ROBUST AND RELIABILITY-BASED DESIGN OPTIMIZATION UNDER EPISTEMIC UNCERTAINTY

By

PRITHBEY RAJ DEY

A Thesis

Submitted to the

Department of Industrial & Production Engineering,

Bangladesh University of Engineering and Technology

in partial fulfillment of the requirements

for the Degree

of

MASTER OF SCIENCE

in

Industrial & Production Engineering

January, 2015

Bangladesh University of Engineering and Technology, Bangladesh

The thesis titled **ROBUST AND RELIABILITY-BASED DESIGN OPTIMIZATION UNDER EPISTEMIC UNCERTAINTY** submitted by Prithbey Raj Dey, Roll No.: 0412082029, Session- April, 2012, has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science (M. Sc.) in Industrial and Production Engineering on January 19, 2015.

BOARD OF EXAMINERS

1.	Dr. AKM Kais Bin Zaman Associate Professor Department of Industrial & Production Engineering (IPE) Bangladesh University of Engineering and Technology, Dhaka	Chairman
2.	Dr. Sultana Parveen Professor & Head Department of Industrial & Production Engineering (IPE) Bangladesh University of Engineering and Technology, Dhaka	Member (Ex-officio)
3.	Dr. M. Ahsan Akhtar Hasin Professor Department of Industrial & Production Engineering (IPE) Bangladesh University of Engineering and Technology, Dhaka	Member
4.	Dr. Shuva Ghosh Assistant Professor Department of Industrial & Production Engineering (IPE) Bangladesh University of Engineering and Technology, Dhaka	Member
5.	Dr. Md. Forkan Uddin Assistant Professor Department of Electrical and Electronic Engineering (EEE) Bangladesh University of Engineering and Technology, Dhaka	(External)

Robust and Reliability-based Design Optimization under Epistemic Uncertainty

ABSTRACT

This thesis proposes formulations and algorithms for robust design optimization with uncertainty representation and propagation considering both aleatory (e.g. produced due to natural variability) and epistemic (e.g. variability due to lack of information or imprecise information) uncertainty arising from interval data. Multiple interval data are treated for uncertainty representation including both overlapping and non-overlapping in characteristics. A general likelihood-based approach for uncertainty representation has been proposed in this research. Uncertainty analysis through the likelihood approach is capable of estimating the uncertainty for different distribution types and parameters. The proposed likelihood-based representation of epistemic uncertainty has been used in the framework for robustness-based design optimization to achieve computational efficiency. A methodology is also outlined for solving reliability-based design optimization (RBDO) under epistemic uncertainty using the proposed likelihood-based uncertainty representation. The proposed robust design optimization methodology is illustrated with two numerical examples including a general mathematical problem and a real engineering problem.

CANDIDATE'S DECLARATION

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Prithbey Raj Dey
Date: 19^{th} January, 2015

ACKNOWLEDGEMENT

I wish to acknowledge my sincere and profound gratitude and pay a lot of thanks to my thesis supervisor, Dr. AKM Kais Bin Zaman, Associate Professor of the Department of Industrial and Production Engineering (IPE), Bangladesh University of Engineering and Technology (BUET), Dhaka, for his valuable suggestions, constant inspiration, careful supervision, appropriate advice and affectionate guidance during the continuation of the research done to prepare the thesis. It would be impossible to carry out thesis work without his dynamic direction under a number of constraints and time limitation in particular. It has been a blessing to have him as the thesis supervisor and mentor.

I would like to express my gratitude and thanks to the committee members including Dr. Sultana Parvin, Professor & Head, Department of Industrial & Production Engineering (IPE), BUET, Dr. M. Ahsan Akhtar Hasin, Professor, Department of Industrial & Production Engineering (IPE), BUET and Dr. Shuva Ghosh, Assistant Professor, Department of Industrial & Production Engineering (IPE), BUET for their valuable critics and suggestions.

I would like to express thanks and appreciation to all my well wishers who inspired me to continue this work. I am grateful to my family for their support and love which was the main source of inspiration. Lastly, I convey my thanks to all of those who have supported me in any respect during the research work.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1. Background

In engineering design optimization, the input design variables and system variables may contain uncertainty. Therefore, uncertainty representation and propagation is required in the systems and machineries design, production, maintenance and other applications. Deterministic design optimization is the traditional way which is bounded to a lot of errors due to reclusion of uncertainty. In deterministic design optimization, it is generally assumed that all design inputs are precisely known and the influence of data or distribution parameter uncertainty on the optimality and feasibility of the models is not explicitly considered. However, real-life engineering problems consist of different types of uncertainty and this deterministic assumption about inputs may lead to infeasibility or poor performance (Sim 2004). In recent years, several methods have been developed for design under uncertainty. Reliability-based design (Chiralaksanakul and Mahadevan 2005) and robust design (Du and Chen 2000; Du and Huang 2007) are two major developments among these. While reliability-based design aims to maintain design feasibility at desired reliability levels, robust design optimization attempts to minimize variability in the system performance due to variations in the inputs (Lee et al. 2008). All these methods developed so far work under aleatory uncertainty (i.e., precise probabilistic information). However, uncertainty increases proportionally with the increase of the complexity in engineering design.

Uncertainty can be divided into two types: aleatory and epistemic. Further, epistemic uncertainty can be defined in two ways. It can be defined with reference to a *poorly known quantity but stochastic in character* (Baudrit and Dubois 2006) or with reference to a *fixed but poorly known quantity* (Helton et al. 2004). There is now an extensive volume of literature available to deal with these two definitions of epistemic uncertainty. However, existing methods can handle these two definitions separately (Helton et al. 2004; Zaman et al. 2011a; Zaman et al. 2011b; Zaman and Mahadevan 2013), although it may be possible that an engineering design contains both types of epistemic uncertainty. Therefore, an approach to engineering system design that addresses both aleatory and epistemic uncertainty of both types is needed.

1.2. Objectives with specific aims

The specific objectives of this research are:

- Development of a methodology for representation of epistemic uncertainty using likelihood-based approach.
- Development of formulations and algorithms for robustness-based design optimization under epistemic uncertainty.
- Development of formulations and algorithms for reliability-based design optimization under epistemic uncertainty.

Therefore, the proposed research develops and demonstrates generalized methodologies and tools for managing uncertainty in engineering systems, which will provide decision support to engineers for robust and reliable design of engineering systems.

1.3. Outline of Methodology

The proposed research methodology is outlined below:

a) A framework for the representation of aleatory uncertainty and stochastic epistemic uncertainty with insufficient multiple interval data has been developed.

b) Formulations and algorithms for robustness-based design optimization under epistemic uncertainty have been proposed based on the developed uncertainty representation framework.

c) Formulations and algorithms for reliability-based design optimization under epistemic uncertainty have been proposed based on the developed uncertainty representation framework.

d) The proposed robustness-based design methodology has been illustrated for several example problems.

1.4. Scopes and limitations

The uncertainty representation methodology developed in this thesis is applicable to any engineering applications under epistemic uncertainty. However, the developed methodology has been used in robust and reliability-based design optimization framework under epistemic uncertainty for solving single discipline engineering problems. The proposed optimization framework can easily be extended to solve multidisciplinary optimization problems under epistemic uncertainty. The proposed method may also contribute in the field of financial engineering problems, such as portfolio optimization and network optimization including supply chain management, warehouse management, inventory management, etc.

The epistemic uncertainty analysis requires extra computational effort to estimate the uncertain parameters of the epistemic variables. However, this is essential to develop a better and efficient methodology for the design optimization.

CHAPTER II

LITERATURE REVIEW

The current research is intended to develop a methodology that provides decision support to engineers for design and analysis of engineering systems for poor amount of data counting stochastic epistemic uncertainty. This uncertainty prevails in the natural characteristics of a design and observed data and this is ignored in the deterministic design optimization. Non-deterministic design optimization has gained increasing attention in last few decades due to this reason. There are now extensive volume of methods and applications available for non-deterministic design optimization problems. Robustness-based design (Parkinson et al. 1993; Du and Chen 2000; Doltsinis and kang 2004; Huang and Du 2007) and reliability-based design (Chiralaksanakul and Mahadevan 2005; Ramu et al. 2006; Agarwal et al. 2007; Du and Huang 2007) are two prominent fields of optimization which consider the uncertainty in the design parameters. Robustness is the performance criteria for a system to operate continuously for a wide range of operational conditions and will be failed outside the conditions (Steven 2001). Taguchi developed the concept of robust design and proposed a method where the product performance or the output remains insensitive to the variation in design variables in manufacturing process (Taguchi 1993). The variation in the design variables was designated as noise which could be created from various factors in the manufacturing process. As, all of the engineering models are becoming more and more complex day by day, application of statistical design tools in Taguchi"s method is not well enough to calculate optimal feasible solution for multiple measurements of performance and design constraints (Wei et al. 2009). Due to application of nonlinear programming in the robust design, it became possible to achieve robustness in both performance outputs and design constraints (Du and Chen 2000).

There exists high volume of researches for robust and reliability-based design optimization. However, most of the methods have been developed considering the physical or natural variability represented by probability distribution. There are normally some significant variables or elements in the system which arise uncertainty in the system. Uncertainty may be aroused from two sources which are aleatory and epistemic (Oberkampf et al. 2004). Aleatory uncertainty cannot be reduced due to including natural phenomena that exhibit natural variation like operating condition, material properties, geometric tolerances, etc. on the other hand, epistemic uncertainty arises from a lack of knowledge about the system, or due to approximations in the system behavior models, or due to limited or subjective data. Epistemic uncertainty can be reduced by gathering as much as information about the system. There will be continuous variation in the manufacturing process due to lack of proper precision in the tools and system. When there is limited knowledge about a system or approximations have to apply for various estimations, epistemic uncertainty arouses. There may be limited data to properly define the distribution parameters of the random variables. This type of uncertainty may be reduced by harnessing more data. There are few studies on robust design optimization which have considered the epistemic uncertainty arising from the lack of information. Probability-based method (Youn et al. 2007) may be used to redefine the performance measure of robust design using the most likely values of fuzzy random variables. Two two-step methods have been developed by Dai and Mourelatos (2003) for robust design optimization which can treat aleatory and epistemic uncertainty separately using a range of method and fuzzy set approach.

There exist a few methods that develop robustness-based design optimization methodology under data uncertainty (i.e., using sparse point data and interval data on input random variables). Zaman et al. (2011b) proposed a decoupled approach for robustness-based design optimization using both sparse point and interval data. They achieve computational efficiency by un-nesting the design optimization from the uncertainty analysis of the epistemic variables. This is a sequential approach, where two optimization formulations are solved until convergence. In this thesis, an optimization approach is proposed where the epistemic analysis is completely eliminated from the design optimization framework by estimating the distribution parameters of the epistemic variables using a likelihood-based uncertainty representation of interval data.

2.1. Interval data uncertainty

There are normally two types of sparsity lies in the data which are controlled and random. Based on the computational method, there are two types of interval data which are single and multiple intervals. Comparing to single interval, multiple intervals require consideration of two additional issues: (1) from the context of computational expense, estimating statistics from multiple intervals can be more challenging, (2) from the context of aggregation of information represented in the multiple intervals, there may be no basis to believe that the "true" value of the variables lies at any particular location of any intervals, such as endpoints or midpoints of the intervals. It is assumed in this research that all the intervals are equally likely to include the true value of the variable, i.e., all the intervals have an equal weight (Ferson et al. 2007). When data is available in multiple intervals, the information contained in the intervals can be broadly categorized as non-overlapping and overlapping intervals. Different types of experiments and procedures are responsible to generate different types of interval data. Therefore, it is important to know about the sources of the interval data when they are collected as the observed data of the random variables.

Sources of interval data

For developing efficient and effective system, interval data are used frequently in practical engineering problems. Interval data are used in some specific situations (Ferson et al. 2007 ;Du et al. 2005), for example: (a) information could be gotten through physical limits and theoretical constraints, which can only provide possible ranges of quantities resulting in interval data, (b) expert opinions are strong sources of interval data , which specify a range of possible values for a variable, (c) Reporting data includes positive or negative uncertainties linked with the calibration of measuring devices also leads to interval data. (d) The results in the experiments like chemical or purity quantification below a certain detection limit, providing an interval or range of the observation for the amount of impurity between zero and the threshold. (e) For temporary or interim observations, there is always a chance to detect faults which can be evolved during two consecutive observations. Hence, the time of failure of a machine or a system is given through intervals. If the interval width cannot be ignored comparing with the magnitude of the variable, then it requires special treatment.

Representation of Interval data

In order to estimate and propagate the uncertainty in the design of a real engineering system, it is important to have proper uncertainty representation through models from which it will be possible to quantify the uncertainty. Therefore, the uncertainty quantification will help to estimate the system response. Two theories are popular to represent and mention the characteristic of interval data for uncertainty representation.

Equi-probability model corresponds to the Laplacian principle of indifference (Howson and Urbach 1993) and considers each interval as a uniform distribution (Bertrand and Groupil 2000). Single probability mass and density are estimated for each possible realization of the interval data for the random variable. Assumption of uniform distribution or any other distribution within a particular interval cannot be valid and could be a limitation of the equi-probability model. The inherent imprecision in the interval data cannot be captured though equi-probability model results in a precise probabilistic representation.

The second interpretation of the interval data representation can be illustrated as interval data efficiently represents incertitude (Ferson et al. 2007). Therefore, the probability of an event will be an interval, unlike a single value for point data or sparse data. Hence, the probabilistic representation of the interval data is different from the uncertainty representation of the single point data. Different possible combinations of information in the interval data can be used for uncertainty representation.

The representation of the interval data of a variable for a specific distribution, (e.g., empirical, normal) could be a probability box or *p*-box (Williamson and Downs 1990). Evidence theory and fuzzy logic are also popular and used in a great variety within the interpretation of incertitude. Various techniques have been developed to represent interval data to represent and propagate the uncertainty.

When a variable is described by interval data and for which the set of all probability distribution of a specific distribution type (e.g., empirical, normal) are feasible, then the set of the probability distribution is known as a probability box or p-box in short. The empirical p-box shows the interval data set graphically. The empirical p-box has an increasing step function with a constant vertical height of $1/n$, where n is the number of intervals. This step height of each interval data for the empirical CDF is identical which represents the intervals are equally weighted. The sorting from the lower and upper bounds for the set of intervals is required to construct the p-box. Empirical cumulative distribution function (CDF) for each bound is plotted for the non-design variables used in the robust design optimization problem. The non-design variables may consist of overlapping, non-overlapping and mixed interval data which can be observed from Figs. 1.1-1.3. There are 5 sets of interval data for each non-design variable.

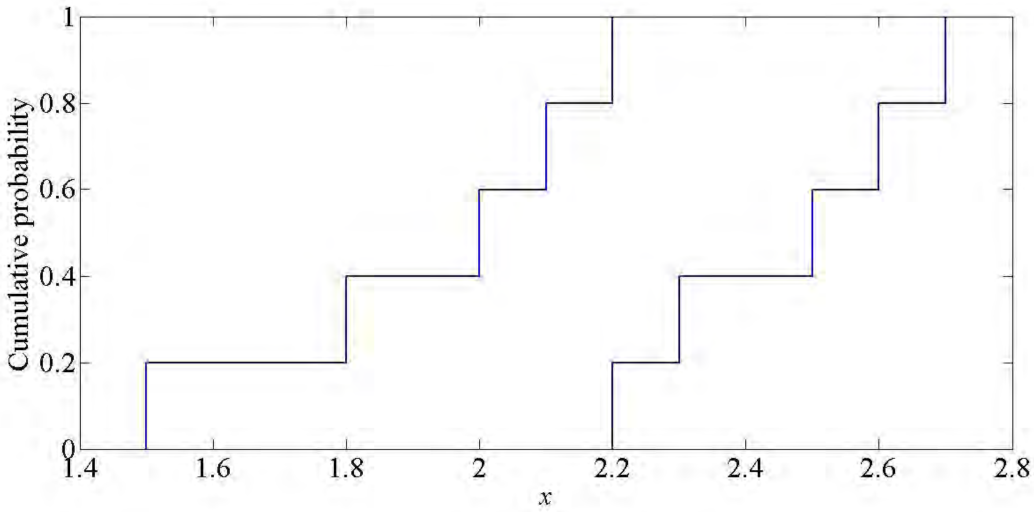

Fig. 1.1 Empirical *p*-box for overlapping interval data

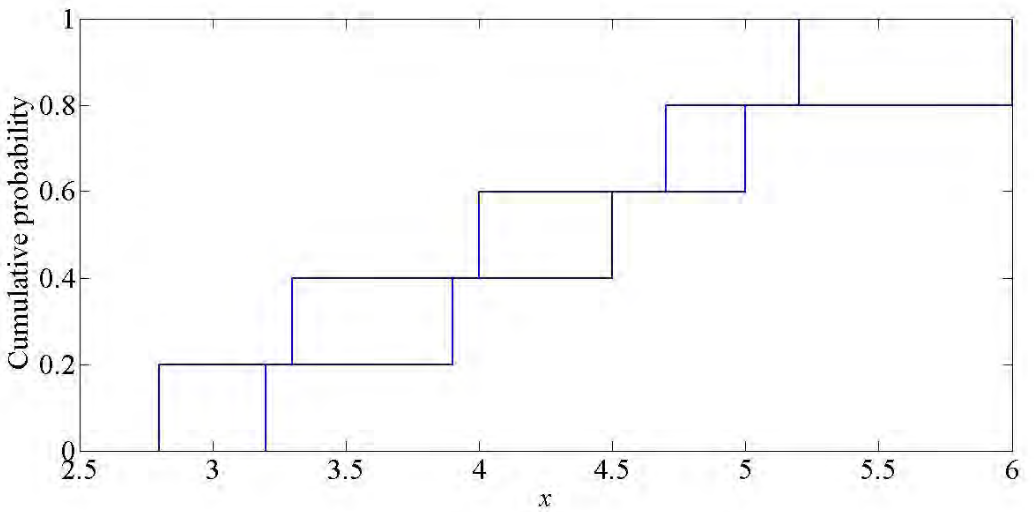

Fig. 1.2 Empirical *p*-box for non-overlapping interval data

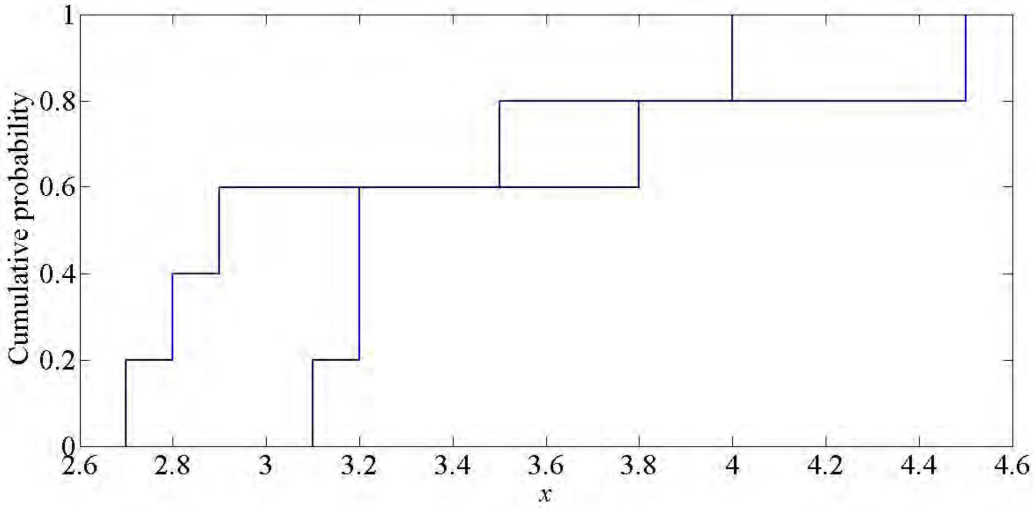

Fig. 1.3 Empirical *p*-box for mixed interval data

Different techniques on quantification and propagation of epistemic data uncertainty for single point and interval data (Ferson et al. 2004) were the major topics in the workshop of the Sandia epistemic uncertainty project (Oberkampf et al. 2004). For representation and propagation of interval uncertainty, the popular uncertainty theories include probability distributions (Helton et al.2004), Dempster-Shafer structures (Helton et al.2004; Klir 2004), p-boxes (Ferson and Hajagos 2004), possibility distributions (Helton et al.2004), subjective probabilities (O"Hagan and Oakley 2004), random intervals (Fetz and Oberguggenberger 2004), set of probability measures (Fetz and Oberguggenberger 2004), fuzzy sets (Fetz and Oberguggenberger 2004), random sets (Berleant and Zhang 2004; Hall and Lawry 2004), imprecise coherent probabilities (Kozine and Utkin 2004), coherent lower provisions (De Cooman and Troffaes 2004), families of polynomial chaos expansions (Red-Horse and Benjamin 2004), info gap models (Ben-Haim 2004) etc.

Different researches including *p*-box method providing a probability theory for interval data has focused on developing bounds mainly on CDFs (Hailperin 1986) with the method of propagation of these probability intervals through uncomplicated expressions. Following this ideas, probabilistic arithmetic expressions in the density domain were developed by Hyman (1982). Williamson and Downs (1990) proposed algorithms to compute arithmetic operations (addition, subtraction, multiplication and division) on pairs of p-boxes. The notion of the convolution between probability distributions was generalized through these methods (Berleant 1993; Berleant 1996; Berleant and Goodman-Strauss 1998). More research works involving bounds on CDFs also developed (Helton et al. 2004; Helton et al. 2008).

Subjective probability is another medium of representing epistemic uncertainty (O"Hagan and Oakley 2004; Apeland et al. 2002; Hofer et al. 2002). Probability is the common used method for representation of both aleatory and epistemic uncertainty but not the only way (O"Hagan and Okaley 2004). There may be lack of precision in probability judgments and in the estimation of the uncertainty, it requires elicitations, not representation. This Bayesian approach (Howson and Urbach 1993) is popular for epistemic uncertainty for former distributions which are updated in the presence of new data. Some researchers also argue that a probabilistic representation for interval data is not enough due to addition of information to the problem (Du et al. 2005; Agarwal et al. 2004a).

