韓國 컴퓨터情報學會 論文誌
第8卷 第4號, 2003. 12.

2003 8 4-3-6

측정결과의 불확도산정을 위한 모델링과 불확도 전파에 관한 연구

김 종 상*, 조 남 호**

A Study on the Modeling and Propagation to Evaluate Uncertainties in Measurement Results

Jong-Sang Kim *, Nam-Ho Cho **

요 약

국제적으로 측정결과의 신뢰성을 판단할 수 있는 척도로서 불확도(Uncertainty)개념이 도입되고 국제표준화기구(ISO)가 여러 국제기구와 합동으로 "측정불확도표현지침(GUM)"을 1993년에 발간하게 되었다. 본 논문에서는 시료의 산포가 존재하는 경우 시료산포를 불확도 인자로 적용하여 측정결과에 대한 불확도를 평가할 수 있는 측정모델을 구축하여 제시하고, GUM에서 제시한 불확도 전파법칙의 문제점을 분석하여 이를 보완할 수 있는 새로운 불확도의 평가방법으로 몬테칼로 시뮬레이션을 이용한 컴퓨터프로그램 활용의 필요성을 논하고자 한다. 또한 이러한 이론적 근거를 바탕으로 하여 불확도를 평가할 수 있는 컴퓨터 프로그램 개발사례를 제시하고자 한다.

Abstract

The concept of measurement uncertainty has been recognised for many years since "Guide to the Expression of Uncertainty in Measurement" was published 1993 by ISO. This study firstly propose the mathematical model to evaluate uncertainty considering the dispersion of samples because the mathematical model of a measurement is an important to evaluate uncertainty, and it must contains every quantity which contribute significantly to uncertainty in the measurement result. Secondly the standard uncertainty of the result of a measurement, namely combined standard uncertainty is evaluated using the law of propagation of uncertainty, what is termed in GUM method. In GUM method, a measurand is usually approximated by a linear function of its variables by the transforming its input quantities. Furthermore central limit theorem is applied to the input quantity. However the mathematical model of a measurement is generally not always a linearity function, and a distribution function of input or output quantity is not necessarily normal distribution. Then, in some cases GUM method is not favorable to evaluate a measurement uncertainty. Therefore this study propose a new method and its algorithm which use the Monte-carlo simulation to evaluate a measurement uncertainty in both case of linearity or non-linearity function.

▸ Keyword : 불확도(Uncertainty), 측정(Measurement), Monte-carlo simulation.

* 한국건자재시험연구원 ** 건국대학교 산업공학과 명예교수

I. 서 론

미지의 대상물에 대한 특성을 정량적으로 이해하고자 하는 인간의 노력이 끊임없이 계속되어 옴에 따라서 각종 사물에 대하여 시험, 실험 등의 측정을 통한 평가가 이루어지고 있다. 계측(Metrology)분야에서는 이러한 측정결과의 신뢰성에 관한 정량적인 방법의 표현으로서 오차(Error)나 오차분산을 오랫동안 사용해 왔으나, 참값을 영원히 알 수 없기 때문에 명시된 측정결과의 신뢰성에는 의심이 남아있을 수밖에 없다. 이에 따라 국제적으로는 측정결과의 신뢰성을 판단할 수 있는 척도로서 불확도(Uncertainty) 개념 [11]이 도입되고, 국제표준화기구(ISO)가 여러 국제기구와 합동으로 "측정불확도 표현지침(이하 'GUM'이라 한다)을 1993년에 발간하게 되었으며, 1994년 우리나라에서도 국가공인시험·검사기관 인정제도가 도입되고, WTO가입 등 국제화시대에 발맞추어 가면서 최근 공인 교정/시험기관을 중심으로 1998년부터 점차 활용되고 있다[1].

현재 국제적으로 각 측정분야별로 널리 보급되고 있는 GUM에서는 불확도 평가를 위한 측정량의 함수표현, 즉, 측정모델의 설정방법을 구체적으로 명시하고 있지 않다. 특히 표준물질(Reference Material)이 없는 경우 시료가 여러 개 또는 여러 로트일 때 각 시료 측정치의 평균으로 결정되는 측정결과치의 불확도는 각 시료간 산포가 측정결과의 불확도 성분에 기여하는 정도는 매우 크다고 할 수 있으나 측정모델을 잘못 설정함으로써 중요한 불확도 인자가 포함되지 못하게 된다. 따라서 측정결과를 보고할 때 시료의 산포가 존재하는 경우 이 시료의 산포에 의한 영향을 배제한 채 불확도를 평가하여 표현하는 것은 측정결과를 이용하여 의사결정을 해야 하는 고객에게 올바른 정보를 제공할 수 없게 되며, 경우에 따라서는 잘못된 의사결정을 유발시키는 치명적인 오류를 초래할 수도 있게 된다.

또한, GUM에서 제시된 불확도 전파법칙에 의한 합성표준불확도의 계산방법은 입력량의 분포함수가 서로 다른 경우에도 중심극한정리를 적용하여 각 입력량의 표준불확도를 숫자적으로 합성함으로써 합성표준불확도를 계산하게 되는데, 서로 다른 분포함수를 갖는 입력량을 단순히 숫자적으로만 합성시켜 출력량의 분포함수를 결정하는 것은 우선, 합성하는 계산방법상에 문제가 있으며, 출력량의 분포가 항상 대칭형의 t-분포 또는 정규분포가 될 수는 없는 것이다.

이와 같이 GUM이 다국적으로 준용되는 유력한 지침이라고 하더라도 모든 측정분야에 적용될 수 있는 기준으로 제작되었기 때문에 전기전자분야의 전자기적합성(EMC : Elec-tromagnetic Compatibility of Appliances)시험 및 미생물 분야에서는 별도의 보완이 요구되고 있다.

따라서 본 연구에서는 시료의 산포가 존재하는 경우 시료산포를 불확도 인자로 적용하여 측정결과에 대한 불확도를 평가할 수 있는 측정모델을 구축하여 제시하고, GUM에서 제시한 불확도 전파법칙의 문제점을 분석하여 이를 보완할 수 있는 새로운 불확도의 평가방법으로 몬테칼로 시뮬레이션을 이용한 컴퓨터프로그램 활용의 필요성을 논하고자 한다. 또한, 이렇게 개발된 측정불확도 평가알고리즘을 응용하여 몬테칼로 시뮬레이션에 의한 컴퓨터프로그램을 개발하고, 그 실행결과를 GUM에 의한 방법의 측정불확도 평가결과와 비교, 분석한다[1].

