REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
08-08-2011	Conference Proceeding	-

4. TITLE AND SUBTITLE
Reliability-based Design Optimization with Confidence Level for Non-Gaussian Distributions Using Bootstrap Method

5a. CONTRACT NUMBER
W911NF-09-1-0250

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Yoojeong Noh, Kyung K. Choi, Ikjin Lee, David Gorsich

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Iowa @ Iowa City
Office of Sponsored Programs
The University of Iowa
Iowa City, IA 52242 -

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
56025-NS.24

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
For reliability-based design optimization (RBDO), generating an input statistical model with confidence level has been recently proposed to offset the inaccurate estimation of the input statistical model with Gaussian distributions. For this, the confidence intervals of mean and standard deviation are calculated using the Gaussian distributions of input random variables. However, if the input random variables are non-Gaussian, the use of the Gaussian distributions of input variables will provide inaccurate confidence intervals, and thus, yield undesirable confidence

15. SUBJECT TERMS
Reliability-based design optimization, input statistical model, confidence level, non-Gaussian distribution, bootstrap method

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Kyung Choi
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 319-335-5684

Report Title

Reliability-based Design Optimization with Confidence Level for Non-Gaussian Distributions Using Bootstrap Method

ABSTRACT

For reliability-based design optimization (RBDO), generating an input statistical model with confidence level has been recently proposed to offset the inaccurate estimation of the input statistical model with Gaussian distributions. For this, the confidence intervals of mean and standard deviation are calculated using the Gaussian distributions of input random variables. However, if the input random variables are non-Gaussian, the use of the Gaussian distributions of input variables will provide inaccurate confidence intervals, and thus, yield undesirable confidence level of the reliability-based optimum design meeting the target reliability bt. In this paper, the RBDO method using the bootstrap method, which does not use the Gaussian distributions of input variables to calculate the confidence intervals of mean and standard deviation, are proposed to obtain the desirable confidence level of output performance for non-Gaussian distributions.

Conference Name: ASME 2010 International Design Engineering Technical Conferences & Computers and Information in En

Conference Date: August 15, 2010

Proceedings of the ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2010
August 15 − 18, 2010, Montreal, Canada

DETC2010-28597

Reliability-based Design Optimization with Confidence Level for Non-Gaussian Distributions Using Bootstrap Method

Yoojeong Noh, Kyung K. Choi*, Ikjin Lee
Department of Mechanical & Industrial Engineering, College of Engineering
The University of Iowa, Iowa City, IA 52242, USA
noh@engineering.uiowa.edu, kkchoi@engineering.uiowa.edu, ilee@engineering.uiowa.edu
*Corresponding author

David Gorsich
US Army RDECOM/TARDEC
Warren, MI 48397-5000, USA
david.gorsich@us.army.mil

ABSTRACT

For reliability-based design optimization (RBDO), generating an input statistical model with confidence level has been recently proposed to offset the inaccurate estimation of the input statistical model with Gaussian distributions. For this, the confidence intervals of mean and standard deviation are calculated using the Gaussian distributions of input random variables. However, if the input random variables are non-Gaussian, the use of the Gaussian distributions of input variables will provide inaccurate confidence intervals, and thus, yield undesirable confidence level of the reliability-based optimum design meeting the target reliability β_t. In this paper, the RBDO method using the bootstrap method, which does not use the Gaussian distributions of input variables to calculate the confidence intervals of mean and standard deviation, are proposed to obtain the desirable confidence level of output performance for non-Gaussian distributions.

KEYWORDS

Reliability-based design optimization, input statistical model, confidence level, non-Gaussian distribution, bootstrap method

1. INTRODUCTION

Obtaining an accurate input statistical model, which includes marginal distributions and a joint distribution of input random variables, is crucial to obtain an accurate reliability-based optimum design. However, for many input random variables such as loading, material properties, and manufacturing geometric variability, only limited data are available due to expensive testing costs. If the input statistical model is obtained from insufficient data, it could yield unreliable design. To offset the inaccurate estimation of the input model, generating an input model with the confidence level has been recently proposed by using adjusted standard deviations and a correlation coefficient that include the effect of inaccurate estimation of mean value and correlation coefficient [1].

The adjusted standard deviation and correlation coefficient are obtained from the confidence intervals of input distribution parameters such as the mean, standard deviation, and correlation coefficient. The confidence intervals of mean and standard deviation are usually calculated using the Gaussian distributions of input variables [2]. If the input variables have marginal Gaussian distributions, the confidence intervals of mean and standard deviation can be explicitly calculated, which are exact for the Gaussian distribution. However, if the input variables have marginal non-Gaussian distributions, the estimated confidence intervals using the assumption of the Gaussian distributions of input variables will be inaccurate. Thus, the bootstrap method [3-5], which does not require the Gaussian distribution of input variables, is used for non-Gaussian distributions.

To validate whether use of the adjusted standard deviation and correlation coefficient obtained using the bootstrap method provides the desirable confidence level of the input model, the confidence level of the input model with adjusted parameters is assessed through simulation tests. The β_t -

contour, for the given target reliability index β_t, is used to measure the confidence level of the input model. A mathematical example and an M1A1 Abrams tank roadarm problem with non-Gaussian correlated variables are used to illustrate how the input model with a target confidence level using the bootstrap method provides a more desirable output confidence level compared to the one without using the proposed method for correlated input distributions.

2. ESTIMATION OF INPUT STATISTICAL MODEL

The input statistical model, which consists of marginal and joint distributions, needs to be identified. If the input random variables are independent, the joint distribution is obtained by multiplication of the marginal distributions, so that only marginal distributions and their associated parameters need to be obtained. However, if the input random variables are correlated, the joint distribution is required. Since only limited data is available in practical engineering applications, it is difficult to obtain the joint distribution directly from limited data. Thus, a copula, which is a function of marginal distributions and correlation parameters, is used to model the input joint distribution. The marginal distributions, copulas, and their associated parameters need to be identified and quantified for the correlated input variables.

2.1 Quantification of Input Model

The two-parameter marginal distributions, which are used in many engineering applications, have their own parameters a and b, which determine the distribution shape. In some distributions, a determines the location of the distributions, while, in other distributions, a determines the scale of the distributions. Thus, it is difficult to choose which parameter needs to be used to enlarge the β_t-contour, which is used to obtain the RBDO result. On the other hand, for many distributions, these parameters (a and b) can be expressed in terms of mean and standard deviation (μ and σ). Since the mean and standard deviation determine the location and variability of the distributions, the standard deviation can be used to enlarge the β_t-contour for these types of distribution. Accordingly, once the mean and standard deviation are calculated from given data, the parameters a and b can be calculated using the explicit functions, which are presented for various marginal distributions in Ref. 6.

According to Sklar's theorem [7], a joint cumulative distribution function (CDF) $F_{X_1 \cdots X_n}(x_1, \cdots, x_n)$ of random variables X_i can be expressed in terms of marginal CDFs $F_{X_i}(x_i)$ of X_i for $i = 1, \cdots, n$ and copula function C as

$$F_{X_1,...,X_n}(x_1,...,x_n) = C\left(F_{X_1}(x_1),...,F_{X_n}(x_n)\middle|\boldsymbol{\theta}\right) \tag{1}$$

where $\boldsymbol{\theta}$ is the matrix of correlation parameters between $X_1,...,X_n$. Most copula applications consider bivariate data because only few copula families have *n*-dimensional generalization. It has been observed that two input variables are correlated in many cases [8-11], so that only bivariate copulas are considered in this paper.

Since copula functions have their own correlation parameters, it is necessary to use a common correlation coefficient. The Kendall's tau, which is a widely used correlation coefficient, is used in Ref. 6 and, accordingly, in this paper. The population version of Kendall's tau is expressed as

$$\tau = 4\int\int_{I^2} C(u,v|\theta)\, dC(u,v|\theta) - 1 \tag{2}$$

where $u = F_{X_1}(x_1)$ and $v = F_{X_2}(x_2)$, and the unit square I^2 is the product $I \times I = [0,1] \times [0,1]$ of the domain of marginal CDFs of X_1 and X_2. The sample version of Kendall's tau is

$$t = \frac{c-d}{c+d} \tag{3}$$

where c and d are the number of concordant and discordant pairs, and *ns* is the number of samples. Once the Kendall's tau is obtained from samples using Eq. (3), the correlation parameter θ can be obtained using Eq. (2) or explicit formulations presented in Ref. 6.