In the context of Dempstar-Shafer evidence theory (Shafer 1976) for interval data, there are many rules to aggregate different sources of information. However, this evidence theory may not be suitable for cases where there is inconsistency in the available evidence (Agarwal et al. 2004a; Oberkampf et al. 2001) e.g., for non-overlapping intervals. In such cases, a mixture of averaging rule can be applied (Oberkampf et al. 2001). Evidence theory has popular application to represent epistemic uncertainty for interval data in multidisciplinary system design (Agarwal et al. 2004a), which is used to develop or formulate the non-deterministic design constraints. Other researchers also used evidence theory for epistemic uncertainty quantification (Guo and Du 2007; Guo and Du 2009).

Uncertain events from patterns can be designed using convex models of uncertainty (Ben-Haim and Elishakoff 1990). Ellipses of any convex sets are the examples of convex models consist of intervals. There is less detailed information in the convex models to represent uncertainties than probabilistic models. The convex model requires a worst case analysis to the design applications which can be developed as a constrained optimization problem. The computation of interval analysis is expensive when complex model has intervals.

Possibility or fuzzy set theory for interval data used a modern technique. The possibility distribution of the membership function for an interval variable with a given possibility distribution can be estimated using Zadeh"s Extension Principle (Dubois and Prade 1988). This process requires combinatorial interval analysis. Therefore, the computational expense increases exponentially with the number of uncertain variables and with the nonlinearity of the function which is a major drawback. Rao and Annamdas (2009) proposed weighted fuzzy theory of intervals where fuzzy set based representations of interval variables from evidences of different credibility are combined to estimate the system margin of failure.

The aggregation methods of multiple sources of information for multiple interval data is an important issue in characterizing input uncertainty. Different aggregation methods treated multiple interval data includes stochastic mixture modeling (Helton et al. 2004; Ferson and Hajagos 2004), posteriori mixture (Red-Horse and Benjamin 2004), Dempster‟s rule (Agarwal et al. 2004a; Rutherford 2004), a natural extension of pointwise maximum (De Cooman and Troffaes 2004), etc. However, the nature of the aggregation method should be consistent with the uncertainty theory specifically used for uncertainty representation (Helton et al. 2004).

Probability theory, evidence theory, possibility theory and inter analysis are used by (Helton et al. 2004) for the representation and propagation of epistemic uncertainty. K. zaman et al. (2011a) proposed a sampling based approach with each of the uncertainty theories. Uniform distribution is assumed over the sets of the possible values of the input variables to define the probability. Different information sources are gathered by simply averaging the distributions by assigning equal weight to each source. Imprecise probabilistic information described by intervals can also be used to formulate different uncertainty approaches, such as probability theory, possibility theory, belief functions etc. (Baudrit and Dubois 2006).

There are many methods developed for uncertainty representation and propagation considering both interval and aleatory uncertainties. Evidence theory or possibility theory are common approaches used to represent interval variables. However, probabilistic representation is generally used to represent aleatory uncertainty. The propagation of an evidence theory through a model for the representation of uncertainty estimating system response is computationally more expensive than that of probability theory (Helton et al. 2007). Helton et al. (2008) discussed the efficiency of different alternatives for the representation and propagation of epistemic uncertainty. Computational effort and the expense are huge for the propagation of epistemic uncertainty using evidence theory and possibility theory than that of probability theory. In uncertainty propagation analysis for the interval data, there are various combinations produced by the computational method. For every combination of interval values, the probabilistic analysis for aleatory variables is repeated, which results in a computationally expensive nested analysis. Therefore, great efforts have been given on managing these computational expenses (Penmetsa and Grandhi 2002; Rao and Cao 2002). Representation and propagation of interval uncertainty are illustrated in the models of structural problems (Langley 2000) and multidisciplinary problems (Du and Chen 2000b). The users used non-probabilistic method to avoid computational complexity for uncertainty analysis due to lack of awareness about various methods. Therefore, huge educational effort is required to make the end users familiar with these non-traditional uncertainty analysis methods (Helton et al. 2008).

There are advantages and limitations in the different approaches for uncertainty analysis for the interval data. Most of the approaches required nested analysis in the presence of the interval variables. Zaman et al. (2011a) proposed a probabilistic representation for interval data using a family of Johnson distributions. An aggregation technique is proposed which enables the use of the moment matching method to represent the uncertainty described by the multiple intervals through a family of probability distributions. An important advantage is that it allows for a unified probabilistic framework which can jointly handle aleatory and epistemic uncertainties. Efficient analytical probabilistic methods such as first-order reliability method (FORM) and second-order reliability method (SORM) are allowed to use for uncertainty propagation with the unified probabilistic framework. The expensive nested analysis is avoided and this method enables the use of an optimization-based strategy for the estimation of the distribution parameters of the input variables that minimizes or maximizes a system response.

Likelihood-based approach has become a popular method to interval uncertainty representation. The likelihood-based uncertainty representation method is capable to deal with any type of interval data. The simplicity of this approach has made it well established method for uncertainty representation. A likelihood-based approach is used in this research to estimate the uncertainty from the stochastic epistemic variables. The simplicity and the versatility of the proposed method could be realized through observing the existing methods.

2.2. Likelihood-based method to uncertainty representation

Likelihood-based methodology has been applied for a probabilistic representation of a stochastic quantity where sparse point data and interval data are available (Sankararaman and Mahadevan 2011). Likelihood function is developed from the probability density function of the sparse point data and the cumulative distribution function of the interval data jointly. A full likelihood function is used to combine the uncertainty of the sparse point and interval data through a joint probability function. Parametric and non-parametric (without considering any particular distribution type) both types of approaches are developed to estimate the properties of the distribution through the full likelihood estimation method.

Two or more coupled component models are usually complex in nature and need to quantify uncertainty through multidisciplinary iterative analysis (Sankararaman and Mahadevan 2012). This methodology is based on computing the probability of satisfying the interdisciplinary compatibility equations, conditioned on specific values of the coupling (or feedback) variables, and this information is used to estimate the probability distributions of the coupling variables. The estimation of the coupling variables is analogous to likelihood-based parameter estimation in statistics and thus leads to the proposed likelihood approach for multidisciplinary analysis.

Likelihood-based methodology can quantify the distribution type uncertainty while fitting probability distributions to sparse and imprecise data (Sankararaman and Mahadevan 2013). In probabilistic representation of uncertainty, it is common to assume a particular type of probability distribution (e.g. normal, lognormal, etc.) while fitting distributions to available data; once this type is chosen, the distribution parameters and the uncertainty in the distribution parameters are estimated. This method analyzes the effect of the choice of the distribution type and quantifies the resulting uncertainty in the probabilistic characterization. Once the distribution type uncertainty in a particular random variable is quantified, the uncertainty in the distribution parameters can also be quantified.

Technique is also developed based on a maximum-likelihood to estimate the state of the pipeline constrained by the underlying physics in the field of fluid mechanics (Modisette 2012). Not only that, advanced modeling and simulation tools used to calibrate a high-fidelity thrust program that predicts steady-state and transient operation of a fighter aircraft turbofan engine. Maximum likelihood parameter estimation is used to estimate engine-model calibration factors (Monaco et. al 2008). The Extended Maximum Likelihood Method (EMLM) is known to make a reliable estimate of the directional wave spectrum from wave measurements at fixed locations in the wave field (Waals et al. 2002). Generally, for accurate and efficient parameter estimation strategy for developing dynamic vehicle models, conventional techniques such as Least Square Estimation (LSE), Maximum Likelihood Estimation (MLE) and Instrumental Variable Methods (IVM) can deliver sufficient estimation results for given models that are linear-in-the-parameter (Xiao and Kulakowski 2003).

Likelihood approach provides flexibility in the simulation which reduces the computational expenses. Krigling model is popular for providing mathematical relationship which can interpolate a set of observations. However, this model is computationally expensive to estimate the model parameters. MLE of kriging model provides the opportunity to estimate the parameters without the computational burden which only requires a continuous numerical optimization (Martin 2007).

The uncertainty representation of the likelihood-based approach can be used for robust and reliability-based design optimization. The likelihood-based approach is popular to estimate the parameters of the distributions. The estimated parameters are used to design robust and reliable systems.

2.3. Robustness-based design optimization

As mentioned in Zaman et al. (2011b), robustness-based design optimization has four key elements: (1) maintaining robustness in the objective function which is called objective robustness; (2) achieve robustness in the constraints which is called feasibility robustness; (3) estimate the mean and the variance of the performance function; and (4) multi-objective optimization. There are various methods of estimating these elements which are necessary for robust design optimization.

Objective Robustness

In robustness-based optimization, objective robustness can be achieved by simultaneously optimizing mean and minimizing variances of the objective function. Two types of robustness measurement processes are popular: one is based on the variance, (Du and Chen 2000; Lee and Park 2001; Doltsinis and Kang 2004) and the other is based on the percentile difference (Du et al. 2004).

Feasibility Robustness

Feasibility robustness can be defined as robustness in the constraints. Robustness in the constraints indicates satisfying the constraints of the design in the presence of uncertainty. Du and Chen (2000) classified the methods of maintaining feasibility robustness into two categories: (i) methods based on stochastic and statistical analysis, e.g., a probabilistic feasibility formulation (Du and Chen 2000; Lee et al. 2008), and a moment matching formulation (Parkinson et al. 1993) and (ii) methods that do not use probabilistic and statistical analysis. Several methods have been developed which are not depending on probabilistic and statistical analysis like worst case analysis (Parkinson et al. 1993), corner space evaluation (Sundaresan et al. 1995) and manufacturing variation patterns (MVP) (Yu and Ishii 1998). The feasible region reduction method (Park et al. 2006) is general in application and normality assumption is not necessary. This is a tolerance design method, where width of the feasible space in each direction is reduced by the amount of $k\sigma$ where k is user defined constant and σ is the standard deviation of the performance function.

Estimating Mean and Variance of the Performance Function

The mean and the standard deviation or variance can be estimated through several methods. There are some existing methods for estimating the mean and standard deviation of performance function which can be classified into three major categories: (1) Taylor series expansion methods, (2) Sampling based methods, and (3) Point estimate methods (Huang and Du 2007). The Taylor series expansion method (Halder and Mahadevan 2000; Du and Chen 2000; Lee and Park 2001) is quite appreciable to measure the mean and variance for nonlinear performance function. However, the approximation may result in huge error if the variances of the random variables are large (Du et al. 2004). Information on distribution of the random variables is required to estimate mean and variance by sampling based methods which has made it expensive. The effective and efficient sampling techniques can be listed as importance sampling, Latin hypercube sampling (Robert and Casella 2004) and surrogate models (Ghanem and Spanos 1991; Bichon et al. 2008; Cheng and Sandu 2009). Point estimate method (Rosenblueth 1975) can be used to ease the computation of the derivatives required in the Taylor series expansion. Several types of point estimate methods (Hong 1998; Zhao and Ono 2000; Zhao and Ang 2003) can be applied to estimate mean and variance. Dimension reduction method (DRM) (Rahman and Xu 2004; Xu and Rahman 2004; Lee et al. 2008) is developed recently which overcomes the troubles associated with Taylor series expansion and sampling method.

Multi-objective Optimization

Multi-objective optimization can be achieved by optimizing the mean of the objective function and minimizing its variation (Marler and Arora 2004). Weighted sum approach has widespread application in multi-objective optimization for robust design problems (Lee and park 2001; Doltsinis and Kang 2004; Zou and Mahadevan 2006). Other optimization methods called ε-constrained method (Mavrotas 2009), goal programming (Zou and Mahadevan 2006), compromise decision support problem (Bras and Mistree 1993, 1995; Chen et al. 1996), compromise programming (CP) (Zeleny 1973; Zhang 2003; Chen et al. 1999) and physical programming (Messac 1996; Messac et al. 2001; Messac and Ismail-Yahaya 2002; Chen et al. 2000) can be enlisted. There are different advantages and shortcomings of all those methods.

2.4. Reliability-based design optimization

Reliability-based design optimization (RBDO) uses numerical optimization algorithms which help to obtain optimal design with reliability (Agarwal et. al 2004b). The reliability optimization usually performed to ensure a safety limit or target reliability for deterministic data. However, without considering the uncertainty may lead to system failure. Therefore, it is necessary to include uncertainty in the design constraints for having a reliable design. The uncertainties can be modeled and represented using the probability theory like likelihood-based approach. In the reliability optimization, it may require to optimize a single or multi-objective function while satisfying the reliability constraints. The reliability constraints consider the probability of failure (P_f) which is related to the failure mood of the system or design. Different simulation methods with high computational ability like Monte Carlo Simulation (MCS), importance sampling etc. can be used to get a reliable solution. In the stochastic design, it is necessary to estimate the uncertainty of the random variables. Likelihood-based approach can have contribution to represent the uncertainty. After estimating the uncertainty, the designer will be able to go for a reliability-based design optimization which may provide an optimal and reliable solution of the problem.

In this thesis, formulation of robustness-based design optimization through likelihood estimation method is proposed for interval data uncertainty. Uncertainty representation and propagation are done for the interval data. This robust design considers both aleatory and epistemic uncertainty through probabilistic representation. The proposed mathematical models are developed by considering multiple interval data of the random variables using normal distribution and bounded Johnson distribution. However, the probabilistic representation is analyzed through the maximum likelihood estimation (MLE) method which helps to find out the moments and the parameters of the distributions. A closed form of the bounded Johnson distribution is proposed to estimate the mean and the standard deviation from the parameters of the bounded Johnson distribution calculated through MLE method. Nonlinear functions create complexity during calculating the mean and the variance. First order Taylor series expansion method is used to estimate the mean and variance of the performance function. Numerical examples are used to show the effectiveness of the formulated methods. One general mathematical problem is given where the coefficients of the objective function are treated as epistemic non-design variables and robust design optimization is performed for both MLE-based approaches for normal and bounded Johnson distribution. Another real engineering problem of a Two-Stage-To-Orbit (TSTO) vehicle is given with six input random variables where some are design variables including epistemic non-design variables with interval data and robust design optimization is done using both uncertainty analysis models.

The performance of a robust design can be measured by the characteristics of the mean and variation of the objective or performance function (Zaman et al. 2011b). In the proposed formulation, we obtain the optimal value of the mean of the performance or objective function (e.g. weight) and simultaneously minimizing its variation or standard deviation. The formulation will provide an optimal target value of the objective function satisfying both the objective function and design constraints which will lead to feasibility robustness.

As mentioned earlier, there exist a few methods for robust and reliability-based design optimization that can handle epistemic uncertainty (e.g., Zaman (2010) and Zaman et al. 2011b). The existing methods either solve a nested optimization formulation or a decoupled approach of optimization, which are computationally expensive. In the current research, a formulation is proposed to completely separate the epistemic analysis from the design optimization framework to achieve computational efficiency.

The specific objectives of this research are the development of formulations and algorithms for uncertainty representation and propagation through robustness-based design optimization under epistemic uncertainty. Therefore, the proposed research develops and demonstrates generalized methodologies and tools for managing uncertainty in engineering systems, which will provide decision support to engineers for robust and reliable design of engineering systems.

This thesis is organized as follows. In chapter III, a nested optimization-based methodology is developed to estimate the distribution parameters under interval uncertainty using maximum likelihood-based approach. The proposed methodology illustrates for normal and bounded Johnson distributions. Chapter IV illustrates the formulation of the nonlinear robustness-based design optimization model for both MLE approach. In chapter V, the opportunities and the applications of the reliability-based design optimization are described. Chapter VI provides the application of the proposed methods with illustration of two examples including a mathematical problem and an engineering problem. Chapter VII describes the summary of the research and recommendation for future research needs.

CHAPTER III

LIKELIHOOD-BASED APPROACH TO EPISTEMIC UNCERTAINTY REPRESENTATION

Likelihood-based approach has been applied particularly for single point data in many researches. In this chapter, a log maximum likelihood-based approach is proposed for variable *x* consists of multiple interval data. Maximum likelihood estimation (MLE) constructs an estimator to estimate the unknown distribution parameters (*P*). The „likelihood" term first introduced by an English mathematical statistician named R. A. Fisher in 1921. The likelihood for a parameter „*P*" is a quantity proportional to the probability of the parameter *P* quantified for a specific population where variable *x* consists the observed data as a sample (Fisher 1921). When there is less opportunity to show our confidence in making decisions, intuitions are used to make inferences which helps to produce the preference list from different results where the populations do not directly follow the rules of probability. R. A. Fisher used this term „likelihood" to present the inference quantity and distinguish it from the probability (Fisher 1925).

According to Edward (1972), Likelihood contains two adjacent hypotheses to estimate a parameter from the observed data. However, all kinds of information regarding the hypotheses are available in the likelihood function to estimate the parameters from the observed data using inferences or intuitions.

The observed data are used as the sample space for likelihood estimation. How the sample space is produced is irrelevant in likelihood estimation. The sample space generally does not depend on the parameter rather the parameter solely depends on the sample space. As mentioned before, evidential significance of the sample space or experimental information about the data *x* to influence the parameter *P* is contained in the likelihood function. In the inference of the parameter *P* for the observed data *x*, all the information are available in the likelihood method. If two likelihood functions are proportional to each other, the likelihood functions contain the same experimental information regarding the parameter.

Likelihood estimation is applied to estimate the parameter *P* from different distributions, e.g., normal, lognormal, bounded Johnson distributions etc. If the sample size is big in quantity, then likelihood is a good estimator of the parameters. The likelihood function for a specific distribution not only depends on the parameters as well as on the variables. If the variable x is a continuous variable, then $f(x \mid \mathbf{P})$ is called probability density function (PDF). However, if x is discrete random variable, then $f(x \mid \mathbf{P})$is called point mass function. If the random variables are independent, then the joint density function can be written as (Halder and Mahadevan 2000):

$$f(x_i \mid \mathbf{P}) = f(x_1, x_2, \ldots \ldots x_n \mid \mathbf{P}) = f(x_1 \mid \mathbf{P}) f(x_2 \mid \mathbf{P}) \ldots\ldots f(x_n \mid \mathbf{P}) \tag{1}$$

where, $i = 1,2,\ldots n$. This is called the likelihood function denoted by L (***P***). In the likelihood estimation, the value of parameter P is selected in such a way so that the estimation of the sample space $x_1, x_2, \ldots x_n$ follows the actually observed value. For a specific value of $P = P^*$, there is high probability that the function f (x_i |**P**) will produce the actual observed value. Therefore, it is reasonable to estimate the value of P in such a way so that it is P^*.

For a specific continuous distribution, the parameter P is estimated when the likelihood function or the probability density function $f(x_i \mid \mathbf{P})$ is very high and this probability value becomes an estimator of P. Therefore, it is very reasonable to estimate the value of parameter P for which the value of the likelihood function will be maximum for the observed data. This is the well known maximum likelihood estimation (MLE). In the MLE, the chosen value of parameter P provides the maximum value of the likelihood function. If there is no actual data for the random variable x_i (where, $i=1\ldots n$) to estimate the likelihood function, then it is called „maximum likelihood estimate". When there are specific numerical values for the random variable x_i (where, $i=1\ldots n$), then it is called „maximum likelihood estimator".

The MLE requires to maximize the likelihood function L(***P***) with respect to the parameter P. However, the likelihood function L(***P***) may be complex in nature for maximization. Hence, natural logarithm of the likelihood function is taken as support (Jeffreys 1934) which can be denoted as *log*(L(***P***)). Due to logarithm being a monotonic increasing function, the maximization of L(***P***) is equal to the maximization of *log*(L(***P***)). The log likelihood function can be expressed as:

$$\log(L(\mathbf{P})) = \log \prod_{i=1}^{n} f(x_i \mid \mathbf{P}) = \sum_{i=1}^{n} \log f(x_i \mid \mathbf{P}) \tag{2}$$

In the paper of Sankararaman and Mahadevan (2011), a probabilistic representation is developed to estimate the parameters from the distributions including single point data and interval data. Sankararaman and Mahadevan (2011) used a full likelihood estimate where the entire likelihood function constructs the PDF of the distribution parameter *P*. Edwards (1972) suggests that the likelihood method is valid up to a proportional constant which allows using a proportional sign to likelihood representation. The probability value for a discrete or specific point is zero, if $f(x \mid \mathbf{P})$ is a continuous density function (Pawitan 2001).

If the data are observed from different experiments or different expert opinions are the sources of the data, then the data can be treated as independent following Eqn. (1). Hence, the joint likelihood function for the parameter *P* of random variable x_i is expressed as:

$$L(\mathbf{P}) \propto \prod_{i=1}^{n} f(x_i \mid \mathbf{P}) \tag{3}$$

Therefore, the parameter *P* can be estimated through maximizing the equation which is popular as Maximum Likelihood Estimation (MLE).

In the following derivation, the likelihood considers interval around the data point x_i. Hence, this derivation can be applied for any general interval [*lb, ub*] and the expression for likelihood function of the parameter *P* is expressed as:

$$\begin{aligned} L(\mathbf{P}) \quad & \propto prob\ (observed\ data \mid \mathbf{P}) \\ & \propto prob\ (x \in [lb, ub] \mid \mathbf{P}) \\ & \propto prob\ (lb \le x \le ub \mid \mathbf{P}) \end{aligned} \tag{4}$$

There can be different types of multiple interval data. For multiple intervals including overlapping and non-overlapping data (e.g., $x_i \in [lb_i, ub_i]$ where, *n = 1* to *n*), the likelihood function can be written as:

$$L(\mathbf{P}) \propto \prod_{i=1}^{n} prob\ (lb_i \le x_i \le ub_i \mid \mathbf{P}) \tag{5}$$

3.1. Maximum likelihood approach for normal distribution

In this research, likelihood function is developed with an illustration of normal distribution from which the parameters can be estimated using the observed interval data.

Likelihood function becomes complex in nature for several variables. Application of the natural logarithm to the likelihood function will make it simple for calculation. The parameters for the normal distribution are estimated by maximizing the log-likelihood function. For the maximum value of the log-likelihood function, the partial differentiation with respect to the parameters becomes zero which helps to estimate the parameters using point data. However, the random variables consist of interval data. Non-linear optimization algorithm is used to estimate the parameters from the maximum log-likelihood function in this research. Hence, the parameters and the moments are identical for the normal distibution. Therefore, maximum log-likelihood estimation is done through optimization of the log-likelihood function which calculate the parameters or moments (e.g. mean and standard deviation) of the normal distribution.

