

Research Article

An Optimization-Based Approach to Calculate Confidence Interval on Mean Value with Interval Data

Kais Zaman and Saraf Anika Kritee

Department of Industrial and Production Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh

Correspondence should be addressed to Kais Zaman; kaiszaman@yahoo.com

Received 5 February 2014; Revised 28 June 2014; Accepted 29 June 2014; Published 13 July 2014

Academic Editor: Ferrante Neri

Copyright © 2014 K. Zaman and S. A. Kritee. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we propose a methodology for construction of confidence interval on mean values with interval data for input variable in uncertainty analysis and design optimization problems. The construction of confidence interval with interval data is known as a combinatorial optimization problem. Finding confidence bounds on the mean with interval data has been generally considered an NP hard problem, because it includes a search among the combinations of multiple values of the variables, including interval endpoints. In this paper, we present efficient algorithms based on continuous optimization to find the confidence interval on mean values with interval data. With numerical experimentation, we show that the proposed confidence bound algorithms are scalable in polynomial time with respect to increasing number of intervals. Several sets of interval data with different numbers of intervals and type of overlap are presented to demonstrate the proposed methods. As against the current practice for the design optimization with interval data that typically implements the constraints on interval variables through the computation of bounds on mean values from the sampled data, the proposed approach of construction of confidence interval enables more complete implementation of design optimization under interval uncertainty.

1. Introduction

Uncertainty quantification plays an increasingly important role in assessing the performance, safety, and reliability of complex physical systems, even in the absence of adequate amount of experimental data for many applications. Uncertainty in engineering analysis and design arises from several different sources and must be quantified with accuracy for further analysis. Sources of uncertainty may be divided into two types: aleatory and epistemic. Aleatory uncertainty is irreducible. Examples include phenomena that exhibit natural variation like environmental conditions (temperature, wind speed, etc.). Manufacturing variations due to limited precision in tools and processes also result in this type of uncertainty. In contrast, epistemic uncertainty results from a lack of knowledge about the system, or due to approximations in the system behavior models, or due to limited or subjective (e.g., expert opinion that results in interval data) data; it can be reduced as more information about the system or the variables is obtained.

Epistemic uncertainty can be viewed in two ways. It can be defined with reference to a stochastic but poorly known quantity [1] or with reference to a fixed but poorly known physical quantity [2]. The term “stochastic but poorly known” refers to the uncertainty about the distribution type and parameters of a random variable, which implies that computing statistics and making inferences from data with this type of uncertainty need separate treatments than aleatory uncertainty. This paper focuses on handling the first definition of epistemic uncertainty, that is, epistemic uncertainty with reference to a stochastic but poorly known quantity resulting from interval data.

Interval data are encountered frequently in practical engineering problems. Several of such situations where interval data arise are discussed in [3–5], for example, a collection of expert opinions, which specify a range of possible values for a random variable. As mentioned in [5], when data is available in multiple intervals (e.g., given by multiple experts), the information contained in one interval could contradict that in the other interval(s) or could be contained by other

interval(s). In this context, intervals can be broadly categorized as *nonoverlapping* and *overlapping* intervals. The methods developed in this paper are equally applicable to any type of multiple interval data, that is, nonoverlapping, overlapping, or mixed intervals (a combination of the former two).

There exists an extensive volume of literature that presents efficient probabilistic as well as nonprobabilistic methods to treat interval data in uncertainty analysis problems (e.g., [1–12]). Zaman et al. [5] proposed efficient probabilistic uncertainty representation and propagation methods for interval data that suggest that a unified framework for representing all forms of uncertainty is necessary to avoid computationally involved nested analysis. Methods for robustness-based design optimization under interval uncertainty are available in the literature (e.g., [13, 14]), which implement the constraints on interval variables through the computation of bounds on mean values from the sampled data. However, for more complete implementation of design optimization under interval data, it is necessary to implement such constraints on interval variables through the construction of confidence intervals.

A confidence interval is a range of values in which an unknown population parameter (e.g., the mean, variance, etc.) may be located. This range is estimated from a given set of sampled data. Confidence interval on the mean is an interval estimate of the population mean. Since it is not practical to collect information of the whole population, the population mean is estimated from the information on the sample mean and variance, which may introduce some error in the prediction of the population mean. Therefore, it is often desirable to calculate an interval for the population mean, the upper limit of which is called the upper confidence limit (UCL) and the lower limit is called the lower confidence limit (LCL). The confidence limits are usually estimated considering the sample size, the confidence level required, and the uncertainty in the random variable under consideration [15]; this uncertainty is quantified through the estimation of variance.