2.2 Identification of Input Model

The joint distribution can be identified by a one-step procedure, which directly tests all candidate joint distributions, or by a two-step procedure, which first identifies marginal distributions and then a copula [6,12-15]. The two-step approach is more efficient and accurate than the one-step approach [16]. For example, if the seven candidate marginal distributions of X_1 and X_2 and nine candidate copulas are used to identify a joint distribution, the one-step approach requires to test $7 \times 7 \times 9 = 441$ cases, whereas the two-step approach requires to test $7 + 7 + 9 = 23$ cases. It is more challenging to identify a correct joint distribution from 441 candidates compared to 23 candidates, so the two-step approach is more preferred. According to the measure of identification, the weight-based method [6] and MCMC-based method [17] can be used. However, the MCMC-based method uses random samples of the posterior distribution, which causes randomness of identification results. Thus, in this paper, the two-step weight-based Bayesian method is used to identify the input model. More detailed information on the two-step weight-based method is presented in Ref. 6.

3. CALCULATION OF CONFIDENCE INTERVALS OF INPUT PARAMETERS USING BOOTSTRAP METHOD

If the input random variables have Gaussian distribution, the confidence intervals of mean and standard deviation can be explicitly and exactly obtained. However, if not, there are no explicit functions for calculation of the confidence intervals of mean and standard deviation for non-Gaussian distribution, so the bootstrap method needs to be introduced to obtain accurate confidence interval of the standard deviation for non-Gaussian distributions. Because the confidence interval of the mean can be accurately estimated even for the non-Gaussian distribution [18,19], only the confidence interval of the standard deviation is tested. Section 3.1 illustrates how to calculate confidence intervals of input parameters using the bootstrap method. To

test performance of the bootstrap method, percentage that the confidence interval of standard deviation includes the true standard deviation (confidence level of standard deviation) should be equivalent to the target confidence level. Thus, in Section 3.2, the confidence level of standard deviation is assessed through simulation tests for non-Gaussian distributions. The confidence interval of the correlation parameter is presented in Section 3.3.

3.1 Quantification of Input Model

If the input variables follow Gaussian distribution with μ and standard deviation σ, the lower and upper bounds of the confidence interval of the mean ($\tilde{\mu}^L$ and $\tilde{\mu}^U$) can be obtained as, [1,2]

$$\tilde{\mu}^L = \tilde{\mu} - t_{\alpha/2,ns-1}\frac{\tilde{\sigma}}{\sqrt{ns}} \text{ and } \tilde{\mu}^U = \tilde{\mu} + t_{\alpha/2,ns-1}\frac{\tilde{\sigma}}{\sqrt{ns}} \quad (4)$$

where *ns* is the number of samples, $\tilde{\mu}$ and $\tilde{\sigma}$ are the sample mean and sample standard deviation, respectively, and $t_{\alpha/2,ns-1}$ is the value of student's *t*-distribution with (*ns*-1) degree of freedom at two-sided confidence level, $100\times(1-\alpha/2)$.

Using a similar procedure of calculating the confidence interval of the mean, the lower and upper bounds of the confidence interval for the standard deviation, $\tilde{\sigma}^L$ and $\tilde{\sigma}^U$, respectively, are calculated as [1,2]

$$\tilde{\sigma}^L = \sqrt{\frac{(ns-1)\tilde{\sigma}^2}{c_{1-\alpha/2,ns-1}}} \text{ and } \tilde{\sigma}^U = \sqrt{\frac{(ns-1)\tilde{\sigma}^2}{c_{\alpha/2,ns-1}}} \quad (5)$$

where $c_{\alpha/2,ns-1}$ and $c_{1-\alpha/2,ns-1}$ are the critical values of the chi-square distribution evaluated at two-sided confidence level $100\times(\alpha/2)$ and $100\times(1-\alpha/2)$ with (*ns*–1) degrees of freedom, respectively.

If input variables follow Gaussian distribution, Eqs. (4) and (5) are exact. However, if not, the estimated confidence intervals are not correct. Thus, the bootstrap method, which does not use the Gaussian distribution of input variables to calculate the confidence interval of the mean and standard deviation, needs to be used.

The bootstrap method calculates the confidence interval of estimated standard deviation $\tilde{\sigma}$ by constructing a distribution of the standard deviation using the frequency distribution of $\tilde{\sigma}^*$ obtained from randomly generated bootstrap samples based on the given data. Table 1 shows how to calculate the confidence interval of the standard deviation using the bootstrap method.

The first step is to construct an empirical distribution $F_{ns}(\mathbf{x})$ or a parametric distribution $\tilde{F}(\mathbf{x}|a,b)$ from given samples, $\mathbf{x}=[x_1,x_2,\cdots,x_{ns}]$. In the second step, bootstrap samples are generated from an empirical distribution or parametric distribution. If a random sample of size *ns* with replacement is drawn from the empirical distribution $F_{ns}(\mathbf{x})$, then this is called a non-parametric approach. If the resample is drawn from the specified model $\tilde{F}(\mathbf{x}|a,b)$ determined from the given samples, this is called a parametric approach. In this study, the distribution type of the parametric model is used for the two-step weight-based Bayesian method. The third step is to calculate $\tilde{\sigma}$ from the resample, drawn from either empirical or parametric distribution, yielding $\tilde{\sigma}^*_{bs}$. In the fourth step, the second and third steps are repeated B times (e.g., B=1000). Then, the fifth step is to construct a probability distribution from $\tilde{\sigma}^*_1,\tilde{\sigma}^*_2,\cdots,\tilde{\sigma}^*_B$. This distribution is the bootstrap sampling distribution of $\tilde{\sigma}$, $\tilde{G}^*(\tilde{\sigma}^*)$, which is used to calculate the confidence interval of $\tilde{\sigma}$. To obtain the bootstrap sampling distribution of $\tilde{\sigma}$, the normal approximation, percentile, bias corrected (BC), percentile-*t*, or bias corrected accelerated (BCa) methods can be used.

Table 1. Bootstrap Procedures

Bootstrap Procedures	
Step 1	From given samples $\mathbf{x}=[x_1,x_2,\cdots,x_{ns}]$, construct an empirical distribution $F_{ns}(\mathbf{x})$ for the non-parametric approach; or parametric distribution $\tilde{F}(\mathbf{x}\|a,b)$ for the parametric approach.
Step 2	Generate bootstrap samples $\mathbf{x}^*=[x^*_1,x^*_2,\cdots,x^*_{ns}]$ from the constructed distribution in Step 1.
Step 3	Calculate a statistic of interest $\tilde{\sigma}$ from bootstrap samples, yielding $\tilde{\sigma}^*_{bs}$, $bs=1,\cdots,B$
Step 4	Repeat Step 2 and 3 B times (e.g., B=1000).
Step 5	Construct a probability distribution $\tilde{G}^*(\tilde{\sigma}^*)$ from $\tilde{\sigma}^*_1,\tilde{\sigma}^*_2,\cdots,\tilde{\sigma}^*_B$, and then calculate confidence interval for estimated parameter, $\tilde{\sigma}$ using $\tilde{G}^*(\tilde{\sigma}^*)$.

3.1.1 Normal Approximation Method

The normal approximation method assumes that the distribution of $\tilde{\sigma}$ is a Gaussian distribution. Using the assumption, the confidence interval for $\tilde{\sigma}$ is obtained as [1,20]

$$\tilde{\sigma} - z_{\alpha/2}\tilde{\sigma}^*_{\tilde{\sigma}} < \sigma < \tilde{\sigma} + z_{\alpha/2}\tilde{\sigma}^*_{\tilde{\sigma}} \quad (6)$$

where $\tilde{\sigma}^*_{\tilde{\sigma}} = \sqrt{\sum_{bs=1}^{B}\left[\tilde{\sigma}^*_{bs}-\tilde{\sigma}^*\right]^2/(B-1)}$, $\tilde{\sigma}^* = \sum_{bs=1}^{B}\tilde{\sigma}^*_{bs}/B$, and $z_{\alpha/2}$ is the value of standard Gaussian distribution CDF at $\alpha/2$.

