Transactions of the Korean Nuclear Society Autumn Meeting
Goyang, Korea, October 24-25, 2019

Development of a Bayesian Analysis Program, MUDAP

*Mina Cho, Byeongmun Ahn, Moosung Jae**
**Hanyang University, 418 HIT, 222 Wangsimni-ro, Seongdong-gu Seoul, 04763*
**Corresponding author: jae@hanyang.ac.kr*

1. Introduction

A risk of nuclear power plant is assessed based on the estimates of failure probabilities, frequencies of initiating event etc. in probabilistic safety analysis (PSA). For more realistic risk assessment, the data that is close to the real and latest data are preferred. In Korea, data from NUREG/CR-6928 or NUREG-1829 etc. were used sometimes in the case that the specific failure rate or initiating event frequency of Korea did not exist. To assess the risk of current domestic nuclear power plant more realistically, the data should be updated by the latest and the actual failure records of Korea.

To obtain new failure probabilities or initiating event frequencies in which latest failure records are reflected, KAERI developed a program (Bayesian Update for reliability Data, BURD) which updates the new failure records and release the data immediately. Failure probabilities and initiating event frequencies could follow various distributions (lognormal, normal, gamma, beta etc.). BURD can use the uniform, Jefferey's non informative, beta, gamma, and lognormal distribution as the prior distribution and for likelihood function, BURD can use beta, gamma, binomial and Poisson distribution. However, failure probabilities or initiating event frequencies can follow other distribution that BURD does not provide.

Therefore, a tool that immediately calculates the new probability which reflect the real and latest failure records and can use various distributions was required for a more realistic risk assessment and as a result, a Bayesian updating program (Multi Unit Data Analysis Program, MUDAP) was developed.

2. Methodologies of the program

MUDAP uses the Bayesian theorem that represent the relation between the prior distribution and the posterior distribution. According to the Bayesian theorem, the posterior probability can be obtained from the prior probability. The Bayesian theorem is as follows.

θ is a continuous parameter with pdf f(θ) and range [a, b], and f(θ) is prior probability density function (pdf). x is random discrete data and together they have likelihood p(x|θ). Then, the posterior pdf can be obtained by equation (1).

$$f(\theta|x) = \frac{p(x|\theta)f(\theta)}{p(x)} = \frac{p(x|\theta)f(\theta)}{\int_a^b p(x|\theta)f(\theta)d\theta} \quad (1)$$

In case of the discrete prior probability distribution, MUDAP calculates posterior probabilities as shown in table I.

Table I: Bayesian updating procedure in MUDAP

θ	Prior	Likelihood	Product	Posterior
θ_1	$f(\theta_1)$	$p(x\|\theta_1)$	$f(\theta_1)p(x\|\theta_1)$	$\frac{f(\theta_1)p(x\|\theta_1)}{\sum_{i=1}^{n} f(\theta_i)p(x\|\theta_i)}$
⋮				
θ_k	$f(\theta_k)$	$p(x\|\theta_k)$	$f(\theta_k)p(x\|\theta_k)$	$\frac{f(\theta_k)p(x\|\theta_k)}{\sum_{i=1}^{n} f(\theta_i)p(x\|\theta_i)}$
⋮				
Total	1			1

In case of the continuous prior pdf, MUDAP selects N samples θ and calculates the posterior probability for the N samples by dealing with the continuous distribution as a discrete distribution. The sampling method is as follows. (N is the number of samples and, it is set to 10,000 by default.)

First, find the cumulative distribution function (cdf) of prior distribution. Next, select the N probability values between 0 and 1 except 0 and 1 evenly. Then, using inverse cdf, obtain the coincident θ values. The sampling procedure is shown in Fig. 1.

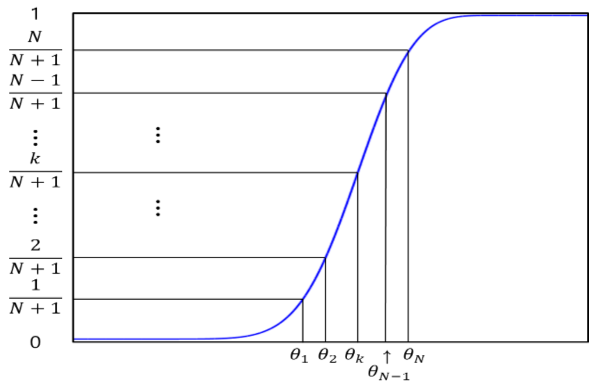

Fig. 1. MUDAP result of the example.

3. Program Description

MUDAP provides distributions in the table II as a prior distribution. For the other distributions, the user can type the prior distribution in form of a table of discrete probability distribution. Also, in MUDAP, user can use distributions in table II as a likelihood.

As a result, MUDAP provides the data table and the graph of prior and posterior probability distribution. Also, MUDAP shows the mean, variance, median etc. of prior and posterior pdf. The result screen is shown is Fig. 2.

Table II: the type of distribution that MUDAP provides as prior and likelihood distribution

Prior	Lognormal, Normal, Beta, Gamma, Exponential, Uniform, Jeffrey's rule (Binomial), Jeffrey's rule (Poisson)
Likelihood	Lognormal, Normal, Poisson, Binomial, Exponential

The result may vary depending on the setting of 'Dist Config' and 'Plot config' parts. In the 'Dist config' part, the number of samples and the sampling range can be set at user's discretion. In the 'Plot config' part, the scale of axis can be set to log scale and the result graph can be provided in form of cdf. Also, at user's discretion the likelihood graph can be provided.

