# 베이지안 접근법을 이용한 입력변수 및 근사모델 불확실성 하에서의 신뢰성 분석

# Reliability Analysis under Input Variable and Metamodel Uncertainty using Bayesian Approach

안 다 운* • 원 준 호** • 최 주 호***
An, Dawn • Won, Junho • Choi, Jooho

요 약

신뢰성 분석은 불확실성으로 인한 제품의 성능 변동을 안전확률이나 파괴확률로 정량화 하여 설계에 이용하기 위해 연구되어 왔다. 불확실성은, 데이터의 양에 따라-물질의 본질적인 특성으로서의 많은 데이터가 주어진 경우의 물리적 불확실성과 부족한 데이터에서의 인식론적 불확실성으로 구분되고, 불확실성을 갖는 대상에 따라-입력변수 및 근사모델 불확실성으로 구분된다. 물리적 불확실성에 대한 연구는 많이 진행되어 왔지만, 실제 산업현장에는 부족한 데이터로 인한 인식론적 불확실성이 지배적이며 이에 대한 연구는 최근에서야 진행되고 있다. 불확실성을 고려하는 신뢰성 기반 설계에는 효율성을 위해 실제모델을 대체하는 근사모델이 이용되는데, 근사모델법 자체에 대한 연구는 많이 진행되어 왔으나, 근사모델 이기 때문에 존재하는 불확실성을 고려한 연구는 최근에서야 연구되기 시작하였다. 본 연구에서는 베이지안 접근법에 기반하여 입력변수 및 근사모델 불확실성을 통합 고려하는 새로운 신뢰성 분석 기법을 제시하고 수치예제를 통해 타당성을 증명한 후, 이를 공학문제에 적용한다.

*keywords* : 신뢰성 분석, 베이지안 접근법, 입력변수 불확실성, 근사모델 불확실성

## 1. 서 론

산업현장에서는 제품의 경쟁력을 높이기 위해, 생산비용과 불량률을 감소시키는 일에 관심을 집중하고 있다. 이를 위해 통계적 품질관리나 QC 분임조 활동의 정착을 통해 품질관리를 실시하고 있으며, 최근에는 사후관리 뿐만 아니라 초기 설계에서부터 제품의 경쟁력을 높이기 위한 노력을 기울이고 있다. 이러한 노력의 일환으로 과거 경험에 의한 '적절한' 안전계수 기반 확정론적 설계에서 탈피하여, 1970년대 초부터 확률적인 기법을 이용하여 설계변수의 불확실성을 고려하고 생산된 제품의 품질을 실패 확률로 정량화 하는 신뢰성 설계에 대한 연구가 활발히 진행되어 왔다(Rahman, 2004). 초기의 신뢰성 설계에 대한 연구는 불확실성을 물질의 본질적인 특성으로서의 물리적(aleatory) 불확실성으로 고려하여, 무수히 많은 데이터를 통해 분포를 확실히 알고 있다는 가정 하에서 진행되었다. 하지만, 산업현장에 실질적으로 존재하는 불확실성은, 데이터의

* 학생회원 • 한국항공대학교 항공우주 및 기계공학과 석사과정 skal34@nate.com
** 한국항공대학교 항공우주 및 기계공학과 박사과정 openworldsm@hanmail.net
*** 정회원 • 한국항공대학교 항공우주 및 기계공학부 교수 jhchoi@kau.ac.kr

부족으로 인한 인식론적(epistemic) 불확실성인 경우가 대부분으로, 최근에는 베이지안 접근법을 통해 인식론적 불확실성과 물리적 불확실성을 통합 고려하는 방법이 연구되고 있다(Cruse, 2007). 불확실성을 고려한 효율적이고 현실적인 신뢰성 기반 설계 최적화를 위해, 실제모델을 대체하는 근사모델이 사용되는데 근사모델(metamodel) 불확실성 역시 베이지안 접근법을 이용하여 고려 될 수 있다(O'Hagan, 2006). 본 연구에서는 실제 현장에 존재하는 불확실성을 현실적인 측면에서 고려하기 위해 입력변수 및 근사모델 불확실성을 통합 고려하여 보다 실용적인 신뢰성 기반 설계 최적화에 대한 방향을 제시하고자 한다.

## 2. 베이지안 접근법

베이지안 접근법은 조건부 확률을 이용하는 다음과 같은 베이의 법칙을 기본이론으로 한다.