3.1.2 Percentile Method

The percentile method calculates the confidence interval for the parameter based on the bootstrap sampling distribution $\tilde{G}^*(\tilde{\sigma}^*)$ approximating the population distribution $G(\tilde{\sigma})$. The basic idea of this method is that the confidence interval for $(1-\alpha)$ level includes all the values of $\tilde{\sigma}^*$ between the

$(\alpha/2\times100)^{\text{th}}$ and $(1-\alpha/2)\times100^{\text{th}}$ percentiles of $\tilde{G}^*(\tilde{\sigma}^*)$. The sorting vector of $\tilde{\sigma}^*_{bs}$ is obtained from each bootstrap sample for $bs=1,\cdots,B$ and the values of $\tilde{\sigma}^*_{bs}$ evaluated at the $(\alpha/2\times100)^{\text{th}}$ and $(1-\alpha/2)\times100^{\text{th}}$ percentiles of $\tilde{G}^*(\tilde{\sigma}^*)$ are used as the lower and upper bounds of $\tilde{\sigma}$,

$$\tilde{\sigma}^*_{\alpha/2} < \sigma < \tilde{\sigma}^*_{1-\alpha/2} \tag{7}$$

Since the percentile method does not assume that the bootstrap sampling distribution follows a Gaussian distribution like the normal approximation method, it allows $\tilde{G}^*(\tilde{\sigma}^*)$ confirming to any shape that the data follow. For this reason, it is the most widely used bootstrap technique among applied statisticians [21]. However, when the number of samples is small, $\tilde{G}^*(\tilde{\sigma}^*)$ might be a biased estimator of $G(\tilde{\sigma})$, i.e., $\tilde{\sigma}^*$ is a biased estimator of $\tilde{\sigma}$. In that case, the percentile method can be inaccurate.

3.1.3 Bias Corrected Method

The bias corrected (BC) method corrects the bias term by introducing an adjusted parameter z_0. Suppose that there exist some monotonic transformations of $\tilde{\sigma}^*$ and $\tilde{\sigma}$, say, $\tilde{\varphi}-\varphi$. Instead of assuming that $\tilde{\sigma}^*-\tilde{\sigma}$ is centered on zero, the BC method assumes that $\varphi(\tilde{\sigma}^*)-\varphi(\tilde{\sigma})+z_0=z$ follows a standard Gaussian distribution. Since $\tilde{\varphi}$ and φ are monotonic functions, it holds that [3]. Accordingly, z_0 is calculated using [20]

$$z_0 = \Phi^{-1}\left(\Pr\left(\tilde{\sigma}^* \le \tilde{\sigma}\right)\right) \tag{8}$$

where z_0 is a biasing constant that compensates for the bias between $\tilde{\sigma}^*$ and $\tilde{\sigma}$. Since $\tilde{G}^*(\tilde{\sigma}^*)$ is invariant to transformation, the transformation does not need to be known. Using z_0, the confidence interval for $\tilde{\sigma}$ is obtained as

$$\tilde{\sigma}^*_{\Phi(2z_0+z_{\alpha/2})} < \sigma < \tilde{\sigma}^*_{\Phi(2z_0+z_{1-\alpha/2})} \tag{9}$$

$\tilde{\sigma}^*_{\Phi(2z_0+z_{\alpha/2})}$ is the value of $\tilde{\sigma}^*$ evaluated at the $\Phi(2z_0+z_{\alpha/2})\times100^{\text{th}}$ percentile and $\tilde{\sigma}^*_{\Phi(2z_0+z_{1-\alpha/2})}$ is the value of $\tilde{\sigma}^*$ evaluated at the $\Phi(2z_0+z_{1-\alpha/2})\times100^{\text{th}}$ percentile. The BC method corrects the bias term, but it still requires the parametric assumption that there exist monotonic transformations of $\tilde{\sigma}^*$ and $\tilde{\sigma}$.

3.1.4 Bias Corrected and Accelerated Method

The bias corrected and accelerated (BCa) method generalizes the BC method. The BC method only corrects the bias, whereas the BCa method corrects both the bias and the skewness. The BCa method assumes that for certain monotone transformations, $\tilde{\varphi}$ and φ, certain bias constant z_0 and acceleration constant A result in [4]

$$\frac{(\tilde{\varphi}-\varphi)}{\sigma_{\tilde{\varphi}}} \sim N\left(-z_0\sigma_h, \sigma_h^2\right), \quad \sigma_h = 1+Ah \tag{10}$$

where $\sigma_{\tilde{\varphi}}$ is the constant standard error of $\tilde{\varphi}$. The acceleration A is defined as

$$A = \frac{\sum_{i=1}^{ns}\left(\tilde{\sigma}-\tilde{\sigma}_{(i)}\right)^3}{6\left\{\sum_{i=1}^{ns}\left(\tilde{\sigma}-\tilde{\sigma}_{(i)}\right)^2\right\}^{3/2}} \tag{11}$$

where $\tilde{\sigma}_{(i)}$ is the estimated parameter of $\mathbf{x}_{(i)}=(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_{ns})$ without the i^{th} point x_i and $\tilde{\sigma}=\sum_{i=1}^{ns}\tilde{\sigma}_{(i)}/ns$. Using Eq. (10), the BCa confidence interval is defined as

$$\tilde{\sigma}^*_{\alpha_1} < \sigma < \tilde{\sigma}^*_{\alpha_2} \tag{12}$$

where $\alpha_1 = \Phi\left(z_0 + \frac{z_0+z_{\alpha/2}}{1-A(z_0+z_{\alpha/2})}\right)$,

$\alpha_2 = \Phi\left(z_0 + \frac{z_0+z_{1-\alpha/2}}{1-A(z_0+z_{1-\alpha/2})}\right)$, and $z_0 = \Phi^{-1}\left(\Pr\left(\tilde{\sigma}^* \le \tilde{\sigma}\right)\right)$.

However, since the BCa method highly depends on the acceleration *A*, if it is not accurate, the BCa also is inaccurate.

3.1.5 Percentile-t Method

The percentile-*t* method uses the distribution of a standardized estimator to calculate the confidence interval. The percentile-*t* interval is expected to be accurate to the extent that standardizing depends less on the boot sampling estimator, $\tilde{\sigma}^*$, than the percentile method. The standardized parameter t^*_{bs} can be defined as [21]

$$t^*_{bs} = \left(\tilde{\sigma}^*_{bs}-\tilde{\sigma}\right)/\tilde{\sigma}_{\tilde{\sigma}^*_{bs}} \tag{13}$$

where $\tilde{\sigma}^*_{bs}$ is the estimated parameter from each re-sampled data, $\mathbf{x}^*_{bs}=\left[x^*_{1bs},x^*_{2bs},\cdots,x^*_{nsbs}\right]$ for $bs=1,\cdots,B$. $\tilde{\sigma}$ is the estimated parameter from the original data, $\mathbf{x}=\left[x_1,x_2,\cdots,x_{ns}\right]$. $\tilde{\sigma}_{\tilde{\sigma}^*_{bs}}$ is the standard deviation of $\tilde{\sigma}$ obtained from a double bootstrap, which is another level of re-sampling. That is, the double bootstrap sample $\mathbf{x}^*_d=\left[x^{**}_{1d},x^{**}_{2d},\cdots,x^{**}_{nsd}\right]$ for $d=1,\cdots,D$ is re-sampled from the bootstrap samples, $\mathbf{x}^*_{bs}=\left[x^*_{1bs},x^*_{2bs},\cdots,x^*_{nsbs}\right]$ for $bs=1,\cdots,B$. Thus, the percentile-*t* method requires a large number of

bootstrap samples $(D \times B)$. Using double bootstrap samples, $\tilde{\sigma}_{\tilde{\sigma}_{bs}^*}$ is obtained as

$$\tilde{\sigma}_{\tilde{\sigma}_{bs}^*} = \sqrt{\frac{\sum_{d=1}^{D}\left[\tilde{\sigma}_d^{**} - \tilde{\sigma}^{**}\right]^2}{D-1}} \tag{14}$$

where $\tilde{\sigma}^{**} = \sum_{d=1}^{D} \tilde{\sigma}_d^{**} / D$.

