Journal of the Korean Nuclear Society
Volume 33, Number 3, pp.270~282, June, 2001

Quantification of Entire Change of Distributions Based on Normalized Metric Distance for Use in PSAs

Seok-Jung Han and Moon-Hyun Chun
Korea Advanced Institute of Science and Technology
373-1 Kusong-dong, Yusong-gu, Taejon 305-701, Korea
hanseok@netian.com
Nam-IL Tak
Korea Atomic Energy Research Institute
150 Dukjin-dong, Yusong-gu, Taejon 305-353, Korea

(Received August 24, 2000)

Abstract

A simple measure of uncertainty importance based on normalized metric distance to quantify the entire change of cumulative distribution functions (CDFs) has been developed for use in probability safety assessments (PSAs). The metric distance measure developed in this study reflects the relative impact of distributional changes of inputs on the change of an output distribution, while most of the existing uncertainty importance measures reflect the magnitude of relative contribution of input uncertainties to the output uncertainty. Normalization is made to make the metric distance measure a dimensionless quantity. The present measure has been evaluated analytically for various analytical distributions to examine its characteristics. To illustrate the applicability and strength of the present measure, two examples are provided. The first example is an application of the present measure to a typical problem of a system fault tree analysis and the second one is for a hypothetical non-linear model. Comparisons of the present result with those obtained by existing uncertainty importance measures show that the metric distance measure is a useful tool to express the measure of uncertainty importance in terms of the relative impact of distributional changes of inputs on the change of an output distribution.

Key Words : uncertainty importance; metric distance; uncertainty analysis; importance measure.

1. Introduction

Currently, the quantification of uncertainty has become an essential part of the overall PSA. Knowing how input uncertainties are related to the output uncertainty it is possible to focus one's attention on the areas where more information is needed. In this respect, the uncertainty importance measure is an important tool to find where future efforts should be directed. There are

Table 1. Uncertainty Importance Measures Proposed in the Recent PSA Study

Proposers	Uncertainty importance measure	Criteria and characteristics
Nakashima and Yamato [1]	$UI(j)=\frac{S_j \mathrm{Var}(X_j)}{\mathrm{Var}(Y)}=\frac{\mathrm{Var}(X_j)\partial \mathrm{Var}(Y)}{\mathrm{Var}(Y)\partial \mathrm{Var}(X_j)}=\frac{\partial \ln[\mathrm{Var}(Y)]}{\partial \ln[\mathrm{Var}(X_j)]}$ where $S_j=\frac{\partial \mathrm{Var}(Y)}{\partial \mathrm{Var}(X_j)}=\mathrm{Var}(Y)\big\|_{\mathrm{Var}(X_j)=1}-\mathrm{Var}(Y)\big\|_{\mathrm{Var}(X_j)=0}$	The variance-importance of component j
Bier [2]	$UI(j)=\frac{\mathrm{Var}(X_j)\partial \mathrm{Var}(Y)}{\mathrm{Var}(Y)\partial \mathrm{Var}(X_j)}$	A reduction in the output variance
Hora and Iman [3]	$UI(j)=[\mathrm{Var}(Y)-E[\mathrm{Var}(Y\|X_j)]]^{1/2}$	A standard deviation measure: square root of expected reduction in the output variance
Iman [4]	$UI(j)=[\mathrm{Var}(X_j)]^{1/2}\frac{\partial Y}{\partial X_j}\cong\left[\frac{(\sum_{i=1}^{n} y_{i,j}^2)-(\sum_{i=1}^{n} y_{i,j})^2}{n-1}\right]^{1/2}$	A standard deviation measure: square root of expected reduction in the output variance
Helton et al. [5]	$UI(j)=\frac{\partial Y}{\partial X_j}\left[\frac{\mathrm{Var}(X_j)}{\mathrm{Var}(Y)}\right]^{1/2}\Bigg\|_{X=x_0}$	*A percentage contribution in the output:* standardized regression coefficient
Andsten and Vaurio [6]	$UI(j)=\frac{B_j^2 \mathrm{Var}(X_j)}{\sum_r B_i^2 \mathrm{Var}(X_i)},\ B_j=\frac{\partial Y}{\partial X}=Y\big\|_{X_j=1}-Y\big\|_{X_j=0}$	The relative contribution of event j to the total variance of output
Iman and Hora [7]	$UI(j)=R^2$ statistic value $UI(j)=\left[\frac{y_{j,\alpha}}{y_\alpha},\frac{y_{j,1-\alpha}}{y_{1-\alpha}}\right]$	(1) A percentage change in the output: determinant coefficient of regression model (2) *A bivariate measure of the change in* the output distributions, α = probability
Khatib-Rahbar et al. [8]	$UI(j)=[\mu,\sigma^2,\xi],\ \xi=\frac{[y_{j,\alpha}/y_{j,1-\alpha}]}{[y_\alpha/y_{1-\alpha}]},\ \alpha=0.95$	A combination of statistical parameters
Park and Ahn [9]	$UI(j)=\int f_{Y_j}(x)\ln\left[\frac{f_{Y_j}(x)}{f_{Y_o}(x)}\right]dx$	Information theoretic entropy measure based on the definition of Kullback-Leibler information discrimination
The present method	$UI(j)=\left[\int_0^1 (y_\alpha^j-y_\alpha^o)^2 d\alpha\right]^{\frac{1}{2}}/E(Y^o)$	Based on the normalized metric distance

a number of existing uncertainty importance measures that have been proposed by earlier investigators as summarized in Table 1.