II. 측정불확도의 기본개념

1. 측정과 오차[5][12]

측정(Measurement)의 목적은 측정량의 값, 즉 측정하고자 하는 특정한 양의 값을 결정하는 것이다. 따라서 측정을 하기 위해서는 먼저 측정량(Measurand)[11], 측정방법, 측정절차를 적절히 정의하고, 명시하여야 한다. 일반적으로 측정결과는 측정량의 값에 대한 근사치 또는 추정치일 뿐이므로 그 값에 대한 불확도가 함께 명시될 때 비로소 완전하다.

한편, 오차(Error)는 측정결과에서 측정량의 참값을 뺀 값으로 정의하는 데, 일반적으로 측정에는 여러 가지 불완전한 요소가 있으므로 측정결과에는 오차가 있게 마련이다. 오차는 우연오차(Random Error)와 계통오차(Systematic Error)로 구분하여 왔다. 우연오차는 영향량이 시간적, 공간적으로 예측할 수 없게 변동하므로 생긴다.

이러한 변동의 영향을 우연효과(Random Effects)라 하며, 이는 측정량을 반복 관측할 때 그 값이 변동하는 원인이 된다.

계통오차도 우연오차와 마찬가지로 제거할 수는 없지만 줄일 수는 있다.

2. 불확도와 오차[5][12]

측정결과의 불확도는 측정량의 값을 정확하게 알 수 없다는 사실을 반영하고 있다. 측정결과는 이미 알고 있는 계통효과를 적절하게 보정하여도 역시 추정치에 불과하다. 왜냐하면 계통효과에 대한 완전한 보정이 불가능하고, 또 우연효과가 있기 때문에 측정결과에는 항상 불확도가 존재하기 때문이다. 적절한 보정을 한 후의 측정결과는 우연히 측정량 값에 매우 가까울 수도 있으며, 따라서 오차는 매우 작을 수 있다. 그러나 측정결과의 불확도는 매우 클 수도 있기 때문에 불확도와 오차는 구별하여야 한다.

측정자는 반복 관측으로 얻은 미지의 확률분포 또는 이용 가능한 정보에 근거한 주관적 또는 선험적 분포로부터, 인지된 계통효과에 대한 보정을 포함한, 입력량의 값과 표준불확도(추정표준편차)를 함께 추정한다. 그 다음 입력량의 추정치로부터 측정결과를 구하고, 그 추정치의 표준불확도로부터 합성표준불확도를 구한다.

이러한 모든 과정이 적정하게 이루어지고 주요한 계통효과를 빠뜨리지 않고 고려하였다면 측정결과는 측정량의 값에 대한 신뢰성있는 추정치가 되며, 합성표준불확도는 측정결과의 가능한 오차(Possible Error)에 대한 신뢰성있는 척도가 된다고 할 수 있다.

실제로 측정과정에는 불확도의 요인이 많이 존재하며, 다음과 같은 것이 포함된다.

① 측정량에 대한 불완전한 정의 (Incomplete Definition)

② 측정량의 정의에 대한 불완전한 실현 (Imperfect Realization)

③ 대표성없는 시료채취 (Non-representa-tive Sampling)

④ 측정 환경효과에 대한 지식부족 및 환경 조건에 대한 불완전한 측정

⑤ 아날로그 기기에서 개인적인 판독차이

⑥ 기기의 분해능력(Resolution)과 검출 (Discrimination Threshold)한계

3. GUM에 의한 측정불확도[1][5][12]

GUM에서 제시한 측정불확도 산정 방법은 〈그림 2-1〉과 같이 우선 측정모델을 구축하고, 구축된 측정모델링으로부터 입력량의 표준불확도를 구한 다음, 불확도 전파의 법칙을 이용하여 합성표준불확도를 계산하고, 신뢰도 및 유효자유도에 따른 포함인자를 결정하여 확장불확도를 구하는 4단계로 구분될 수 있다.

3.1 측정모델링

측정모델은 측정량(Measurand)을 출력량(Output Quantity)으로 하고, 측정량에 영향을 미치는 모든 영향량을 입력량으로 하여 형성되는 함수관계의 측정방정식을 말한다. 따라서 측정모델의 설정은 불확도 평가시 가장 먼저 결정되어야 할 중요한 과정으로서 측정의 목적, 측정결과의 용도 및 경제성을 고려하여 영향량이 최대한 포함될 수 있도록 모델링이 이루어져야 한다.

대부분의 경우, 측정량 Y는 직접 측정되지 않고, 측정방정식이라는 함수관계 f를 통하여 N개의 다른 양 X_1, X_2, …, X_N으로부터 결정된다.

$$Y=f(X_1,\ X_2,\ \cdots,\ X_N)$$

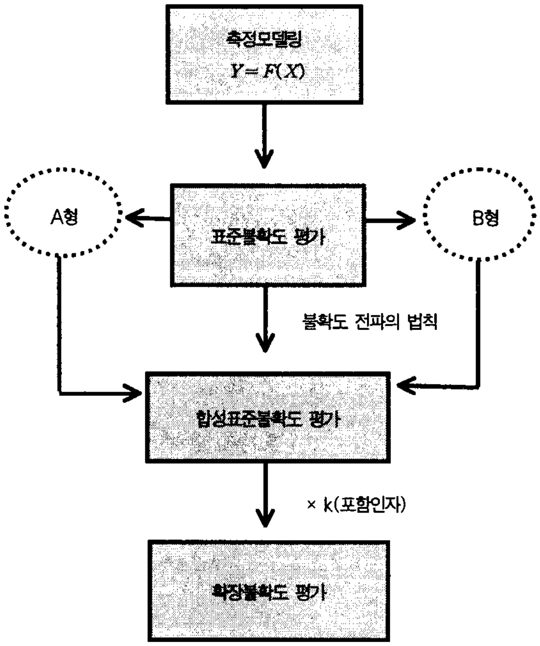

그림 2-1. 측정불확도 산정절차

3.2 표준불확도 평가(A형 및 B형 평가)

각 입력량의 추정치 x_i와 이에 대한 표준불확도 u, 즉 $u(x_i)$는 입력량 X_i가 가질 수 있는 값들의 분포로부터 얻어진다. A형 불확도 평가에서 분산의 추정치 u^2은 반복 측정된 값으로부터 계산되며, 이는 통계학에서 사용되는 시

료분산 s^2에 해당한다. 추정 표준편차 u는 u^2의 양(+)의 제곱근으로서 $u = s$이며, 이를 편의상 'A형 표준불확도(Type A Standard Uncertainty)'라고 하고, B형에 의해 산출되는 분산의 추정치 u^2은 이미 알려진 정보를 이용하여 구하며, 이때 추정 표준편차 u를 'B형 표준불확도(Type B Standard Uncertainty)'라고 한다. 즉, A형 표준불확도는 반복 측정치의 도수분포에 근거한 확률밀도함수에서 구하는 반면, B형 표준불확도는 기존의 정보 또는 문헌을 통해 측정치가 가질 수 있는 확률밀도함수를 가정하여 구한다. 두 가지 방법은 모두 확률에 근거를 둔 것이다.