The confidence interval of the sample standard deviation $\tilde{\sigma}$ is obtained as

$$\tilde{\sigma} + t_{\alpha/2}^* \tilde{\sigma}_{\tilde{\sigma}} < \sigma < \tilde{\sigma} + t_{1-\alpha/2}^* \tilde{\sigma}_{\tilde{\sigma}} \tag{15}$$

where $t_{\alpha/2}^*$ and $t_{1-\alpha/2}^*$ are the values of t^* evaluated at $\alpha/2 \times 100^{\text{th}}$ and $(1-\alpha/2) \times 100^{\text{th}}$ percentiles. In Eq. (15), $\tilde{\sigma}_{\tilde{\sigma}}$ is the population standard deviation, which is calculated as $\tilde{\sigma}_{\tilde{\sigma}} = \sqrt{\sum_{d=1}^{D}\left[\tilde{\sigma}_d^* - \tilde{\sigma}^*\right]^2 / (D-1)}$, where $\tilde{\sigma}_d^*$ is the estimator obtained from re-sampled data, $\mathbf{x}^* = \left[x_1^*, x_2^*, \cdots, x_D^*\right]$, randomly generated from the original data.

3.2 Confidence Levels of Standard Deviation

To test the performance of the bootstrap methods, a lognormal marginal distribution is considered as the true model with $\mu = 5.0$ and $\sigma = 5.0$. The confidence level of the standard deviation is assessed by calculating the probability that the upper bound of the confidence interval of the standard deviation is larger than the true standard deviation over 1000 data sets with *ns*=30, 100, 300 samples generated from the true model. In this paper, for 95% of the two-sided target confidence level, the target confidence level is 97.5% for the upper side of the confidence interval. Thus, if the estimated confidence level is close to the target confidence level, it is the most desirable. Even if the estimated confidence level is close to the target confidence level, the unnecessarily large upper bound for standard deviation is not desirable because it yields unnecessarily conservative RBDO design. Therefore, the most desirable confidence interval should just include the true standard deviation with the target confidence level.

Table 2 shows the obtained confidence levels using the five bootstrap methods where the bootstrap distribution is obtained from the empirical distribution (nonparametric approach) and the parametric distribution (parametric approach). For the parametric approach, the true marginal and identified CDFs are used. Since the distribution of the standard deviation is highly skewed due to the lognormal distribution with large coefficient of variation (COV), which is the ratio of standard deviation to mean, all methods have poor confidence levels, especially for *ns*=30. As the number of samples increases, the obtained confidence levels are mostly closer to the target confidence level, whereas the method using the Gaussian distribution of input variable in Eq. (5) provides nearly only 65% regardless of the number of samples. The parametric approach using the true marginal CDF has the best performance, and the parametric approach with identified CDF is not as good as the one with true CDF due to the identification and quantification error. However, it is better than the nonparametric approach because it yields more desirable confidence level than the nonparametric approach when the identified CDFs are correct.

As shown in Table 2, when the nonparametric approach is used, the percentile-*t* has the highest confidence level, followed by the BCa method among five bootstrap methods. For the parametric approach using identified CDF, the percentile-*t* method is the best for *ns*=30 and 100, and the percentile method is the best for *ns*=300. When the parametric approach using the true CDF is used, the percentile method has the best performance for *ns*=30, 100, and 300.

Table 2. Estimated Confidence Levels of Standard Deviation

ns	Approach	Nor. Approx.	Percentile	BC	BCa	Percentile-*t*
30	Nonpar.	64.9	61.3	64.8	70.4	83.7
	Par. (Iden.)	74.9	79.0	85.0	88.9	89.6
	Par. (True)	89.3	95.2	88.3	92.4	88.3
100	Nonpar.	76.0	75.1	78.1	83.1	88.8
	Par. (Iden.)	87.2	87.5	87.4	89.9	92.0
	Par. (True)	95.2	97.7	86.9	89.9	87.1
300	Nonpar.	82.3	82.7	84.5	87.9	91.3
	Par. (Iden.)	95.9	97.5	86.8	88.4	85.7
	Par. (True)	96.7	99.0	86.3	88.2	85.3

* Method using Gaussian distribution of input variable yields 65.8% for *ns*=30, 66.0% for *ns*=100, and 65.0% for *ns*=300

Even though the percentile-*t* method provides more desirable confidence levels than other methods for the nonparametric approach, it provides unnecessarily large upper bounds of standard deviation. For example, the mean value of the upper bound of standard deviation using the nonparametric percentile-*t* method is 20.01 as shown in Table 3, which is significantly larger than true standard deviation 5.0. In addition, it has a large standard deviation, 32.78, of the upper bounds of confidence intervals of standard deviation as shown in Table 4. That is, the upper bounds of confidence intervals of the standard deviation are overestimated and widely spread over the wide range of the standard deviation. As the number of samples increases, the upper bound of the standard deviation approaches the true standard deviation and its variation is reduced, but it converges very slowly to the true standard deviation compared to other methods. Thus, the BCa method, which has second highest confidence level and adequate value of upper bound, is preferred for the nonparametric approach.

Likewise, for the parametric approach, the percentile method yields a desirable confidence level, and has adequate values of upper bounds for standard deviation and small variations as shown in Table 3 and 4.

Figure 1 shows the histograms of the upper bounds of confidence interval of the standard deviation using the parametric percentile bootstrap method with identified CDF. The estimated upper bounds of standard deviations mostly centered at the true standard deviation (5.0) even for *ns*=30. As the number of samples increases, a large amount of the upper bounds of standard deviation tends to be very close to the true standard deviation with a small variation.

Table 3. Mean Values of Upper Bound of Confidence Interval of Standard Deviation

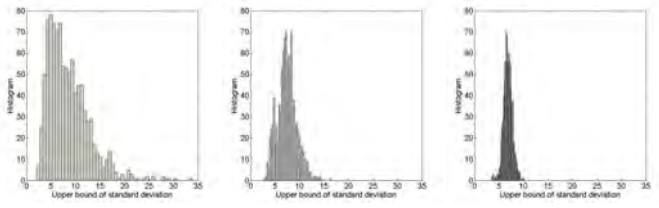
ns	Approach	Nor. Approx.	Percentile	BC	BCa	Percentile-*t*
30	Nonpar.	6.900	6.458	6.803	7.379	20.01
	Par. (Iden.)	7.830	8.396	12.72	15.22	16.62
	Par. (True)	8.742	9.998	12.57	15.83	16.04
100	Nonpar.	6.482	6.367	6.604	7.144	10.78
	Par. (Iden.)	7.167	7.596	9.204	10.43	10.83
	Par. (True)	7.367	7.978	9.342	10.60	10.77
300	Nonpar.	6.088	6.078	6.219	6.571	7.810
	Par. (Iden.)	6.416	6.763	7.324	7.877	7.642
	Par. (True)	6.491	6.824	8.034	7.422	7.984

* Method using Gaussian distribution of input variable yields 6.451 for *ns*=30, 5.708 for *ns*=100, and 5.401 for *ns*=300

Table 4. Standard Deviations of Upper Bound of Confidence Interval of Standard Deviation

ns	Approach	Nor. Approx.	Percentile	BC	BCa	Percentile-*t*
30	Nonpar.	3.629	3.130	3.491	3.990	32.78
	Par. (Iden.)	3.873	4.316	11.62	12.96	20.66
	Par. (True)	3.587	3.958	10.18	11.81	21.21
100	Nonpar.	2.272	2.132	2.365	2.813	11.71
	Par. (Iden.)	1.977	2.019	4.725	5.373	12.34
	Par. (True)	1.917	1.663	4.790	5.520	12.77
300	Nonpar.	1.219	1.200	1.308	1.594	3.930
	Par. (Iden.)	0.971	0.873	2.728	2.978	4.692
	Par. (True)	0.929	0.815	2.598	2.946	4.695

* Method using Gaussian distribution of input variable yields 2.484 for *ns*=30, 1.488 for *ns*=100, and 0.834 for *ns*=300



(a) *ns*=10 (b) *ns*=100 (c) *ns*=300

Figure 1. Histograms of Upper Bound of Confidence Interval of Standard Deviation

In summary, when the distribution is not Gaussian and COV is relatively small, e.g., COV<0.2, the performances of five bootstrap methods are quite identical to each other and are very good. However, when COV is large as shown in the above example, the parametric percentile method has the best performance among five bootstrap methods. Even though the bootstrap methods do not achieve the target confidence level for a small number of samples, as the number of samples increases, the obtained confidence levels tend to converge to the target confidence level while the method using Gaussian distribution of input variable does not.