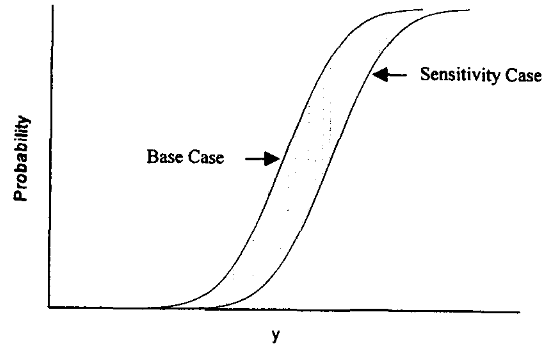
The objective of the uncertainty importance measure may be different depending on the specific goal or the purpose of its application. For example, if one wants to reduce the uncertainty of an output estimate, the uncertainty importance measure that gives more information about the relative impact of the input uncertainty on the output uncertainty should be preferred. However, if an accurate quantification of the output distribution is needed, one should use the uncertainty importance measure that shows the relative impact of distributional changes of inputs on the change of an output distribution. As can be seen in Table 1, while Nakashima and Yamato [1], Bier [2], Hora and Iman [3], Iman [4], Helton et al. [5], Andsten and Vaurio [6] focused on the reduction of the output uncertainty, Iman and Hora [7], Khatib-Rahbar et al. [8], Park and Ahn [9] focused on the precise quantification of an output distribution. One of the key questions that arise from the previous studies is how can we quantitatively compare and rank two results when one result has a high mean value with a narrow uncertainty range while the other result has a lower mean value with a broad uncertainty range. With respect to the reduction of an output uncertainty, the range of uncertainty, which is expressed usually in terms of a variance (or a standard deviation), is more meaningful. With respect to the precise quantification of an output distribution, on the other hand, the relative impact on the change of an output distribution is more important. These uncertainty importance measures are very useful to express the uncertainties that are related to rare events and/or less known phenomena. In this case, input uncertainties used to quantify the output distribution are usually quantified based on the opinions of experts. For this type of problem, in particular, it is necessary to assess the relative impact of different subjective assumptions on the change of the output distribution. In the present work, a simple measure of the uncertainty importance that uses the entire change of CDFs has been developed. The entire change of CDFs is quantified by means of the normalized metric distance between two CDFs. The present uncertainty importance measure focuses on the relative impact of the change of distribution of inputs on the change of an output distribution.

2. Metric Distance Measure

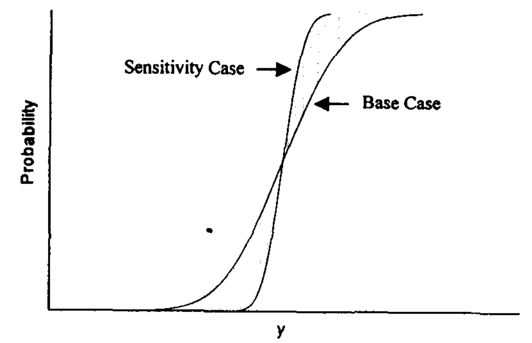
The metric distance measure was originally developed as a measure of fuzziness [10], but its concept can also be used to measure the uncertainty importance expressed by a probabilistic manner. A general form of the metric distance D, called the Minkowski class of distance, is as follows:

$$D = \left(\sum_{x \in X} | f_1(x) - f_2(x) |^w\right)^{1/w} \tag{1}$$

where $f_1(x)$ and $f_2(x)$ are functions of x, and w is a number greater than 1. The Hamming and Euclidean distances are special cases of the Minkowski class of distances for w = 1 and w = 2, respectively. Figure 1 shows the characteristics of



(a) The Hamming distance (dashed area) when output distribution is shifted.



(b) The Hamming distance (dashed area) when uncertainty of output is reduced.

Fig. 1. Characteristics of the Hamming Distance Between Two CDFs

the Hamming distance between two CDFs. The Hamming distance is equivalent to the dashed area shown in Fig. 1.