The most commonly used method for the construction of confidence interval on mean is the standard t interval, which is based on an asymptotic normality assumption [16]. This standard t interval results in large errors in the presence of skewness in the distribution. Johnson [17] proposed a modified t -statistic to reduce the effect of population skewness on the distribution of the t variable. It is less affected by asymmetry in the population (when the population is non-normal) when compared to the commonly used t statistic. Therefore, the resulting approximation in confidence bounds is a better approximation than what is available now, which is simply the use of the t -statistic. However, Hall [18] claimed that Johnson's transformation is not monotone and not invertible. He proposed empirical transformations which are able to correct both bias and skewness. Although there is now an extensive volume of different transformation-based approximation methods and their bootstrapped versions [16, 18, 19] available to obtain reliable estimates of confidence bounds on mean, all these methods have only been developed with respect to physical or natural variability described by point

data. An efficient approach to construction of confidence interval on mean for interval data is yet to be developed.

A few studies are reported in the literature to deal with the methods for constructing confidence interval on mean with interval data. Ferson et al. [3] discussed the methods of computing two confidence intervals for the mean of the interval data based on the assumption that the data come from a normal population, one for the lower bound on the mean called the lower confidence limit and the other for the upper bound on mean called the upper confidence limit. Although this concept of two confidence limits is useful for outlier detection problems, a single confidence interval that contains whole range of possible values for a variable described by interval data is necessary for many engineering problems (e.g., robust design optimization under interval uncertainty). It has been reported in the literature that the general problem of computing the outer bounds on the endpoints of confidence interval, that is, the upper bound on the upper confidence limit and the lower bound on the lower confidence limit, is an NP-hard problem [20]. Kreinovich et al. [20] developed feasible algorithms to compute the lower bound on the upper confidence limit and upper bound on the lower confidence limit, which are then termed as the inner bounds on the endpoints of the confidence interval. Kreinovich et al. [20] also developed efficient algorithms to find the upper bound on the upper confidence limit and the lower bound on the lower confidence limit for some special types of interval data. These approaches that calculate confidence bounds on mean combinatorially search for points within the intervals that minimize or maximize the confidence bounds. A major contribution of this paper is the development of algorithms based on continuous optimization methods that are valid for any type of interval data, that is, overlapping, nonoverlapping, or mixed intervals. This paper develops efficient methods for the construction of confidence interval on mean with interval data that is computationally tractable and also can ensure rigorous confidence bounds on mean values.

In this paper, we develop two different approaches for calculating confidence intervals for interval data. The first approach assumes that the moments of data are independent of each other. This approach is based on the methods of calculating the bounds on moments for interval data developed in [5]. The second approach preserves the dependence among moments of interval data. As opposed to computationally expensive heuristic approach discussed in [3], we have used an optimization-based approach to calculate the confidence interval. With numerical examples, we have shown that the proposed confidence bound algorithms are scalable in polynomial time with respect to increasing number of intervals. Unlike [3], we have used Johnson's modified t -statistic to calculate the outer confidence interval on mean. As mentioned earlier, this modified statistic can reduce the effect of population skewness on the distribution of the t -variable. Therefore, the resulting approximation in confidence bounds is a better approximation than simply the use of the t -statistic.

The remainder of the paper is organized as follows. Section 2 describes the proposed methodology for construction of confidence interval with interval data. Section 3 illustrates

the proposed developments using different examples of interval data, where comparisons with alternate approaches, such as the one developed in [3], are made. Section 4 concludes the paper with summary and future work.

2. Construction of Confidence Interval with Interval Data

This section discusses the proposed algorithms that estimate the upper and lower bounds of confidence interval on mean for interval data. We propose two approaches to find the confidence interval on mean. The first approach uses the moment bounding methods developed in [5] and is able to produce rigorous confidence bounds. This approach does not consider the dependence among moments of interval data. The second approach preserves the dependence among moments of interval data and is able to produce optimal confidence bounds. Both approaches use Johnson's modified t -statistic [17] to construct confidence interval on mean values.