When the input variable has the Gaussian distribution, the method using Gaussian distribution of the input variable needs to be used because it has an exact formulation of calculating the confidence interval of the standard deviation. The bootstrap method can be used, but it might not be as accurate as the method using the Gaussian distribution of input variable because of the randomness of the bootstrap samples. The bootstrap method can be applied to any types of distribution, and the test results for various types of distributions are presented in Ref. 19.

The upper bound of the confidence interval calculated from the method using the Gaussian distribution of input variable and the bootstrap method will be used to calculate the adjusted standard deviation, which is introduced in Section 4.

3.3 Confidence Interval of Correlation Coefficient

It is known that as *ns* goes to infinity, the sample correlation parameter follows a Gaussian distribution [22] as

$$\tilde{\theta} \sim N\left(\theta, \frac{1}{ns}\left\{4w\frac{dg^{-1}(\tilde{\tau})}{d\tilde{\tau}}\right\}^2\right) \tag{16}$$

where $\theta = g^{-1}(\tau)$, $w = \sqrt{\frac{1}{ns}\sum_{i=1}^{ns}\left(w_i + \tilde{w}_i - 2\bar{w}\right)^2}$, $w_i = \frac{1}{ns}\sum_{j=1}^{ns} I_{ij}$, $\tilde{w}_i = \frac{1}{ns}\sum_{j=1}^{ns} I_{ji}$, and $\bar{w} = \frac{1}{ns}\sum_{i=1}^{ns} w_i$. If $x_{1j} < x_{1i}$ or $x_{2j} < x_{2i}$ (x_{1i} and x_{2i} are i^{th} sample point for X_1 and X_2), then $I_{ij} = 1$, otherwise, $I_{ij} = 0$. Thus, the confidence interval for the correlation parameter for $100\times(1-\alpha)$ of the confidence level is obtained as

$$\Pr\left[\tilde{\theta} - h \le \theta \le \tilde{\theta} + h\right] = 1 - \alpha \tag{17}$$

where $z_{\alpha/2}$ is the CDF value of Gaussian distribution evaluated at $\alpha/2$ and $h = z_{\alpha/2}\frac{1}{\sqrt{ns}}4w\left|\frac{dg^{-1}(\tilde{\tau})}{d\tilde{\tau}}\right|$.

Using the lower and upper bounds of the confidence interval for the correlation parameter θ, the upper and lower bounds of the confidence interval for the correlation coefficient τ are calculated from $\tau = g(\theta)$ using Eq. (2) or explicit functions in Ref. 6. The confidence interval of correlation parameter is accurately estimated regardless of copula function types; the bootstrap method is not necessary to obtain the confidence interval of the correlation parameter in this paper.

4. ASSESSMENT OF CONFIDENCE LEVEL OF INPUT MODEL

When an input model is estimated from given limited data, we are concerned how much confidence level of output performance the estimated input model provides. However, it is difficult to predict the accurate confidence level of output performance of the RBDO result because the confidence level of output performance can be different for different RBDO problems according to the location of most probable point (MPP) on the β_t-contour at the optimum design even though the same input model is used. Thus, the confidence level of the input model needs to be first estimated before stepping into the estimation of the confidence of output performance. Even though the confidence level of the input model is not necessarily equivalent to the output confidence level, if a conservative measure for estimating the input confidence

level, i.e., a β_t-contour is used, then it can be assured that the confidence level of output performance is at least larger than the confidence level of input model. Section 4.1 explains the β_t-contour as a measure of input confidence level.

To have an input model with a target confidence level, confidence intervals of the input parameters need to be used to offset the prediction error of the estimated input parameters. However, neither the upper nor lower bounds of the confidence intervals of mean and correlation yield reliable design. In Section 4.2, the adjusted parameters obtained from the confidence intervals of the input parameters are introduced. Section 4.3 shows simulation test results for the confidence levels of input models obtained using the proposed method.

4.1 Measure of Input Confidence Model

The reliability β_t-contour acts as a safety barrier that locates the optimum design point away from constraint boundary with the target probability of failure. That is, the MPP search is carried out on the β_t-contour for the inverse reliability analysis. Therefore, if the β_t-contour is large enough for the optimum design point to be away from the constraint boundary with the target probability of failure, it means that the obtained optimum design is reliable.

Suppose X_1 and X_2 have marginal distributions, $u = F_{X_1}(x_1)$ and $v = F_{X_2}(x_2)$ correlated with the Frank copula. The β_t-contour in the standard normal space can be obtained as

$$u_1^2 + u_2^2 = \beta_t^2 \tag{18}$$

where u_1 and u_2 are the independent standard normal variables, and β_t is the target reliability index, $\beta_t = 2.0$ for 2-σ design.

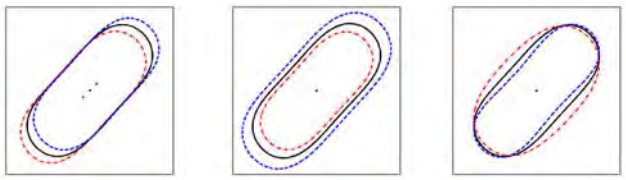
Once the joint distribution is obtained from the marginal distributions and the copula, u_1 and u_2 can be transformed to the original space using the Rosenblatt transformation [23] as

$$\begin{aligned} u_1 &= \Phi^{-1}\left(F_{X_1}(x_1)\right) = \Phi^{-1}(u) \\ u_2 &= \Phi^{-1}\left(F_{X_2}(x_2|x_1)\right) \\ &= \Phi^{-1}\left(\frac{e^{-\theta u}\left(e^{-\theta v}-1\right)}{e^{-\theta}-1+\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}\right) \end{aligned} \tag{19}$$

where the marginal distributions, $u = F_{X_1}(x_1)$ and $v = F_{X_2}(x_2)$, involve the mean and standard deviations of X_1 and X_2. Substituting Eq. (19) into Eq. (18), the β_t-contour for the Frank copula can be obtained. Once the explicit copula function is available, the β_t-contour can be obtained for any type of copulas using Eq. (18).

To observe how the input parameters affect the β_t-contour shape, the β_t-contour using the lower and upper bound of confidence intervals of input parameters and true parameters are drawn as a dashed line, dash dotted line, and solid line, respectively. It is assumed that the distribution types are correctly identified. As shown in Figs. 2 (a) and (c), neither the lower nor upper bounds of confidence intervals of the mean and correlation coefficient fully cover the true β_t-contour, which means that the β_t-contour may or may not yield reliable design. On the other hand, the upper bound of the confidence interval of the standard deviation fully covers the β_t-contour, and thus, it can be readily used to obtain the β_t-contour covering the true β_t-contour, which will lead to the reliable optimum design.

However, the prediction errors in the mean and correlation coefficient still exist in RBDO problems when the available data is insufficient. Therefore, instead of using the estimated mean and correlation coefficient, adjusted parameters using the confidence intervals of the mean and correlation coefficient are proposed to ensure that the estimated β_t-contour with the adjusted parameters covers the true β_t-contour, which will lead to a desirable confidence level of the input model.



(a) Mean (b) Standard Deviation (c) Correlation Coefficient

Figure 2. β_t-contours Using True, Lower and Upper bound of Each Parameter [1]

4.2 Adjusted Parameters Using Confidence Intervals of Input parameters

Since the confidence interval of the mean cannot be used to enlarge β_t-contour, the change in the sample standard deviation ($\Delta\tilde{\sigma}$) caused by the change in the sample mean $\Delta\tilde{\mu}$ is added to the confidence interval of the standard deviation $\tilde{\sigma}^U$ to obtain the adjusted standard deviation $\tilde{\sigma}^A$ as

$$\tilde{\sigma}^A = \tilde{\sigma}^U + \Delta\tilde{\sigma} \approx \tilde{\sigma}^U + \frac{\tilde{\sigma}}{\tilde{\mu}} \times \Delta\tilde{\mu} \tag{20}$$

where it is assumed that $\frac{\Delta\tilde{\sigma}}{\Delta\tilde{\mu}} \approx \frac{\tilde{\sigma}}{\tilde{\mu}}$, and $\Delta\tilde{\mu} = \left|\tilde{\mu}^U - \tilde{\mu}\right| = \left|\tilde{\mu} - \tilde{\mu}^L\right|$ since the sample mean is the middle point of the confidence interval for the Gaussian distribution. In Eq. (20), the COV ($\tilde{\sigma}/\tilde{\mu}$) functions as a scale factor such that the effect of the confidence interval of the mean on the adjusted standard deviation is proportional to COV. Using the estimated mean and the adjusted standard deviation, parameters *a* and *b* of the identified distribution can be estimated using explicit functions, which are expressed as mean and standard deviation in Ref. 6.