In the present work, the Euclidean metric distance between two CDFs normalized with the mean of the base distribution is proposed for the measure of uncertainty importance

$$MD\,(i:o) = \frac{\left(\int_0^1 [y_p^i - y_p^o]^2 dp\right)^{\frac{1}{2}}}{E(Y^o)} \qquad (2)$$

where $MD(i{:}o)$ is the metric distance measure in terms of quantiles between the base case and its sensitivity case, y_p^o is the pth quantile of a CDF for the base case, y_p^i is the pth quantile of a CDF for its sensitivity case, and $E(Y^o)$ is the mean of output distribution for the base case. The base case refers to the case where an output distribution is obtained with all the input distributions set to their nominal distribution, whereas the sensitivity case refers to the case where an output distribution is obtained with a change in only one of the input distributions. Normalization is made to make the metric distance measure a dimensionless quantity. The metric distance measure $MD(i{:}o)$ represents an integrated distance of quantiles between two CDFs generated by the base case and its sensitivity case. The quantile of CDF is mathematically defined as the inverse function of CDF. Thus, the metric distance measure $MD(i{:}o)$ defined in the present study has a unique merit in the sense that it can not only be derived for analytical distributions but also it can easily be calculated for analytical and empirical distributions, which will be shown in the next paragraph. Certainly, the metric distance measure $MD(i{:}o)$ can provide information on how much a given input parameter impacts on the output distribution when its input distribution is changed. If the value of y_p^o is equal to that of y_p^i over all ranges, the two distributions become identical and $MD(i{:}o)$ goes to zero. A larger $MD(i{:}o)$ means that there has been a larger distributional change. Thus, the input parameter that gives relatively a large value of $MD(i{:}o)$ means that the input parameter is more important than the other input parameters.

3. Analytical Evaluation of Metric Distance Measure for Various Distributions

The normalized metric distance measure $MD(i{:}o)$ developed in the present work has a unique merit in the sense that it can easily be derived or calculated for analytical distributions as well as for empirical distributions. An analytical derivation of the metric distance measure for analytical distributions and a simple representation of the metric distance measure for empirical distributions are shown here. To examine the characteristics of the metric distance measure, three different cases are considered: In the first two cases, analytical distributions are treated, and an empirical type of distribution is considered in the third case. The first case corresponds to the case when all the analytical distributions have symmetric distributions. The second case consists of two asymmetric distributions (i.e., lognormal and two-parameter Weibull distributions). The detailed derivations can be found in reference 11.

3.1. Case 1: Symmetric Distributions

Typical examples of symmetric distributions are uniform and normal distributions. Let Y be a random variable representing an output and let F be the CDF of Y. Since F is monotonically increasing function, the pth quantile of the random variable Y is expressed by the standardized random variable Z as follows:

$$y_p = \mu + \sigma z_p \qquad (3)$$

where y_p is the pth quantile of the random variable

Y, z_p is the pth quantile of the standardized random variable Z, μ is the mean of Y, and σ is the standard deviation of Y.

For symmetric distributions, the metric distance measure *MD(i:o)* normalized with the mean of the base distribution can analytically be derived by substituting Eq. (3) into Eq. (2):

$$MD(i:o) = \frac{\sqrt{(\mu_i - \mu_o)^2 + (\sigma_i - \sigma_o)^2}}{\mu_o} \tag{4}$$

The subscripts o and i refer to the base case and its distributional sensitivity case, respectively. Equation (4) shows that the magnitude of *MD(i:o)* obtained from symmetric distribution is determined by differences of the mean and standard deviation of Y.

3.2. Case 2: Two Asymmetric Distributions

The metric distance *MD(i:o)* between two symmetric distributions given by Eq. (4) is not valid for asymmetric distributions because of their skewness. However, the magnitude of *MD(i:o)* for the lognormal distribution and that of the Weibull distribution can be determined from their unique features such as their shape and scale factors.

3.2.1. Lognormal Distribution:

The random variable Y has a lognormal distribution when logY is a normal random variable. In reliability and risk analyses of nuclear power plants, a lognormal distribution has been widely used when a given event occurs very infrequently or the range of the given event is very large. Its functional form is given by

$$f(y) = \frac{1}{\sqrt{2\pi}\alpha y} \exp\left[-\frac{\{\ln(y/\beta)\}^2}{2\alpha^2}\right] \tag{5}$$

where α and β are the shape and scale factors of Y, respectively. The pth quantile of lognormal random variable Y is as follows:

$$y_p = \beta \exp(\sqrt{2}\alpha k_p) \tag{6}$$

Here, k_p is expressed by the inverse error function as follows:

$$k_p = erf^{-1}(2p-1) \tag{7}$$

The normalized metric distance measure *MD(i:o)* can analytically be derived by substituting Eq. (6) into Eq. (2) as follows:

$$MD(i:o) = \frac{\sqrt{\beta_i^2 \exp(2\alpha_i^2) + \beta_o^2 \exp(2\alpha_o^2) - 2\beta_i\beta_o \exp(\frac{(\alpha_i + \alpha_o)^2}{2})}}{\beta_o \exp(\frac{\alpha_o^2}{2})} \tag{8}$$

Equation (8) shows that the metric distance measure *MD(i:o)* for the lognormal distribution is determined by unique features such as its shape and scale factors. The metric distance measure *MD(i:o)* for the lognormal distribution is different from that of the symmetric distribution due to its skewness (See Eq. (4) and Eq. (11)).