The likelihood function for n observations of random variable x_i for normal distribution is:

$$L(\mu,\sigma)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \tag{6}$$

Therefore, to make the calculation simple, the natural logarithm of the likelihood function is taken which provides the log-likelihood function:

$$\begin{aligned}\log\left(L(\mu,\sigma)\right)&=\sum_{i=1}^{n}\log\left(\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}\right)\\&=-n\log\sqrt{2\pi}-n\log\sigma-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\end{aligned} \tag{7}$$

From this Eqn. (7), using multiple interval data, the parameter mean and standard deviation of the random variables can be estimated through non-linear optimization.

3.2. Maximum likelihood approach for bounded Johnson distribution

According to the formula of likelihood function, the likelihood function for n observations of random variable x_i for bounded Johnson distribution is:

$$L(\delta,\gamma,\xi,\lambda)=\prod_{i=1}^{n}\frac{\delta}{\sqrt{2\pi}}\times\frac{1}{\left[\left(\frac{x_i-\xi}{\lambda}\right)\left\{1-\left(\frac{x_i-\xi}{\lambda}\right)\right\}\right]}\times\exp\left\{-\frac{1}{2}\left[\gamma+\delta\log\left(\frac{x_i-\xi}{\lambda-x_i+\xi}\right)\right]^2\right\} \tag{8}$$

Taking natural logarithm of the likelihood function will make it simpler:

$$\log\left(L(\delta,\gamma,\xi,\lambda)\right)=n\log\delta-n\log\sqrt{2\pi}-\sum_{i=1}^{n}\log\left(\frac{x_i-\xi}{\lambda}\right)-\sum_{i=1}^{n}\log\left(\frac{\lambda-x_i+\xi}{\lambda}\right)-\frac{1}{2}\sum_{i=1}^{n}\left[\gamma+\delta\log\left(\frac{x_i-\xi}{\lambda-x_i+\xi}\right)\right]^2 \qquad (9)$$

From Eqn. (9), using multiple interval data, the parameters (δ, γ, ξ and λ) can be calculated through optimization. According to the condition of bounded Johnson distribution, the values of δ and λ must be greater than zero. The value of ξ is assumed as the lowest value of the multiple intervals for the input random variables where λ is the difference of the lowest and highest bounds of the multiple intervals for each individual random variable. However, maximum log likelihood estimation method will be used to estimate the parameters from Eqn. (9).

3.3. Proposed approach for likelihood-based estimation with interval data

Nested optimization problem is designed and applied for maximum log likelihood estimation. The variables of the likelihood-based function are consisted of multiple interval data and the parameters of the distributions are estimated from the likelihood function by applying a nested optimization.

The formulation of nested optimization for the maximum likelihood function:

$$\begin{aligned}&\min_{x}\left(\max_{p}\left(f(x\,|\,\mathbf{P})=\log\left(L(x\,|\,\mathbf{P})\right)\right)\right)\\&s.t.\quad lb_i\le x_i\le ub_i\quad for\; i=1,2,\ldots,n\end{aligned} \qquad (10)$$

However, using the Eqns. (7) and (9) in the nested optimization model formulated in Eqn. (10), the parameters of the normal and bounded Johnson distribution are estimated, respectively. The mean (μ) and the standard deviation (σ) are the parameters for the normal distribution. The parameters (δ, γ, ξ and λ) of the bounded Johnson distribution are also estimated from the multiple interval data using the maximum log likelihood estimation method through the proposed nested optimization.

Zaman et al. (2011a) formulated a method proposing nonlinear programming for calculating the moments from the different combination of multiple interval data rather than calculating the discrete point data. Nonlinear programming is used with minimization and maximization of the objective functions to calculate the bounds of the moments for their respective multiple interval data. Overlapping and non-overlapping both types of interval data are treated with this method. The bounds completely enclosed all the possible moments that can be generated from various combinations of interval data which support the rigorous characteristics of the method. Methods of calculating moment bounds from multiple interval data are shown in Table 2.1.

Table 2.1: Methods of calculating moment bounds from multiple interval data

Moment	Formula
1	$$[\underline{M},\overline{M}]=\left[\frac{1}{n}\sum_{i}^{n} lb_i,\frac{1}{n}\sum_{i}^{n} ub_i\right]$$ where, $lb_i \le x_i \le ub_i$, $i = \{1,2,\ldots,n\}$ and $[\underline{M},\overline{M}]$ are the lower and upper bounds on the mean, respectively.
2	$$\min_{x_1,\ldots,x_n}/\max M_2=\frac{1}{n}\left(\sum_{i=1}^{n} x_i^{\,2}\right)-\frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2$$ for $lb_i \le x_i \le ub_i$, $i = \{1,2,\ldots,n\}$
3 and 4	$$\min_{x_1,\ldots,x_n}/\max M_k=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\frac{1}{n}\sum_{j=1}^{n} x_j\right)^k$$ where, $lb_i \le x_i \le ub_i$, $i = \{1,2,\ldots,n\}$ and minimizing or maximizing for $k = 3$ provides the lower or upper bound on the third moment and for $k = 4$ yields the or upper bound on the fourth moment.

The bounds of the moments help to verify the parameters estimated from the maximum likelihood estimation method. The empirical cumulative density function (CDF) for the parameters of the normal density function and the bounded Johnson density function should be remained or lied within the bounds of empirical cumulative density function (CDF) constructed from the bounds of the parameters of normal distribution and from the bounds of the parameters of bounded Johnson density function, respectively. The bounds of the parameters of bounded Johnson density function are estimated through the moment matching method. The validation of the parameters calculated from the maximum likelihood estimation depends on remaining the empirical CDF of the parameters in the bounds of empirical CDFs produced from the moment bounds.

CHAPTER IV

ROBSUTNESS-BASED DESIGN OPTIMIZATION UNDER EPISTEMIC UNCERTAINTY

4.1. Existing methods for robustness-based design optimization

The variables in the designs and processes are considered as fixed in the deterministic optimization formulation without considering any stochastic characteristics or data uncertainty in the variables. The deterministic optimization formulation can be written as:

$$\begin{aligned} &\min_{x} \quad f(x) \\ &s.t. \quad LB \le g_i(x) \le UB \quad for\ all\ i \\ &\qquad lb \le x \le ub \end{aligned} \tag{11}$$

where, $f(x)$ is the objective function, x is expressed as a vector for design variables, $g_i(x)$ is the ith constraint, LB and UB are the vectors of lower and upper bound of the constraints $g_i(x)$ and lb and ub are the vectors of lower and upper bounds of the design variables.

However, in the real situation, the input variables are normally uncertain and solution obtained from the deterministic design optimization could be sensitive to the variations of the input variables. Uncertainty analysis could be an important issue in the robustness-based design optimization. The deterministic design optimization could be used to get an optimal point which might be applied as an initial guess in the robustness-based optimization problem.

In practice, the robust design optimization is complex in nature which requires nonlinear optimization. In the robust design optimization method, objective robustness is achieved by measuring the variation in the objective function through the variance or standard deviation. Feasibility robustness is achieved by feasible region reduction method. First order Taylor series expansion is used to estimate the mean and variance of the objective function. Weighted sum method is used to trade off the multiple objectives in the performance function of the robust design optimization. The formulation of robustness-based design optimization considering only aleatory uncertainty is (Zaman et al. 2011b):

$$
\begin{aligned}
&\min_{x} \quad f(\mu,\sigma) = w\times\mu_f + v\times\sigma_f \\
&s.t. \quad LB + k\sigma(g_i(d,z)) \le E(g_i(d,z)) \le UB - k\sigma(g_i(d,z)) \quad for\ all\ i \\
&\qquad lb_i + k\sigma(x_i) \le d_i \le ub_i - k\sigma(x_i) \quad for\ i = 1,2,\ldots,nrdv \\
&\qquad lb_i \le d_i \le ub_i \qquad\qquad\qquad for\ i = 1,2,\ldots,nddv
\end{aligned} \tag{12}
$$

where, μ_f and σ_f are the mean value and standard deviation of the objective function, respectively; *d* is the vector deterministic variable which can be the mean values of the uncertain variables *x*; *nrdv* is the number of random design variables and *nddv* is the number of deterministic design variables; *z* is the vector of non-design input random variables. $w \ge 0$ and $v \ge 0$ are the weighting coefficients that represent the relative importance of the objectives μ_f and σ_f is the *i*th constraint*;* $E(g_i(d, z))$ is the mean and $\sigma(g_i(d, z))$ is the standard deviation of the *i*th constraint. *LB* and *UB* are the vectors of lower and upper bounds of the constraints g_i's; *lb* and *ub* are the vectors of lower and upper bounds of the design variables. ζ(*x*) is the vector of standard deviations of the random variables and *k* is some constant. *k* is used here to adjust the robustness of the method against the level of conservatism of the solution. Considering the variations in the design variables, *k* reduces the feasible region and is related to the probability of constraint satisfaction.

The design methodology includes complexity in the robust design due to uncertainty. Therefore, the design variables *d* and the input random variables *z* in (12) have a high chance for having epistemic uncertainty in the form of multiple interval data. Normally, the designer has no control on the non-design epistemic variables *z*. Therefore, the optimization method has to employ a search among the possible values of the epistemic variables in order to find an optimal solution. The design becomes conservative due to this approach. The formulation of the robustness-based design due to the epistemic data uncertainty can be written as:

$$
\begin{aligned}
&\min_{d}\left(\max_{\mu_z} f(\mu,\sigma)\right) = \left(w\times\mu_f + v\times\sigma_f\right) \\
&s.t. \quad LB + k\sigma(g_i(d,\mu_z)) \le E(g_i(d,z)) \le UB - k\sigma(g_i(d,\mu_z)) \quad for\ all\ i \\
&\qquad lb + k\sigma(x) \le d \le ub - k\sigma(x) \\
&\qquad Z_l \le \mu_z \le Z_u
\end{aligned} \tag{13}
$$

where, Z_l and Z_u are the vectors of lower and upper bounds of the decision variables μ_z of the inner loop optimization problem (Zaman et al. 2011 b). Here, *d* can be expressed as stochastic design variables as well as epistemic design variables. The outer loop optimization is the design optimization problem where a robust design optimization may be carried out for a fixed set of non-design epistemic variables. The inner loop optimization is used to analyze the non-design epistemic variables. The inner loop optimizer searches the possible values of non-design epistemic variables to calculate the upper bound of the objective function value.

This nested formulation in Eqn. (13) is a very expensive formulation and does not provide any guarantee to converge. This nested optimization analyzes the aleatory uncertainty of the design variables with every epistemic analysis which makes it costly. Therefore, another design optimization algorithm is developed which is un-nested from the epistemic analysis with the computational accuracy.

The optimization problem is decoupled and expressed (Zaman et al. 2011b) as:

$$\begin{aligned} d^* &= \arg\min_{d}\left(w\times\mu_f\left(d,\mu_z^*\right)+v\times\sigma_f\left(d,\mu_z^*\right)\right) \\ s.t.\quad & LB+k\sigma\left(g_i\left(d,\mu_z^*\right)\right)\le E\left(g_i(d,z)\right)\le UB-k\sigma\left(g_i\left(d,\mu_z^*\right)\right)\quad for\ all\ i \\ & lb+k\sigma(x)\le d\le ub-k\sigma(x) \end{aligned} \tag{14}$$

$$\begin{aligned} \mu_z^* &= \arg\max_{\mu_z}\left(w\times\mu_f\left(d^*,\mu_f\right)+v\times\sigma_f\left(d^*,\mu_f\right)\right) \\ s.t.\quad & LB+k\sigma\left(g_i\left(d^*,\mu_z\right)\right)\le E\left(g_i(d,z)\right)\le UB-k\sigma\left(g_i\left(d^*,\mu_z\right)\right)\quad for\ all\ i \\ & Z_l\le\mu_z\le Z_u \end{aligned} \tag{15}$$

The optimization problem in Eqns. (14) and (15) are solved iteratively until convergence. In Eqn. (15), the optimization is continued by all non-design epistemic variables where in Eqn. (14), the optimization is done for only design variables. In the optimization, d^* are the fixed quantities in the optimization in Eqn. (15) and μ_z^* are the fixed quantities in the optimization in Eqn. (14).

The existing methods for robustness-based design optimization solve a nested formulation as shown in Eqn. (13), where the optimizers searches among the possible values of epistemic variables to find a robust design. The decoupled approach given in Eqns. (14) and (15) can achieve computational efficiency by un-nesting the design analysis from the epistemic analysis. However, this is an iterative approach. The formulations in Eqns. (14) and (15) are solved iteratively until convergence to obtain the robust design in the presence of data uncertainty. We can achieve further computational efficiency if the uncertainty analysis for the epistemic variable is carried out outside the design optimization framework. In the following subsection, we propose an efficient approach for robustness-based design optimization, where epistemic analysis is based on the likelihood-based approach described in chapter 2.

4.2. Proposed MLE-based Robust design optimization with interval data

In this research, the uncertainty analysis of the epistemic variables is done outside the optimization framework which helps to increase the computational efficiency and reduce cost. The uncertainty is represented by likelihood-based approach mentioned in chapter 2 and used in the robust optimization model as fixed values. Therefore, it is not required to search in the epistemic variables for uncertainty representation with respect to design variables which actually makes it costly in the nested design optimization. Neither any convergence is needed for the uncertainty representation as mentioned in the decoupled approach.

The proposed formulation of nonlinear MLE-based robust design optimization can be expressed as:

$$
\begin{aligned}
& d^{*} = \arg\min_{d} \left(w \times \mu_f \left(d, \mu_z^{*}, \sigma_z^{*} \right) + v \times \sigma_f \left(d, \mu_z^{*}, \sigma_z^{*} \right) \right) \\
& s.t. \quad LB + k\sigma\left(g_i \left(d, \mu_z^{*}, \sigma_z^{*} \right) \right) \leq E\left(g_i (d, z) \right) \leq UB - k\sigma\left(g_i \left(d, \mu_z^{*}, \sigma_z^{*} \right) \right) \\
& \qquad lb_i + k\sigma(x_i) \leq d_i \leq ub_i - k\sigma(x_i) \quad for\ all\ i
\end{aligned}
\tag{16}
$$

In the Eqn. (16), the standard deviation (σ_d) or variance of the design variables and the mean (μ_z) and standard deviation (σ_z) of the non-design epistemic variables are provided as fixed values for robust design optimization. The estimated uncertainty of the non-design variables can be provided through the likelihood-based method. The uncertainty representation method is proposed with the illustrations of normal and bounded Johnson distribution in chapter 2. In the decoupled approach, Eqn. (15) is supplying the mean values of epistemic variables in the Eqn. (14) until convergence. Therefore, the Eqn. (16) is equivalent to Eqn. (14). The proposed MLE-based approach estimated the parameters which can be used to provide the mean (μ_z) and standard deviation (σ_z) of the non-design epistemic variables. This moment estimation method based on MLE converts the decoupled approach into a robust design optimization formulation as shown in Eqn. (16).

The mean (μ) and standard deviation (σ) of the performance function in (16) can be obtained by first order Taylor series expansion method. If there is a response variable Y

which is represented by a non-linear function f consists a set of random variables $(x_1, x_2, \ldots, x_n)$ then the response variable can be represented as,

$$Y = f(x_1, x_2, \ldots, x_n) \tag{17}$$

If the mean and variance of each x_i are known but the distribution is unknown, the approximate and variance can be estimated by expanding the function $f(x_1, x_2, \ldots\ldots, x_n)$ in a Taylor series about the mean values $(\mu_{x_1}, \mu_{x_2}, \ldots\ldots, \mu_{x_n})$ (Halder and Mahadevan 2000).

$$Y = f(\mu_{x_1}, \mu_{x_2}, \ldots\ldots, \mu_{x_n}) + \sum_{i=1}^{n} (x_i - \mu_{x_i}) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \mu_{x_i})(x_j - \mu_{x_j}) \frac{\partial^2 f}{\partial x_i \partial x_j} + \ldots \tag{18}$$

However, the derivatives are evaluated at the mean values of the x_i*'s*.

So, the first-order approximate mean of Y:

$$E(Y) = f(\mu_{x_1}, \mu_{x_2}, \ldots\ldots, \mu_{x_n}) \tag{19}$$

The first-order variance of Y can be written as:

$$Var(Y) \approx \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 Var(x_i) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} Cov(x_i, x_j) \tag{20}$$

If x_i"s are uncorrelated, then the equation becomes:

$$Var(Y) \approx \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 Var(x_i) \tag{21}$$

Using Eqns. (19) and (21) based on approximation method, the moments of the performance function can be estimated. These procedures are followed to estimate the mean and the standard deviation for the robust design optimization.

The uncertainty representation by likelihood approach has made the robust design optimization simple. The MLE method is illustrated for normal and bounded Johnson distribution. The moments and the parameters are identical for normal distribution. Hence, the mean (μ_z) and standard deviation (σ_z) of the non-design epistemic variables for the normal distribution can be directly calculated from the log likelihood estimation. However, the moments from the parameters of the bounded Johnson distribution are estimated through a new proposed method.

4.3. Estimation of moments from the bounded Johnson distribution parameters

There are extensive researches to estimate the moments of the bounded Johnson distribution (mean and standard deviation). J. Draper (1952) suggested a method to estimate the moment of the transformed variable *y* which helps to get a generalized standard form to estimate the moments of input epistemic variable *x* for the bounded Johnson distribution function. The general standard form of the bounded Johnson distribution is:

$$Z = \gamma + \delta \log\left\{\frac{x-\xi}{\lambda+\xi-x}\right\}, \quad \xi < x < \xi + \lambda \tag{22}$$

There can be a substitution in (22) as $y = \frac{x-\xi}{\lambda}$,

$$Z = \gamma + \delta \log\left\{\frac{y}{1-y}\right\} \tag{23}$$

where, *Z* is the unit normal variable. The „*r*"th moment of *y* about zero is:

$$\mu_r(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} \left(1 + e^{-\frac{(z-\gamma)}{\delta}}\right)^{-r} dz \tag{24}$$

For $r = 1$, the first central moment can be estimated directly from the equation through integration which will provide the expected value of „*y*" as:

$$\mu_1(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} \left(1 + e^{-\frac{(z-\gamma)}{\delta}}\right)^{-1} dz \tag{25}$$

For $r = 2$, the second central moment can be estimated through integration as:

$$\mu_2(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} \left(1 + e^{-\frac{(z-\gamma)}{\delta}}\right)^{-2} dz \tag{26}$$

Therefore, the mean of the input random variables *x* can be estimated as:

$$\mu(x) = \xi + \lambda\mu_1(y) \tag{27}$$

If, y is a random variable, then from the general statistical formula of variance,

$$\sigma^2 = E[(y - E(y))^2] = E[y^2] - (E[y])^2$$

Therefore, the standard deviation,

$$\sigma = \sqrt{E[y^2] - (E[y])^2} \tag{28}$$

Hence, the standard deviation of the bounded Johnson distribution from the central moments can be estimated as:

$$\sigma(x) = \lambda \sigma(y) \tag{29}$$

where, $\sigma(y) = \sqrt{\mu_2(y) - (\mu_1(y))^2}$

After calculating the parameters of the bounded Johnson distribution through maximum log likelihood estimation, the mean (μ_z) and standard deviation (σ_z) of the non-design epistemic variables are estimated from Eqns. (27) and (29), respectively. Monte Carlo Simulation (MCS) method is used to verify the mean and the standard deviation. For that, the cumulative density function of the bounded Johnson distribution is expressed for input random variable or for the coefficients x of a function. Ten thousand standard normal random numbers are generated in MATLAB for Monte Carlo Simulation (MCS) technique which produce ten thousand values of each input variable or coefficients. The moments (mean and standard deviation) are calculated then from this enormous number of input variables.

The bounded system of Johnson distribution in Eqn. (22) can be converted to:

$$x = \xi + \lambda \frac{1}{1 + e^{-\frac{Z-\gamma}{\delta}}}, \quad \xi < x < \xi + \lambda \tag{30}$$

Using Monte Carlo Simulation in Eqn. (30), the mean and the standard deviation are estimated for the parameters of the bounded Johnson distribution which helps to prove that the proposed method to estimate the mean (μ_z) and standard deviation (σ_z) of the epistemic variables from Eqns. (27) and (29), respectively is valid.

The Design Structure Matrixes (DSM) are drawn to understand the robustness-based design optimization method easily. In Fig. 3.1, the MLE-based approach for robustness-based design optimization for normal distribution is presented. The parameters of the distributions can be estimated using the maximum likelihood estimation method from the non-design epistemic variables in the form of interval data. The parameters and the moments are same for normal distribution. However, the moments are used in the robust design optimization which will provide the estimated mean values of the design variables for the standard deviations which consist deterministic uncertainty. This is a robust design optimization framework to optimize the mean and minimize the standard deviation of a performance function.

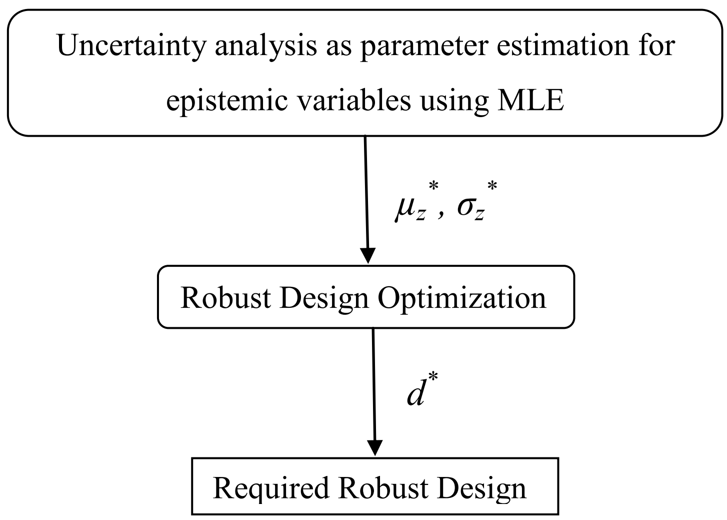

Fig. 3.1 MLE-based approach for robustness-based design optimization considering normal distribution

Fig. 3.2 illlustrates the DSM of the MLE-based approach for robustness-based design optimization for bounded Johnson distribution. The four parameters δ, γ, ξ and λ) are estimated form the likelihood-based approach. A method is developed to estimate the moments (mean and standard deviation) of the bounded Johnson distribution is from the estimated parameters. The estimated moments of the non-design epistemic variables are used in the robust design optimization to calculate the mean values of the design variables with standard deviations consists of fixed uncertainty. This robust optimization will optimize the performance function and minimize the standard deviation or variance. This single loop robust design optimization can be used to optimize the mean and minimize the variation.