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{1}$$

$\theta$는 추정하고자 하는 입력변수의 파라미터(정규분포를 예로 들면, $\theta=\mu,\sigma$), $y$는 측정된 데이터, $p(y|\theta)$는 $\theta$ 하에서 측정된 데이터의 우도(likelihood), $p(\theta)$는 $\theta$의 사전분포, $p(\theta|y)$는 측정된 데이터 $y$하에서 업데이트 되는 사후분포이다. 새로운 데이터가 추가되면 사후분포 $p(\theta|y)$는 다시 사전분포 $p(\theta)$가 되며, 이러한 과정이 반복되면서 $\theta$에 대한 신뢰도를 점점 더 향상 시키게 되는데 이를 베이지안 학습효과라 한다.

## 3. 입력변수 및 근사모델 불확실성의 통합

### 3.1. 입력변수 불확실성

물리적 불확실성의 경우에는 확정값으로 주어지는 입력변수의 파라미터 $\theta$를 통해 확률분포를 알 수 있고 이로부터 하나의 반응치 확률값을 구하게 된다. 하지만 산업현장에서 실질적으로 존재하는 인식론적 불확실성의 경우에는, 주어진 데이터 하에서 베이지안 접근법을 이용하여 입력변수의 사후분포를 예측하게 된다. 이를 위한 방법에는, 다음 식(2)를 통해 입력변수의 사후 예측분포를 직접 구하여 반응치의 확률'값'을 구하는 방법(이하 prediction)과 (여기서 $\bar{x}$는 측정 샘플 평균, $n$은 샘플 수, $t(n-1)$은 자유도가 $(n-1)$인 t-분포, $s$는 샘플 표준편차, $\tilde{x}$는 예측된 사후분포) $\theta$를 분포로 예측하고, 이로부터 샘플 $\theta_i(i=1,...,N)\rightarrow N$개의 입력변수의 분포→확률 $p_i(i=1,...,N)\rightarrow p_i$의 확률'분포'를 구하는 방법(이하 double loop)의 두 가지가 있다.

$$\tilde{x} \sim \bar{x}+t(n-1)s\sqrt{1+\frac{1}{n}} \tag{2}$$

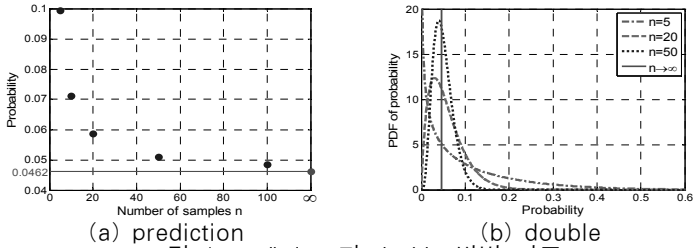
두 방법에 대한 타당성을 논하기 위해 다음과 같은 수치 예제를 고려한다(O'Hagan, 2006).

$$g(X)=X+3\sin(X/2) \tag{3}$$

위 식(3)의 입력변수는 샘플로부터 $X \sim N(3.5,1.2)$인 분포를 따른다. 앞서 언급한 두 가지 방법에 따라 신뢰성 분석을 실시하여 표 1, 그림 1에 그 결과를 나타내었다. prediction에서는 $n$의 증가에 따라 물리적 불확실

표 1 prediction과 double 방법 비교

| aleatory p=0.0462 | | 5% | p/ 95% | prctile/ interval |
|---|---|---|---|---|
| predic-tion | n=5 | | 0.0994 | 65.05 |
| | n=10 | | 0.0712 | 60.89 |
| | n=20 | | 0.0586 | 58.74 |
| | n=50 | | 0.0509 | 56.40 |
| double | n=5 | 0.0024 | 0.3329 | 0.3305 |
| | n=10 | 0.0066 | 0.2108 | 0.2042 |
| | n=20 | 0.0118 | 0.1350 | 0.1232 |
| | n=50 | 0.0206 | 0.0935 | 0.0729 |



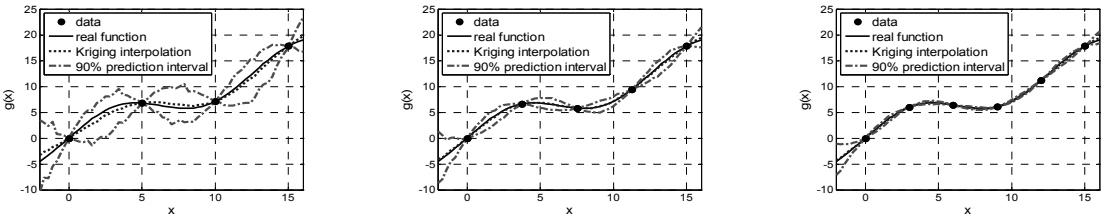
(a) prediction (b) double

그림 1 prediction 과 double 방법 비교

성에서의 확률값 0.0462에 수렴하며, double loop를 통해 구한 확률의 확률분포에서 50% 확률값에 가까워 짐을 알 수 있다. double loop를 통한 결과는 $n$의 증가에 따라 신뢰구간이 줄어들면서 역시 확률값 0.0462에 수렴하며, 95%에서의 확률이 prediction을 통한 확률값 보다 높은 파괴확률로 더 보수적인 경향을 보인다.