3.2.2. Two-parameter Weibull Distribution:

The two-parameter Weibull distribution covers a wide range of parameter space and can completely be quantified by specifying the shape factor α and the scale factor β. These two parameters are defined in the probability density function as follows:

$$f(y) = \frac{\alpha}{\beta}\left(\frac{y}{\beta}\right)^{\alpha-1} \exp\left[-(\frac{y}{\beta})^{\alpha}\right] \tag{9}$$

The pth quantile of Y can be derived as follows:

$$y_p = \beta\left(\log\frac{1}{1-p}\right)^{\frac{1}{\alpha}} \tag{10}$$

By substituting Eq. (10) into Eq. (2), the normalized metric distance measure $MD(i{:}o)$ for the Weibull distribution can analytically be derived as:

$$MD(i:o)=\frac{\sqrt{\beta_i^2\Gamma(\frac{2}{\alpha_i}+1)+\beta_o^2\Gamma(\frac{2}{\alpha_0}+1)-2\beta_i\beta_o\Gamma(\frac{1}{\alpha_i}+\frac{1}{\alpha_0}+1)}}{\beta_o\Gamma(\frac{1}{\alpha_o}+1)} \quad (11)$$

where Γ is the gamma function. The metric distance measure $MD(i{:}o)$ for the Weibull distribution is determined by its unique features as in the case of the lognormal distribution.

3.3. Case 3: Empirical Distribution

The output distribution of the PSA results, where simulations such as Monte Carlo are widely used to obtain an output distribution from the complex PSA models, may be considered as one of the most representative empirical distribution. An empirical distribution function S(y) can be obtained directly from the Monte Carlo simulation because of the probabilistic nature of the Monte Carlo simulation:

$$S(y)=\frac{1}{N}\sum_{n=1}^{N}\delta(y>y_n) \quad (12)$$

$$\delta(y>y_n)=\begin{cases}1, \text{ if } y>y_n\\ 0, \text{ otherwise}\end{cases} \quad (13)$$

where N is the sample size and n is the sample index. The quantiles can easily be obtained from the inverse function of S. If the sample sizes of M and N used in the Monte Carlo simulation are the same (i.e., if the sample size of the base case is equal to that of its sensitivity case), the normalized metric distance measure $MD(i{:}o)$ between the two empirical distributions can be expressed as follows:

$$MD(i:o)=\frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}[y_{n/N}^{i}-y_{n/N}^{o}]^2}}{\frac{1}{N}\sum_{n=1}^{N}y_n^o} \quad (14)$$

where $y^{o}_{n/N}$ is the (n/N)th quantile for the base case $(0<n<N)$, and $y^{i}_{n/N}$ is the (n/N)th quantile for its sensitivity case.

4. Examples

Two examples are selected to show the applicability and strength of the metric distance measure developed in the present work. In the first example, a general application of the present uncertainty importance measure for PSA is illustrated. In the second example, on the other hand, an effort has been made to show a more clear advantage of the present importance measure over existing measures to which it was compared.

4.1. Example 1 : System Fault Tree Analysis

To examine the general applicability of the present measure, an uncertainty importance analysis has been performed for the typical example of a system fault tree analysis used in references 4 and 9. The example selected here is the uncertainty analysis associated with estimations of the system unavailability or reliability obtained from a Boolean representation of a system fault tree. The mathematical relationship of the top event is expressed as follows:

$$\begin{aligned}Top(X)=&X_1X_3X_5+X_1X_3X_6+X_1X_4X_5+X_1X_4X_6\\&+X_2X_3X_4+X_2X_3X_5+X_2X_4X_5\\&+X_2X_5X_6+X_2X_4X_7+X_2X_6X_7\end{aligned} \quad (15)$$

In Eq. (15), X_1 and X_2 are initiating events expressed as the number of occurrences per year, and X_3 ~ X_7 are basic events which represent the component failure rate. The same assumptions of input uncertainties used in references 4 and 9 are adopted and shown in Table 2. It is also assumed