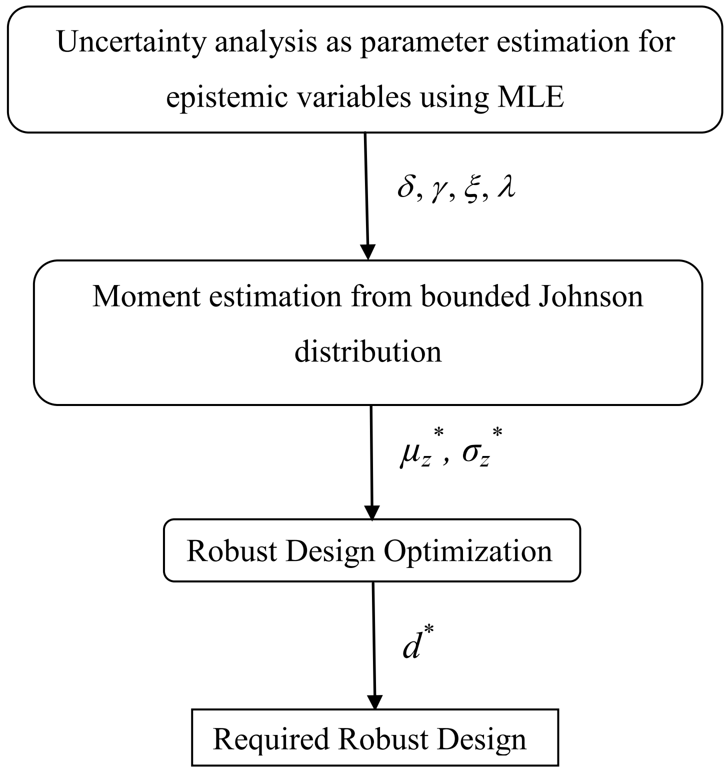

Fig. 3.2 MLE-based approach for robustness-based design optimization considering bounded Johnson distribution

CHAPTER V

RELIABILITY-BASED DESIGN OPTIMIZATION UNDER EPISTEMIC UNCERTAINTY

Reliability-based design optimization (RBDO) has become a well-established method to ensure the quality and safety in the automotive, aerospace, various consumer products and especially in the structural designs and applications. It is used to optimize the quality in the product with minimizing the cost. The quality of a product is judged by the characteristics of withstand the failure. Reliability-based design optimization is such a method which minimizes the cost considering the design constraints for the probability of failure. However, the reliability optimization requires a huge computational effort which may become expensive. The expensive computational effort for the complex engineering problem may not allow it become an efficient tool. At the same time, there may be single or multiple objectives in reliability-based design optimization. Due to the complexity in the design constraints, the objective function becomes non-linear. In the non-linear multi-objective reliability-based design optimization, the goal is to establish the best tradeoff between the objectives. The summation of the amount of reliability and the quantity of the probability of failure is expressed as unity. Hence, it is mandatory to estimate the amount of the probability of failure to estimate the reliability. The estimation of the probability of failure (P_f) for large number of random variables is difficult and impossible in some cases. First Order Reliability Method (FORM), Second Order Reliability Method (SORM), Monte Carlo Simulation (MCS), Importance Sampling etc. are used to estimate the uncertainty from the complex engineering models consist of many random variables. There is huge influence of the optimization algorithms for reliability analysis in the complex engineering designs (Rackwitz 2001).

For a reliable design, the resistance or the capacity must be greater than the load or demand. For RBDO, there must be a relationship or performance function consists of both the capacity and load. For a deterministic design, there are no uncertainty lies in the resistance or the load. However, in a stochastic design, it is natural to have uncertainty in the variables used in expressing the objectives in the performance function. Hence, it is required to estimate the uncertainty of the random variables. The proposed likelihood-based approach can be used to represent the epistemic uncertainty from the interval data for different probability distributions. The estimated parameters of different distributions using likelihood approach can be used for the reliability-based design optimization.

In the worldwide competitive market, it is important to design a product which will be optimal in quality and reliability. The manufacturing cost of the product is trying to reduce simultaneously with achieving performance and quality which requires reliability-based design optimization. Different types of optimization frameworks have become popular for efficient reliability-based design optimization. In the simulation based environment, there are three generalized methods which are well established for the reliability-based design. These methods are nested optimization, decoupled approach and single loop optimization for reliability-based design optimization.

Nested optimization is the well known traditional method for reliability optimization. Nested optimization is also known as the double loop method. However, nested optimization is computationally expensive to execute. The decoupled reliability-based design optimization is popular for reducing the computational cost less than the nested optimization. This is also popular as the sequential reliability design optimization method. In sequential optimization, a deterministic and a reliability analysis are decoupled for convergence (Chen et al. 1997; Wang and Kodiyalam 2002). The deterministic optimization is updated based on information from the reliability analysis and the updated design is used for the next cycle. The decoupled process is continued until the convergence is obtained. However, this method may not be convergent and the optimal design may possess errors. Different methodologies are developed based on the sequential or decoupled method. Chen et al. (1997) proposed a sequential RBDO method for normally distributed random variables. Wang and Kodiyalam (2002) developed a more advanced model for non-normal random variables with huge computational expenses. Chen and Du (2002) proposed a method called SORA (Sequential Optimization and Reliability Assessment).

Unilevel reliability optimization has similarity to the nested optimization but cost effective. Kuschel and Rackwitz (2000) developed unilevel formulation where the FORM problem is replaced by corresponding first order Karush-Kuhn-Tucker (KKT) optimality conditions of the first order reliability problem. The unilevel method developed by Agarwal et al. (2004b) can reduce the computational cost around fifty percent with comparison to the nested optimization. It can deal with the equality constraints through continuation method.

This method follows the results produced from the nested optimization which proved its validity. It avoids the singularities associated with zero probability of failure.

In the reliability-based design optimization, there is tradeoff between getting higher reliability and lowering the cost. There must be achieved two important characteristics in reliability-based design optimization which are efficiency and robustness. The computation method should be efficient with respect to time and expense while reliability optimization must achieve robustness in the result without having any local optimal solution. The reliability optimization requires characterizing the uncertain variables and all kinds of failure moods to estimate the most probable point (MPP) of failure. The uncertainty can be represented by probability theory, Dempster-Shafer theory, convex models, possibility or fuzzy set theory and many others. The probability theory requires sufficient data for uncertainty representation. The uncertainty of the different distributions of the random variables usually estimated through the statistical models (Klir and Wierman 1998; Parsons 2001). There may be multiple failure moods in a design or system. Hence, a reliable design has to satisfy all the failure moods to gain overall reliability of the system. The probability of failure (P_f) is also denoted as reliability index which can be estimated using simulation-based probabilistic reliability analysis (Enevoldsen and Sorensen 1994).

In the deterministic design, there are fixed limits for the design constraints which do not allow any kind of incertitude. Therefore, deterministic design has high chance of failure due to lack of uncertainty. The uncertainties may be controllable (physical dimension) or uncontrollable (material property). Uncertainties can be responsible to produce huge amount of variations in the performance function and this may lead to failure. Therefore, the deterministic design optimization may lead to unreliable solution.

A typical reliability formulation only considering the aleatory uncertainty with component level reliability constraints can be written as (Zaman 2010):

$$\begin{aligned} &\min \quad f(d,Z) \\ &s.t. \quad p_{f_i} = P(g_i(X,Z) \le 0) \le p_i \quad for\ i = 1,2,\ldots,k \end{aligned} \tag{31}$$

where, $f(d,Z)$ is the objective function, d is a set of design variables, Z is a set of input random variables and p_i could be ith threshold failure probability. The vector d may include both deterministic design variables as well as distribution parameters of random design variables

x. In RBDO, the objective function value is estimated at the mean values of the random variables x and z.

The inclusion of epistemic uncertainty in reliability-based design optimization leads to more complexity in the design methodology. The design variable d and/or the input random variable Z in Eqn. (31) might have epistemic uncertainty. Since the non-design epistemic variables could not be controlled, the RBDO methodology has to employ a search among the possible values of such epistemic variables in order to find an optimal solution. This may provide a conservative reliability-based design. The nested RBDO considering the epistemic uncertainty can be formulated as:

$$\begin{aligned} &\min_{d}\left(\max_{\mu_z} f(d,Z)\right) \\ &s.t.\quad p_{f_i} = P(g_i(X,Z)\le 0)\le p_i \quad for\ i=1,2,\ldots,k \\ &\qquad Z_l \le \mu_z \le Z_u \end{aligned} \tag{32}$$

Where, Z_l and Z_u are the vectors of the lower and upper bound of the non-design variables μ_z of the inner loop optimization problem.

In Eqn. (32), the outer loop decision variables d may consist of stochastic design variables as well as epistemic design variables. In the RBDO, the outer loop optimization is carried out for a fixed set of non-design epistemic variables. The inner loop optimization is the analysis of the non-design epistemic variables where the optimizer searches among the possible values of the non-design epistemic variables within the interval data for a conservative solution of the RBDO.

The nested optimization is expensive than the decoupled approach while there is probability of failure in convergence in the decoupled reliability optimization. Therefore, it would be better to find a single loop optimization which will be efficient in computation and cost effective. At the same time, the performance of the design methodology would be better than these methods.

A single loop reliability-based design optimization is proposed which can be written as:

$$\begin{aligned} &d^* = \arg\min_{d}\left(f\left(d,\mu_z^*,\sigma_z^*\right)\right) \\ &s.t.\quad p_{f_i} = P\left(g_i\left(d,\mu_z^*,\sigma_z^*\right)\le 0\right)\le p_i \quad for\ i=1,2,\ldots,k \end{aligned} \tag{33}$$

The single loop optimization can be solved using Eqn. (33). However, this requires the uncertainty estimation of the non-design epistemic variables. The proposed likelihood-based method can be applied for the epistemic uncertainty representation as mentioned in chapter 2. The parameters of the non-design epistemic variables z are estimated using maximum likelihood method. In the uncertainty representation, the parameters and the moments may be identical. Either, methods can be developed to estimate the moments from the parameters as mentioned in chapter 3. The fixed values of the moments of the non-design random variables are used in Eqn. (36) for the reliability optimization.

The proposed maximum likelihood approach can estimate specific values of the parameters which can be used for moment estimation. Therefore, the estimated uncertainty from the likelihood method can provide the most probable reliability-based design optimization. Hence, application of this MLE-based single loop optimization method will be efficient and reduce the computational expenses. This proposed MLE-based reliability design optimization can be illustrated with numerical examples in future.

In robust design optimization, the main objective is to minimize the variation in the performance function where the optimization is focused to minimize the failure probability in the reliable design. Both methods required the uncertainty representation of the random variables which can be done easily and efficiently by the proposed maximum likelihood based approach in chapter 2. This thesis mainly focused on the robust design optimization considering the epistemic uncertainty representation. Nonlinear robustness-based design optimization is illustrated with numerical examples to establish the proposed method.

CHAPTER VI

NUMERICAL EXAMPLES

In this research, Maximum Likelihood Estimation-based (MLE) approach has been proposed for robustness-based design optimization. This is a generalized method which is used to estimate the parameters of the specified distributions. The MLE-based robust design optimization is illustrated with normal distribution and bounded Johnson distribution. A established method is used to estimate the moments of the bounded Johnson distribution from the parameters.

Two examples are illustrated in this study where one of them is a mathematical design problem which includes a objective function of three random variables and four coefficients. The observed data of the coefficients are multiple intervals which are overlapping or non-overlapping in nature. Another is a real engineering problem describing the upper stage design problem for a Two Stage To Orbit (TSTO) vehicle. The proposed methodology is applied for this two examples to develop robust design models.

6.1. Mathematical Example

Robust design optimization is done for a mathematical example using the MLE-based liklihood approach proposed previously. There is a performmance function consisted of two nonlinear constraints. There are three variables of *x* and four coefficients (*a, b, c* and *l*) in the performance function. The variables x_i (where, i = 1,2,3) are treated as design variables where the coefficients are treated as the non-design epistemic variables and denoted as z_i.

The performance function with the nonlinear constraints of the mathematical model:

$$\begin{aligned} &\min f(x) = ax_1^2 - bx_1x_2 + cx_1x_2^2 + lx_1x_3 \\ &s.t. \quad -2x_1^2 - x_1x_2 + x_3 \le -1 \\ &\qquad -x_1x_2 + x_2^2 - x_1x_3 \le -3 \end{aligned} \tag{34}$$

From the nonlinear function *f*(*x*), applying first order Taylor series expansion method, the mean and variance can be obtained from Eqns. (19) and (21), respectively.

Hence, the mean of the objective function:

$$\mu_f = ax_1^2 - bx_1x_2 + cx_1x_2^2 + lx_1x_3 \tag{35}$$

There are 7 variables in the objective function where some are design variables with non-design epistemic variables. Therefore, the variance of the objective function:

$$\sigma_f^2 = \sigma_{x_1}^2\left(\frac{\partial(f(x))}{\partial x_1}\right)^2 + \sigma_{x_2}^2\left(\frac{\partial(f(x))}{\partial x_2}\right)^2 + \sigma_{x_3}^2\left(\frac{\partial(f(x))}{\partial x_3}\right)^2 + \sigma_a^2\left(\frac{\partial(f(x))}{\partial a}\right)^2$$
$$+ \sigma_b^2\left(\frac{\partial(f(x))}{\partial b}\right)^2 + \sigma_c^2\left(\frac{\partial(f(x))}{\partial c}\right)^2 + \sigma_l^2\left(\frac{\partial(f(x))}{\partial l}\right)^2$$
$$or, \sigma_f^2 = \sigma_{x_1}^2\left(2ax_1 - bx_2 + cx_2^2 + lx_3\right)^2 + \sigma_{x_2}^2\left(-bx_2 + 2cx_1x_2\right)^2 + \sigma_{x_3}^2\left(lx_1\right)^2 + \sigma_a^2\left(x_1^2\right)^2$$
$$+ \sigma_b^2\left(-x_1x_2\right)^2 + \sigma_c^2\left(x_1x_2^2\right)^2 + \sigma_l^2\left(x_1x_3\right)^2 \quad (36)$$

The coefficients of the objective function consist of multiple interval data. Maximum log likelihood method is used in Eqn. (10) to find out the parameters of the coefficients of the objective function considering the coefficients follow normal distribution. The interval data including overlapping, non-overlapping and mixed of the coefficients are presented in Table 5.1.

Table 5.1: Multiple interval data of the coefficients

Coefficient	Multiple Interval Data [lb_i-ub_i]
a	[1.5 - 2.2; 1.8 – 2.3; 2.0 – 2.5; 2.1 – 2.6; 2.2 – 2.7]
b	[2.8 – 3.2; 3.3 – 3.9; 4.0 – 4.5; 4.7 – 5.0; 5.2 – 6.0]
c	[2.7 – 3.1; 2.9 – 3.2; 2.8 – 3.2; 3.5 – 3.8; 4.0 – 4.5]
l	[4.8 – 5.1; 5.3 – 5.9; 6.0 – 6.8; 6.5 – 6.9; 6.6 – 7.0]

The parameters of the coefficients (*a, b, c* and *l*) are estimated using maximum log likelihood method from multiple interval data through optimization by using MATLAB solvers called „fminunc" and „fmincon". The optimization is applied for each coefficient individually to find out the optimum values of the parameters from the multiple interval values. The mean (μ) and the standard deviations (σ) are the parameters for normal distribution. The parameters are obtained through nested optimizations which are given in Table 5.2.

Table 5.2: Parameters of normal distribution for the coefficients

Coefficient	μ	σ
a	2.2200	0.4792
b	4.2200	1.1565
c	3.3400	0.7003
l	6.1600	0.9222

For the log likelihood function of bounded Johnson distribution, the parameters (δ, γ, ξ and λ) of the coefficients (*a, b, c* and *l*) are estimated from Eqn. (9) through the nested optimization method proposed earlier from multiple interval. The nested optimization is applied for each coefficient individually to find out the parameters δ, γ, ξ and λ from the multiple intervals. The estimated values of the parameters (δ, γ, ξ and λ) are given in Table 5.3.

Table 5.3: Parameters of bounded Johnson distribution for the coefficients

Parameters / Coefficient	δ	γ	ξ	λ
a	0.4901	0.2548	1.5000	1.2000
b	0.5110	0.3101	2.8000	3.2000
c	0.4926	-0.0827	2.7000	1.8000
l	0.5038	0.1073	4.8000	2.2000

The empirical CDF curve of the parameters estimated from nested optimization for both normal and bounded Johnson distribution are plotted in the graphs Figs. 5.1-5.8. The MLE-based curves are well within the bounds of the empirical CDFs produced from the parameters obtained through moment matching method from the moment bounds method developed in Zaman et al. (2011b).