Like the mean, the β_t-contours for the lower and upper bounds of the correlation coefficient do not yield reliable design when the sample standard deviation is not large enough to cover the true contour. To resolve this problem, the adjusted correlation coefficient is used. As shown in Fig. 2, neither the lower nor upper bounds of the confidence interval of correlation coefficient cover the true β_t-contour. However, when the lower bound of the correlation coefficient with the adjusted standard deviation is used, it yields more reliable design than when the upper bound with the adjusted standard deviation is used. Thus, when the true correlation coefficient is small, the adjusted correlation coefficient needs to be close to the estimated one. Otherwise, the adjusted correlation coefficient needs to be underestimated. Thus, the adjusted correlation coefficient is proposed as

$$\tilde{\tau}^A = \tilde{\tau} - \tilde{\tau} \times \max\left(\left|\tilde{\tau}^U - \tilde{\tau}\right|, \left|\tilde{\tau} - \tilde{\tau}^L\right|\right) \tag{21}$$

such that it can be applied to both small and large correlation coefficients. As the number of samples increases, the adjusted correlation coefficient converges to the true correlation coefficient.

4.3 Confidence Levels of Input Model

Let X_1 have a lognormal distribution with $\mu_{X_1} = 5.0$ and $\sigma_{X_1} = 5.0$ (COV=1.0); X_2 have a Gaussian distribution with $\mu_{X_2} = 5.0$ and $\sigma_{X_2} = 1.0$ (COV=0.2). X_1 and X_2 are correlated with the Frank copula. From the true input model, a different number of samples, *ns*=30, 100, and 300, are randomly generated for a sufficient number of trials, 300. Using the generated samples, the marginal distributions and copulas are identified and their parameters are quantified. If the marginal Gaussian distributions are identified as correct ones, the method using Gaussian distribution of input variable is used to calculate the confidence interval of the standard deviation. If not, the bootstrap method is used. Once the marginal distributions and copulas are obtained, the β_t-contour can be obtained. The input confidence level is assessed by calculating the probability that the obtained β_t-contour is larger than the true β_t-contour over 300 trials. The target confidence level is 97.5%.

Table 5 shows the obtained input confidence levels using the identified marginal distributions and copulas. Since the bootstrap more accurately calculates the upper bound of the confidence interval of standard deviation for a non-Gaussian distribution, it yields a more desirable confidence level than the method using the Gaussian distribution of input variable. Likewise, when the true marginal distribution types are used, the input confidence levels using the bootstrap method are more accurate than the method using the Gaussian distribution of input variable as shown in Table 6. When the distribution types are correct, the performance of the bootstrap method is improved, so that the results using the true CDFs are better than those using the identified CDFs as shown in Tables 5 and 6.

Table 5. Input Confidence Levels Using Identified CDF Types

ns	$\tau = 0.2$		$\tau = 0.5$		$\tau = 0.8$	
	Norm.	Bootstrap	Norm.	Bootstrap	Norm.	Bootstrap
30	83	87	78	84	77	83
100	85	90	85	89	83	87
300	93	96	89	93	90	92

Table 6. Input Confidence Levels Using True CDF Types

ns	$\tau = 0.2$		$\tau = 0.5$		$\tau = 0.8$	
	Norm.	Bootstrap	Norm.	Bootstrap	Norm.	Bootstrap
30	86	92	82	90	82	89
100	88	96	90	96	86	92
300	94	97	92	97	92	94

5. NUMERICAL EXAMPLES

In this section, a mathematical example and an M1A1 tank roadarm with correlated non-Gaussian input variables are used to show how the bootstrap method yields more reliable design than the method using the Gaussian distribution of input variable. To carry out RBDO, the MPP-based DRM [24] is used to more accurately calculate the probability of failure than the FORM.

5.1 Mathematical Example

Let X_1 and X_2 have lognormal and Gaussian distributions, $X_1 \sim LN\left(3, 1.5^2\right)$ and $X_2 \sim N\left(3, 0.3^2\right)$, which are correlated with the Frank copula and $\tau = 0.7$. For the comparison study, three types of input models – one model with estimated parameters, another model with adjusted parameters obtained from the method using the Gaussian distribution of input variable, and the other model with adjusted parameters obtained from the bootstrap method – are tested. The output confidence levels are assessed using 100 data sets with *ns*=30, 100, and 300, which are randomly generated from the true input model. The marginal distribution, the copula type, and their parameters are determined from each data set. Using the estimated input models from 100 data sets, the RBDO is carried out. The target probability of failure is given as 2.275%.

An RBDO problem is formulated to

$$\begin{aligned}
&\text{minimize} \quad \text{cost}(\mathbf{d}) = d_1 + d_2 \\
&\text{subject to} \quad P\left(G_i\left(\mathbf{X}\right) \ge 0\right) \le P_{F_i}^{Tar} (= 2.275\%),\ i = 1,2,3 \\
&\qquad \mathbf{d} = \boldsymbol{\mu}\left(\mathbf{X}\right),\ 0 \le d_1, d_2 \le 10 \\
&G_1(\mathbf{X}) = 1 - \left(0.9010X_1 - 0.4339X_2 + 1.5\right)^2 \\
&\qquad \times \left(0.4339X_1 + 0.9010X_2 + 2\right)/20 \\
&G_2(\mathbf{X}) = 1 - \left(X_1 + X_2 - 2.8\right)^2/30 - \left(X_1 - X_2 - 12\right)^2/120 \\
&G_3(\mathbf{X}) = 1 - 200/\{2.5(0.9010X_1 - 0.4339X_2 - 3)^2 \\
&\qquad + 8\left(0.4339X_1 + 0.9010X_2\right) + 5\}
\end{aligned} \tag{22}$$

Table 7 shows the minimum, mean, and maximum values of the probabilities of failure P_{F_1} and P_{F_2} for two active constraints evaluated at the optimum designs using the Monte Carlo simulation (MCS). The output confidence levels are

estimated by calculating the probability that the obtained probability of failure is smaller than the target probability of failure. As shown in Table 7, when the input model with the estimated parameters is used for *ns*=30, the mean value of P_{F_1} (4.313%) is larger than the target probability of failure, 2.275%. The maximum value of P_{F_1} (20.842%) is not even close to 2.275% due to the wrong identification and quantification of the input model. This means that even if the mean values of the probability of failure using the estimated parameters are closer to the true target probability of failure, it is not desirable to trust the RBDO results due to the lack of statistical information. Thus, the output confidence level (54%) is significantly smaller than the target confidence level of 97.5%.

On the other hand, when the input model with the adjusted parameters is used, the mean values of P_{F_1} and P_{F_2} are smaller than 2.275%, which gives us more confidence than the case of estimated parameter. Accordingly, the obtained output confidence levels using the adjusted parameters become much closer to the target confidence level, 97.5%. When the number of samples is small, e.g., *ns*=30, the estimated output confidence levels using the bootstrap method are still smaller than the target confidence level due to the incorrect identification of marginal distribution and the copula. However, as the number of samples increases, the output confidence levels using the bootstrap method are closer to the target confidence level.

Table 7. Probabilities of Failure and Output Confidence Levels

ns	Par.	Estimated		Adjusted (Bootstrap)	
		P_{F_1}	P_{F_2}	P_{F_1}	P_{F_2}
30	Min	0.079	0.317	0.000	0.050
	Mean	4.313	2.127	0.600	0.995
	Max	20.842	9.577	10.633	6.176
	Conf.	54	73	94	93
100	Min	0.190	0.454	0.004	0.394
	Mean	2.668	1.725	0.562	1.467
	Max	16.71	4.833	8.277	3.727
	Conf.	48	82	97	93
300	Min	0.700	0.684	0.067	0.865
	Mean	2.393	1.695	0.707	1.675
	Max	4.647	3.275	2.300	2.960
	Conf.	49	91	99	96

*Optimum design using true input model is (3.623, 1.770) ; P_{F_1} = 2.250%, P_{F_2} =1.922%

5.2 M1A1 Abrams Tank Roadarm

The roadarm in the M1A1 tank is modeled using 1572 eight-node isoparametric finite elements (SOLID45) and four beam elements (BEAM44) of a commercial program, ANSYS [25], as shown in Fig. 3. The material of the roadarm is S4340 steel with Young's modulus E=3.0×10^7 psi and Poisson's ratio ν=0.3. The durability analysis of the roadarm is carried out to obtain the fatigue life contour using Durability and Reliability Analysis Workspace (DRAW) [26,27]. The fatigue lives at the 13 critical nodes are selected for design constraints of the RBDO in Fig. 4.