3.3 합성표준불확도(Combined Standard Uncertainty)

측정결과가 여러 개의 다른 입력량(Input Quantity)으로부터 구할 때 측정결과에 대한 표준불확도를 '합성표준불확도'라 하며, 이를 u_c로 표시한다. 이는 측정결과에 대한 추정 표준편차이고, 합성표준불확도는 여러 입력량의 분산(Variance)과 공분산(Covariance) 성분으로부터 알아내는 합성분산(Combined Variance)의 양의 제곱근으로서 불확도 전파의 법칙(The Law of Propagation of Uncertainty)에 의해 구할 수 있다.

3.4 확장불확도(Expanded Uncertainty)

보건, 안전분야는 물론 산업과 통상분야에서도 그 필요성이 대두됨에 따라 확장불확도가 도입되었으며, 확장불확도 U는 합성표준불확도 u_c에 포함인자(Coverage Factor) k를 곱하여 얻는다. 이러한 확장불확도를 도입하는 목적은 측정량의 합리적인 추정치가 이루는 분포의 대부분을 포함할 것으로 기대되는 측정결과 주위의 구간을 제공하자는 것이다. 포함인자 k의 값은 그 구간에 대해 요구되는 포함확률 또는 신뢰수준에 따라 정해지는 데, 보통 2와 3 사이의 값을 갖는다[5][12].

III. 시료의 산포를 고려한 측정모델링 고찰

불확도 산정에 있어서 측정모델의 설정은 곧 출력량(측정결과)에 영향을 미치는 인자를 결정하게 되며, 출력량과의 관계를 함수식으로 표현하게 된다. 일단 측정모델링이 이루어지면 그 다음은 모두 통계학과 수학적인 지식을 이용하여 계산(Calculation)하는 과정이라 할 수 있다. 이렇게 측정모델의 설정이 매우 중요함에도 불구하고 GUM에서는 측정모델의 설정방법을 구체적으로 명시하고 있지 않음으로서 많은 시험자는 측정방법이나 규격에서 명시된 모델을 불확도 산정을 위한 측정모델로 그대로 사용함으로써 측정결과의 불확도 성분에 기여하는 모든 양을 포함하지 못하고 있다.

특히 시험분야의 측정에서는 표준물질이 없는 경우 시료가 여러 개 또는 여러 로트일 때 각 시료 측정치의 평균으로 결정되는 측정결과치의 불확도는 각 시료간 산포가 측정결과의 불확도 성분에 기여하는 정도는 매우 크다고 할 수 있으나 측정모델을 잘못 설정함으로써 중요한 불확도 인자가 포함되지 못하게 된다.

따라서 측정결과의 불확도를 산정할 때는 시료산포가 측정결과에 미치는 영향의 정도를 항상 검토하여야 하며, 시료의 영향이 존재하는 경우 이 시료산포의 영향을 고려한 적절한 측정모델을 설정하여야 한다.

일반적으로 시험 또는 실험을 하는 경우 측정결과의 표현은 모집단으로부터 시료를 랜덤추출하여 그 시료를 측정방법에 의거 측정을 실시하는 측정분야와 측정결과를 통계적으로 처리하여 대표치(최종측정치)을 설정하고, 그 대표치를 근거로 모집단을 추정하는 통계적 추정부문의 두 가지 측면으로 구분할 수 있다.

지금까지 측정학부문에서의 측정불확도 산정은 측정시스템내에서 즉, 측정시스템부문에서 불확도를 추정하는 데 있다. 따라서 측정과정에서 발생하는 측정치에 영향을 미치는 인자의 파악 및 그 영향정도를 계량화함으로써 불확도를 추정하는 것으로 여기에 표현된 측정결과치 및 불확도는 개개의 시료 그 자체의 특성치와 시험기관의 측정시스템에 기인한 불확실한 정도를 나타내는 것이다.

이에 반하여 2개 이상의 시료측정치를 평균 또는 기타 통계적인 방법을 적용하여 대표치로 표현하여야 하는 경우는 이 대표치를 근거로 하여 모집단의 특성치를 추정하고자 하는 고객의 의지가 내재되어 있으므로 시험자는 시험성적서에 불확도 보고시 고객이 모집단에 대한 특성치를 올바르게 판단할 수 있도록 시료간 산포를 포함하여 불확도를 표현하여야 한다.

1. 시료의 산포를 고려한 측정모델의 제안

1.1 측정데이터의 구조

어떤 모집단으로부터 n개(일반적으로 시험분야에서는 10개 미만임)의 시료를 랜덤하게 추출하여 n개의 시료를 대상으로 정해진 소정의 동일한 시험규격(방법)에 의거 동일한 시험자가 시험을 실시하여 산술평균에 의한 대표치를 최종결과치($\overline{Y}$)로 설정하고, 최종결과치에 대한 불확도를 표현하고자 하는 경우, n개의 측정치 및 최종결과치의 데이터구조는 (표 3-1)과 같이 표현할 수 있다[1][3][4][6][9].