2.1. Rigorous Confidence Interval Formulations. The t -distribution was proposed by William S. Gossett under the pseudonym of "Student" [15], which assumes that all possible samples are drawn from a normal population. Johnson [17] proposed a modified t -statistic to reduce the effect of population skewness on the distribution of the t -variable. It is less affected by asymmetry in the population (when the population is non-normal) when compared to the commonly used t -statistic. As the input variables are described by the interval data, it is possible that the underlying distributions of the variables might have major deviations from normality. This modified statistic takes into account the skewness of the distribution and thus provides a better estimate of the confidence bounds in the presence of interval data.

Johnson's modified t -statistic [17] is used to construct the confidence bounds on means of the input variables described by point data as follows:

$$\begin{aligned} Z_l &= \bar{z} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} + \frac{\mu_3}{6s^2n}, \\ Z_u &= \bar{z} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} + \frac{\mu_3}{6s^2n}, \end{aligned} \quad (1)$$

where \bar{z} is the vector of means of the epistemic variables, s is the vector of standard deviations, n is the sample size of the point data, μ_3 is the third central moment, and $t_{\alpha/2, n-1}$ is obtained from the Student's t -distribution at $(n-1)$ degrees of freedom and α significance level.

As mentioned earlier, the moments of interval data are obtained only as bounds. Therefore, we cannot use (1) to find the confidence bounds on mean for interval data. Zaman et al. [5] proposed methods to compute the bounds of moments for both single and multiple interval data. The methods for computing bounds of the first three moments for interval data are given later in this section. Once the bounds on the first three moments of interval data are estimated, we search for the configuration of moments, m_i constrained to lie within the

TABLE 1: Methods for calculating moment bounds for multiple interval data.

Moment	Formula
1	$[\underline{M} \quad \overline{M}] = \left[\frac{1}{n} \sum_i lb_i, \frac{1}{n} \sum_i ub_i \right]$
2	$\min_{x_1, \dots, x_n} / \max_{x_1, \dots, x_n} M_2 = \frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2$ s.t. $lb_i \leq x_i \leq ub_i \quad i = \{1, \dots, n\}$
3	$\min_{x_1, \dots, x_n} / \max_{x_1, \dots, x_n} M_3 = \frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^3$ s.t. $lb_i \leq x_i \leq ub_i \quad i = \{1, \dots, n\}$

Note: $[lb_i \quad ub_i]$: set of intervals; n : number of intervals

respective moment bounds of interval data that minimizes (or maximizes) the function shown below to yield the lower (or upper) bound of the confidence interval. Therefore, we construct a multivariable optimization problem as follows.

For the lower bound of confidence interval on mean,

$$\begin{aligned} \min_{\bar{z}, s, \mu_3} \quad & Z = \bar{z} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} + \frac{\mu_3}{6s^2n} \\ \text{s.t.} \quad & a \leq \bar{z} \leq b \\ & c \leq s \leq d \\ & e \leq \mu_3 \leq f. \end{aligned} \quad (2)$$

For the upper bound of confidence interval on mean,

$$\begin{aligned} \max_{\bar{z}, s, \mu_3} \quad & Z = \bar{z} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} + \frac{\mu_3}{6s^2n} \\ \text{s.t.} \quad & a \leq \bar{z} \leq b \\ & c \leq s \leq d \\ & e \leq \mu_3 \leq f, \end{aligned} \quad (3)$$

where \bar{z} , s , and μ_3 represent the first, second, and third moments, respectively which are constrained by their respective lower and upper bounds.

The following discussions briefly summarize the methods to estimate the bounds on the first three moments for multiple interval data.

Bounds on Moments with Multiple Interval Data. The methods for calculating bounds on the first three moments for multiple interval data are summarized in Table 1 below.

Once the bounds on the mean, variance, and third central moment of interval data are estimated by the methods described in Table 1, we can now use these bounds to solve the formulations in (2)-(3) to find the confidence bounds on mean with interval data.

The implementation of the algorithm for calculating the rigorous confidence bounds is as follows:

- (1) calculate the bounds on the first four moments of multiple interval data by the methods outlined above;

- (2) solve the optimization problems in (2) and (3) to obtain confidence bounds on mean. Minimizing the objective function gives the lower bound on the mean and maximizing the objective function gives the upper bound on the mean.