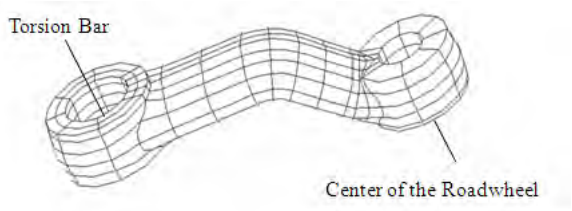


Figure 3. Finite Element Model of Roadarm

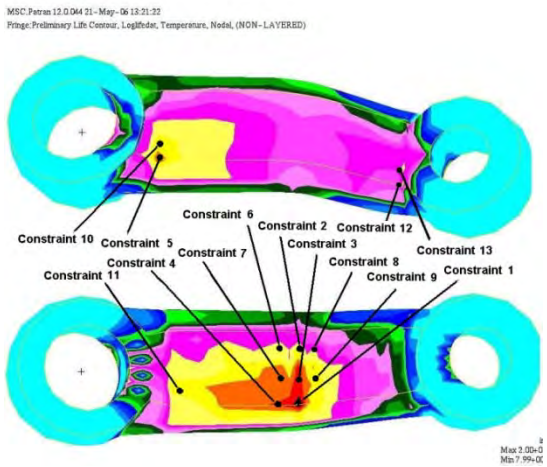


Figure 4. Fatigue Life Contour and Critical Nodes of Roadarm

In Fig. 5, the widths (x_1-direction) of the cross-sectional shapes are defined as design variables, d_1, d_3, d_5, and d_7, at intersections 1, 2, 3, and 4, respectively, and the heights (x_3-direction) of the cross-sectional shapes are defined as design variables, d_2, d_4, d_6, and d_8. Table 8 shows the initial design point, the lower and upper bounds of eight design variables with their standard deviations and distribution types, and the four material parameters with their means and standard deviations.

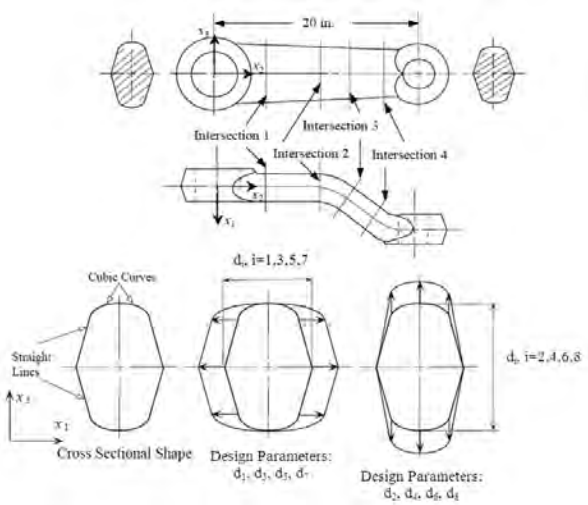


Figure 5. Shape Design Variables for Roadarm

Table 8. Random Variables and Fatigue Material Properties

Random Variables	Lower Bound	Initial Design	Upper Bound	Standard Deviation	Distribution Type

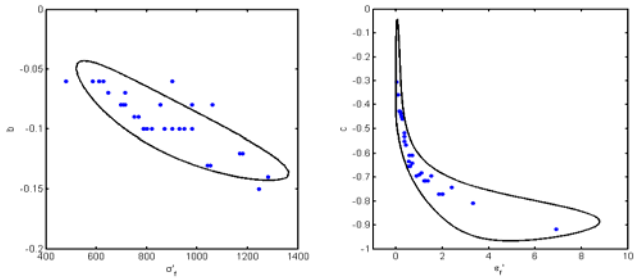
	$\mathbf{d}^L$	$\mathbf{d}^0$	$\mathbf{d}^U$		
d_1	1.3500	1.7500	2.1500	0.0875	Gaussian
d_2	2.6496	3.2496	3.7496	0.1625	Gaussian
d_3	1.3500	1.7500	2.1500	0.0875	Gaussian
d_4	2.5703	3.1703	3.6703	0.1585	Gaussian
d_5	1.3563	1.7563	2.1563	0.0878	Gaussian
d_6	2.4377	3.0377	3.5377	0.1519	Gaussian
d_7	1.3517	1.7517	2.1517	0.0876	Gaussian
d_8	2.5085	2.9085	3.4085	0.1454	Gaussian

Fatigue Material Properties

Non-design Uncertainties	Mean	Standard Deviation	Distribution Type
Fatigue Strength Coefficient, σ'_f	177000	44250	Lognormal
Fatigue Strength Exponent, b	-0.0730	0.018	Gaussian
Fatigue Ductility Coefficient, ε'_f	0.4100	0.205	Lognormal
Fatigue Ductility Exponent, c	-0.6000	0.150	Gaussian

To test the input model with the confidence level, experimental data need to be used to obtain the adjusted standard deviation and correlation coefficient. However, the experimental data of S4340, which is used in the roadarm, is not available. Thus, in this paper, 30 paired data are generated from an assumed true input model.

First, it is assumed that Frank copula ($\tau = -0.683$) for σ'_f and b, and the Gaussian copula ($\tau = -0.906$) for ε'_f and c, respectively, are the true copulas. As the two copulas well describe the experimental data of SAE 950X [8] as shown in Fig. 6, it seems to be reasonable to select these two copulas to model the joint CDFs of the four correlated random parameters of S4340. The marginal distribution types of S4340 are assumed to be the same as those of SAE 950X.



(a) σ'_f and b (b) ε'_f and c

Figure 6. Joint PDF Contours of Gaussian and Frank Copula Identified from 29 Paired Data of 950X Steel

Second, once the copula and marginal distribution types are obtained, the mean and standard deviation of S4340 need to be determined. The mean values of four fatigue material properties of S4340 are known, but the standard deviations are unknown. Therefore, the standard deviations are assumed using COV of SAE 950X. The coefficient of variation of SAE 950X is 115% for ε'_f and 25% for other material properties [8]. Since S4340 is a stronger material than SAE 950X, in this paper, it is assumed that the COV of S4340 is 50% for ε'_f and 25% for other material properties to estimate the standard deviation as shown in Table 8.

Assuming that a true input model has the above statistical information on S4340, 30 paired data are randomly generated, and RBDO is carried out using the estimated input model with and without confidence level. Table 9 shows the estimated and adjusted parameters, and the target confidence level is specified as 95% in this roadarm example.

Table 9. Estimated and Adjusted Parameters

	σ'_f	b	ε'_f	c
$\tilde{\mu}$	176738	–0.073	0.395	–0.594
$\tilde{\sigma}$	35141	0.015	0.143	0.110
$\tilde{\sigma}^A$	45356	0.020	0.223	0.155
$\tilde{\tau}$	–0.701		–0.921	
$\tilde{\tau}^A$	–0.596		–0.866	
Copula	Gaussian		Frank	

The RBDO is formulated to

$$\begin{aligned} &\text{minimize} \quad \text{cost}(\mathbf{d}) \\ &\text{subject to} \quad P(G_i(\mathbf{X}) \ge 0) \le P_{F_i}^{Tar},\ i = 1 \sim 13 \\ &\mathbf{d} = \boldsymbol{\mu}(\mathbf{X}),\ \mathbf{d}_L \le \mathbf{d} \le \mathbf{d}_U,\ \mathbf{d} \in R^n,\ P_{F_i}^{\text{Tar}} = 2.275\% \\ &G_i(\mathbf{X}) = 1 - \frac{L(\mathbf{X})}{L_t},\ i = 1 \sim 13 \end{aligned} \tag{23}$$

$\text{cost}(\mathbf{d})$: Weight of Roadarm

$L(\mathbf{X})$: Crack Initiation Fatigue Life,

L_t : Crack Initiation Target Fatigue Life (=5 years)

Table 10 shows a comparison of RBDO results for various input models where the true input model is assumed to have Frank and Gaussian copulas with the fatigue material properties in Table 9. First, the RBDO results are compared for the independent and correlated input fatigue material properties. As shown in the table, when the correlation between material properties is considered, the optimized weight of the roadarm is significantly reduced from 592.22 to 514.02 for the same target reliability. This is because the material properties are highly and negatively correlated. Thus, it is very important to correctly model the correlation between material properties to carry out the RBDO.