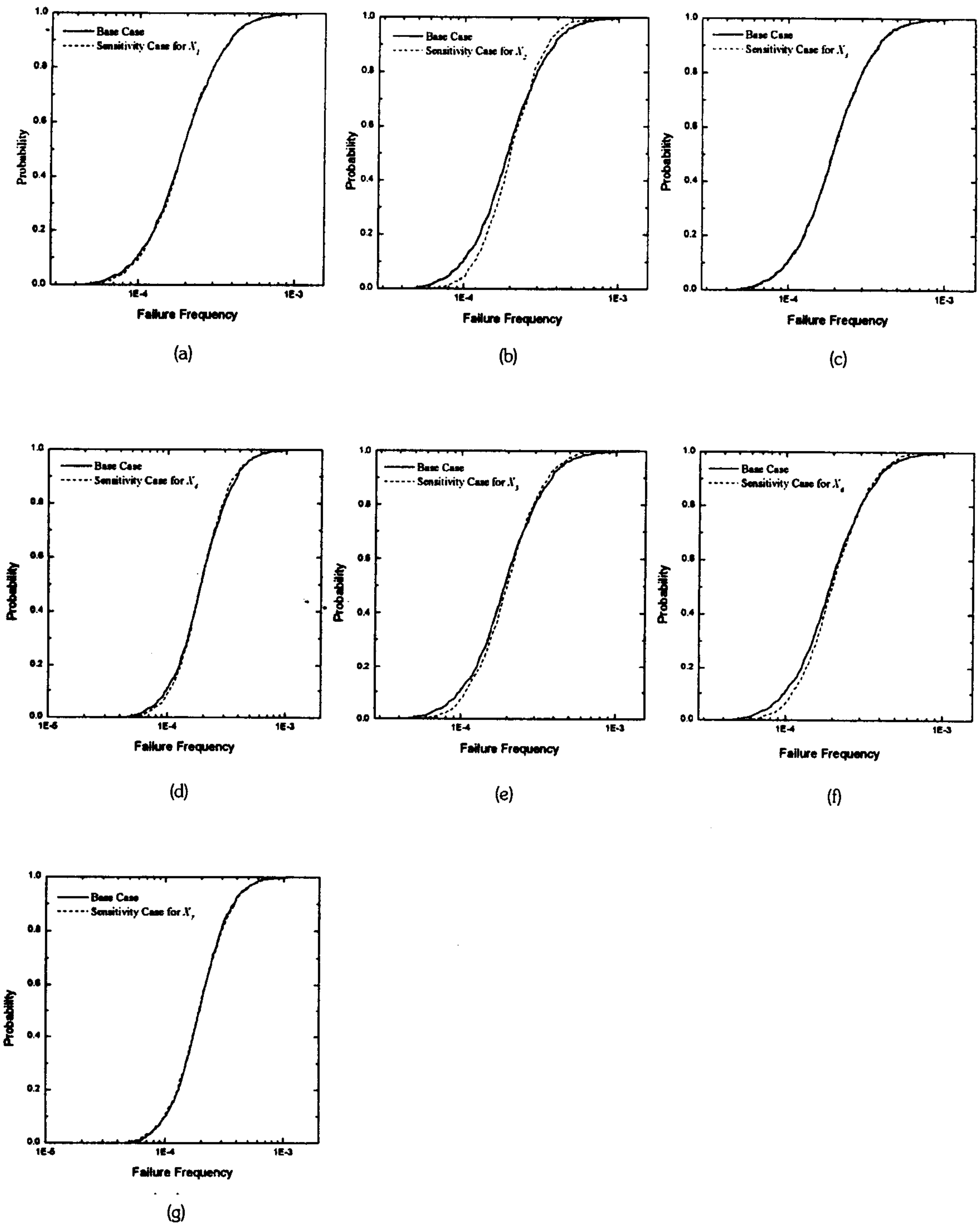


Fig. 2. Empirical Distributions of the Top Event Frequency for the Base Case and the Sensitivity Case of 7 Input Parameters in Example 1

Table 2. Assumptions of Inputs for Example 1

Variable	Event description	Nominal value	Distribution	Error factor
X_1	Initiating event	2	Lognormal	2
X_2	Initiating event	3	Lognormal	2
X_3	Component failure rate	0.001	Lognormal	2
X_4	Component failure rate	0.002	Lognormal	2
X_5	Component failure rate	0.004	Lognormal	2
X_6	Component failure rate	0.005	Lognormal	2
X_7	Component failure rate	0.003	Lognormal	2

that all events are independent of one another.

There are three methods of change in input distribution for the uncertainty importance analysis, i.e., (1) the uncertainty is completely eliminated; (2) the uncertainty range is changed; (3) the type of distribution is changed [9]. In the present example, the uncertainty importance analysis is performed using the first method (1) where one of the input uncertainties is completely eliminated one after another.

The crude Monte Carlo simulations with 1000 sampling for one simulation are implemented to obtain the empirical distributions of the top event frequency for the base case and its sensitivity cases. The base case is the case where the empirical distribution of the top event frequency is calculated using all input distributions which are equal to their original distributions shown in Table 2. The sensitivity case, on the other hand, is the case where the empirical distribution of the top event frequency is obtained while replacing only one of the input distributions with its nominal value. In this case, the metric distance measure can easily be obtained from empirical distributions of the base case and its sensitivity case using Eq. (14).