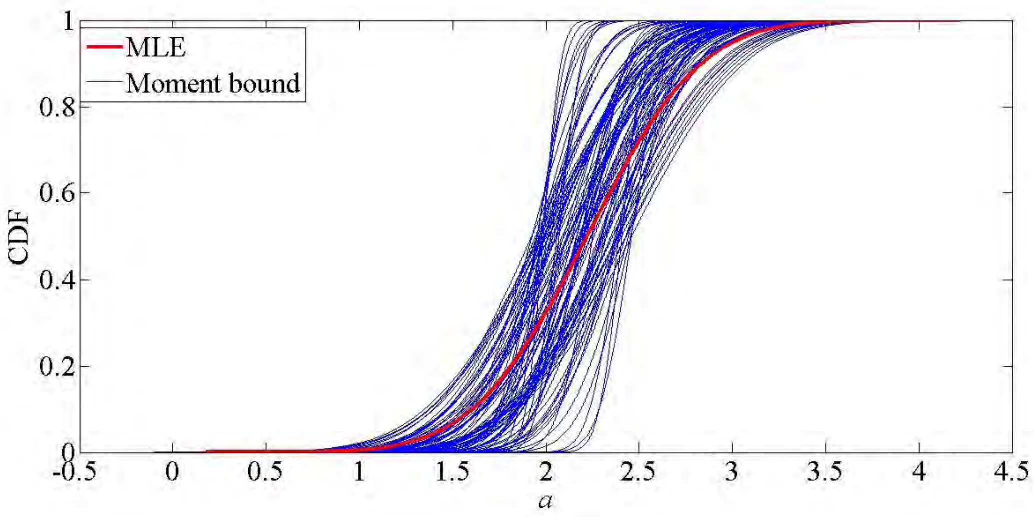

Fig. 5.1 MLE-based curve fitting of coefficient „*a*" for normal distribution

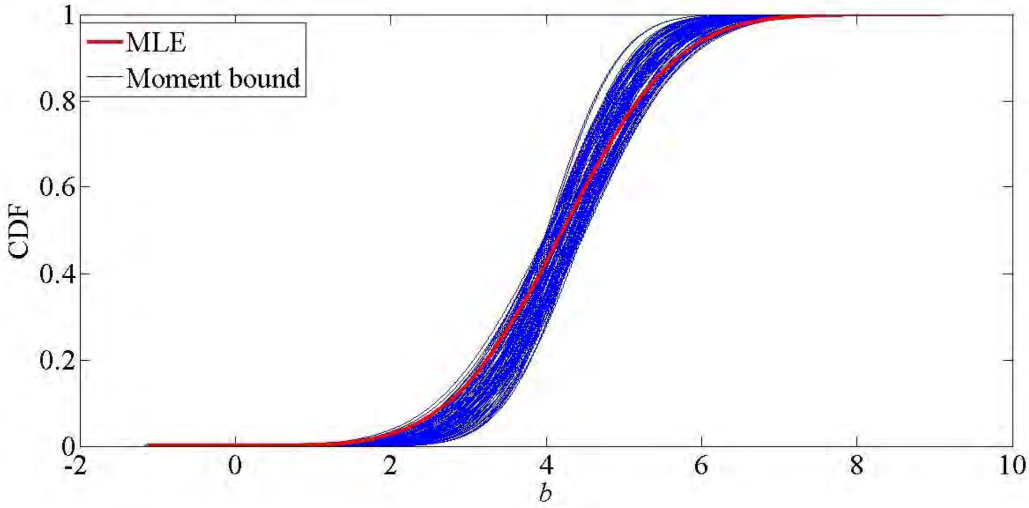

Fig. 5.2 MLE-based curve fitting of coefficient „*b*" for normal distribution

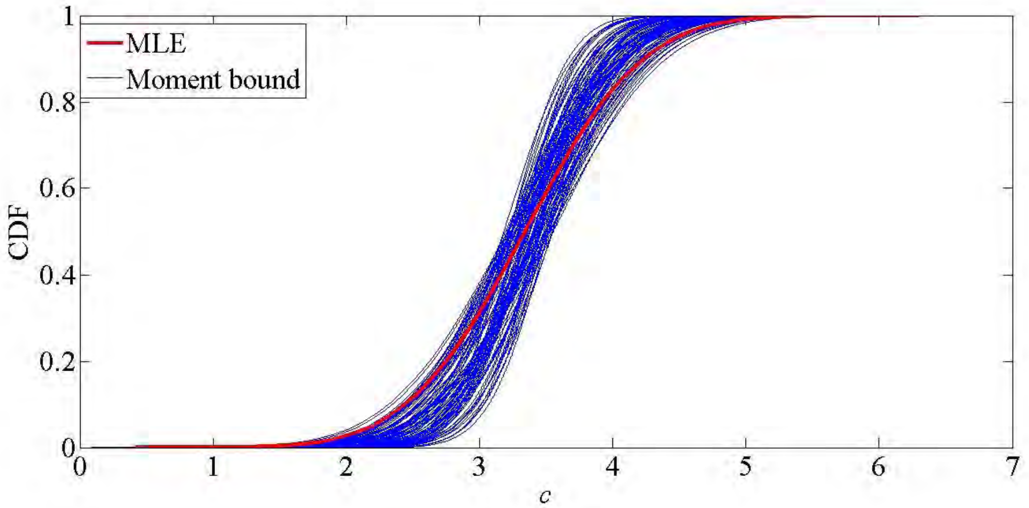

Fig. 5.3 MLE-based curve fitting of coefficient „*c*" for normal distribution

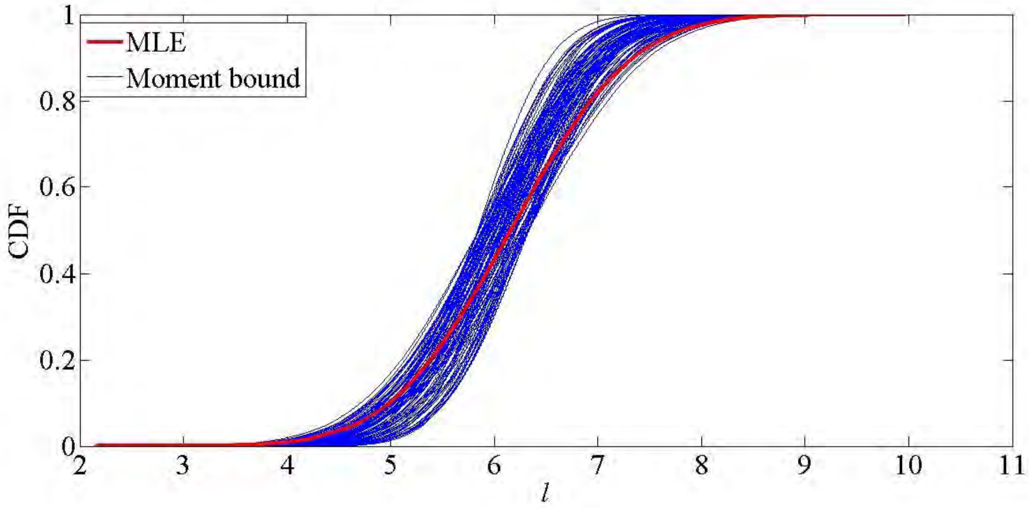

Fig. 5.4 MLE-based curve fitting of coefficient „l" for normal distribution

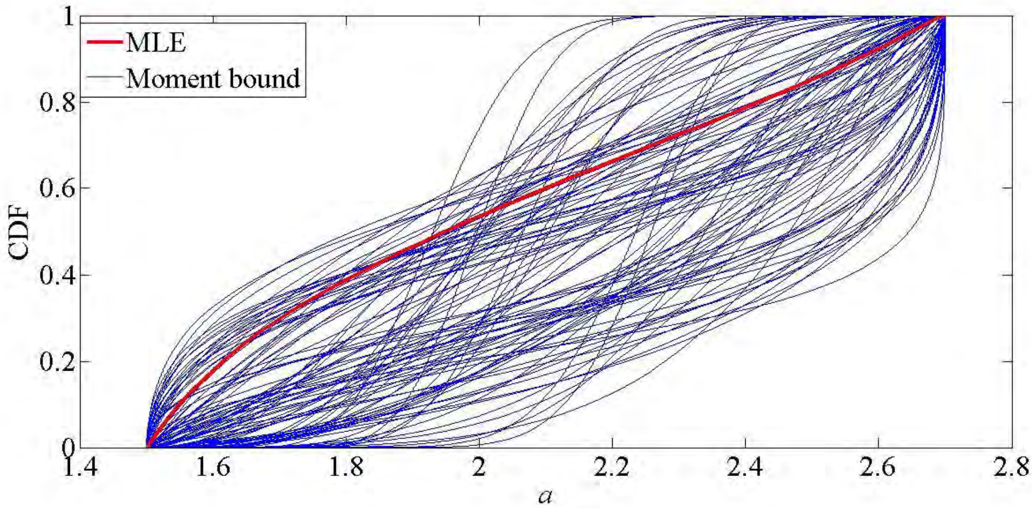

Fig. 5.5 MLE-based curve fitting of coefficient „a" for bounded Johnson distribution

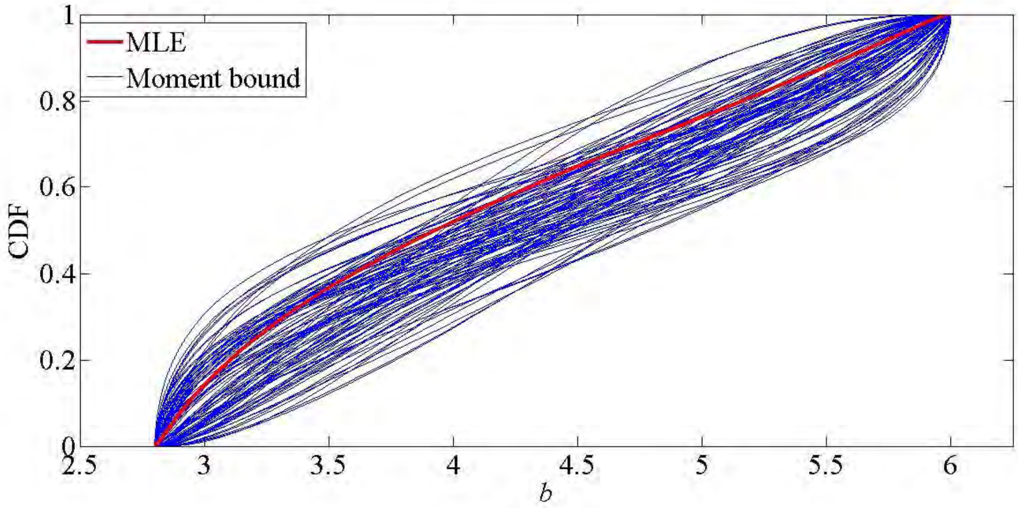

Fig. 5.6 MLE-based curve fitting of coefficient „*b*" for bounded Johnson distribution

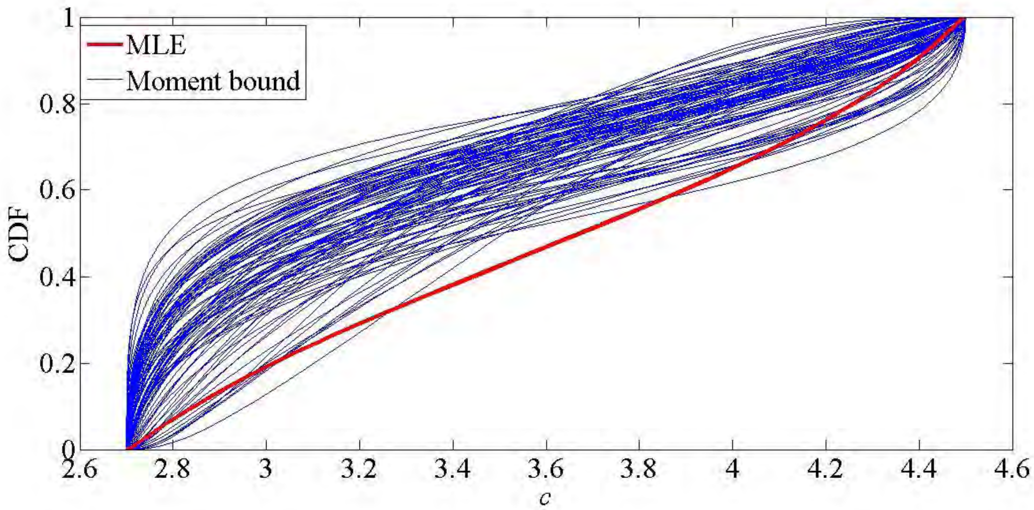

Fig. 5.7 MLE-based curve fitting of coefficient „*c*" for bounded Johnson distribution

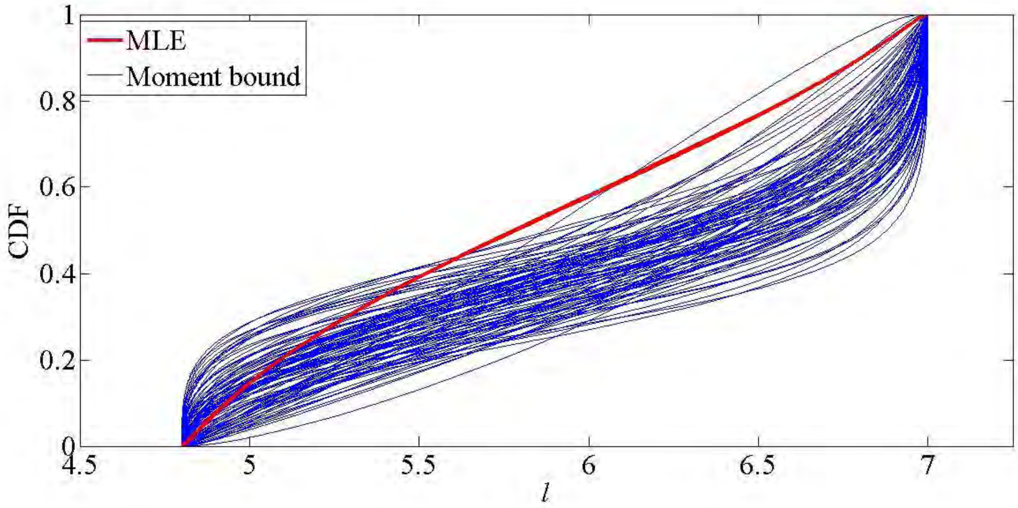

Fig. 5.8 MLE-based curve fitting of coefficient „l" for bounded Johnson distribution

In the proposed robust design optimization, it requires the standard deviations (σ_x) of the design variables and the moments (mean (μ_z) and standard deviation (σ_z)) of the epistemic non-design variable.

Let"s assume, the standard deviation of the random variables as,

$$\sigma_{x_1} = \sigma_{x_2} = \sigma_{x_3} = 0.5$$

The moments of the epistemic variables (z_i) for the bounded Johnson distribution from the parameters can be estimated from Eqns. (27) and (29). The proposed equations for bounded Johnson distribution provide the values of the mean (μ_z) and the standard deviation (σ_z) of the coefficients in Table 5.4.

Table 5.4: Mean values (μ_z) and standard deviations (σ_z) of the bounded Johnson distribution in mathematical problem

Coefficient	Mean (μ_z)	Std. Dev (σ_z)
a	2.0073	0.3744
b	4.1049	0.9727
c	3.6453	0.5681
l	5.8288	0.6865

The robust design formulation from Eqn. (16) for the mathematical example can be written as:

$$
\begin{aligned}
d^* &= \arg\min_d \left(w \times \mu_f\left(d, \mu_z^*, \sigma_z^*\right) + (1-w) \times \sigma_f\left(d, \mu_z^*, \sigma_z^*\right) \right) \\
s.t. \quad & LB + k\sigma\left(g_i\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E\left(g_i(d, z)\right) \le UB - k\sigma\left(g_i\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& lb_i + k\sigma(x_i) \le d_i \le ub_i - k\sigma(x_i) \quad for\ i = 1,2,3
\end{aligned} \tag{37}
$$

The mean (μ) and standard deviation (σ) of the performance function are estimated form first order Taylor series expansion method using Eqns. (35) and (36), respectively. As mentioned previously, „$w \geq 0$" is the weighting coefficient which ranges from *0* to *1* that provides the relative importance of the objectives μ_f and σ_f. The value of „k" is assumed to be unity which adjusts the robustness of the method against any types of conservatism in the solution. The value of „k" ensures feasibility robustness in the robust design optimization.

However, the Matlab solver „fmincon" is used for robust design optimization using Eqn. (37) which will provide the mean (μ) and standard deviation (σ) of the performance function. Here, „fmincon" uses „Active set algorithm" for optimization of the performance function. From the optimization, mean (μ_f) and standard deviation (σ_f) of the performance function are obtained for different values of „w". The robustness-based design optimization is obtained using maximum likelihood-based approaches (MLE) with the illustration of normal and bounded Johnson distribution. Zaman et al. (2011b) developed a decoupled method for robust design optimization which is illustrated here to compare with the methods proposed in this research.

The mean (μ_f) and standard deviation (σ_f) are estimated from the Robustness-based design optimization using MLE for Normal distribution, MLE for bounded Johnson distribution and decoupled approach (Zaman et al. 2011b). The estimated values of the means (μ_f) and standard deviations (σ_f) are given in Table 5.6, respectively.

From Table 5.6, it is easy to compare the robust design optimizations using different methods for different values of „w". For $w = 0$, the MLE-based robust optimization using bounded Johnson distribution is better than the other two methods with minimum standard deviation as it is the requirement to minimize the standard deviation with optimizing the mean value. For $w = 0$, the MLE-based robust optimization using normal distribution is better method than the decoupled approach for robust design due to contain less deviation. For $w =$

1, the MLE-based robust optimization using bounded Johnson distribution with minimum standard deviation is better than the MLE-based robust optimization using normal distribution and the MLE-based robust optimization using normal distribution is better than decoupled approach which has maximum standard deviation. For *w = 0.2, 0.4, 0.6 and 0.8*, the MLE based robust optimizations are better than the decoupled approach for robust design optimization which has bigger standard deviations. It is easier to observe the differences in Figs. 16-18.

Table 5.5: Mean values (μ_f) and standard deviations (σ_f) from robust optimization using MLE and decoupled approaches in mathematical problem

w	Moments of performance function (μ_f, σ_f)		
	MLE (Normal distribution)	MLE (bounded Johnson distribution)	Decoupled approach (Zaman et al. 2011)
0.0	(29.2744, 12.1637)	(27.7160, 11.1894)	(31.3929, 13.1637)
0.2	(29.2146, 12.1663)	(27.6411, 11.1927)	(31.2753, 13.1663)
0.4	(29.1709, 12.1739)	(27.5873, 11.2021)	(31.1848, 13.1739)
0.6	(29.1397, 12.1866)	(27.5500, 11.2173)	(31.1159, 13.1866)
0.8	(29.1195, 12.2061)	(27.5266, 11.2397)	(31.0676, 13.2061)
1.0	(29.1116, 12.2366)	(27.5178, 11.2732)	(31.0476, 13.2366)

Robustness-based design optimization using maximum likelihood approach for Normal distribution and bounded Johnson distribution are shown in Figs. 5.9 and 5.10, respectively. In Figs. 5.11-5.13, comparisons are shown among the MLE-based robust design optimization method using normal distribution, MLE-based robust design optimization method using bounded Johnson distribution and decoupled approach for robust design optimization (Zaman et al. 2011b).

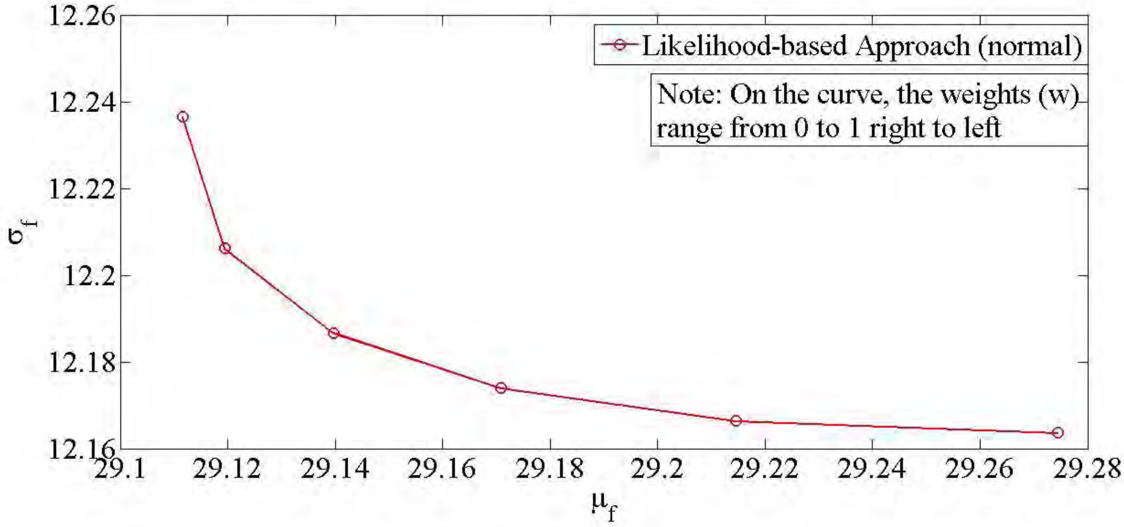

Fig. 5.9 MLE-based robust design optimization using Normal distribution

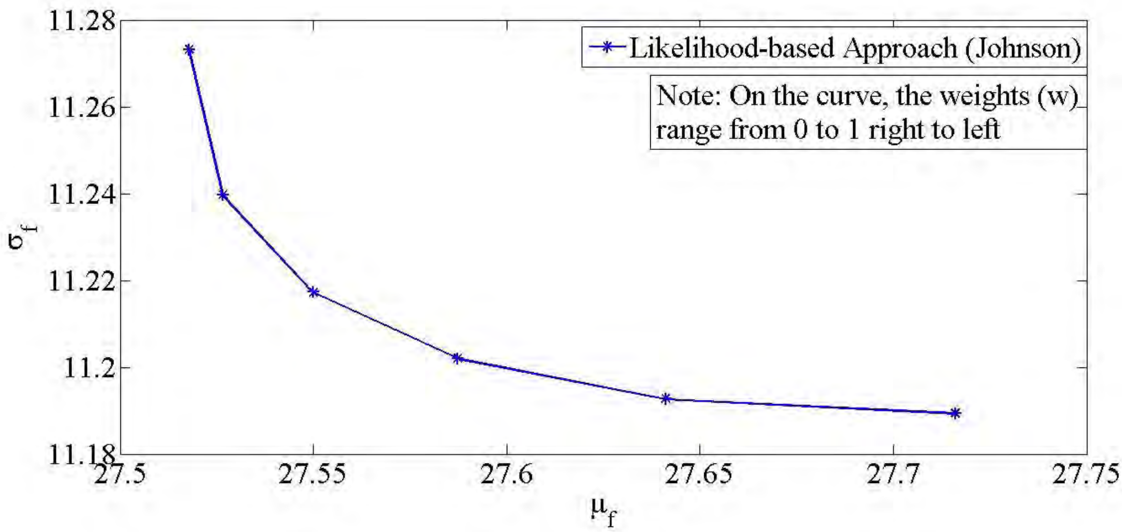

Fig. 5.10 MLE-based robust design optimization using bounded Johnson distribution

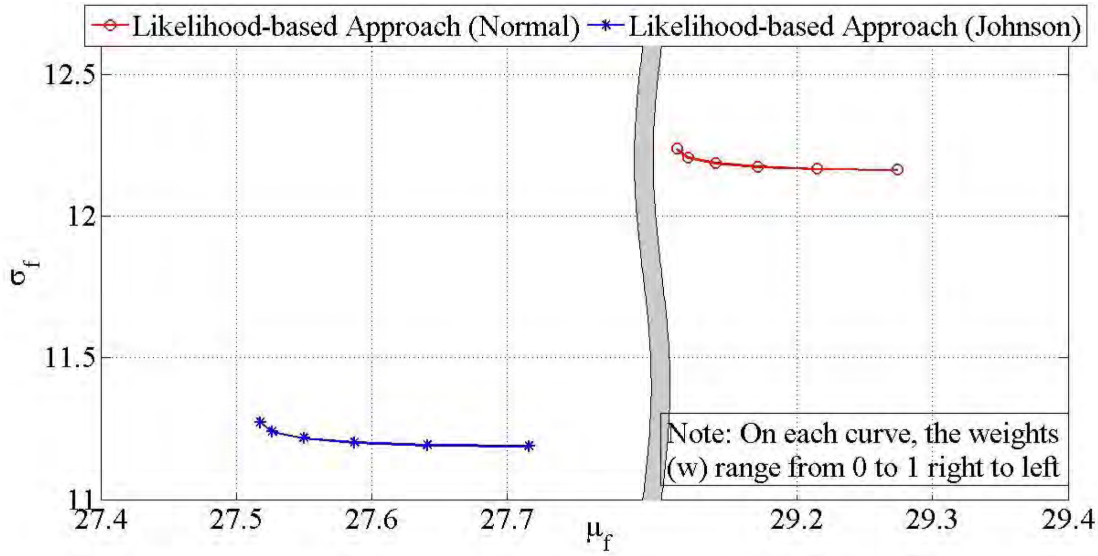

Fig. 5.11 Robust design optimization using MLE-based (Normal distribution) and MLE (bounded Johnson distribution)

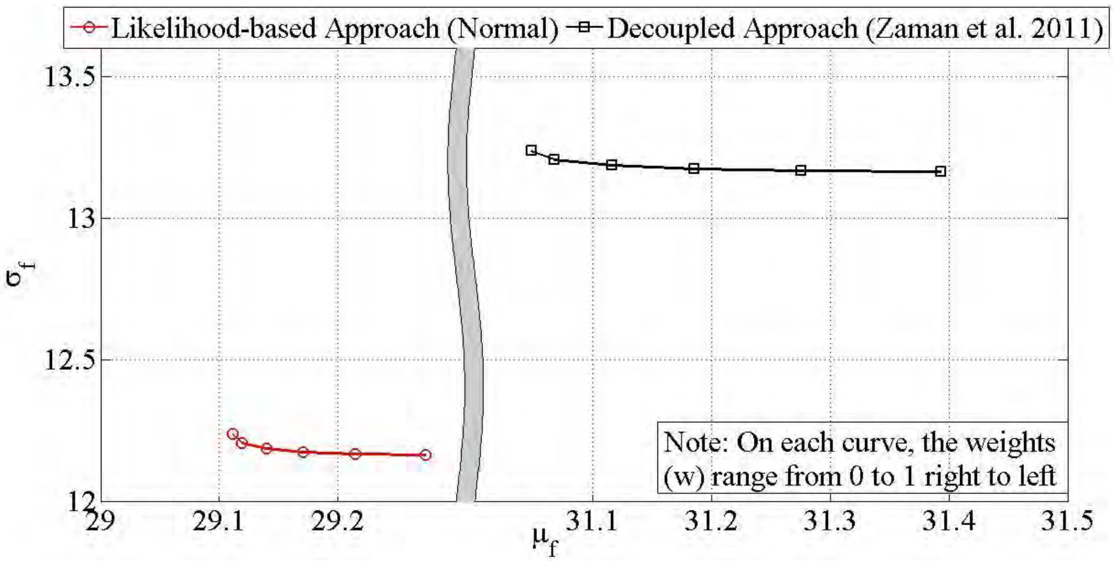

Fig. 5.12 Robust design optimization using MLE-based (Normal distribution) and decoupled approach

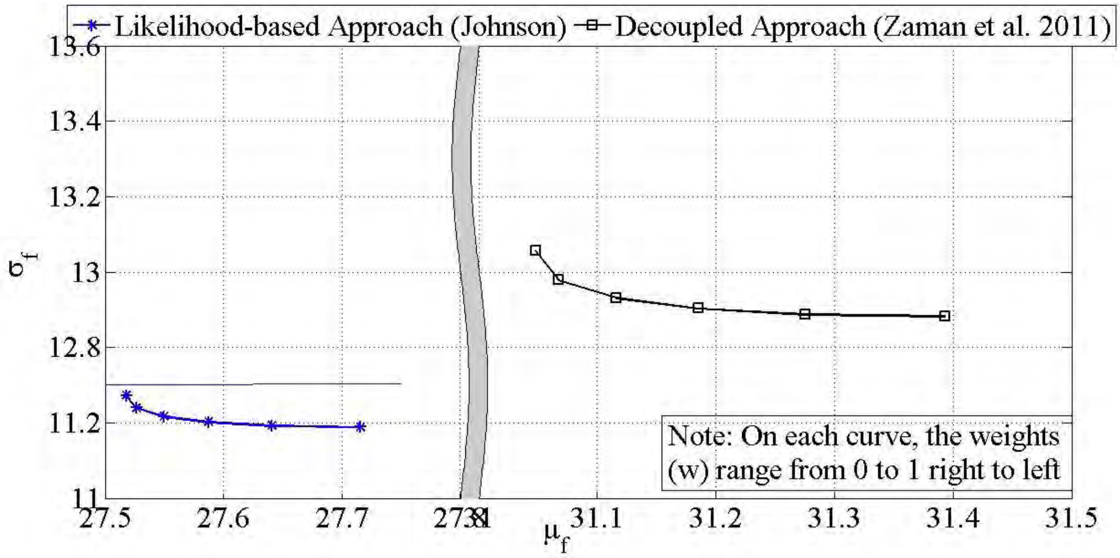

Fig. 5.13 Robust design optimization using MLE-based (bounded Johnson distribution) and decoupled approach

6.2. Engineering Example

In this section, the proposed methods are illustrated for the conceptual level design process of a TSTO vehicle. This problem has been adapted from Zaman et al (2001). The engineering system analysis may consist of geometric modeling, aerodynamics, aerothermodynamics, engine performance analysis, trajectory analysis, mass property analysis and cost modeling (Stevenson et al. 2002). In this example, a simplified version of the upper stage design process of a TSTO vehicle is used to examine the proposed methods.