Note that the moment bounding methods described above are scalable in polynomial time with respect to increasing number of intervals [5]. The proposed confidence interval estimation method in (2) and (3) uses these bounds on moments to construct the confidence interval on the mean value. It is seen in (2) and (3) that the number of decision variables in these optimization formulations is always three, the first three moments of interval data. Therefore, the computational efficiency of the optimization formulations in (2) and (3) does not explicitly depend on the number of intervals; rather it depends on the computational complexity of the moment bounding algorithms that are used to compute the bounds on moments for interval data. This implies that the proposed confidence interval estimation method is scalable in polynomial time with respect to increasing number of intervals.

2.2. Optimal Confidence Interval Formulations. Note that although the method presented above is able to give rigorous confidence bounds on mean it does not consider dependence among moments. However, it is more helpful to evaluate bounds in terms of both “rigor” and “optimality” as conceptually sketched in Figure 1. By rigorous, it is meant that the true interval of the possible values lies within the computed bounds. By optimal, it is meant that the bounds are the narrowest possible, while still being rigorous. The optimal bounds preserve the dependence among moments of interval data. Consider the general formulation of confidence interval as shown below:

$$\begin{aligned} \min \quad & Z_l(\theta) \\ \text{s.t.} \quad & \theta \in \Theta, \end{aligned} \quad (4)$$

where θ is the set of moments, selected from a set of admissible values Θ .

The proposed confidence bounds are rigorous, provided that the set Θ encompasses all admissible values of moments. If the set of all admissible values of moments is equal to Θ , then the bounds obtained by (2) and (3) are optimal. Again, if the set Θ is a superset of all actually admissible values of moments, the bounds will still be rigorous, as the search over Θ includes a search over the set of all actually admissible values of moments; however, the bounds will not be optimal because Θ is larger than the set of all admissible moment values. In (2) and (3), the optimizer independently selects a set of moments for the input variable to estimate the confidence bounds. However, for a random variable, the moments are not independent of each other. For example, when the first moment is selected from a configuration of multiple interval data, it is obvious that the other moments will be estimated using the same configuration of multiple interval data. Therefore, if the moments are selected independently as in (2) and (3) presented in this section, the set Θ becomes a superset

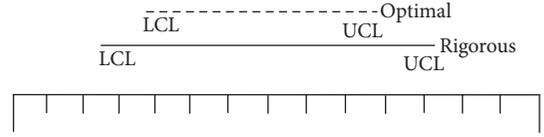


FIGURE 1: Rigorous versus optimal confidence bounds.

of all actually admissible moment values resulting rigorous confidence bounds, which may be wider than the actual ones. In the following discussion, we propose formulations that result in optimal confidence bounds on mean for multiple interval data.

Optimal Confidence Interval. The approach is the same as in (2) and (3) presented earlier in this section, which minimizes and maximizes the confidence bounds $Z(m)$ conditioned on a set of moments (m_i) for the input variables. The optimal confidence bounds problem solves the following optimization formulations to obtain bounds on the mean.

For the lower bound of confidence interval on mean,

$$\begin{aligned} \min_{x_1, \dots, x_n} \quad & Z = \bar{z}(x) - t_{\alpha/2, n-1} \frac{s(x)}{\sqrt{n}} + \frac{\mu_3(x)}{6[s(x)]^2 n} \\ \text{s.t.} \quad & P_i \leq x \leq Q_i \quad i = \{1, \dots, n\}. \end{aligned} \quad (5)$$

For the upper bound of confidence interval on mean,

$$\begin{aligned} \max_{x_1, \dots, x_n} \quad & Z = \bar{z}(x) + t_{\alpha/2, n-1} \frac{s(x)}{\sqrt{n}} + \frac{\mu_3(x)}{6[s(x)]^2 n} \\ \text{s.t.} \quad & P_i \leq x \leq Q_i \quad i = \{1, \dots, n\}, \end{aligned} \quad (6)$$

where $\bar{z} = (1/n) \sum_i x_i$, $s = \sqrt{((1/n)(\sum_{i=1}^n x_i^2) - (1/n^2)(\sum_{i=1}^n x_i)^2)}$, $\mu_3 = (1/n) \sum_{i=1}^n (x_i - \sum_{i=1}^n x_j)^3$, and n is the number of intervals, which are constrained by their respective lower and upper bounds. Note that, in (2) and (3), the decision variables were the set of moments ($m = [m_1 \ m_2 \ m_3]$); however, in (5) and (6), the decision variables are configurations of multiple interval data ($x = [x_1 \ x_2 \ x_3 \ \dots \ x_n]$). The set of moments m are estimated using this configuration x of interval data inside the optimizer and thus the dependency relationships among the moments are preserved resulting in optimal confidence bounds on the mean.