(표 3-1) 측정데이터의 구조

	측정부문 (각 시료별 측정결과치)			
시 료	No. 1	No. 2	•••	No. n
	Y_1	Y_2	•••	Y_n
측정치 Y_i를 결정하기 위한 재현 (l회 반복측정)	y_{11}	y_{21}	•••	y_{n1}
	y_{12}	y_{22}	•••	y_{n2}
	⋮	⋮		⋮
	y_{1l}	y_{2l}	•••	x_{nl}
측 정 치	$y_{1.}$	$y_{2.}$	•••	$y_{n.}$

단, $\overline{Y} = F(Y_1, Y_2, Y_3, \cdots, Y_n) = \dfrac{\sum_{i=1}^{n} Y_i}{n}$

$Y_i = G(X_{i,1}, X_{i,2}, X_{i,3}, \cdots, X_{i,m})$

$i = 1, \cdots, n$

1.2 측정량에 대한 정의

측정자(시험자)는 측정된 시료가 모집단을 대표하도록 추출된 시료라는 것을 전제로 시험시료 n개를 측정하여 평균치 $\bar{y}$ 을 추정하고, 이에 따른 불확도 $\pm U(k=k', 95.45\%)$를 추정하여 결과보고서(시험성적서)에 표현하고자 하는 경우이다. 즉, 측정자가 모집단으로부터 다시 랜덤하게 추출하여 시료 n개를 시험한다면 그 결과치(평균)은 95.45%의 신뢰도로 평균을 중심으로 $\pm U$ $(k=k', 95.45\%)$ 이내에 존재하게 되리라는 것을 고객(시험의뢰자)에게 제시하는 것을 의미하는 것이다. 따라서 측정량이 시료 각각의 특성치를 나타내는 것이 아니라 시료로부터 얻은 특성치의 평균치이며, 측정결과는 산술평균에 의한 대표치를 추정하여 시험결과보고서에 최종결과치를 표현하는 것이 목표인 것이다.

1.3 입력량의 분류

측정량의 정의를 최대한 실현시킬 수 있도록 측정량의 함수표현시 모든 보정치와 보정인자뿐만 아니라 측정결과의 불확도 성분에 기여하는 모든 양을 포함하여야 한다. 따라서 실험적으로 결정되거나 혹은 수치로 계산되는 알고리즘뿐만 아니라 대표치(산술평균) 추정으로 나타낼 수 있는 시료간의 산포를 불확도 기여인자에 추가 설정한다. 통계학적으로 모집단에서 추출된 시료를 조사하여 시료로부터 시료의 평균치를 구하여 대표치를 설정하면 그 평균치를 중심으로 시료값 또는 평균치는 특정한 산포를 나타내게 된다. 따라서 입력량 Y_i 간의 산포에 의한 오차 $\varepsilon = Y - \mu_Y$가 존재하며, 이러한 산포는 측정결과의 불확도 성분에 중요한 기여인자가 된다.

1.4 입력량 Y_i간의 산포에 의한 오차 ε 의 정의

시료간 산포에 의한 오차 ε_i는 시료 Y_i가 전체의 모평균 μ_Y로부터 어느 정도 치우침을 가지고 있는가를 나타내는 수치를 말하며, 다음과 같이 표현할 수 있다.

$$\varepsilon_i = Y_i - \mu_Y$$

전체의 모평균 μ_Y는 일정한 상수이나 각 시료가 변량으로서 랜덤하게 정해지므로 Y_i는 샘플링 때마다 변하는 값이며, 따라서 $\varepsilon_i = Y_i - \mu_Y$도 샘플링 때마다 변하는 변량이다. 그 성질을 다음과 같이 정리하여 볼 수 있다 [1][3][4][8]

$\varepsilon = Y - \mu_Y$ 라 하면,

① ε 는 랜덤으로 변하는 값이며, 기대치를 μ_ε라 하면

$$\mu_\varepsilon = E(\varepsilon) = E(Y - \overline{Y}) = E(Y) - E(\overline{Y}) = 0$$

② ε 의 분산을 σ_ε^2 , Y의 분산을 σ_Y^2이라하면,

$$\begin{aligned}\widehat{\sigma_\varepsilon}^2 = V_\varepsilon &= E[(\varepsilon - \mu_\varepsilon)^2] \\ &= E[\varepsilon^2 - 2\varepsilon\mu_\varepsilon + \mu_\varepsilon^2] \\ &= E(\varepsilon^2) = E[(Y - \mu_Y)^2] \\ &= \widehat{\sigma_Y}^2\end{aligned}$$

이며, 여기서 $\widehat{\sigma_Y}^2$ 는 다음과 같다.

$$\begin{aligned}\widehat{\sigma_Y}^2 = V_Y &= E[(Y - \mu_Y)^2] \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu_y)^2\end{aligned}$$

③ $\bar{\varepsilon}$ 의 분산을 $\sigma_{\bar{\varepsilon}}^2$라 하면 $\sigma_{\bar{\varepsilon}}^2$의 불편추정량은 다음과 같다.

$$\begin{aligned}\widehat{\sigma_{\bar{\varepsilon}}}^2 = V_{\bar{\varepsilon}} &= V[\frac{\sum\varepsilon_i}{n}] = \frac{1}{n^2} V[\sum_{i=1}^{n}\varepsilon_i] \\ &= \frac{1}{n^2}[n V_\varepsilon] = \frac{1}{n} V_\varepsilon \\ &= \frac{1}{n} V_Y\end{aligned}$$

1.5 측정량의 함수표현

입력량은 시료 각각의 측정치 Y_i와 시료간 산포에 의한 오차 $\varepsilon = Y - \mu$ 를 추가한 즉, Y_1 , Y_2 , … Y_n , ε 이 되며 , 최종출력량 $\overline{Y}$는 다음과 같은 함수로 표현할 수 있다.

$$\overline{Y} = f(Y_1, Y_2, \cdots, Y_n, \varepsilon) = \frac{\sum_{i=1}^{n} Y_i}{n} + \varepsilon$$

여기서 $\varepsilon = Y - \mu$로서 $\varepsilon_i = y_i - \mu_Y$

2. 시료의 산포를 고려한 측정모델의 불확도 산정방법 제안

2.1 표준불확도 계산

입력량 Y_i의 표준불확도 $u(Y_i)$를 불확도전파법칙에 의거 추정하면 다음과 같다.

$$Y_i = G(X_{i,1}, X_{i,2}, X_{i,3}, \cdots X_{i,m}),\quad i = 1, \cdots, n$$

서브입력량 $X_{i,1}$, $X_{i,2}$, $X_{i,3}$, … $X_{i,m}$ 이 서로 상관관계가 없다고 가정하면,

$$u^2(Y_i) = \sum_{j=1}^{m}\left[\frac{\partial g}{\partial x_{i,j}}\right]^2 u^2(x_{i,j}) = \sum_{j=1}^{m}[c_{i,j}u(x_{i,j})]^2$$

단, $i = 1, 2, \cdots, n$

이므로 입력량 Y_i의 표준불확도 $u(Y_i)$는 다음과 같이 표현할 수 있다.