The analysis outputs (performance function) are Gross Weight (*GW*), Engine Weight (*EW*), Propellant Fraction Required (*PFR*), Vehicle Length (*VL*), Vehicle Volume (*VV*) and Body Wetted Area (*BWA*). Each of the analysis outputs is approximated by a second order response surface and is a function of the random design variables Nozzle Expansion Ratio (*ExpRatio*), Payload Weight (*Payload*), Separation Mach (*SepMach*), Separation Dynamic Pressure (*SepQ*), Separation Flight Path Angle (*SepAngle*) and Body Fineness Ratio (*Fineness*).

The objective is to optimize an individual analysis output (e.g., Gross Weight) while satisfying the constraints imposed by each of the design variables as well as the analysis output. The numerical values of the design bounds for the input variables and analysis outputs are given in Tables 5.7 and 5.8, respectively.

Table 5.6: Design bounds for the input variables

Design Variables	lower bound (lb)	upper bound (ub)
ExpRatio	40	150
Payload	8000	40000
SepMach	7	12
SepQ	40	200
SepAngle	7	12
Fineness	4	6

Table 5.7: Design bounds for the analysis outputs

Analysis Outputs	Lower Bound (LB)	Upper Bound (UB)
GW	0	100e+005
EW	0	100e+005
PFR	0.4	0.95
VL	0	100e+002
VV	0	100e+003
BWA	0	100e+005

The performance function of Gross Weight (*GW*) is expressed in Eqn. (38) as:

$$
\begin{aligned}
GW = {} & 1189320.749 - 3028.704706\, x_1 + 18.41485\, x_2 - 176044.4712\, x_3 + 908.5213095\, x_4 \\
& - 41769.8467\, x_5 + 14727.08053\, x_6 - 0.007369851\, x_1 x_2 + 118.2472034\, x_1 x_3 \\
& + 0.006718916\, x_1 x_4 + 78.93978363\, x_1 x_5 + 7.645949064\, x_1 x_6 - 1.003462593\, x_2 x_3 \\
& - 0.004146963\, x_2 x_4 - 0.24342992\, x_2 x_5 - 0.30960956\, x_2 x_6 - 45.02435999\, x_3 x_4 \\
& + 2296.000898\, x_3 x_5 - 4111.544204\, x_3 x_6 - 3.727848286\, x_4 x_5 - 72.68000687\, x_4 x_6 \\
& + 264.2316861\, x_5 x_6 + 6.401261905\, x_1^2 + 2.98956\, e06\, x_2^2 + 7986.357268\, x_3^2 \\
& - 0.210020773\, x_4^2 + 916.7392297\, x_5^2 + 4231.6528\, x_6^2
\end{aligned}
\tag{38}
$$

The performance function of Engine Weight (*EW*) is expressed in Eqn. (39) as:

$$\begin{aligned} EW = {} & 15374.58847 - 27.702079\, x_1 + 0.26187849\, x_2 - 2425.361607\, x_3 + 12.37410577\, x_4 \\ & - 389.3583447\, x_5 + 537.286141\, x_6 + 0.000104819\, x_1x_2 + 0.095950445\, x_1x_3 \\ & - 0.003058586\, x_1x_4 + 0.498878041\, x_1x_5 + 0.837557912\, x_1x_6 - 0.016045124\, x_2x_3 \\ & - 3.53203\, e-05\, x_2x_4 - 0.003595284\, x_2x_5 - 0.003359929\, x_2x_6 - 0.68176651\, x_3x_4 \\ & + 29.88198246\, x_3x_5 - 63.99813459\, x_3x_6 - 0.061435395\, x_4x_5 - 0.950124874\, x_4x_6 \\ & + 5.613278991\, x_5x_6 + 0.095697511\, x_1^2 + 2.24406\, e-07\, x_2^2 + 115.508073\, x_3^2 \\ & - 7.06903\, e-05\, x_4^2 + 2.726825794\, x_5^2 + 27.5074007\, x_6^2 \end{aligned} \tag{39}$$

The performance function of Propellant Fraction Required (*PFR*) is expressed in Eqn. (40) as:

$$\begin{aligned} PFR = {} & 0.893093233 - 0.0002417266\, x_1 - 4.75121\, e-07\, x_2 - 0.005464932\, x_3 + 0.000101483\, x_4 \\ & - 0.00322987\, x_5 - 0.00928557\, x_6 + 5.13829\, e-10\, x_1x_2 - 5.46612\, e-06\, x_1x_3 \\ & + 6.76376\, e-08\, x_1x_4 - 4.53026\, e-06\, x_1x_5 + 3.30224\, e-06\, x_1x_6 - 3.75202\, e-08\, x_2x_3 \\ & + 6.43768\, e-10\, x_2x_4 - 9.55568\, e-09\, x_2x_5 + 5.77463\, e-08\, x_2x_6 - 1.43973\, e-05\, x_3x_4 \\ & - 9.88908\, e-05\, x_3x_5 - 0.0004773\, x_3x_6 + 1.13559\, e-06\, x_4x_5 - 5.42466\, e-06\, x_4x_6 \\ & + 8.51662\, e-05\, x_5x_6 + 6.80047\, e-07\, x_1^2 + 8.9559\, e-12\, x_2^2 - 0.00098585\, x_3^2 \\ & + 1.88595\, e-07\, x_4^2 + 0.000117169\, x_5^2 + 0.000959377\, x_6^2 \end{aligned} \tag{40}$$

The performance function of Vehicle Length (VL) is expressed in Eqn. (41) as:

$$\begin{aligned} VL = {} & 823.1048661 - 1.411069734\, x_1 + 0.006369636\, x_2 - 90.3567667\, x_3 + 0.569733596\, x_4 \\ & - 25.55147283\, x_5 + 261.6412745\, x_6 - 1.06184\, e-06\, x_1x_2 + 0.066102241\, x_1x_3 \\ & - 7.43962\, e-05\, x_1x_4 + 0.013128748\, x_1x_5 - 0.074329589\, x_1x_6 - 0.000214937\, x_2x_3 \\ & - 9.37031\, e-07\, x_2x_4 - 6.43232\, e-05\, x_2x_5 + 0.001046781\, x_2x_6 - 0.040195575\, x_3x_4 \\ & + 1.323122789\, x_3x_5 - 14.70655866\, x_3x_6 - 0.004019241\, x_4x_5 - 0.01496916\, x_4x_6 \\ & - 1.387289781\, x_5x_6 + 0.003662151\, x_1^2 - 5.09359\, e-08\, x_2^2 + 4.70505342\, x_3^2 \\ & - 7.94011\, e05\, x_4^2 + 0.93934255\, x_5^2 + 0.009053289\, x_6^2 \end{aligned} \tag{41}$$

The performance function of Vehicle Volume (*VV*) is expressed in Eqn. (42) as:

$$\begin{aligned} VV = {} & 54689.41793 - 71.0056374\, x_1 + 0.63271800\, x_2 - 6583.263146\, x_3 + 29.16000643\, x_4 \\ & - 1213.035573\, x_5 - 1813.643318\, x_6 - 0.000246467\, x_1 x_2 + 3.774590948\, x_1 x_3 \\ & - 0.01858828\, x_1 x_4 - 0.142544958\, x_1 x_5 - 0.043939802\, x_1 x_6 - 0.034727527\, x_2 x_3 \\ & + 0.00027002\, x_2 x_4 - 0.004930386\, x_2 x_5 + 0.002161457\, x_2 x_6 - 1.819045749\, x_3 x_4 \\ & + 64.33859515\, x_3 x_5 - 102.110372\, x_3 x_6 - 0.283472503\, x_4 x_5 - 1.482888979\, x_4 x_6 \\ & - 16.35284892\, x_5 x_6 + 0.148782073\, x_1^2 - 1.26129\, e - 06\, x_2^2 + 2.66.7127867\, x_3^2 \\ & - 0.009071094\, x_4^2 + 37.3538488\, x_5^2 + 281.5212617\, x_6^2 \end{aligned} \quad (42)$$

The performance function of Body Wetted Area (*BWA*) is expressed in Eqn. (43) as:

$$\begin{aligned} BWA = {} & 1894320.177 - 2535.255106\, x_1 + 18.24634963\, x_2 - 209077.2388\, x_3 + 1002.2601\, x_4 \\ & - 43645.41464\, x_5 + 39116.04318\, x_6 - 0.006531596\, x_1 x_2 + 123.9268282\, x_1 x_3 \\ & - 0.364359382\, x_1 x_4 - 8.855865687\, x_1 x_5 - 38.42451662\, x_1 x_6 - 0.860599695\, x_2 x_3 \\ & + 0.00322796\, x_2 x_4 - 0.157168092\, x_2 x_5 + 0.653315508\, x_2 x_6 - 64.3040588\, x_3 x_4 \\ & + 2275.808742\, x_3 x_5 - 10585.43583\, x_3 x_6 - 8.305159162\, x_4 x_5 - 38.77262826\, x_4 x_6 \\ & - 1106.758605\, x_5 x_6 + 5.682243503\, x_1^2 - 6.48464\, e - 05\, x_2^2 + 8896.931233\, x_3^2 \\ & - 0.292543211\, x_4^2 + 1404.757323\, x_5^2 + 8985.576363\, x_6^2 \end{aligned} \quad (43)$$

The proposed robust design optimization method is illustrated here for the TSTO problem. It is assumed that four input variables (*ExpRatio*, *Payload*, *SepMach* and *SepQ*) are design variables denoted by d_i and another two variables (*SepAngle* and *Fineness*) are the epistemic non-design variables denoted by z_i which are expressed by interval data as given in Table 5.9.

Table 5.8: Interval data of epistemic variables of the TSTO system

Sample	SepAngle	Fineness
i.	8.0 – 8.5	4.00 – 4.50
ii.	8.2 – 8.6	4.25 – 4.75
iii.	8.5 – 9.0	4.50 – 5.25
iv.	8.2 – 9.2	4.50 – 5.50
v.	8.8 – 9.5	5.00 – 6.00

In chapter 2, a maximum log likelihood method is proposed with the illustration of normal and bounded Johnson distribution for the uncertainty representation of the interval data. Using Eqns. (7) and (9), the log likelihood density functions are formed for normal and bounded Johnson distribution. The proposed nested optimization for the likelihood estimation in Eqn. (10) provides the parameters of the distributions. The estimated parameters of normal and bounded Johnson distribution are given in Tables 5.10 and 5.11 respectively.

Table 5.9: Parameters of normal distribution estimated in TSTO problem

	Mean (μ)	Std. Dev (σ)
SepAngle	8.7800	0.5810
Fineness	4.8500	0.7681

Table 5.10: Parameters of bounded Johnson distribution estimated in TSTO problem

Parameters / Variable	δ	γ	ξ	λ
SepAngle	0.4902	-0.2411	8	1.5
Fineness	0.4923	-0.2008	4	2

The empirical CDF curves constructed for the parameters estimated from nested optimization for both normal and bounded Johnson distribution are plotted in the graphs (Figs. 5.14-5.17). The MLE-based curves are well within the bounds for the empirical CDF produced from the parameters obtained through moment matching method from the bounds of moment estimation methodology developed in Zaman et al. (2011a).

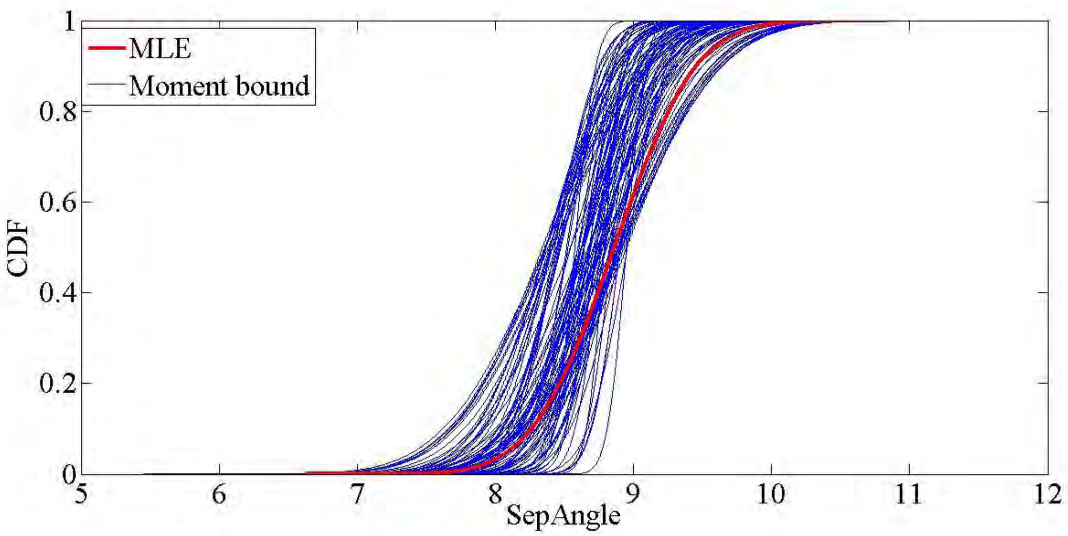

Fig. 5.14 MLE-based curve fitting of *SepAngle* for normal distribution

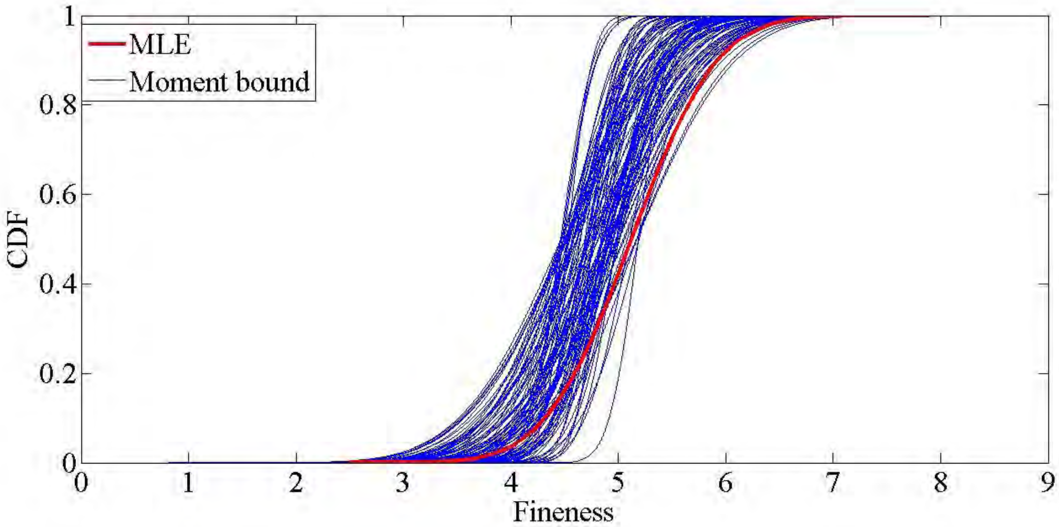

Fig. 5.15 MLE-based curve fitting of *Fineness* for normal distribution

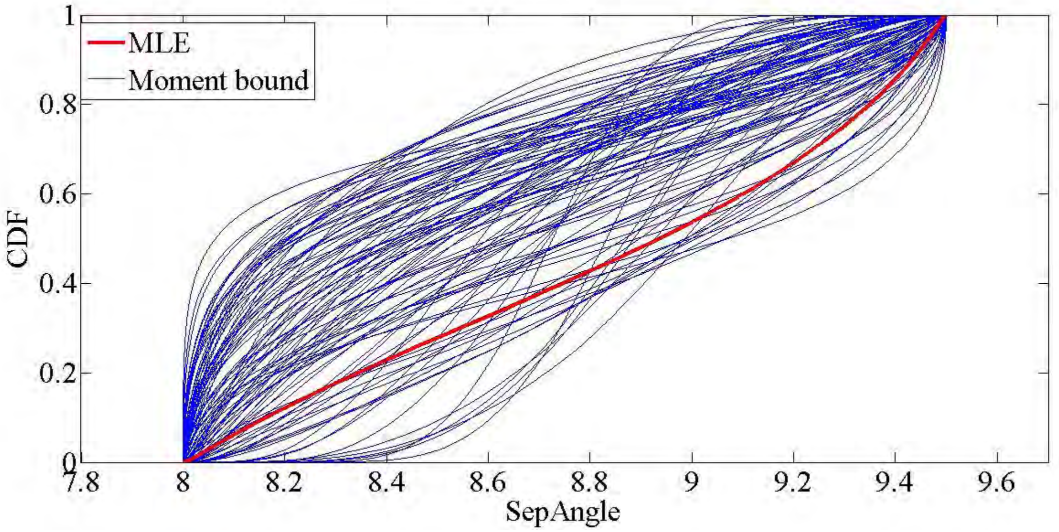

Fig. 5.16 MLE-based curve fitting of *SepAngle* for bounded Johnson distribution

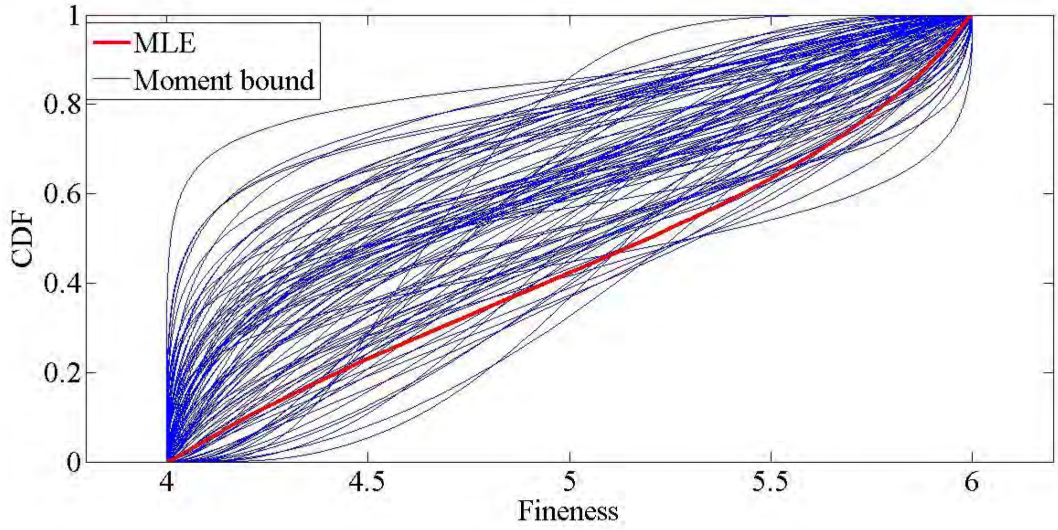

Fig. 5.17 MLE-based curve fitting of *Fineness* for bounded Johnson distribution

The moments (mean (μ_z) and the standard deviations (σ_z)) of the epistemic variables (z_i) for the bounded Johnson distribution from the parameters can be estimated from Eqns. (27) and (29). Estimated mean (μ_z) and standard deviations (σ_z) of the epistemic variables are given in Table 5.12.

Table 5.11: Mean values (μ_z) and standard deviations (σ_z) of bounded Johnson distribution in TSTO problem

	Mean (μ_z)	Std. Dev (σ_z)
SepAngle	8.8597	0.4687
Fineness	5.1219	0.6264

Let‟s assume, the standard deviations of the random variables are calculated from the coefficient of variation (C_v) which is asssumed to be 0.1. hence, the standard deviation of the decision variable can be calculated from the following equation:

$$C_v = \frac{\sigma_d}{\mu_d} \tag{44}$$

Using (44), the standard deviations (σ_d) of the decision variables are calculated from the average value of the bounds as shown in Table 5.14.