Note that unlike the algorithm for rigorous confidence bounds the implementation of algorithm for optimal confidence bounds does not require computing the bounds on moments. At each iteration, the optimizer selects a set of decision variables, that is, a configuration of multiple interval data, which are then used to calculate moments inside the optimizer. As in rigorous confidence bounds, minimizing the objective function gives the lower bound on the mean and maximizing the objective function gives the upper bound on the mean.

We have implemented the formulations in (5) and (6) to calculate the lower and upper confidence bounds on the mean value for various test cases with increasing number of intervals. We considered nonoverlapping, overlapping, and

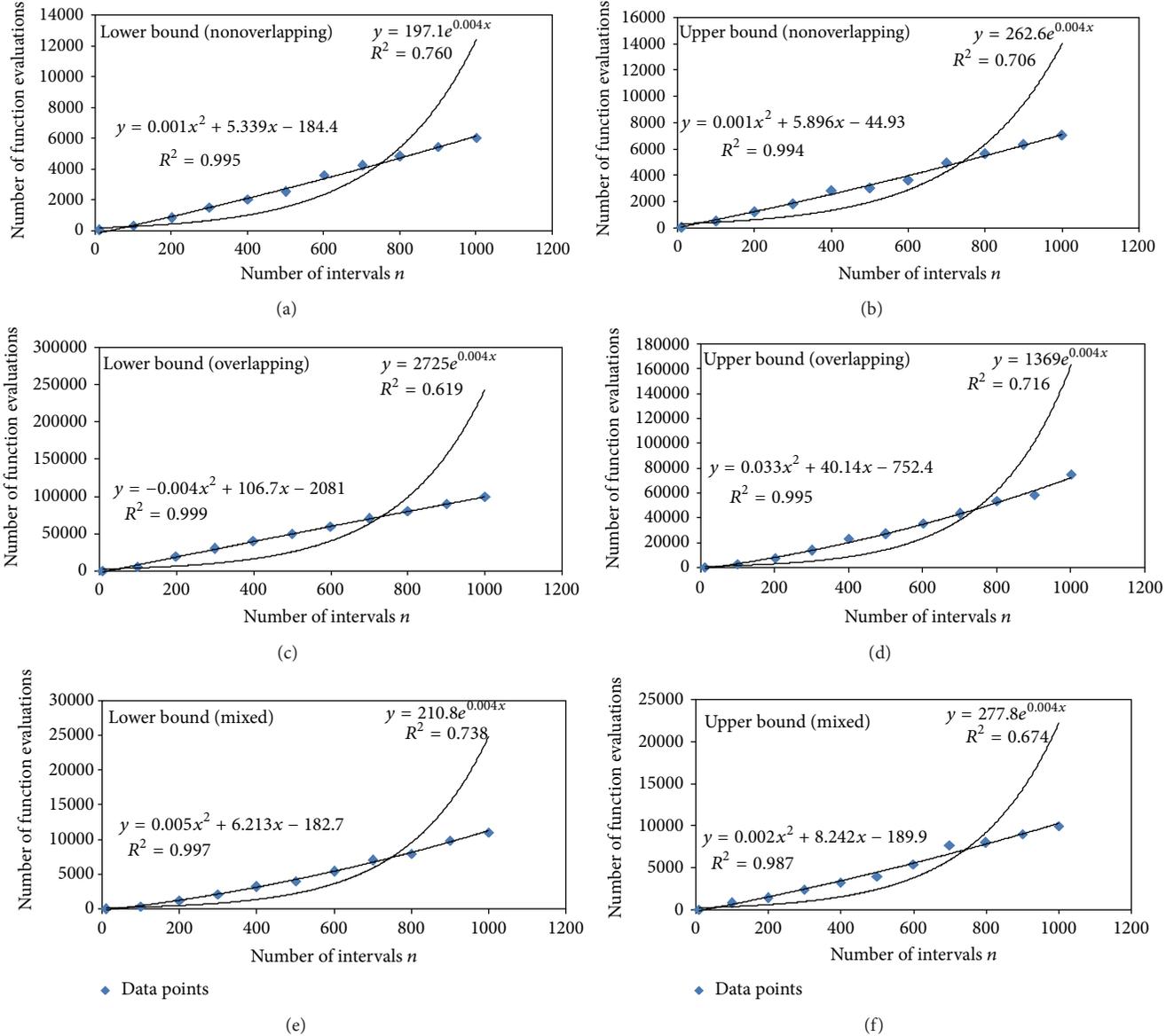


FIGURE 2: Computational effort for the estimation of confidence bounds on mean value for nonoverlapping, overlapping, and mixed intervals.