### 3.2. 근사모델 불확실성

근사모델에는 반응표면모델, 이동최소자승 및 크리깅 모델 등이 있으며, 본 연구에서는 보다 정확한 근사가 가능한 크리깅 모델을 이용하여 근사모델 불확실성의 특성을 파악하여 신뢰성 설계에 적용코자 한다. 간단한 예제로써, 식(3) $g$의 사후분포를 예측하여, 5%와 95%의 예측구간을 그림 2에 나타내었다. 쉽게 예상할 수 있었던 결과대로, 실험점(data)의 증가에 따라 예측구간은 점차 좁아진다.




(a) data=4 (b) data=5 (a) data=6

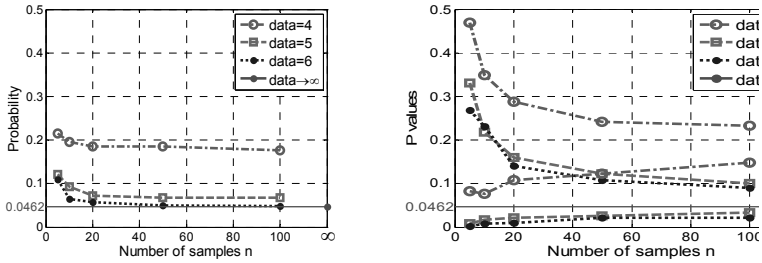
그림 2 크리깅모델과 예측구간

### 3.3. 불확실성의 통합

실제 존재하는 불확실성을 제대로 고려한 신뢰성 설계를 위해서는 입력변수 및 근사모델 불확실성에 대한 통합과정이 필요하다. 두 가지 불확실성의 통합 과정을 통한 신뢰성 분석 결과를 표 2와 그림 3에 나타내었다. $n$의 증가에 따라 신뢰구간이 줄어들고, 데이터수의 증가에 따라서는 $n$의 증가에 따른 수렴의 양상 그대로 물리적 불확실성 확률값 쪽으로 이동되는 것을 알 수 있다. 데이터 수가 적어서 근사모델의 불확실성이 커지면 그 만큼 높은 파괴확률로 불확정성을 보정한 확률값을 내어주는 결과양상을 통해 입력변수 및 근사모델 불확실성의 통합과정에 대한 논리성이 뒷받침 된다.

표 2 입력변수 및 근사모델 통합

| real p=0.0462 | | predic. p | double 5% | 95% | interval |
|---|---|---|---|---|---|
| data =4 | n=5 | 0.2146 | 0.0831 | 0.4700 | 0.3870 |
| | n=10 | 0.1954 | 0.0766 | 0.3488 | 0.2722 |
| | n=20 | 0.1859 | 0.1075 | 0.2882 | 0.1808 |
| | n=50 | 0.1853 | 0.1232 | 0.2412 | 0.1180 |
| data =5 | n=5 | 0.1210 | 0.0077 | 0.3316 | 0.3239 |
| | n=10 | 0.0924 | 0.0169 | 0.2176 | 0.2007 |
| | n=20 | 0.0722 | 0.0211 | 0.1599 | 0.1388 |
| | n=50 | 0.0683 | 0.0263 | 0.1229 | 0.0966 |
| data =6 | n=5 | 0.1094 | 0.0024 | 0.2682 | 0.2658 |
| | n=10 | 0.0648 | 0.0084 | 0.2321 | 0.2237 |
| | n=20 | 0.0574 | 0.0101 | 0.1403 | 0.1302 |
| | n=50 | 0.0502 | 0.0215 | 0.1074 | 0.0859 |




(a) prediction (b) double

그림 3 입력변수 및 근사모델 불확실성 통합 결과

## 4. 설계 문제에의 적용

### 4.1. I-beam 예제

$$g(X_1, X_2) = 1172.12 - \frac{27000 X_2 (X_1 - X_2)}{0.0822 \times 59^3 X_1} \tag{4}$$