Figure 2 shows empirical distributions of the top event frequency obtained by crude Monte Carlo simulations for the base case and its sensitivity case with 7 inputs. The metric distance measures for each 7 input parameters have been calculated and their relative impacts on the distribution of the top event frequency have been ranked according to the magnitude of the normalized metric distance measure obtained for each input parameter as shown in Table 2.

In an effort to assess the applicability of the metric distance measure, the results of the present method are compared with two existing results calculated by other typical uncertainty importance measures, i.e., Iman' s standard deviation [4] and Iman and Hora' s bivariate [7] measures. The results of calculation by Iman' s standard deviation can be found in Tables 1 of reference 4. For Iman and Hora' s bivariate measure, Park and Ahn' s results in Tables 3 of reference 9 are used for the comparison. The summary of the results is shown in Table 3. Table 3 shows that the results obtained by the metric distance measure proposed here and two existing measures agree that X_2 is the most important parameter with respect to the uncertainty importance. They are also in agreement about the ranks of top three high rankers (i.e., the top events of X_2, X_5, and X_6). The rankings of the remaining parameters except X_2, X_5, and X_6 seem to be unimportant because

Table 3. Uncertainty Importance Rankings Obtained by the Present and Two Existing Measures for Example 1

Changed variable	Uncertainty importance measure and their rankings: Present Measure		Iman's Standard deviation measure[b]		Iman & Hora's bivariate measure[c]: $R_{0.05}$	$R_{0.95}$	
X_1	1.98×10^{-2}	(6)[a]	2.54×10^{-5}	(6)	1.07	0.99	(5)
X_2	1.31×10^{-1}	(1)	7.23×10^{-5}	(1)	1.25	0.84	(1)
X_3	1.68×10^{-2}	(7)	1.51×10^{-5}	(7)	1.01	0.98	(6)
X_4	5.22×10^{-2}	(4)	4.46×10^{-5}	(4)	1.06	0.98	(7)
X_5	1.18×10^{-1}	(2)	7.20×10^{-5}	(2)	1.16	0.94	(2)
X_6	1.11×10^{-1}	(3)	6.05×10^{-5}	(3)	1.13	0.91	(3)
X_7	3.27×10^{-2}	(5)	2.60×10^{-5}	(5)	1.07	0.98	(4)

[a]Ranking.
[b]See the second column of Table 1 in reference 4.
[c]See Case 1 of Table 3 in reference 9.

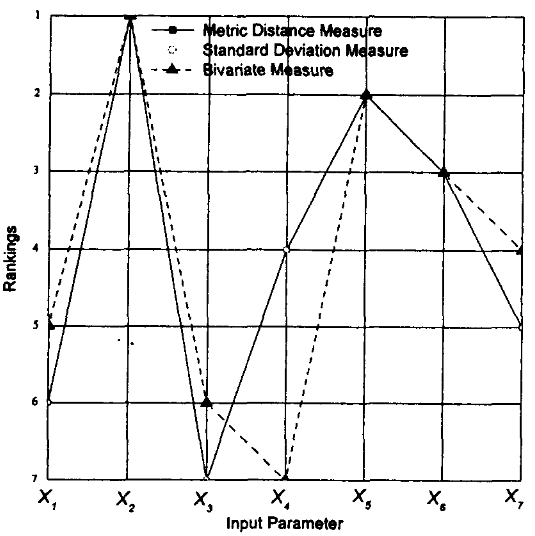


Fig. 3. A Summary Plot of the Uncertainty Importance Rankings Given in Table 3*

*Only two lines can be seen since the results of metric distance and standard deviation measures are the same.

distributional changes of the remaining parameters are negligibly small.

Figure 3 is a summary plot of the uncertainty importance rankings given in Table 3. The summary plot is very useful when different measures are applied to the same problem. The overall trend of the results by different measures can be seen clearly.

From Table 3 and Fig. 3, it can be concluded that the present measure as well as two existing measures provides good information about the uncertainty importance for this example in spite of their different approaches.

4.2. Example 2 : Hypothetical Analytical Model

In the above example, a typical application of the present measure for PSA is made. From this example, however, one could not see a clear advantage of the present importance measure over the other measures to which it was compared. Therefore, a hypothetical non-linear analytical model with 3 input variables is used in the present measure to examine the strength and the advantage of the present measure as follows:

$$g(X_1, X_2, X_3) = \sin X_1 + a\sin^2 X_2 + bX_3^4 \sin X_1 \quad (16)$$

Table 4. Uncertainty Importance Rankings Obtained by the Present and Two Existing Measures for Example 2

Changed variable	Uncertainty importance measure and their rankings						
	Present Measure		Hora & Iman's Standard deviation measure		Iman & Hora's bivariate measure		
					$R_{0.05}$	$R_{0.95}$	
X_1	0.6676	(3)[a]	10.5989[b]	(1)	0.0608	0.6915	(1)
X_2	0.9809	(1)	9.3750	(2)	1.5614	0.6766	(3)
X_3	0.6740	(2)	6.2500	(3)	0.1777	0.7191	(2)

[a] Ranking.
[b] U_j value defined in reference 13.