mixed intervals examples to demonstrate the performance of the proposed formulations. The following procedure was used to generate the intervals for overlapping interval test cases. The interval extremes (lowest of the lower bound and the highest of the upper bound) were arbitrarily assumed. In order to generate a desired number of intervals for each test case, a uniform random number generator was used to generate overlapping intervals between interval extremes. To generate nonoverlapping interval data with n intervals for the test problems, we used the following procedure. First, a sequence of monotonically increasing random numbers is generated, $\{1, \dots, 2 \times n\}$. The i th interval is generated by collecting the $(2i - 1)$ th and $(2i)$ th random number. Thus the interval widths and the end points are generated randomly. We generated mixed interval data for the test case by combining both overlapping and nonoverlapping data sets.

We solved the above optimization formulations in (5) and (6) using the Matlab solver *fmincon*, which implements a sequential quadratic programming algorithm. The plot in Figure 2 illustrates the scalability of the proposed formulations with increasing number of intervals for nonoverlapping (Figures 2(a) and 2(b)), overlapping (Figures 2(c) and 2(d)), and mixed intervals cases (Figures 2(e) and 2(f)), respectively. For each plot shown in Figure 2, we fit a quadratic function as well as an exponential function (solely for comparison purposes). The regression coefficients (i.e., the values of R^2) indicate a strong quadratic trend for the scalability of the algorithms.

Observe that the computational effort for estimating both the lower and upper bounds on mean value increases polynomially with increasing number of intervals for overlapping, nonoverlapping, and mixed interval data. Therefore, for all test cases, the computational effort to estimate the lower and

TABLE 2: Interval data for the four numerical examples.

Example	Data
Example 1 with 5 overlapping intervals	[5, 6; 5.5, 6.1; 6, 6.5; 5.4, 6.2; 5.6, 6.6]
Example 2 with 9 overlapping intervals [3]	[3.5, 6.4; 6.9, 8.8; 6.1, 8.4; 2.8, 6.7; 3.5, 9.7; 6.5, 9.9; 0.15, 3.8; 4.5, 4.9; 7.1, 7.9]
Example 3 with 4 nonoverlapping intervals	[5, 6; 6.1, 6.7; 6.9, 7.8; 8, 9]
Example 4 with 6 nonoverlapping intervals [3]	[1, 1.52; 2.68, 2.98; 7.52, 7.67; 7.73, 8.35; 9.44, 9.99; 3.66, 4.58]

TABLE 3: Comparison of confidence bounds on the mean.

Example	Proposed approaches		Ferson et al. [3]	
	Independent moments	Dependent moments	Normal (UCLs)	Distribution free
1	[4.6936, 7.0766]	[5.0004, 6.7671]	—	—
2	[2.0013, 9.9159]	[2.8281, 9.1108]	[5.958, 8.863]	[1.706, 13.787]
3	[4.0351, 9.8494]	[4.4489, 9.4540]	—	—
4	[1.8780, 9.2973]	[2.0914, 9.0655]	[8.070, 8.613]	[1.160, 14.694]

the upper confidence bounds on mean value with increasing number of intervals is observed to be $O(n^2)$, making this a computationally affordable procedure, even for relatively large data sets. These plots show the best fitting polynomial and exponential trend lines to show that the trend is indeed polynomial in the number of intervals.

In the following section, we illustrate the proposed methods with numerical example problems to compare our approaches with the existing ones.

3. Numerical Examples

In this section, we apply the proposed approaches to four example problems. We consider four multiple interval examples, each with different numbers of intervals and overlaps. The example problems are adapted from Zaman et al. [5]. Examples 2 and 4 are also used in Ferson et al. [3].

We consider two examples each for overlapping and nonoverlapping multiple interval data, each with different numbers of intervals (Table 2). We follow the methods developed in Section 2 to find the confidence interval on the mean for each multiple interval data set in Table 2.

3.1. Independent Value of Moments. In this case, the first three moments of interval data are first estimated as bounds using the method described in Section 2. Once the moment bounds are obtained, we solve the formulations in (2) and (3) using Matlab's *fmincon* solver to get the confidence interval on mean for each of the data set given in Table 2. We have used a level of significance, $\alpha = 0.05$. Note that this significance level is chosen arbitrarily for the sake of illustration. The results are presented in Table 3.