각 변수는 독립적인 관계이고 $X_1(l)$은 인식론적 불확실성을 갖고 10개의 샘플로부터의 평균이 $3100mm$,

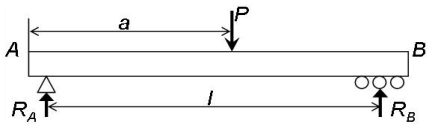


그림 4 단순 지지 I-beam

표 3 I-beam 신뢰성 분석 결과

| data=10, sample n=10 | Input uncertainty | | | Input+Modeling uncertainty | | |
|---|---|---|---|---|---|---|
| | p/5% | 95% | interval | p/5% | 95% | interval |
| aleatory | 0.9504 | – | – | 0.9069 | – | – |
| prediction | 0.9300 | – | – | 0.8628 | – | – |
| double | 0.8390 | 0.9880 | 0.1490 | 0.7890 | 0.9596 | 0.1707 |

표준편차 $25mm$인 정규분포를 따르고, $X_2(a)$는 물리적 불확실성으로 $N(1830,25)\,m$인 분포이다. 10개의 데이터로 구축한 근사모델을 통한 신뢰성 분석결과가 표 3과 같으며 물리적에서 인식론적으로, 입력변수에서 입력변수+근사모델 불확실성으로 갈수록 높아지는 불확실성으로 인해 신뢰도가 떨어짐을 알 수 있다.

4.2. Fortini's clutch 예제

$$g(X_1,X_2,X_3,X_4)=\arccos\left(\frac{X_1+1/2(X_2+X_3)}{X_4-1/2(X_2+X_3)}\right) \tag{5}$$

접촉각도 $g$는 서로 독립적인 설계변수 $X_1 \sim X_4$에 의해 위 식(5)와 같이 표현되며 설계사양으로 $g$는 5°~9°에 놓여야 한다. 각 설계변수의 분포는 표 4와 같으며 표 5에 신뢰성 분석 결과가 나타나 있다. 30개의 데이터를 통한 근사모델에서 입력변수의 불확실성을 함께 고려한 신뢰성 분석 결과는 네 개의 설계 설계변수를 포함하기 때문에 앞서 실시한, 변수가 한두 개일 경우보다 신뢰도가 많이 떨어지는 것이 확인된다.

표 4 설계변수의 분포

| | $X_1 \sim Bet$ | | $X_2 \sim N$ | $X_3 \sim N$ | $X_4 \sim Ray$ |
|---|---|---|---|---|---|
| n | $\bar{x}$ | 55.29 | $\mu$ 22.86 | 22.86 | 101.6 |
| =10 | $s$ | 0.0793 | $\sigma$ 0.0043 | 0.0043 | 0.0793 |

표 5 clutch 신뢰성 분석 결과

| data=30, sample n=10 | Input uncertainty | | | Input+Modeling uncertainty | | |
|---|---|---|---|---|---|---|
| | p/5% | 95% | interval | p/5% | 95% | interval |
| aleatory | 0.9981 | – | – | 0.4914 | – | – |
| prediction | 0.9880 | – | – | 0.4903 | – | – |
| double | 0.9720 | 1.0000 | 0.0280 | 0.4856 | 0.4938 | 0.0082 |

## 5. 결론

본 연구에서는 입력변수 및 근사모델의 불확실성을 통합 고려하는 신뢰성 분석 방법을 연구·제시하였고, 두 불확실성을 통합 고려할 경우에 대해 그 결과양상을 설명하였다. 입력변수에서의 샘플수의 증가에 따라 반응치 확률의 신뢰구간이 감소되면서 물리적 불확실성에서의 확률값에 가까워지고, 근사모델에서의 데이터수의 증가에 따른 반응치 확률의 변화는 샘플수에 따른 수렴 양상 그대로 물리적 불확실성에서의 확률값 쪽으로 이동됨을 확인하였다. 이를 토대로 공학예제에 적용하여 각각의 불확실성하에서의 결과양상을 비교하였으며, 산업현장에 실제 존재하는 불확실성을 제대로 고려하기 위한 새로운 신뢰성 설계 방법을 제시하였다.

## 감사의 글

이 논문은 한국학술진흥재단(과제번호 2008-02-010)의 지원에 의해 수행되었고 이에 감사합니다.

## 참고문헌

Rahman, S. and Xu, H. (2004) A Univariate Dimension-Reduction Method for Multi-Dimensional Integration in Stochastic Mechanics, *Probabilistic Engineering Mechanics*, 19, pp.393~408.

Cruse, T.A. and Brown, J.M. (2007) Confidence interval simulation for systems of random variables, *Journal of Engineering for Gas Turbines and Power ASME*, 129, pp.836~842.

O'Hagan, A. (2006) Bayesian analysis of computer code outputs: A Tutorial, *Reliability Engineering and System Safety*, 91, pp.1290~1300.