$$u(Y_i) = \sqrt{\sum_{j=1}^{m}[c_{i,j}u(x_{i,j})]^2}$$

단, $i = 1, 2, \cdots, n$

한편, 시료간의 산포에 의한 오차 ε 의 불확도 $u(\varepsilon)$는

$$u(\varepsilon) = s_{\overline{Y}} = \frac{s_Y}{\sqrt{n}}$$

이다. 또한, 유효자유도 ν_{y_i}, ν_ε는 각각 Satterthwaite의 유효자유도 공식에 의하여 다음과 같이 구할 수 있다[8].

$$\nu_{Y_i} = \nu_{eff} = \frac{u(y_i)^4}{\sum_{j=1}^{m}\frac{[c_j u(x_j)]^4}{\nu_{x_j}}}$$

단, $i = 1, 2, \cdots, n$ 이고,

ν_{x_j} ; x_j의 자유도

$$\nu_\varepsilon = \phi^* = n - 1$$

2.2 합성표준불확도 산정

측정모델식으로부터 감도계수 c_i, c_ε는

$$c_i = \frac{\partial f}{\partial y_i} = \frac{1}{n},\ i = 1, 2, \cdots, n \text{ 이고,}$$

$c_\varepsilon = \frac{\partial f}{\partial \varepsilon} = 1$ 이므로 $\overline{Y}$ 의 불확도 $u_{\overline{Y}}$는 입력량 Y_1 , Y_2 , Y_3 , … Y_n , ε의 합성표준불확도로서 불확도전파법칙에 의해 구할 수 있으며 그 결과는 다음과 같다[1].

① 측정치 y_i가 서로 상관관계가 없는 경우, 즉 $r(y_i, y_j) = 0$

$$u_{\overline{Y}} = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} u(y_i)^2 + u(\varepsilon)^2}$$

$$\nu_{eff} = \nu_{\overline{Y}} = \frac{u_y^{\ 4}}{\sum_{i=1}^{n}\frac{[c_i u(y_i)]^4}{\nu_{Y_i}} + \frac{[c_\varepsilon u(y_\varepsilon)]^4}{\nu_\varepsilon}} = \frac{u_y^{\ 4}}{\sum_{i=1}^{n}\frac{[(\frac{1}{n})^2 u(y_i)]^4}{\nu_{Y_i}} + \frac{u(y_\varepsilon)^4}{\nu_\varepsilon}}$$

만일 $u_c = u_{Y_1} = u_{Y_2} = u_{Y_3} = \cdots = u_{Y_n}$ 라 하면

$$u_{\overline{Y}} = \sqrt{\frac{\sum_{i=1}^{n} u(Y_i)^2}{n^2} + u(\varepsilon)^2} = \sqrt{\frac{u_c^{\ 2}}{n} + u(\varepsilon)^2}$$

$$\nu_{eff} = \frac{\frac{u_c}{\sqrt{n}}^{\ 4}}{\frac{n[\frac{1}{n^2}\frac{u_c}{\sqrt{n}}]^4}{\nu_{Y_i}} + \frac{u(y_\varepsilon)^4}{\nu_\varepsilon}} = \frac{\frac{u_c}{\sqrt{n}}^{\ 4}}{\frac{\frac{1}{n^7}(\frac{u_c}{\sqrt{n}})^4}{\nu_{Y_i}} + \frac{u(y_\varepsilon)^4}{\nu_\varepsilon}}$$

이다.

② 측정치 y_i가 서로 상관관계가 있는 경우, 즉 $r(y_i, y_j) \neq 0$

$$u_{\overline{Y}} = \sqrt{\frac{1}{n^2}\left[\sum_{i=1}^{n} u(y_i)^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} u(y_i, y_j)\right] + u(\varepsilon)^2}$$

$$= \sqrt{\frac{1}{n^2}\left[\sum_{i=1}^{n} u(y_i)^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} u(y_i)u(y_j)r(y_i, y_j)\right] + u(\varepsilon)^2}$$

$$\nu_{eff} = \nu_{\overline{Y}} = \frac{u_y^4}{\sum_{i=1}^{n}\frac{[c_i u(y_i)]^4}{\nu_{Y_i}} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{[2c_ic_ju(y_i)u(y_j)r(y_i, y_j)]^2}{\nu_{(Y_{i,j})}} + \frac{[c_\varepsilon u(y_\varepsilon)]^4}{\nu_\varepsilon}}$$

단, Y_i, Y_j의 공분산 $u(y_i, y_j)$에 대한 자유도 $\nu_{(Y_{i,j})}$는 $\nu_{(Y_{i,j})} = 2\times\nu_{Y_i}$ 이다.

만일 $r(y_i, y_j) = 1$ 이라 하면,

$$u_{\overline{Y}} = \sqrt{\frac{1}{n^2}\left[\sum_{i=1}^{n} u(y_i)\right]^2 + u(\varepsilon)^2}$$

이고, 또한 $u_c = u_{Y_1} = u_{Y_2} = u_{Y_3} = \cdots = u_{Y_n}$ 이면, 합성표준불확도 u_Y 및 유효자유도 ν_{eff}는 다음과 같다.

$$u_{\overline{Y}} = \sqrt{\frac{1}{n^2}(n\times u_c)^2 + u(\varepsilon)^2} = \sqrt{u_c^2 + u(\varepsilon)^2}$$

$$\nu_{eff} = \frac{u_y^4}{\sum_{i=1}^{n}\frac{[(\frac{1}{n})^2 u(y_i)]^4}{\nu_{Y_i}} + \sum_{i=1}^{n}\frac{[2(\frac{1}{n})^2 u(y_i)^2]^2}{\nu_{(Y_{i,j})}} + \frac{[u(y_\varepsilon)]^4}{\nu_\varepsilon}}$$

2.3 확장불확도 및 불확도 보고

유효자유도 ν_{eff}와 신뢰수준(또는 유의수준)에 따라 t분포에서 t값을 설정하여 다음과 같이 포함계수 k를 결정한다.

$k = t(\nu_{eff}$, 신뢰수준%)

예) $k' = t(\nu_{eff}, 95.45\%)$

따라서 확장불확도 $U = k' \times u_Y$ 이며, 불확도보고는 다음과 같이 표현할 수 있다.