Table 5.12: Standard deviations (σ_d) of the decision variables

Decision variable	Standard deviation (σ_d)
ExpRatio	9.50
Payload	2400
SepMach	0.95
SepQ	12

For the TSTO vehicle, the robust design problem becomes:

$$
\begin{aligned}
d^* = \arg\min_d & \left(w \times E\left(GW\left(d, \mu_z^*, \sigma_z^*\right)\right) + (1-w) \times \sigma\left(GW\left(d, \mu_z^*, \sigma_z^*\right)\right)\right) \\
s.t. \quad & LB_1 + k\sigma\left(GW\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(GW(d,z)) \le UB_1 - k\sigma\left(GW\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& LB_2 + k\sigma\left(EW\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(EW(d,z)) \le UB_2 - k\sigma\left(EW\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& LB_3 + k\sigma\left(PFR\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(PFR(d,z)) \le UB_3 - k\sigma\left(PFR\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& LB_4 + k\sigma\left(VL\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(VL(d,z)) \le UB_4 - k\sigma\left(VL\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& LB_5 + k\sigma\left(VV\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(VV(d,z)) \le UB_5 - k\sigma\left(VV\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& LB_6 + k\sigma\left(BWA\left(d, \mu_z^*, \sigma_z^*\right)\right) \le E(BWA(d,z)) \le UB_6 - k\sigma\left(BWA\left(d, \mu_z^*, \sigma_z^*\right)\right) \\
& lb_i + k\sigma(x_i) \le d_i \le ub_i - k\sigma(x_i) \quad for\, i = 1,2,3,4
\end{aligned} \tag{45}
$$

The mean (μ_{GW}) and standard deviation (σ_{GW}) from the variance of the performance function are estimated form first order Taylor series expansion method using Eqns. (19) and (21) respectively. As mentioned previously, „$w \geq 0$" is the weighting coefficient which ranges from 0 to 1 and that „k" is assumed to be unity.

The mean of Gross Weight (GW) can be calculated as expressed in Eqn. (46):

$$
\begin{aligned}
E(GW) = & 1189320.749 - 3028.704706\, x_1 + 18.41485\, x_2 - 176044.4712\, x_3 + 908.5213095\, x_4 \\
& - 41769.8467\, x_5 + 14727.08053\, x_6 - 0.007369851\, x_1x_2 + 118.2472034\, x_1x_3 \\
& + 0.006718916\, x_1x_4 + 78.93978363\, x_1x_5 + 7.645949064\, x_1x_6 - 1.003462593\, x_2x_3 \\
& - 0.004146963\, x_2x_4 - 0.24342992\, x_2x_5 - 0.30960956\, x_2x_6 - 45.02435999\, x_3x_4 \\
& + 2296.000898\, x_3x_5 - 4111.544204\, x_3x_6 - 3.727848286\, x_4x_5 - 72.68000687\, x_4x_6 \\
& + 264.2316861\, x_5x_6 + 6.401261905\, x_1^2 + 2.98956\, e06\, x_2^2 + 7986.357268\, x_3^2 \\
& - 0.210020773\, x_4^2 + 916.7392297\, x_5^2 + 4231.6528\, x_6^2
\end{aligned} \tag{46}
$$

The variance of the Gross Weight (*GW*) can be calculated as expressed in Eqn. (47):

$$Var(GW) = \sum_{i=1}^{6} \left(\frac{\partial (GW)}{\partial x_i} \right)^2 Var(x_i)$$

$$= \begin{aligned} &\left(\frac{\partial (GW)}{\partial x_1} \right)^2 Var(x_1) + \left(\frac{\partial (GW)}{\partial x_2} \right)^2 Var(x_2) + \left(\frac{\partial (GW)}{\partial x_3} \right)^2 Var(x_3) \\ &+ \left(\frac{\partial (GW)}{\partial x_4} \right)^2 Var(x_4) + \left(\frac{\partial (GW)}{\partial x_5} \right)^2 Var(x_5) + \left(\frac{\partial (GW)}{\partial x_6} \right)^2 Var(x_6) \end{aligned} \tag{47}$$

The mean of Engine Weight (*EW*) can be calculated as expressed in Eqn. (48):

$$\begin{aligned} E(EW) = &\ 15374.58847 - 27.702079\ x_1 + 0.26187849\ x_2 - 2425.361607\ x_3 + 12.37410577\ x_4 \\ &- 389.3583447\ x_5 + 537.286141\ x_6 + 0.000104819\ x_1x_2 + 0.095950445\ x_1x_3 \\ &- 0.003058586\ x_1x_4 + 0.498878041\ x_1x_5 + 0.837557912\ x_1x_6 - 0.016045124\ x_2x_3 \\ &- 3.53203\ e - 05\ x_2x_4 - 0.003595284\ x_2x_5 - 0.003359929\ x_2x_6 - 0.68176651\ x_3x_4 \\ &+ 29.88198246\ x_3x_5 - 63.99813459\ x_3x_6 - 0.061435395\ x_4x_5 - 0.950124874\ x_4x_6 \\ &+ 5.613278991\ x_5x_6 + 0.095697511\ x_1^{\ 2} + 2.24406\ e - 07\ x_2^{\ 2} + 115.508073\ x_3^{\ 2} \\ &- 7.06903\ e - 05\ x_4^{\ 2} + 2.726825794\ x_5^{\ 2} + 27.5074007\ x_6^{\ 2} \end{aligned} \tag{48}$$

The variance of the Engine Weight (*EW*) can be calculated as expressed in Eqn. (49):

$$Var(EW) = \sum_{i=1}^{6} \left(\frac{\partial (EW)}{\partial x_i} \right)^2 Var(x_i) \tag{49}$$

The mean of Propellant Fraction Required (*PFR*) can be calculated as expressed in Eqn. (50):

$$\begin{aligned} E(PFR) = &\ 0.893093233 - 0.0002417266\ x_1 - 4.75121\ e - 07\ x_2 - 0.005464932\ x_3 + 0.000101483\ x_4 \\ &- 0.00322987\ x_5 - 0.00928557\ x_6 + 5.13829\ e - 10\ x_1x_2 - 5.46612\ e - 06\ x_1x_3 \\ &+ 6.76376\ e - 08\ x_1x_4 - 4.53026\ e - 06\ x_1x_5 + 3.30224\ e - 06\ x_1x_6 - 3.75202\ e - 08\ x_2x_3 \\ &+ 6.43768\ e - 10\ x_2x_4 - 9.55568\ e - 09\ x_2x_5 + 5.77463\ e - 08\ x_2x_6 - 1.43973\ e - 05\ x_3x_4 \\ &- 9.88908\ e - 05\ x_3x_5 - 0.0004773\ x_3x_6 + 1.13559\ e - 06\ x_4x_5 - 5.42466\ e - 06\ x_4x_6 \\ &+ 8.51662\ e - 05\ x_5x_6 + 6.80047\ e - 07\ x_1^{\ 2} + 8.9559\ e - 12\ x_2^{\ 2} - 0.00098585\ x_3^{\ 2} \\ &+ 1.88595\ e - 07\ x_4^{\ 2} + 0.000117169\ x_5^{\ 2} + 0.000959377\ x_6^{\ 2} \end{aligned} \tag{50}$$

The variance of the Propellant Fraction Required (*PFR*) can be calculated as expressed in Eqn. (51):

$$Var(PFR) = \sum_{i=1}^{6} \left(\frac{\partial (PFR)}{\partial x_i} \right)^2 Var(x_i) \tag{51}$$

The mean of Vehicle Length (*VL*) can be calculated as expressed in Eqn. (52):

$$\begin{aligned} E(VL) = {} & 823.1048661 - 1.411069734\, x_1 + 0.006369636\, x_2 - 90.3567667\, x_3 + 0.569733596\, x_4 \\ & - 25.55147283\, x_5 + 261.6412745\, x_6 - 1.06184\, e-06\, x_1x_2 + 0.066102241\, x_1x_3 \\ & - 7.43962\, e-05\, x_1x_4 + 0.013128748\, x_1x_5 - 0.074329589\, x_1x_6 - 0.000214937\, x_2x_3 \\ & - 9.37031\, e-07\, x_2x_4 - 6.43232\, e-05\, x_2x_5 + 0.001046781\, x_2x_6 - 0.040195575\, x_3x_4 \\ & + 1.323122789\, x_3x_5 - 14.70655866\, x_3x_6 - 0.004019241\, x_4x_5 - 0.01496916\, x_4x_6 \\ & - 1.387289781\, x_5x_6 + 0.003662151\, x_1^2 - 5.09359\, e-08\, x_2^2 + 4.70505342\, x_3^2 \\ & - 7.94011\, e05\, x_4^2 + 0.93934255\, x_5^2 + 0.009053289\, x_6^2 \end{aligned} \tag{52}$$

The variance of the Vehicle Length (*VL*) can be calculated as expressed in Eqn. (53):

$$Var(VL) = \sum_{i=1}^{6} \left(\frac{\partial (VL)}{\partial x_i} \right)^2 Var(x_i) \tag{53}$$

The mean of Vehicle Volume (*VV*) can be calculated as expressed in Eqn. (54):

$$\begin{aligned} E(VV) = {} & 54689.41793 - 71.0056374\, x_1 + 0.63271800\, x_2 - 6583.263146\, x_3 + 29.16000643\, x_4 \\ & - 1213.035573\, x_5 - 1813.643318\, x_6 - 0.000246467\, x_1x_2 + 3.774590948\, x_1x_3 \\ & - 0.01858828\, x_1x_4 - 0.142544958\, x_1x_5 - 0.043939802\, x_1x_6 - 0.034727527\, x_2x_3 \\ & + 0.00027002\, x_2x_4 - 0.004930386\, x_2x_5 + 0.002161457\, x_2x_6 - 1.819045749\, x_3x_4 \\ & + 64.33859515\, x_3x_5 - 102.110372\, x_3x_6 - 0.283472503\, x_4x_5 - 1.482888979\, x_4x_6 \\ & - 16.35284892\, x_5x_6 + 0.148782073\, x_1^2 - 1.26129\, e-06\, x_2^2 + 2.66.7127867\, x_3^2 \\ & - 0.009071094\, x_4^2 + 37.3538488\, x_5^2 + 281.5212617\, x_6^2 \end{aligned} \tag{54}$$

The variance of the Vehicle Volume (*VV*) can be calculated as expressed in Eqn. (55):

$$Var(VV) = \sum_{i=1}^{6} \left(\frac{\partial (VV)}{\partial x_i} \right)^2 Var(x_i) \tag{55}$$

The mean of Body Wetted Area (*BWA*) can be calculated as expressed in Eqn. (56):

$$\begin{aligned}E(BWA) = {} & 1894320.177 - 2535.255106\ x_1 + 18.24634963\ x_2 - 209077.2388\ x_3 + 1002.2601\ x_4 \\ & - 43645.41464\ x_5 + 39116.04318\ x_6 - 0.006531596\ x_1x_2 + 123.9268282\ x_1x_3 \\ & - 0.364359382\ x_1x_4 - 8.855865687\ x_1x_5 - 38.42451662\ x_1x_6 - 0.860599695\ x_2x_3 \\ & + 0.00322796\ x_2x_4 - 0.157168092\ x_2x_5 + 0.653315508\ x_2x_6 - 64.3040588\ x_3x_4 \\ & + 2275.808742\ x_3x_5 - 10585.43583\ x_3x_6 - 8.305159162\ x_4x_5 - 38.77262826\ x_4x_6 \\ & - 1106.758605\ x_5x_6 + 5.682243503\ x_1^2 - 6.48464e-05\ x_2^2 + 8896.931233\ x_3^2 \\ & - 0.292543211\ x_4^2 + 1404.757323\ x_5^2 + 8985.576363\ x_6^2\end{aligned} \tag{56}$$

The variance of the Body Wetted Area (*BWA*) can be calculated as expressed in Eqn. (57):

$$Var(BWA) = \sum_{i=1}^{6} \left(\frac{\partial (BWA)}{\partial x_i} \right)^2 Var(x_i) \tag{57}$$

Matlab solver „fmincon" is used for robust optimization using Eqn. (45) which will optimize the performance function by estimating the mean values of the design variables d_i. Here, „fmincon" uses „Active set algorithm" for optimization of the performance function. From the optimization, mean and standard deviation of the performance function are obtained for the values *0* to *1* of „*w*". The nonlinear robustness-based design optimization is obtained using maximum likelihood-based approaches (MLE) with the illustration of normal and bounded Johnson distribution which are compared with a decoupled method proposed by Zaman et al. (2011b) for robust design optimization.

The mean (μ_f) and standard deviation (σ_f) are estimated from the robustness-based design optimization using MLE for Normal distribution, MLE for bounded Johnson distribution and based on decoupled approach (Zaman et al. 2011b) given in Table 5.15. Form Table 5.15, it is easy to compare the robust design optimization for the different values of „*w*". For $w = 0$, the robust design for MLE-based robust optimization using normal distribution is better than the other two methods with minimum standard deviation as it is the target to minimize the standard deviation with optimizing the mean value. For $w = 0$, the MLE-based robust optimization using bounded Johnson distribution is better method than the decoupled approach based robust design. The MLE-based robust optimization using bounded Johnson distribution contains smaller deviation than the decoupled approach based robust design. For $w = 1$, the decoupled-based robust design is better than the both MLE-based robust optimization with minimum standard deviation. For *w* = *1*, the MLE-based robust optimization using normal distribution is better than the MLE-based robust optimization using bounded Johnson distribution which has maximum standard deviation. However, for *w* = *0.2, 0.4, 0.6 and 0.8*, the MLE-based robust optimizations are better than the decoupled approach for robust design optimization which has bigger standard deviations. It is easier to observe the differences in the Fig. 5.18.

Table 5.13: Mean values (μ_f) and standard deviations (σ_f) from robust optimization using MLE and decoupled approaches in TSTO problem

w	Moments of performance function (μ_f, σ_f)		
	MLE (Normal distribution)	MLE (bounded Johnson distribution)	Decoupled approach (Zaman et al. 2011b)
0.0	1.3322e+05, 1.0616e+04	1.3751e+05, 1.1091e+04	1.3771e+05, 1.2335e+04
0.2	1.3153e+05, 1.0636e+04	1.3520e+05, 1.1118e+04	1.3448e+05, 1.2373e+04
0.4	1.2955e+05, 1.0723e+04	1.3003e+05, 1.1871e+04	1.2905e+05, 1.2904e+04
0.6	1.2417e+05, 1.1879e+04	1.2460e+05, 1.2449e+04	1.2331e+05, 1.3518e+04
0.8	1.2002e+05, 1.2839e+04	1.2031e+05, 1.3431e+04	1.1834e+05, 1.4776e+04
1.0	1.1086e+05, 1.7156e+04	1.0973e+05, 1.7489e+04	1.1447e+05, 1.6532e+04

Robustness-based design optimization using MLE for Normal distribution and MLE for bounded Johnson distribution and decoupled approach for robust design optimization (Zaman et al. 2011b) for upper stage design problem of TSTO vehicle are shown in Fig. 5.18.

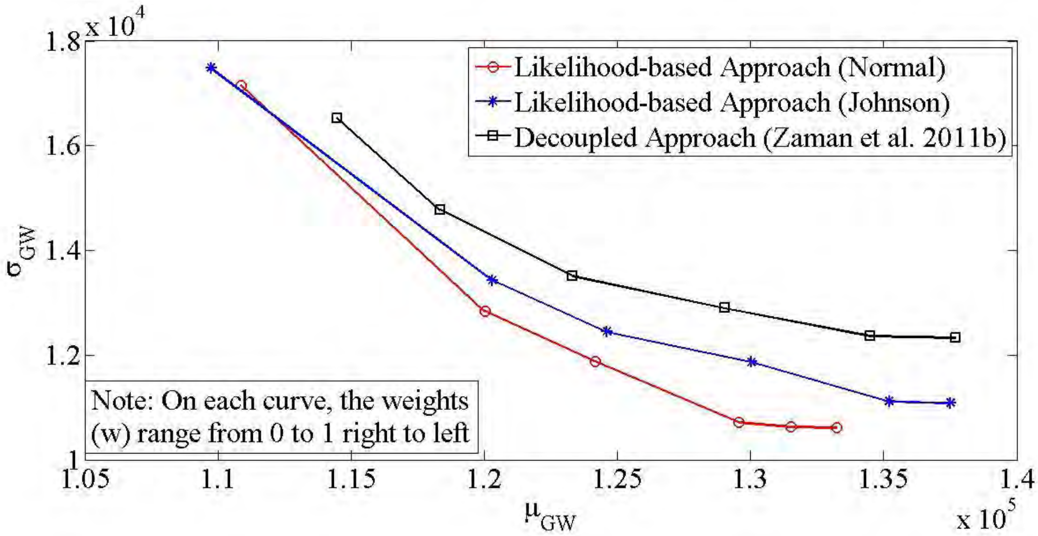

Fig. 5.18 Robust design optimization (RBDO) for MLE (Normal distribution), MLE (bounded Johnson distribution) and decoupled approaches for TSTO problem

CHAPTER VII

CONCLUSION AND RECOMMENDATION

7.1. Conclusion

This research proposed a likelihood-based approach to representation of epistemic uncertainty arising from multiple interval data. The proposed epistemic uncertainty representation methodology is then used in the design optimization framework to solve robustness-based design optimization problem under epistemic uncertainty. The proposed methodology is illustrated for numerical examples: a general mathematical problem and the upper level conceptual design of a TSTO vehicle. The outcomes of this research are mentioned below:

i. A maximum likelihood estimation-based methodology is developed for uncertainty representation with multiple interval data. This methodology is applicable to any probability distributions.

ii. A single loop robustness-based design optimization framework is developed using the proposed likelihood-based uncertainty representation for the non-design epistemic variables.

The proposed likelihood-based approach is verified using empirical cumulative distribution functions (CDF). Empirical CDFs are plotted for the parameters estimated from both normal and bounded Johnson distributions using the proposed likelihood-based approach and moment bounding algorithms developed in Zaman et al (2011a). Considerable agreement is found in the results obtained using the both approaches.

The proposed likelihood-based approach of epistemic uncertainty representation is able to estimate the parameters of any distributions including normal and bounded Johnson. The estimated parameters are used to calculate moments of the distributions. The information about the moments of the non-design epistemic variables has provided us the opportunity to go for robust design optimization. This likelihood-based approach extends the opportunity to obtain specific values of the parameters or moments from interval data rather than taking upper bound value from the moment for conservative robustness-based design optimization.

The proposed methodology of robust design optimizes the mean of the performance function and minimizes the variance or standard deviation of the performance function. The proposed MLE-based robust design optimization models are better than the established robust optimizations methods with respect to efficiency and computational expense. The proposed MLE-based robust optimization can provide most probable solution of the robust design where decoupled approach provides the worst case scenario due to use of upper bound of the moment (standard deviation). This robust optimization reduces the computational cost through minimizing the computational complexity.

From the MLE-based robust design optimization of normal distribution and bounded Johnson distribution, it can be observed for the mathematical example in Figs. 5.9 and 5.10, respectively that the mean of the performance function is optimized with the minimization of the standard deviation. The robust design optimization based on MLE approach of bounded Johnson distribution is better than the MLE-based approach of normal distribution which is shown in Fig. 5.11. The both proposed robust design optimizations on MLE-based approaches are better than a well established method called decoupled approach with respect to variation. In Figs. 5.12 and 5.13, comparisons are shown between the decoupled method and the MLE-based robust design optimizations. The MLE-based robust optimizations are also applied to an engineering design problem of TSTO vehicle. The mean of the performance function is optimized with the minimization of the standard deviation in the robust design optimization for both distributions which are shown in Figs. 5.18. The comparison of the MLE-based approaches with well established decoupled method is show in Fig. 5.18. It can be deduced easily that the MLE-based robust optimization for normal distribution is better than the MLE-based robust optimization for bounded Johnson distribution. The both MLE-based robust optimization methods are better in most cases than the decoupled method for robustness-based design optimization.

This proposed formulation is an approach of robust optimization which can also be applied to develop more multi-objective optimization techniques or methods. It is always difficult to gather precise data from the experiment. Point data has a severe fatality of loosing data due to slight deviation of a system or little fault in the system. To accumulate all the data and information, it is rather safe to collect the data in the form of intervals. It is also possible to gather more and more data through different experiments. However, it will increase the cost though reduce uncertainty. This proposed robust design optimization for multiple intervals assists to reduce the cost associated with system design and analysis compare to other methods.

7.2. Recommendation for future research needs

The method developed in this study is a single disciplinary system, can be used in the multidisciplinary system. The design variables and the input random variables of a problem might have epistemic uncertainty with interval uncertainty. The designer does not have any control on the non-design epistemic variables. Therefore, the design methodology has to employ a search among the possible values of such epistemic variables to find an optimal solution. It became a conservative design. Multidisciplinary system analysis is a system which leads to use of individual disciplinary analysis codes that interact with each other through shared input and output data. A feasible multidisciplinary optimization solution simultaneously satisfies all individual disciplinary constraints. Therefore, it becomes more complex to have a multidisciplinary solution based on interval uncertainty. This MLE-based robust optimization may extend in application for multi-disciplinary robustness-based design optimization which may ensure to reduce the design cost. The proposed uncertainty representation and design optimization methodology can easily be extended to solve problems in other domains such as financial engineering and network optimization.

REFERENCES

Agarwal H, Renaud JE, Preston EL, Padmanabhan D. (2004) *Uncertainty quantification using evidence theory in multidisciplinary design optimization*. Reliability Engineering and System Safety; 85(1–3):281–94.

Agarwal H, Renaud JE, Lee J., Watson LT (2004) *A Unilevel Method for Reliability Based Design optimization,* 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Material Conference, Palm Spring, California.

Agarwal H, Mozumder CK, Renaud JE, Watson LT (2007) *An inversemeasure- based unilevel architecture for reliability-based design optimization.* Struct Multidisc Optim 33:217–227

Apeland S, Aven T, Nilsen T. (2002) *Quantifying uncertainty under a predictive, epistemic approach to risk analysis.* Reliability Engineering and System Safety; 75:93–102.

Baudrit C, Dubois D. (2006) *Practical representations of incomplete probabilistic knowledge.* Computational Statistics & Data Analysis;51:86–108.

Ben-Haim Y, Elishakoff I. (1990) *Convex models of uncertainty in applied mechanics*. Studies in Applied Mechanics:25.

Ben-Haim Y. (2004) *Uncertainty, probability and information-gaps.* Reliability Engineering and System Safety ;85:249–66.

Berleant D. (1993) *Automatically verified reasoning with both intervals and probability density functions.* Interval Computations :48–70.

Berleant D. (1996) *Automatically verified arithmetic on probability distributions and intervals.* In: Kearfott B, Kreinovich V, editors. Applications of interval computations. Kluwer Academic Publishers; pp. 227–44.

Berleant D, Goodman-Strauss C. (1998) *Bounding the results of arithmetic operations on random variables of unknown dependency using intervals*. Reliable Computing ;4:147–65.