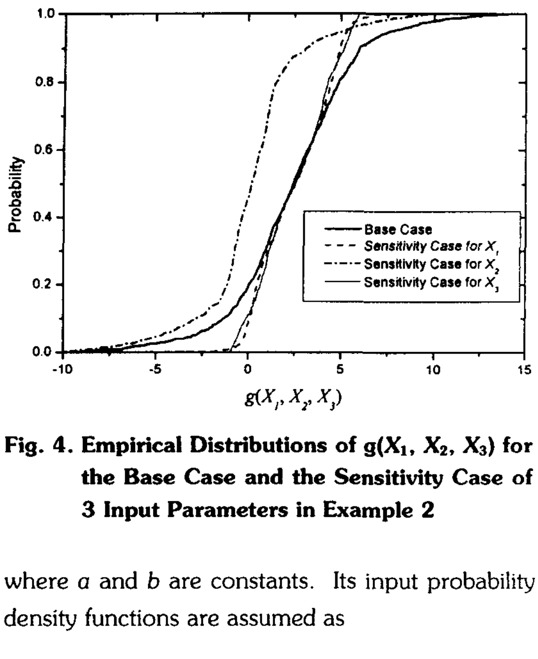


Fig. 4. Empirical Distributions of $g(X_1, X_2, X_3)$ for the Base Case and the Sensitivity Case of 3 Input Parameters in Example 2

where a and b are constants. Its input probability density functions are assumed as

$$h_i(x_i) = \begin{cases} \frac{1}{2\pi}, \text{ when } -\pi \le x_i \le \pi \\ 0, \text{ when } x_i < -\pi, \ x_i > \pi \end{cases} \quad \text{for } i = 1,2,3 \tag{17}$$

The same function was used to investigate the performance of importance measure and the effect of the sampling strategy [12-14]. An uncertainty importance analysis has been carried out for the above hypothetical analytical model. In the sensitivity case of this example, the uncertainty range of input distribution is reduced by a factor of 10. The empirical distributions of the output $g(X_1, X_2, X_3)$ are obtained using the Monte Carlo simulations with 1000 sample size for the base case and its sensitivity cases. The results are shown in Fig. 4. The constants appearing in Eq. (16) are a=5 and b=0.1.

From Fig. 4, it can be observed clearly that the change of X_2's distribution gives the greatest impact on the change of the output distribution. Therefore, this example can be used to check the usefulness of the present measure. The present and two existing measures adopted in example 1 are used in the calculation of example 2 and the results are summarized in Table 4. Since Iman's standard deviation measure is only valid for Boolean representation of a fault tree, more general Hora and Iman's standard deviation measure is considered for the example 2. In Table 4, Homma and Satelli's U_j value defined in reference 13 is used to represent Hora and Iman's standard deviation measure. The analytical expression of U_j for the analytical model considered in the example 2 can be seen in Eqs. (19) ~ (21) of reference 13. The value of U_j has the same information with the Hora and Iman's standard deviation measure about the rank.

Table 4 shows that only the present measure adequately predicts the highest important ranker X_2 in terms of the relative impact on the change of the output distribution. Hora and Iman's standard deviation and Iman and Hora's bivariate measures predict X_1 as the top ranker. In addition, the second rank of the standard deviation measure disagrees with that of the bivariate measure. This problem seems to be due to the fact that the bivariate measure do not depends on the overall output distribution and the standard deviation measure focuses only on the variance reduction. This example clearly shows that variance based importance measures may poorly predict the entire change of output distribution induced from the change of input distribution. Park and Ahn's entropy measure may be useful for this kind of example since their measure also focuses on the relative impact on the change of the output distribution. However, their entropy measure has some disadvantages: For example, the entropy of normal distribution depends only on its variance. In addition, fitting procedures are required for real applications. Data fitting may be another source of uncertainty.

5. Summary and Conclusions

A simple measure of the uncertainty importance focusing on the entire change of CDFs has been developed in the present work. The entire change of CDFs is quantified using the definition of the normalized metric distance between two CDFs. The metric distance measure developed in this study reflects the relative impact of distributional changes of inputs on the change of an output distribution, while most of the existing measures such as the standard deviation measure reflect the relative contribution of input uncertainties to an output uncertainty. To examine the characteristics of the metric distance measure, analytical evaluations of four different distributions have been performed.