$y \pm U$(단위) ($k = k'$, 95.45% 신뢰도)

3. 시료의 산포를 고려한 측정모델 응용사례 (시료수가 3개인 경우)

예를 들어 어떤 모집단으로부터 3개의 시료 Y_1, Y_2, Y_3 를 랜덤하게 추출하여 동일한 시험방법(반복 없는 1회 측정 실시)에 의거 동일한 시험자가 시험을 실시하여 다음과 같은 결과를 얻었다[1].

Y_1; 측정치 $y_1 = 9.95$, 측정불확도 $u_1 = 0.05$

Y_2; 측정치 $y_2 = 10.00$, 측정불확도 $u_2 = 0.05$

Y_3; 측정치 $y_3 = 10.05$, 측정불확도 $u_3 = 0.05$

최종 측정결과는 3개 시료의 산술평균에 의해 얻은 값으로 결정하는 경우 측정 결과치 및 합성불확도를 구하면 다음과 같다.

3.1 측정모델링

$$Y = \frac{Y_1 + Y_2 + Y_3}{3} + \varepsilon$$

3.2 측정결과 Y= 10.00 ,

3.3 합성표준불확도

$u(Y_1) = u(Y_2) = u(Y_3) = 0.05$ 이고,

$$V_{\overline{Y}} = \frac{V_Y}{n} = \frac{\sum_{i=1}^{3}(y_i - \overline{Y})^2/(3-1)}{3} = 0.0008333$$

$u(\varepsilon) = \sqrt{V_{\overline{Y}}} = \sqrt{0.0008333} = 0.02887$ 이므로

$u(Y)^2 = u(Y_1)^2 + u(\varepsilon)^2$이다.

따라서

$u(Y) = \sqrt{0.05^2 + 0.02887^2} = 0.05774$ 이다.

Ⅳ. 몬테칼로법을 이용한 불확도 산정의 고찰.

1. GUM에 의한 불확도 전파법칙 적용의 한계

측정불확도 산정방법에서 합성표준불확도를 계산하는 방법으로 GUM에서 제시한 불확도 전파법칙은 측정모델식을 1차 Taylor 전개에 의한 근사적인 선형모형으로부터 유도된 것으로 측정모델식(함수)이 비선형성이 클 때에는 Taylor 시리즈의 고차 전개항이 포함되어야 한다. 그러나 고차 전개항을 적용하는 것이 일반적으로 쉽지 않으며, GUM에서도 2차항까지만 제시하고 있고, 합성표준불확도의

계산을 모두 1차 항 전개에 의한 방법으로 설명하고 있다. 따라서 실제로 비선형성이 큰 경우 합성표준불확도의 계산은 GUM에서 제시한 방법을 사용하여서는 구할 수 없다[1].

또한 불확도 전파의 법칙은 입력량의 분포함수가 서로 다른 경우에도 중심극한정리를 적용하여 각 입력량의 표준불확도를 숫자적으로 합성함으로써 합성표준불확도를 계산하게 되며, 출력량의 분포는 모두 t-분포 또는 정규분포로 가정하고, 신뢰구간을 설정하여 확장불확도를 구하게 된다. 그런데 서로 다른 분포함수를 갖는 입력량을 단순히 숫자적으로만 합성시켜 구한 점추정치(추정표준편차)만으로는 출력량의 분포의 모양이나 형태를 알 수 없으며, 분산특성을 모두 파악하는 것은 불가능하다.

그리고 입력량의 분포함수를 정규형의 t분포 또는 정규분포에 대응하여 표준편차를 추정하는 것은 시료크기 n이 충분히 큰 경우에 적용되는 중심극한정리를 이용한 것으로서 입력량이 정규형 분포가 되기 위해서 통계적으로 모집단의 비대칭도(왜도 : Skewness)가 심한 경우는 비대칭도가 작은 경우에 비하여 상대적으로 많은 시료크기가 필요하나, 시험분야에서 시료크기를 많이 할 수 없는 경우가 더 일반적이라 할 수 있다. 또한 출력량의 분포도 정규형의 t분포 또는 정규분포로 가정하고, 확장불확도를 계산하고 있으나 출력량의 분포는 입력량의 분포형태와 측정모델식이 어떻게 이루어졌는가에 따라 달라지므로 가정에 대한 명확함이 더욱 더 저하될 수 있다.

따라서 합성표준불확도를 계산할 때 입력량의 분포형태를 고려하여 합성시켜야 할 필요가 있으며, 측정결과에 대한 신뢰구간을 추정하는 확장불확도 계산의 경우에도 출력량의 분포형태를 반영시킴으로서 측정불확도 산정에 대한 명확함을 더욱 높일 수 있다. 그러므로 본 논문에서는 GUM에서 제시한 불확도 전파법칙의 문제점을 보완할 수 있는 새로운 불확도 계산방법으로 몬테칼로 시뮬레이션(Monte-carlo Simulation)[7]에 의한 계산방법을 연구·제시한다.

2. 몬테칼로법을 이용한 불확도 계산 알고리즘의 제안

2.1 측정불확도 계산을 위한 알고리즘 구축

GUM에서 제시된 측정불확도 산정방법과 본 논문에서 검토된 몇 가지 문제점을 고려하면 측정불확도산정 방법은 <그림 4-1>과 같이 크게 두 가지 부문으로 나눌 수 있다.

측정량의 공식화(Formulation)		계산(Calculating)
1. 측정모델의 설정 2. 입력량의 확률밀도함수 설정	➡	1. 출력량의 확률분포 생성 2. 측정결과(추정치)의 분산(표준편차)에 대한 점추정 3. 측정결과(추정치)에 대한 신뢰구간추정 (신뢰구간 95%, 99.% 등)

그림 4-1

측정량의 공식화(Formulation) 부문과 계산(Calculating) 부문의 두 그룹으로 구분될 수 있으므로 몬테칼로 시뮬레이션에 의한 불확도의 계산 알고리즘은 다음과 같은 5 단계로 구성할 수 있다[1].

단계 1 측정모델링

특정량에 대한 명확한 이해와 요구정확도에 맞는 입력량을 선정하여 입력량과 출력량의 함수관계식을 만든다.

$$Y = F(X_1, X_2, X_3, \ .\ .\ .\ .\ ,\ X_n)$$

- 입력량 ; $X_1, X_2, X_3, \ .\ .\ .\ .\ ,\ X_n$
- 출력량 ; 측정결과 Y

단계 2 입력량의 확률밀도함수를 설정한다.

A형 평가 및 B형 평가방법에 의해 각 입력량의 확률밀도함수를 추정하여 설정.