Table 10. RBDO Comparison

	Initial	Independent	Correlated		
			True	Estimated	Adjusted
d_1	1.750	2.194	1.928	1.954	2.052
d_2	3.250	2.650	2.650	2.650	2.650
d_3	1.750	2.602	2.067	2.030	2.038
d_4	3.170	3.010	2.577	2.623	2.702
d_5	1.756	2.656	1.776	1.684	1.804
d_6	3.038	2.538	3.535	3.538	3.500
d_7	1.752	2.422	2.075	2.025	2.152
d_8	2.908	2.895	2.512	2.508	2.754
Cost	515.09	592.22	514.02	509.44	531.64

Second, when the input model with the estimated standard deviation is used, the underestimated standard deviations (see Tables 8 and 9) yield an unreliable optimum design with an optimum cost that is smaller than the optimum cost obtained using the true input model (509.44 vs. 514.02). Since the MCS cannot be used for the benchmark test for this problem due to computational cost, the comparison of optimum costs is used as a measure to check whether the obtained optimum design is reliable or not. On the other hand, when the input model with the adjusted parameters is used, the obtained optimum cost is higher than the optimum cost obtained from the true input model (531.64 vs. 514.02), which indicates the obtained optimum design is reliable. Even though the optimum cost using the estimated input model is closer to the true optimum cost than the one using the adjusted parameters, the estimated input model does not provide the confidence level of output performance whereas the input model with the adjusted parameters does. Accordingly, the input model with the adjusted standard deviation and correlation coefficient is indeed necessary to obtain a reliable optimum design.

6. CONCLUSIONS

In many engineering applications, only insufficient test data are available for input variables, and thus, the input statistical model obtained from the insufficient data could be inaccurate. Thus, the RBDO with confidence level is proposed to offset the inaccurate estimation of the input model by using the adjusted standard deviation and correlation coefficient. The adjusted standard deviation is obtained from the confidence intervals of mean and standard deviation. If the input variables have a Gaussian distribution, the method using the Gaussian distribution of input variables is exact. If not, it yields an inaccurate estimation of the confidence interval of standard deviation. Thus, in this paper, the bootstrap method is proposed to be used to calculate the confidence interval of standard deviation and, thus, the adjusted standard deviation. The input model with the adjusted parameters obtained from the bootstrap method is used to assess the input and output confidence levels for non-Gaussian distributions. Numerical test shows that the percentile method has the most desirable performance out of 5 candidate bootstrap methods for the parametric approach. Numerical examples also show that the input model using the parametric percentile bootstrap method yields more reliable design than the one using other method for non-Gaussian distributions.

7. ACKNOWLEDGEMENT

Research is jointly supported by the Automotive Research Center, which is sponsored by the U.S. Army TARDEC and ARO Project W911NF-09-1-0250. These supports are greatly appreciated.

8. REFERENCES

1. Noh, Y., Choi, K.K., and Lee, I., "Reliability-Based Design Optimization with Confidence Level under Input Uncertainty," *35th ASME Design Automation Conference*, San Diego, California, August 30- September 2, 2009.
2. Haldar, A., and Mahadevan, S., *Probability, Reliability, and Statistical Methods in Engineering Design*, John Wiley & Sons, New York, 2000.
3. Efron, B., *The Jackknife, the Bootstrap, and Other Resampling Plans*, SIAM, Philadelphia, 1982.
4. Efron, B., and Tibshirani, R., *An Introduction to the Bootstrap*, Chapman&Hall/CRC, 1993.
5. Picheny, V., Kim, N.H., and Haftka, R.T., "Application of Bootstrap Method in Conservative Estimation of Reliability with Limited Samples," *Structural and Multidisciplinary Optimization*, Vol. 41, No. 2, pp. 205-217, 2010.
6. Noh, Y., Choi, K.K., and Lee, I., "Identification of Marginal and Joint CDFs Using the Bayesian Method for RBDO," *Structural and Multidisciplinary Optimization*, Vol. 40, No. 1, pp. 35-51, 2010.
7. Nelsen, R.B., *An Introduction to Copulas*, Springer, New York, 1999.
8. Socie, D.F., Seminar notes: "Probabilistic Aspects of Fatigue", 2003, URL: http://www.fatiguecaculator.com [cited May, 8 2008].
9. Annis, C., "Probabilistic Life Prediction Isn't as Easy as It Looks," *Journal of ASTM International*, Vol. 1, No.2, pp. 3-14, 2004.
10. Efstratios, N., Ghiocel, D., and Singhal, S., *Engineering design reliability handbook*, CRC press, New York, 2004.
11. Pham, H., *Springer Handbook of Engineering Statistics*, Springer, London, 2006.
12. Huard, D., Évin, G., and A.-C. Favre, "Bayesian Copula Selection," *Computational Statistics and Data Analysis*, COMSTA3137, Vol. 51, No. 2, pp. 809-822, 2006.
13. Genest, C., Ghoudi, K., and Rivest, L.P., "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distribution," *Biometrika*, Vol. 82, No. 3, pp. 543-552, 1995.
14. Hürliman, W., "Fitting Bivariate Cumulative Returns with Copulas," *Computational Statistics & Data Analysis*, Vol. 45, No. 2, pp. 355–372, 2004.
15. Roch, O., and Alegre, A. "Testing the Bivariate Distribution of Daily Equity Returns Using Copulas. An Application to the Spanish Stockmarket," *Computational Statistics & Data Analysis*, Vol. 51, No. 2, pp. 1312–1329, 2006.
16. Noh, Y., Choi, K.K., and Lee, I., "Comparison Study between MCMC-based and Weight-based Bayesian Methods for Identification of Joint Distribution." *Structural and Multidisciplinary Optimization*, Submitted, January 2010.
17. Silva, R.S., and Lopes, H.F., "Copula, Marginal Distributions and Model Selection: a Bayesian Note," *Statistics and Computing*, Vol. 18, No. 3, pp. 313-320, 2008.
18. Shao, J., and Tu, D., *The Jackknife and Bootstrap*, Springer, 1995.
19. Bonate, B.L., *Phamacokinetic-pharmacodynamic Modeling and Simulation*, Springer, 2005.
20. Mooney, C.Z., and Duval, R.D., *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Sage Publication, 1993.
21. Liu, R.Y., and Singh, K., "Invited Discussion on Theoretical Comparison of Bootstrap Confidence Intervals", *Annals of Statistics*, Vol. 16, pp.978-979, 1988.
22. Genest, C., and Favre, A-C., "Everything You Always Wanted to Know about Copula Modeling but Were Afraid

to Ask." *Journal of Hydrologic Engineering*, Vol. 12, No. 4, pp. 347-368, 2007.
23. Rosenblatt, M., "Remarks on A Multivariate Transformation," *Annal of Mathematical Statistics*, Vol. 23, pp. 470-472, 1952.
24. Lee, I., Choi, K.K., Du, L., and Gorsich, D., "Inverse Analysis Method Using MPP-Based Dimension Reduction for Reliability-Based Design Optimization of Nonlinear and Multi-Dimensional Systems," *Special Issue of Computer Methods in Applied Mechanics and Engineering: Computational Methods in Optimization Considering Uncertainties*, Vol. 198, No. 9-12, pp. 14-27, 2008.
25. Swanson Analysis System Inc., *ANSYS Engineering Analysis System User's Manual*, Vol. I & II, Houston, PA, 1989.
26. Center for Computer-Aided Design, College of Engineering, *DRAW Concept Manual*, The University of Iowa, Iowa City, IA, 1999a.
27. Center for Computer-Aided Design, College of Engineering, *DRAW User Reference*, The University of Iowa, Iowa City, IA, 1999b.