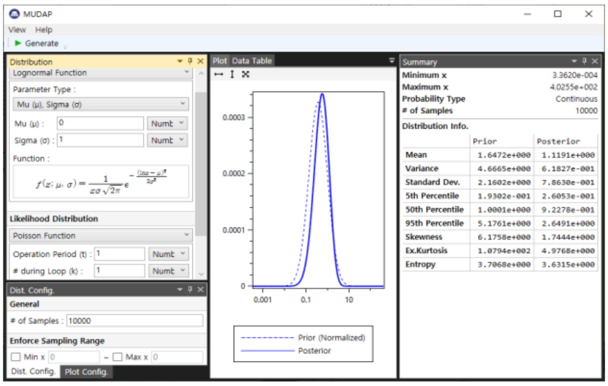

Fig. 2. MUDAP result screen.

4. Verification Example

MUDAP performance was verified by a following example. There is a pump that has the average failure rate, λ. Suppose that from previous experiences with same pump, the expected pump failure rate was estimated to follow exponential distribution, $f(\lambda)=e^{-\lambda}$ with range $\lambda \geq 0$. Occurrence of failure is assumed to be a Poisson process. During the first year after installation of the pump, two accidents occurred. Using these observations, find the posterior probability for λ.

Solution:

The analytic solution of this example is shown in table III and the result comparison between analytic solution and MUDAP solution is shown in table IV. Also, MUDAP result screen for this example is represented in Fig. 3.

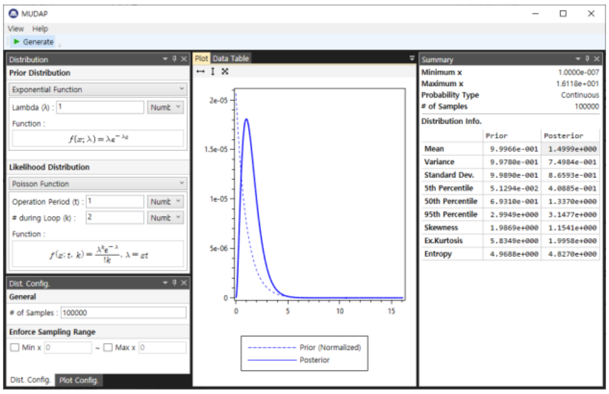

Fig. 3. MUDAP result of the example.

Table III: The analytic solution of the example

Prior	Likelihood	Product	Posterior
$f(\lambda)$	$p(k=2\|\lambda)$	$f(\lambda)p(k=2\|\lambda)$	$f(\lambda\|k=2)$
$e^{-\lambda}$	$\frac{e^{-\lambda\times 1}(\lambda\times 1)^2}{2!}$	$e^{-\lambda}\frac{e^{-\lambda}(\lambda)^2}{2!}$	$\frac{e^{-\lambda}\frac{e^{-\lambda}(\lambda)^2}{2!}}{\int_0^\infty e^{-\lambda}\frac{e^{-\lambda}(\lambda)^2}{2!}d\lambda}$

Table IV: The result comparison between analytic solution and MUDAP result

	Analytic solution	MUDAP (N=10,000)		MUDAP (N=100,000)	
		Result	Error [%]	Result	Error [%]
Mean	1.5	1.4997	0.0200	1.4999	0.006%
Variance	0.75	0.7493	0.0880	0.7498	0.0213
Median	1.337	1.3369	0.0075	1.337	0.0000

The errors are less than 0.1% with 10,000 and 100,000 samples and it shows that MUDAP calculation result is almost accurate.

5. Conclusions

For a realistic domestic risk assessment, using probability or frequency values reflecting up-to-date and Korean actual failure records is necessary.

The program (MUDAP) which calculates posterior probability from prior probability based on the Bayesian theorem was developed for calculation of the new probability which is close to the real. The performance of MUDAP was verified by an example above and the example result showed that MUDAP calculated posterior distribution with acceptable error which is less than 0.1%. Also, MUDAP can use more various distributions than BURD as shown in Table II.

However, using MUDAP only one probability or frequency of an event can be updated at a time, which may cause inconvenience when updating the probability or frequency of multiple events is necessary. Also for a certain distribution, optimization is not perfectly done, which can result in slightly lower performance (it takes a long time). MUDAP does not use a direct method, but an iterative method (as numerical method), because the combination of prior and likelihood distribution is too various to solve it with direct method. As a result, MUDAP does not present the formula of posterior distribution (cdf and pdf) and does not calculate the error, which make determining the accuracy of the posterior distribution data difficult.

MUDAP could produce immediately new probability or frequency reflecting the latest and actual failure records of Korea with acceptable error with various distributions. By using MUDAP and producing the posterior distribution, more realistic probability values could be used in the risk analysis and as a result, it is possible to assess the nuclear power plant risk more realistically.

Acknowledgements

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KOFONS), granted financial resource from the Multi-Unit Risk Research Group(MURRG), Republic of Korea (No. 1705001).

REFERENCES

[1] M. Modarres, What every engineer should know about reliability and risk analysis, CRC Press, 1992

[2] Stan Brown, Measures of Shape: Skewness and Kurtosis, Oak Road Systems, 2008-2016

[3] E.T. Jaynes, Gibbs vs Boltzmann Entropies, American Journal of Physics, 1965

[4] US NRC, Industry-Average Performance for Components and initiating Events at U.S. Commercial nuclear Power plants, NUREG/CR-6928, Initiating Data Sheets, update 2010.

[5] US NRC, Estimating Loss-of-Coolant Accident (LOCA) Frequencies Through the Elicitation process, NUREG-1829, Vol 1, 2008.