

IV. When $x_0 - x_0' \geq D$ and $x_m - x_m' \leq D$:

This case is similar to Case II, where $\pi|\ x_B - x_A\ |_{<D} = 1$.

$$\pi_{d_0 \geq D} = \frac{2(x_1 - x_0' - D)}{(x_1' - x_0') + (x_1 - x_0)} \tag{A.13}$$

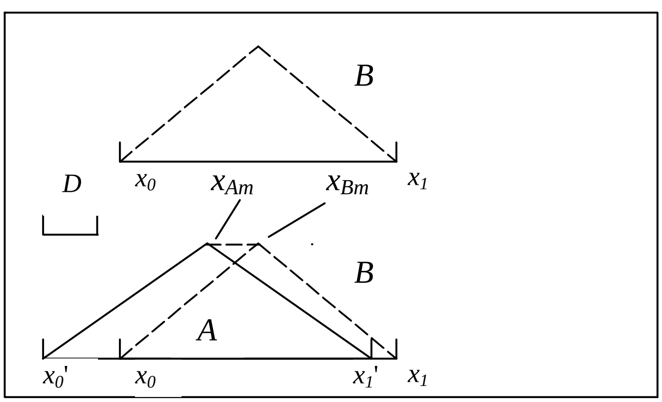

Figure A.5 Illustrations for calculating $\pi|\ x_B - x_A\ |_{<D}$ and $\pi|\ x_B - x_A\ |_{\geq D}$ for case IV

VITA

Sophie Chen was born in Wuxi, China on September 28, 1968. She received her Bachelor of Science degree from Shanghai Jiaotong University in July 1990 and her Master of Science degree from Huazhong Science and Technology University in July 1993. During 1993 and 1997, she worked for China Classification Society, in the offshore engineering area.   She started her Ph.D program in the Aerospace and Ocean Engineering Department at Virginia Tech from August 1997. She will join Novellus Systems at San Jose, California upon graduation from Virginia Tech.