- 실험계획을 수립하여 직접실험을 실시하거나, 과거 경험 및 외부자료로부터 알고 있는 정보를 이용

단계 3 입력량의 확률분포를 생성(프로그램 작성)

설정된 각 입력량의 확률밀도함수로부터 몬테칼로 시뮬레이션에 의한 확률분포를 생성시킨다.

단계 4 출력량의 확률분포의 생성(프로그램 작성)

몬테칼로 시뮬레이션에 의한 입력량의 확률분포를 측정모델식에 의거 전파시킴(Propagation). 이때 확률분포의 전파를 위한 프로그램은 입력량간의 상관관계를 고려할 수 있도록 작성되어야 한다.

단계 5 출력량의 점추정 및 구간추정(프로그램에 의한 시뮬레이션 결과의 출력)

몬테칼로 시뮬레이션의 결과로부터 평균, 표준편차 및 포함구간(Coverage Interval)의 한계치(Endpoint) 즉, 신뢰구간을 계산함으로써 측정결과, 표준불확도 및 확장불확도를 구한다(신뢰수준, 95%, 99.7% 등)

3. 몬테칼로 시뮬레이션에 의한 확률분포의 생성사례

몬테칼로 시뮬레이션을 이용하여 GUM에서 불확도 산정시 입력량의 분포로 많이 사용하고 있는 직사각형분포(Uniform or Rectangular Distribution), t-분포, 삼각형 분포(Triangular Distribution), 사다리꼴분포

(Trapezoidal Distribution), 정규분포 등의 확률분포를 따르는 랜덤수를 만들고, 확률분포를 생성한 결과는 〈그림 4-2〉와 같다. 프로그램언어는 MATLAB[2]을 사용하였으며, 컴퓨터 시뮬레이션 횟수는 20만회를 수행하였다[1].

3.1 각 확률분포의 특성치 입력데이터

① 대칭형 직사각형분포(Rectangular Distribution)
a: 반나비 1, 평균: 0

② 삼각형 분포(Triangular Distribution)
a: 반나비 1, 평균: 0

③ 대칭형 사다리꼴분포(Trapezoidal Distribution)
a: 윗나비 1 , b: 밑나비 2 , 평균: 0

④ t-분포(t-Distribution)
표준편차: 0.1 , 평균: 0 , 자유도 9

⑤ 정규분포(Normal Distribution)
표준편차: 0.1 , 평균 : 0

3.2 몬테칼로 시뮬레이션에 의한 확률분포 생성결과

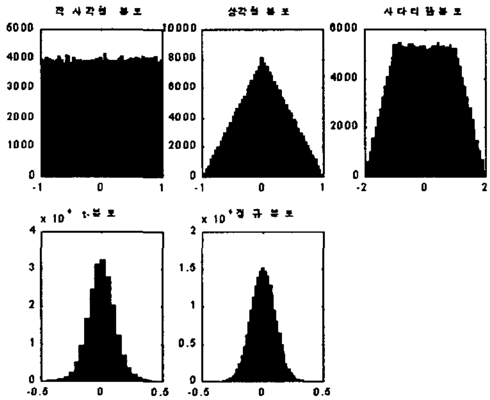

그림 4-2. 몬테칼로 시뮬레이션에 의한 확률분포 생성결과

4. 몬테칼로법을 이용한 불확도 산정의 고찰

몬테칼로법을 이용한 확률분포의 합성은 측정모델식(함수식)에 따라 프로그램이 다양하게 구성될 수 있으며, 본 논문에서는 합성표준불확도를 구하기 위한 방법을 고찰하기 위해 다음과 같은 수학모델(덧셈)의 확률분포 합성사례에 따른 결과만을 제시한다.

$Y = X_1 + X_2$

단, X_1 : 직사각형 분포를 따름.

- 평균100, 반나비 5.
(표준편차 : 2.88765)
변동계수 CV(X_1) ≒ 2.89 %

X_2 : 정규분포를 따름.

- 평균 10, 표준편차(s)

입력량 X_1, X_2의 확률분포를 발생시키고, 확률분포를 합성시키는 프로그램을 작성하였으며, 입력변수 X_2의 표준편차 값을 다음과 같이 6가지 경우로 변화시키면서 몬테칼로 시뮬레이션을 실시(10만회 실행)하고, 출력량의 확률분포의 경향을 분석한 결과는 〈표 4-1〉과 같다.

① X_2의 표준편차가 0.1인 경우
: CV(X_1) 〉 CV(X_2) = 1%

② X_2의 표준편차가 0.29인 경우
: CV(X_1) = CV(X_2) ≒ 2.89%

③ X_2의 표준편차가 0.5인 경우
: CV(X_1) 〈 CV(X_2) = 5%

④ X_2의 표준편차가 1인 경우
: CV(X_1) 〈 CV(X_2) = 10%

⑤ X_2의 표준편차가 2인 경우
: CV(X_1) 〈 CV(X_2) = 20%

⑥ X_2의 표준편차가 2.9 인 경우
: CV(X_1) 〈 CV(X_2) = 29%

상기 (6)의 시뮬레이션 결과는 〈그림 4-3〉과 같다.

시뮬레이션 결과 〈표 4-1〉과 같은 수학모델에서 입력량 X_2의 표준편차(분포의 특성을 결정짓는 인자)가 증가함에 따라 출력량 Y의 평균과 왜도(Skewness)는 각각 110.0, 0.01 미만으로 큰 변화가 없으나 표준편차와 첨도(kurtosis)는 증가하는 경향을 보여주고 있다.

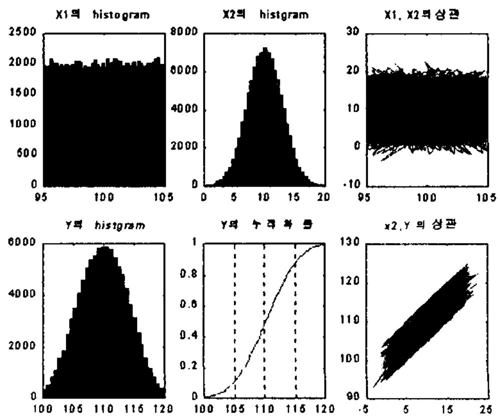

그림 4-3. 몬테칼로 시뮬레이션 결과(6)

즉, 입력량의 표준편차 값이 달라지면 결과치의 표준편차는 물론이고, 분포의 형태가 직사각형 분포(일양분포)에서 정규형 분포로 달라질 수 있음을 알 수 있다.