Berleant D, Zhang J. (2004) *Representation and problem solving with distribution envelope determination (DEnv).* Reliability Engineering and System Safety ;85:153–68

Bertrand P, Groupil F. (2000) *Descriptive statistics for symbolic data*. In: Bock HH, Diday E, editors. Analysis of symbolic data. Berlin: Springer.

Bichon BJ, McFarland JM, Mahadevan S (2008) *Using Bayesian inference and efficient global reliability analysis to explore distribution uncertainty*. In: 49th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, Schaumburg, IL

Bras BA, Mistree F (1993) *Robust design using compromise decision support problems.* Eng Optim 21:213–239

Bras BA, Mistree F (1995) *A compromise decision support problem for robust and axiomatic design.* ASME J Mech Des 117(1): 10–19

Chen W, Allen JK, Mistree F, Tsui K-L (1996) *A procedure for robust design: minimizing variations caused by noise factors and control factors*. ASME J Mech Des 118:478–485

Chen X. C., Hasselman T. K., Neill D. J. (1997) *Reliability based structural design optimization for practical applications*. In Proceedings of the 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, number AIAA-97-1403, pages: 2724-2732.

Chen W, Wiecek MM, Zhang J (1999) *Quality utility—a compromise programming approach to robust design.* J Mech Des (ASME) 121:179–187

Chen W, Sahai A, Messac A, Sundararaj GJ (2000) *Exploration of the effectiveness of physical programming in robust design.* J Mech Des 122:155–163

Chen W., Du X. (2002) *Sequential optimization and reliability assesment method for efficient probabilistic design*. In ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, number DETC2002/DAC-34127, Montreal, Canada.

Cheng H, Sandu A (2009) *Efficient uncertainty quantification with the polynomial chaos method for stiff systems.* Math Comput Simul 79:3278–3295

Chiralaksanakul A, Mahadevan S (2005) *First-order approximation methods in reliability-based design optimization.* J Mech Des 127(5):851–857

Dai Z, Mourelatos ZP (2003) *Incorporating epistemic uncertainty in robust design*. In: Proceedings of DETC, 2003 ASME design engineering technical conferences, 2–6 September 2003, Chicago, Illinois, USA

DeBrota DJ, Swain JJ, Roberts SD, Venkataraman S (1988) *Input modeling with the Johnson system of distributions.* In: Proceedings of the 1988 winter simulation conference

De Cooman G, Troffaes MCM. (2004) *Coherent lower previsions in systems modeling: products and aggregation rules.* Reliability Engineering and System Safety 2004;85:113–34

Doltsinis I, Kang Z (2004) *Robust design of structures using optimization methods*. Comput Methods ApplMechEng 193:2221– 2237

Draper, J. (1952) *Properties of distributions resulting from certain simple transformations of the normal distribution*. *Biometrika*, *39*, 290-301.

Dubois D, Prade H. (1988) *Possibility theory: an approach to computerized processing of uncertainty.* 1st ed. Plenum Press.

Du X, Chen W (2000) *Towards a better understanding of modeling feasibility robustness in engineering.* ASME J Meach Des 122(4):385–394

Du X, Chen W. (2000) *An integrated methodology for uncertainty propagation and management in simulation-based systems design*. AIAA Journal; 38(8):1471–8.

Du X, Sudjianto A, Chen W (2004) *An integrated framework for optimization under uncertainty using inverse reliability strategy*. ASME

Du X, Sudjianto A, Huang B. (2005) *Reliability based design with mixture of random and interval variables*. Journal of Mechanical Design;127:1068–76.

Du X, Beiqing Huang B (2007) *Reliability-based design optimization with equality constraints*. Int J Numer Methods Eng 72:1314– 1331

Edwards A W F (1972) *Likelihood*, 1st edition, Cambridge: Cambridge University Press.

Enevoldsen I., Sorensen J. D. (1994) *Reliability-based optimization in structural engineering*. Structural Safety, 15(3): 169-196.

Ferson S, Joslyn CA, Helton JC, Oberkampf WL, Sentz K. (2004) *Summary from the epistemic uncertainty workshop: consensus amid diversity*. Reliability Engineering and System Safety;85:355–69.

Ferson S, Hajagos J. (2004) *Arithmetic with uncertain numbers: rigorous and (often) best-possible answers.* Reliability Engineering and System Safety;85:135–52.

Ferson, S, Kreinovich, V, Hajagos, J, Oberkampf W, Ginzburg, L. (2007) *Experimental uncertainty estimation and statistics for data having interval uncertainty*. Sandia National Laboratories Technical report SAND2007-0939, Albuquerque, NM.

Fetz T, Oberguggenberger M. (2004) *Propagation of uncertainty through multi-variate functions in the framework of sets of probability measures*. Reliability Engineering and System Safety;85:73–87.

Fisher R A (1921) *On the probable error of a coefficient of correlation deduced from a small sample*, Metron, 1:3-32

Fisher R A (1925) *Theory of statistical estimation,* Proceedings of the Cambridge Philosophical Society, 22:700-725

George, F., Ramachandran, K. M., Lihua, L. (2009) *Gene selection with Johnson distribution*. *Journal of Statistical Research*, *43*, 117-125.

George F., Ramachandran K. M. (2011) *Estimation of parameters of Johnson's system of distributions*, Journal of Modern Applied Statistical Methods (JMASM), Volume 10, Issue 2, Article 9, pp: 494-504

Ghanem R, Spanos P (1991) *Stochastic finite elements: a spectral approach*. Springer, New York

Guo J, Du X. (2007) *Sensitivity analysis with mixture of epistemic and aleatory variables*. AIAA Journal;45:9.

Guo J, Du X. (2009) *Reliability sensitivity analysis with random and interval variables*. International Journal for Numerical Methods in Engineering.

Hailperin T. (1986) *Boole's logic and probability*. Amsterdam: North-Holland

Haldar A, Mahadevan S (2000) *Probability, reliability and statistical methods in engineering design*. Wiley

Hall JW, Lawry J. (2004) *Generation, combination and extension of random set approximations to coherent lower and upper probabilities*. Reliability Engineering and System Safety;85:89–101.

Hahn, J. G., & Shapiro S. S. (1967) *Statistical models in engineering*. John Wiley & Sons, New York.

Hofer E, Kloos M, Hausmann BK, Peschke J, Woltereck M. (2002) *An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties*. Reliability Engineering and System Safety2002;77:229–38.

Helton JC, Johnson JD, Oberkampf WL. (2004) *An exploration of alternative approaches to the representation of uncertainty in model predictions*. Reliability Engineering and System Safety;85:39–71.

Helton JC, Johnson JD, Oberkampf WL, Storlie CB. (2007) *A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory*. Comput. Methods Appl. Mech. Eng.;196(37–40):3980–98.

Helton, JC, Johnson JD, Oberkampf WL, Sallaberry CJ. (2008) *Representation of analysis results involving aleatory and epistemic uncertainty*, Sandia Report, SAND2008-4379.

Hill, I. D., Hill, R., & Holder, R. L. (1976) *Fitting Johnson curvesby moments*. *Applied Statistics*, *25* 180-189.

Hong HP (1998) *An efficient point estimate method for probabilistic analysis*. ReliabEngSystSaf 59:261–267

Howson C, Urbach P. (1993) *Scientific reasoning: the Bayesian approach*. 2nd ed. Chicago, IL: Open Court.

Huang B, Du X (2007) *Analytical robustness assessment for robust design*. Struct Multidisc Optim 34:123–137

Hyman, (1982) *JM.FOR SIG: annex tension of FORTRAN with significance arithmetic*. Report LA-9448-MS, Los Alamos National Laboratory, Los Alamos, NM.

Jeffreys H (1934) *Probability and Scientific Method*, Proceedings of the Royal Society, Ser. A, 146:9-16

Johnson N. L., Kitchen J. O. (1971) *Tables to facilitate fitting S_B curves II: both terminals known*. Biometrica, 58, pp. 657-668.

Klir G. J., Wierman M. J. (1998) *Uncertainty Based Information: Elements of Generalized Information Theory*. Physica-Verlag.

Klir GJ. (2004) *Generalized information theory: aims, results, and open problems*. Reliability Engineering and System Safety;85:21–38.

Kozine IO, Utkin LV. (2004) *An approach to combining unreliable pieces of evidence and their propagation in a system response analysis*. Reliability Engineering and System Safety;85:103–12.

Kuschel N. and Rackwitz R. (2000) A new approach for structural optimization of series systems. *Applications of Statistics and Probability*, 2(8): 987-994.

Langley RS. (2000) *Unified approach to probabilistic and possibilistic analysis of uncertain systems*. Journal of Engineering Mechanics; 126(11):1163–72.

Lee K-H, Park G-J (2001) *Robust optimization considering tolerances of design variables*. ComputStruct 79:77–86

Lee I, Choi KK, Du L, Gorsich D (2008) *Dimension reduction method for reliability-based robust design optimization*.ComputStruct 86(13–14):1550–1562

Marler RT, Arora JS (2004) *Survey of multi-objective optimization methods for engineering*. Struct Multidisc Optim 26:369– 395

Martin J. D. (2007) *Using Maximum Likelihood Estimation to Estimate Kriging Model Parameters*, ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Volume 6, Paper No. DETC2007-34662, pp. 663-673; 11 pages, doi:10.1115/DETC2007-34662

Mavrotas G (2009) *Effective implementation of the e-constraint method in multi objective mathematical programming problems*. Appl Math Comput 213:455–465

Messac A (1996) *Physical programming effective optimization for computational design*. AIAA J 34(1):149–158

Messac A, Melachrinoudis E, Sukam CP (2001) *Mathematical and pragmatic perspectives of physical programming.* AIAA J 39(5):885–893

Messac A, Ismail-Yahaya A (2002) *Multiobjective robust design using physical programming*. Struct Multidisc Optim 23:357– 371

Modisette J. P. (2012) *Maximum Likelihood Approach to State Estimation in Online Pipeline Models*, 9th International Pipeline Conference, Paper No. IPC2012-90664, pp. 813-824; 12 pages, doi:10.1115/IPC2012-90664

Monaco J. F., Kidman D. S., Malloy D. J., Ward D. G., Gist J. F. IV (2008) *Automated Methods to Calibrate a High-Fidelity Thrust Deck to Aid Aeropropulsion Test and Evaluation*, ASME Turbo Expo 2008: Power for Land, Sea, and Air, Paper No. GT2008-50213, pp. 41-54; 14 pages, doi:10.1115/GT2008-50213

Oberkampf, WL, Helton, JC, Sentz, K, (2001) *Mathematical representation of uncertainty*. In: Proceedings of the 42nd AIAA/ASME/ASCE/AHS/ASC struc- tures, structural dynamics, and materials conference & exhibit, Seattle, WA, Paper number AIAA 2001-1645.

Oberkampf WL, Helton JC, Joslyn CA, Wojtkiewicz SF, Ferson S (2004) *Challenge problems: uncertainty in system response given uncertain parameters*. Reliab Eng Syst Saf 85:11–19

O'Hagan A, Oakley JE. (2004) *Probability is perfect, but we can't elicit it perfectly*. Reliability Engineering and System Safety;85:239–48.

Park G-J, Lee T-H, Lee KH, Hwang K-H (2006) *Robust design: an overview*. AIAA J 44(1):181–191

Parkinson A, Sorensen C, PourhassanN (1993) *A general approach for robust optimal design*. Trans ASME 115:74–80

Parsons S. (2001) *Qualitative Methods for Reasoning under Uncertainty*, The MIT Press.

Pawitan Y (2001) *In all likelihood: statistical modeling and inference using likelihood*, New York: Oxford Science Publications.

Penmetsa RC, Grandhi RV. (2002) *Efficient estimation of reliability for problems with uncertain intervals*. Computers and Structures; 80(12):1103–112.

Rackwitz R. (2001) *Reliability analysis-a review and some perspectives.* Structural Safety, 23(4): pp. 365-395.

Rahman S, Xu H (2004) *A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics*. Probab Eng Mech 19:393–408

Ramu P, Qu X, Youn BD, Haftka RT, Choi KK (2006) *Inverse reliability measures and reliability-based design optimization*. Int J ReliabSaf 1(1/2):187–205

Rao SS, Annamdas KK. (2009) *An evidence-based fuzzy approach for the safety analysis of uncertain systems*. In: 50th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, Palm Springs, California, 2009, Paper number AIAA-2009-2263.

Rao SS, Cao L. (2002) *Optimum design of mechanical systems involving interval parameters*. Journal of Mechanical Design;124(3):465–72.

Red-Horse JR, Benjamin AS. (2004) *A probabilistic approach to uncertainty quantification with limited information*. Reliability Engineering and System Safety;85:183–90.

Robert CP, Casella G (2004) *Monte Carlo statistical methods*, 2nd edition. Springer, New York

Rosenblueth E (1975) *Point estimates for probability moment*. Proc Natl Acad Sci U S A 72(10):3812–3814

Rutherford B. (2004) *A response-modeling approach to characterization and propagation of uncertainty specified over intervals*. Reliability Engineering and System Safety 2004;85:211–22.

Shafer Glenn (1976) *A mathematical theory of evidence*, Princeton University Press; 0-608-02508-9.

Sankararaman S, Mahadevan S (2011) *Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data*, Reliability Engineering and System Safety, 96 (2011) pp. 814-824

Sankararaman S, Mahadevan S (2012) *Likelihood-Based Approach to Multidisciplinary Analysis Under Uncertainty*, Reliability Engineering & System Safety, Volume 134, Issue 3, March 2012

Sankararaman S, Mahadevan S (2013) *Distribution type uncertainty due to sparse and imprecise data*, Mechanical Systems and Signal Processing, Volume 37, Issues 1–2, May–June 2013, Pages 182–198

Sim M (2004) *Robust optimization*. PhD dissertation submitted to the Sloan School of Management, Massachusetts Institute of Technology.

Slifker, J., & Shapiro, S. (1980) *The Johnson system: Selection and parameter estimation*. Technometrics, *22*, 239-247.

Steven D. Gribble (2001) *Robustness in Complex Systems*. Proceedings of the 8th Workshop on Hot Topics in Operating Systems (HotOS-VIII)

Stevenson MD, Hartong AR, Zweber JV, Bhungalia AA, Grandhi RV (2002) *Collaborative design environment for space launch vehicle design and optimization*. In: Paper presented at the RTO AVT symposium on "reduction of military vehicle acquisition time and cost through advanced modeling and virtual simulation", held in Paris, France

Sundaresan S, Ishii K, Houser DR (1995) *A robust optimization procedure with variations on design variables and constraints*. Eng Optim 24(2):101–117

Taguchi G (1993) *Taguchi on robust technology development: bringing quality engineering upstream*. ASME, New York

Waals O. J., Aalbers A. B., Pinkster J. A. (2002) *Maximum Likelihood Method as a Means to Estimate the Directional Wave Spectrum and the Mean Wave Drift Force on a Dynamically Positioned Vessel*, ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering, Volume 4, Paper No. OMAE2002-28560, pp. 605-613; 9 pages, doi:10.1115/OMAE2002-28560

Wang L., Kodiyalam S. (2002) An efficient method for probabilistic and robust design with non-normal distribution. In Proceedings of the 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, number AIAA 2002-1754, Denver, Colorado.

Wei DL, Cui ZS, Chen J (2009) *Robust optimization based on a polynomial expansion of chaos constructed with integration pointn rules*. J MechEngSci 223(5):1263–1282 (Part C)

Wheeler, R. (1980). Quantile Estimators of Johnson curve Parameters. *Biometrika*, *67*, 725-728.

Williamson RC, Downs T. (1990) *Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds*. International Journal of Approximate Reasoning 1990; 4: 89–158.

Xiao J., Bohdan Kulakowski (2003) *Hybrid Genetic Algorithm: A Robust Parameter Estimation Technique and Its Application to Heavy Duty Vehicles*, ASME 2003 International Mechanical Engineering Congress and Exposition, Paper No. IMECE:2003-41934, pp. 1087-1096; 10 pages, doi:10.1115/IMECE2003-41934

Xu H, Rahman S (2004) *A generalized dimension-reduction method for multidimensional integration in stochastic mechanics*. Int J Numer Methods Eng 61(12):1992–2019

Youn BD, Choi KK, Du L (2007) *Integration of possibility-based optimization and robust design for epistemic uncertainty*. ASME J Mech Des 129(8):876–882

Yu J-C, Ishii K (1998) *Design for robustness based on manufacturing variation patterns*. Trans ASME 120:196–202

Zaman K (2010) *Modeling and Management of Epistemic Uncertainty for Multidisciplinary system Analysis and Design*, PhD thesis, Vanderbilt University.

Zaman K, Rangavajhala, S, McDonald, M., and Mahadevan, S, (2011) *A probabilistic approach for representation of interval uncertainty*, Reliability Engineering & System Safety - RELIAB ENG SYST SAFETY , vol. 96, no. 1, pp. 117-130

Zaman K, McDonald, M., Mahadevan, S., Green, L. (2011) *Robustness-based design optimization under data uncertainty*. Structural and Multidisciplinary Optimization-STRUCT MULTIDISCIP OPTIM, vol. 44, no. 2, pp. 183-197

Zaman K, Mahadevan S (2013) *Robustness-Based Design Optimization of Multidisciplinary System Under Epistemic Uncertainty*, AIAA Journal, Vol. 51, No. 5, pp. 1021-1031.

Zeleny M (1973) *Compromise programming. In: Cochrane JL, Zeleny M (eds) Multiple criteria decision making.* University of South Carolina Press, Columbia, SC, pp 262–301

Zhang WH (2003) *A compromise programming method using multibounds formulation and dual approach for multicriteria structural optimization*. Int J Numer Methods Eng 58:661–678

Zhao Y-G, Ono T (2000) *New point estimates for probability moments*. J EngMech 126(4):433–436

Zhao Y-G, Ang AH-S (2003) *System reliability assessment by method of Moments*. J StructEng 129(10):1341–1349

Zou T, Mahadevan S (2006) *Versatile formulation for multiobjective reliability-based design optimization*. J Mech Des 128:1217

APPENDIX

A.1. *Johnson family of distributions*

Any data set with finite moments can be fitted by a member of the Johnson families such as S_B, S_U or S_L. The most commonly used methods to estimate the parameters of the Johnson distribution are the Percentile approach (Slifker and Shapiro 1980) and Quantile method (Wheeler 1980). Large volume of researches have been done to estimate the parameters of the Johnson distribution which can be find in Drapper (1952), Hill et al. (1976), Hahn and Shapiro (1967), George et al. (2009).

If, there is a continuous random variable *X* whose distribution is unknown and is to be estimated, Johnson proposed three normalizing transformations (George and Ramachandran 2011) having the general form:

$$Z = \gamma + \delta f\left(\frac{X-\xi}{\lambda}\right)$$

Where *f(.)* denotes the transformation function, *Z* is a standard normal variable, γ and δ are shape parameters, λ is the scale parameter and ξ is a location parameter. Without the loss of generality, it is assumed that $\delta>0$ and $\lambda>0$. The first transformation proposed by Johnson defined the lognormal system of distributions denoted by S_L:

$$\begin{aligned} Z &= \gamma + \delta \log\left(\frac{X-\xi}{\lambda}\right), \quad X > \xi \\ &= \gamma^* + \delta \log(X-\xi), \quad X > \xi \end{aligned}$$

The S_L curves cover the lognormal family.

The bounded system of distributions S_B is defined by

$$Z = \gamma + \delta \log\left(\frac{X-\xi}{\xi+\lambda-X}\right), \quad \xi < X < \xi + \lambda$$

S_B curves cover bounded distribution. The distributions can be bounded on the lower end, the upper end or the both ends. This family covers Gamma distributions, Beta distributions and many others.

The unbounded system of distributions S_U is defined by

$$Z = \gamma + \delta \log\left[\left(\frac{X-\xi}{\lambda}\right) + \left\{\left(\frac{X-\xi}{\lambda}\right)^2 + 1\right\}^{1/2}\right], \quad -\infty < X < \infty$$

$$= \gamma + \delta \sinh^{-1}\left(\frac{X-\xi}{\lambda}\right)$$

The S_U curves are unbounded and cover the t and normal distributions, among others. After the transmission, Z follows standard normal distribution, the probability density function (pdf) of each of the family of the Johnson distribution can be derived. If X follows the Johnson distribution and $Y = \left(\frac{X-\xi}{\lambda}\right)$ then, for S_L family, the probability density function (pdf) is:

$$p(y) = \frac{\delta}{\sqrt{2\pi}} \times \frac{1}{y} \times \exp\left\{-\frac{1}{2}[\gamma + \delta \log(y)]^2\right\}, \quad \xi < X < +\infty$$

Similarly, for the S_B family, the pdf is:

$$p(y) = \frac{\delta}{\sqrt{2\pi}} \times \frac{1}{[y(1-y)]} \times \exp\left\{-\frac{1}{2}\left[\gamma + \delta \log\left(\frac{y}{1-y}\right)\right]^2\right\}, \quad \xi < X < +\xi + \lambda$$

The pdf for the S_U family is

$$p(y) = \frac{\delta}{\sqrt{2\pi}} \times \frac{1}{\sqrt{y^2+1}} \times \exp\left\{-\frac{1}{2}\left[\gamma + \delta \log(y + \sqrt{y^2+1})\right]^2\right\}, \quad -\infty < X < +\infty$$

DeBrota et al. (1988) presents four methods to estimate the Johnson"s parameters. The first method is the moment matching method. This method involves solving a set of four nonlinear equations that equate the first four moments calculated from the given data with those of a Johnson distribution, which are calculated as the function of unknown parameters. The second method is known as the percentile matching, where the parameters are estimated by solving a system of nonlinear equations equating four percentile points of the data and the Johnson distribution. The third method involves a least square estimation of the parameters obtained by minimizing the sum of the squared errors in the percentile values from the data and those from the Johnson distribution. The fourth approach involves minimizing the error norm of the Johnson distribution CDF when compared with the empirical CDF constructed from the data.

Hill et. al (1976) designed an algorithm to estimate γ, δ, λ and ξ matching the first four moments of *X*. Johnson and Kitchen (1971) developed a table which helps to find γ and δ. This method provides the opportunity to match the first two given moments. The four moments provide us better fitting than using two moments. However, during the use of sample values, the two moment method is considered as better process.