3.2. Dependent Value of Moments. In this case, the moments of interval data are considered dependent on each other. This method does not require computing the moments for the data. We solve the formulations in (5) and (6) using Matlab's *fmincon* solver to get the confidence interval on mean for each of the data set given in Table 2. We have used a level of significance, $\alpha = 0.05$. The results are presented in Table 3.

Note that *fmincon* solver can implement four different algorithms: interior point, sequential quadratic programming (SQP), active set, and trust region reflective. Here, *fmincon* uses an SQP algorithm. The estimate of the Hessian of the Lagrangian is updated using the BFGS formula at each iteration. The convergence properties of SQP have been extensively discussed in the literature (e.g., [21, 22]).

3.3. Comparison among Different Approaches. The solutions for two different cases are listed in Table 3: independent value of moments and dependent value of moments. It is seen that the approach with the dependent value of moments produces narrower confidence bounds compared to the approach with independent value of moments, which is quite intuitive. The results obtained by the proposed optimization-based approaches are also compared with earlier solutions [3]. The solutions to Examples 2 and 4 discussed in [3] are presented in Table 3. Ferson et al. [3] calculated the upper confidence limits (UCL) for these sets of data based on the assumption that the data come from a normal population. They also estimated distribution free confidence limits for these data sets. Note that unlike [3] the proposed approaches estimate the outer confidence limit on mean for interval data.

Since Ferson et al. [3] calculated only upper confidence limits (UCL), we compare the upper limit of their UCL with the upper bound of confidence limits obtained using the proposed approaches. It is seen in Table 3 that the results with the proposed approaches for Examples 2 and 4 show some overlaps with the results of the earlier study. We observe a little disagreement in results between our dependent moment approach and their normal distribution approach due to the fact that we have used Johnson's modified *t*-statistic to account for any deviation of the underlying distribution from normality. The disagreements in results among different studies may be also due to the approaches by which the uncertainty described by multiple intervals is aggregated as well as due to different representations of independence [23].

4. Conclusions

This paper proposed several formulations and algorithms for construction of confidence interval on mean for interval

data, which are illustrated through numerical examples with different numbers of intervals and type of overlap. The major contribution of this paper is to develop an efficient methodology for the construction of confidence interval on mean, which does not depend on any particular type of interval data. This paper also discussed the concepts of rigor and optimality with regard to the confidence bounds on the mean and proposed optimization formulations that give optimal confidence bounds. Two approaches are proposed: the first approach assumes that the moments of data are independent of each other. In order to calculate the confidence bounds on mean value, this approach estimates the moment bounds on the first three moments for the interval data, which are computed using polynomial time algorithms. Therefore, this approach is able to find the confidence interval for interval data in polynomial time. The computational complexity of the second approach, which assumes that the moments of data are dependent on each other, is also found to be polynomial with increasing number of intervals through numerical experimentation. This is important because these problems have been generally considered earlier to be NP-hard.

Note that there exist some performance gaps between the rigorous and optimal confidence bounds, which are expected. The computation of rigorous bounds ignores any dependence that may exist among the moments of interval data. Consequently, the estimated bounds on moments are the widest possible when such dependence is ignored. Therefore, the resulting confidence bounds are also the widest. Although we do not calculate moment bounds explicitly for the dependent moment cases, it is intuitive that the feasible sets of moments that preserve the dependence among moments have narrower bounds than the independent moment case. However, no analytical result is available for this performance gap. It would be interesting if we could show analytically that rigorous bounds are wider than the optimal ones, which we would like to pursue in future.