〈표 4-1〉의 결과는 표준불확도가 서로 같거나 유사함에도 불구하고 GUM방법의 결과에 의한 신뢰구간이 시뮬레이션에 의한 신뢰구간과 모두 차이를 나타내고 있으며, 순서 6의 경우는 두 가지 방법에 의한 표준불확도의 차이가 크게 나타나고 있다. 또한 GUM방법에 의한 결과의 표준불확도는 몬테칼로법에 의한 결과의 표준불확도보다 대부분 작은 값을 나타냄에 따라 같은 신뢰도로서 몬테칼로법에 의한 불확도 표현시보다 오류를 범할 수 있는 확률이 더 커질 수 있으며, 신뢰구간의 표현도 출력량의 분포형태에 따라 왜곡되어 표현될 수 있음을 보여주고 있다[1].

따라서 분포의 형태나 모양에 관계없이 측정결과의 신뢰구간을 추정할 수 있고, 복잡한 계산을 쉽고, 빠르고, 좀더 정확하게 계산할 수 있는 새로운 불확도 산정방법으로 확률적 수학모델에 대한 시뮬레이션기법인 몬테칼로법을 도입하여 적용할 수 있음을 연구, 제시하였다.

표 4-1. 수학모델 $Y = X_1 + X_2$ 의 시뮬레이션 결과

순서	구분	시뮬레이션 결과치 y (단위 생략)					GUM방법의 경우 (평균 ; 110)	
		평균	표준불확도	왜도	첨도	신뢰구간	신뢰구간	표준불확도
1	s = 0.10	109.9960	2.8912	-0.0061	1.8021	114.55~ 105.45	115.78~ 104.22	±2.8894
2	s = 0.29	109.9942	2.9043	-0.0008	1.8218	114.56~ 105.44	115.80~ 104.20	±2.9022
3	s = 0.50	110.0009	2.9344	0.0028	1.8645	114.61~ 105.40	115.90~ 104.10	±2.9519
4	s = 1.00	110.0001	3.0581	0.0073	2.0469	114.92~ 105.11	116.11~ 103.89	±3.0560
5	s = 2.00	110.0088	3.5076	-0.0041	2.4547	115.82~ 104.21	116.22~ 103.78	±3.1091
6	s = 2.90	110.0127	4.0842	-0.0101	2.7058	116.86~ 103.12	116.67~ 103.33	±3.3372

- 사용프로그램 ; MATLAB.
- 시뮬레이션 실행회수 ; 100만회실행
- 입력량의 상관계수 $r(x_1, x_2) = 0$, 신뢰구간 ; 95.45 %

V. 결론

결과적으로 시험 대상모집단의 특성을 파악하기 위해 표본으로 추출된 시료를 측정하여 그 결과를 표현할 때 시험기관에서는 고객이 올바른 판단을 할 수 있도록 시료의 산포를 포함한 불확도를 제시하고, 시료의 산포가 불확도 추정시 포함되지 않으면 불확도 표현지침에 의거 그 사실을 고객이 인지할 수 있도록 측정결과 보고시 표현해야 할 필요가 있다.

이러한 경우 본 연구에서 제시한 불확도추정방법을 적용하면 효과적으로 처리할 수 있을 것으로 사려되며, 복잡하고 비선형성이 큰 경우, 다양한 입력량 및 출력량의 확률분포 또는 출력량의 확률분포가 정규분포라 할 수 없는 경우 등에는 본 연구에서 제시한 몬테칼로법을 이용하면 좀더 정도가 좋고, 효과적으로 측정결과의 불확도를 산정할 수 있을 것으로 사료된다.

또한 향후 시험기관에서 샘플링을 포함한 시험결과의 측정불확도를 추정하고저 할 경우 본 논문에서 제시한 불확도 측정모델링 및 몬테칼로 시뮬레이션에 의한 산정방법을 응용하면 좀더 쉽고 확실하게 불확도를 추정할 수 있을 것으로 사료된다. 다만 컴퓨터에 의한 시뮬레이션은 항상 그렇듯이 그 처리속도가 문제이므로 프로그램언어와 컴퓨터의 하드웨어적 성능을 향상시킴으로써 처리속도를 단축

시키고, 많은 분야에서의 측정모델을 시뮬레이션할 수 있으면서 다양한 통계자료를 출력할 수 있는 시뮬레이션 프로그램의 개발이 보강되면 측정불확도 산정의 품질을 더욱 향상시킬 수 있을 것으로 생각된다.

참고문헌

[1] 김종상, "측정결과의 불확도산정을 위한 모델링과 불확도전파에 관한 연구", 2002.
[2] 김용수, Matlab 입문과 활용, 높이깊이, 2000.
[3] 박성현, 현대실험계획법, 민영사, 1995.
[4] 윤기중, 수리통계학, 박영사, 1983.
[5] 기술표준원, "측정결과의 불확도 산정 및 표현을 위한 지침", KOLAS-R-005, 2000.
[6] Graybill, F.A.: An Introduction to Linear Statistical Models, McGraw Hill Book Co., Inc., New York, 1961.
[7] David C. Joy, Monte Carlo Modeling for Electron Microscopy and Microanalysis, Oxford Univ. 1995.
[8] Satterthwaite, F. E.: "An Approximate Distribution of Estimates of Variance Components, "Biometrics Bulletin, 1946, Vol. 2, pp.110~114.
[9] Searle, S.R.: Linear Models, John Wiley & Sons, Inc., New York, 1971.
[10] ISO/IEC 17025, "General Requirement for the Competence of Testing and Calibration Laboratories" ISO, 1999.
[11] ISO, International Vocabulary of Basic and General Terms in Metrology, 2nd edition, ISO, 1995.
[12] ISO, Guide to the Expression of Uncertainty In Measurement. ISO, Geneva, Switzerland 1993.

저 자 소 개

김 종 상

1987년 건국대학교 산업공학과
1993년 건국대학교 산업대학원 산업공학(공학석사)
2002년 건국대학교 대학원 산업공학(공학박사)
현 재 안양과학대학 테크노경영학부 외래강사, (재)한국건자재시험연구원 품질경영부 선임연구원 재직 중

조 남 호

1962년 한양대학교 공업경영학
1978년 건국대학교 공과대학
2003년 산업공학과 교수역임.
현 재 건국대학교 산업공학과 명예교수, 한국우수기술인증협회 이사