In addition, two examples have been chosen to assess its applicability and strength. In the first example, a typical application of the present uncertainty importance measure for PSA has been made. To show the strength of the present measure, a hypothetical non-linear analytical model has been used in the second example. Two typical existing measures, the standard deviation and bivariate measures, are used for the comparison of the results. In general, as shown in example 1, the present as well as existing measures may provide good information about the uncertainty importance. However, the results of uncertainty importance may be different from one another depending on the specific problem as can be seen in example 2. Therefore, using an appropriate measure is very important in uncertainty importance analysis.

The standard deviation and bivariate measures have the limitation of using simple statistics of output distributions rather than the entire output distribution. However, the present measure uses the entire output distributions in the uncertainty importance analysis. This will be a great advantage when the given problem is such that more precise output distribution is needed. The existing measures as well as the present measure possess unique strengths and weakness depending on the objective of their applications. If one wants to know the information on the reduction of uncertainty, the existing measures using the variance (or standard deviation) is useful. However, when the precise output distribution is needed, the normalized metric distance measure developed in the present work will be more useful to express the uncertainty importance of PSA results since it expresses the relative impact of distributional changes of inputs on the change of an output distribution. In addition, it is conceptually simple and easy to calculate.

Nomenclature

a,b	constants used in Eq. (16)
D	metric distance
$E(Y^0)$	mean of output distribution for the base case
F	cumulative distribution function of Y
$f_1(x)$, $f_2(x)$	arbitrary functions of x
g	hypothetical non-linear function defined in Eq. (16)
h	input probability density function defined in Eq. (17)
k_p	variable defined in Eq. (7)
M	sample size of Monte Carlo simulation
$MD(i{:}o)$	metric distance measure defined in the present study
N	sample size of Monte Carlo simulation
n	sample index of Monte Carlo simulation
p	probability
S	empirical distribution function of Y
w	number greater than 1
X	random variable of input
x	input
Y	random variable of output
y	output
y_p	the pth quantile of Y
y_p^i	the pth quantile of Y for the sensitivity case
y_p^o	the pth quantile of Y for the base case
Z	standardized random variable of Y
z_p	the pth quantile of Z
α	shape factor of Y
β	scale factor of Y
Γ	gamma function
δ	delta function defined in Eq. (13)
μ	mean value of Y
σ	standard deviation of Y

Superscripts/Subscripts

i	sensitivity case
o	base case

References

1. K. Nakashima and K. Yamato, "Variance-importance of system components," *IEEE Transactions on Reliability*, **R-31**(1), (1982).
2. V. M. Bier, "A measure of uncertainty importance for components in fault trees," *Transactions of the American Nuclear Society*, **45(1)**, 384 (1983).
3. S. C. Hora and R. L. Iman, "A Comparison of Maxiums/Bounding and Bayes/Monte Carlo for Fault Tree Uncertainty Analysis", SAND85-2839, Sandia National Laboratories (1986).
4. R. L. Iman, "A matrix-based approach to uncertainty and sensitivity analysis for fault trees," *Risk Analysis*, **7**(1), 21 (1987).
5. J. C. Helton, R. L. Iman, J. D. Johnson, and C. D. Leigh, "Uncertainty and sensitivity analysis of a model for multicomponent aerosol dynamics," *Nuclear Technology*, **73**, 320 (1986).
6. R. S. Andsten and J. K. Vaurio, "Sensitivity, uncertainty, and importance analysis of a risk assessment," *Nuclear Technology*, **98**, 160 (1992).
7. R. L. Iman and S. C. Hora, "A robust measure of uncertainty importance for use in fault tree system analysis," *Risk Analysis*, **10**(3), 401 (1990).
8. M. Khatib-Rahbar, E. Cassoli, M. Lee, H. Nourbakhsh, R. Davis, and E. Schmidt, "A probabilistic approach to quantifying

uncertainties in the progression of severe accidents," Nuclear Science and Engineering, **102**, 219 (1989).

9. C. K. Park and K. I. Ahn, "A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment," *Reliability Engineering and System Safety*, **46**, 253 (1994).
10. G. J. Klir and T. A. Folger, Fuzzy Sets, Uncertainty and Information, Prentice Hall, New Jersey, (1988).
11. S. J. Han, Uncertainty analysis methods for quantification of source terms using a large computer code, Ph. D. dissertation, KAIST, (1996).
12. T. Ishigami and T. Homma, "An importance quantification technique in uncertainty analysis for computer models," Proceedings of the ISUMA '90, First International Symposium on Uncertainty Modeling and Analysis, University of Maryland, USA, 3-5, pp.398-403 (December 1990).
13. T. Homma and A. Saltelli, "Use of Sobol' s Quasirandom Sequence Generator for Integration of Modified Uncertainty Importance Measure," *Journal of Nuclear Science and Technology*, **32**, 1164 (1995).
14. T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of nonlinear models," *Reliability Engineering and System Safety*, **52**, 1 (1996).