In the presence of interval uncertainty, the results regarding the confidence interval are valuable to the decision maker as it facilitates more complete implementation of design optimization under interval uncertainty by implementing the constraints on interval variables through the construction of confidence bounds on mean values.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] C. Baudrit and D. Dubois, "Practical representations of incomplete probabilistic knowledge," *Computational Statistics and Data Analysis*, vol. 51, no. 1, pp. 86–108, 2006.
- [2] J. C. Helton, J. D. Johnson, and W. L. Oberkampf, "An exploration of alternative approaches to the representation of uncertainty in model predictions," *Reliability Engineering and System Safety*, vol. 85, no. 1–3, pp. 39–71, 2004.
- [3] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf, and L. Ginzburg, "Experimental uncertainty estimation and statistics for data having interval uncertainty," Sandia National Laboratories Technical Report SAND 2007-0939, Albuquerque, NM, USA, 2007.
- [4] X. Du, A. Sudjianto, and B. Huang, "Reliability-based design with the mixture of random and interval variables," *Journal of Mechanical Design*, vol. 127, no. 6, pp. 1068–1076, 2005.
- [5] K. Zaman, S. Rangavajhala, M. P. McDonald, and S. Mahadevan, "A probabilistic approach for representation of interval uncertainty," *Reliability Engineering and System Safety*, vol. 96, no. 1, pp. 117–130, 2011.
- [6] W. L. Oberkampf, J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson, "Challenge problems: uncertainty in system response given uncertain parameters," *Reliability Engineering & System Safety*, vol. 85, no. 1–3, pp. 11–19, 2004.
- [7] S. Ferson and J. G. Hajagos, "Arithmetic with uncertain numbers: rigorous and (often) best possible answers," *Reliability Engineering and System Safety*, vol. 85, no. 1–3, pp. 135–152, 2004.
- [8] X. Du, "Unified uncertainty analysis by the first order reliability method," *Journal of Mechanical Design*, vol. 130, no. 9, Article ID 091401, 10 pages, 2008.
- [9] J. Guo and X. Du, "Sensitivity analysis with mixture of epistemic and aleatory uncertainties," *AIAA Journal*, vol. 45, no. 9, pp. 2337–2349, 2007.
- [10] J. Guo and X. Du, "Reliability sensitivity analysis with random and interval variables," *International Journal for Numerical Methods in Engineering*, vol. 78, no. 13, pp. 1585–1617, 2009.
- [11] M. McDonald and S. Mahadevan, "Uncertainty quantification and propagation for multidisciplinary system analysis," in *Proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (MAO '08)*, Victoria, Canada, September 2008.
- [12] K. Zaman, M. McDonald, and S. Mahadevan, "Probabilistic framework for uncertainty propagation with both probabilistic and interval variables," *Journal of Mechanical Design*, vol. 133, no. 2, Article ID 021010, 14 pages, 2011.
- [13] K. Zaman, M. McDonald, S. Mahadevan, and L. Green, "Robustness-based design optimization under data uncertainty," *Structural and Multidisciplinary Optimization*, vol. 44, no. 2, pp. 183–197, 2011.
- [14] K. Zaman and S. Mahadevan, "Robustness-based design optimization of multidisciplinary system under epistemic uncertainty," *AIAA Journal*, vol. 51, no. 5, pp. 1021–1031, 2013.
- [15] A. Haldar and S. Mahadevan, *Probability, Reliability and Statistical Methods in Engineering Design*, John Wiley & Sons, New York, NY, USA, 2000.
- [16] X. Zhou and S. Gao, "One-sided confidence intervals for means of positively skewed distributions," *The American Statistician*, vol. 54, no. 2, pp. 100–104, 2000.
- [17] N. J. Johnson, "Modified t tests and confidence intervals for asymmetrical populations," *Journal of the American Statistical Association*, vol. 73, no. 363, pp. 536–544, 1978.
- [18] P. Hall, "On the removal of skewness by transformation," *Journal of the Royal Statistical Society B: Methodological*, vol. 54, no. 1, pp. 221–228, 1992.
- [19] X. H. Zhou and P. Dinh, "Nonparametric confidence intervals for the one-and two-sample problems," *Biostatistics*, vol. 6, no. 2, pp. 187–200, 2005.
- [20] V. Kreinovich, L. Longpré, P. Patangay, S. Ferson, and L. Ginzburg, "Outlier detection under interval uncertainty: algorithmic solvability and computational complexity," *Reliable Computing*, vol. 11, no. 1, pp. 59–76, 2005.

- [21] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1987.
- [22] E. R. Panier and A. L. Tits, "On combining feasibility, descent and superlinear convergence in inequality constrained optimization," *Mathematical Programming*, vol. 59, no. 1-3, pp. 261-276, 1993.
- [23] S. Ferson, C. A. Joslyn, J. C. Helton, W. L. Oberkampf, and K. Sentz, "Summary from the epistemic uncertainty workshop: consensus amid diversity," *Reliability Engineering and System Safety*, vol. 85, no. 1-3, pp. 355-369, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

