Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 17-18, 2018

An Analytical SOKC Solver for Log-normal Distribution

Jongsoo Choi
Korea Institute of Nuclear Safety, 62 Gwahak-ro, Yuseong-gu, Daejeon 34142
Corresponding author: *k209cjs@kins.re.kr*

1. Introduction

The state of knowledge correlation (SOKC) arises because, for identical or similar components, the state-of-knowledge about their failure parameters is the same. In other words, the data used to obtain mean values and uncertainties of the parameters in the basic event models of these components may come from a common source and, therefore, are not independent, but are fully correlated.

When the basic event mean values and uncertainty distributions are propagated in the PSA model without accounting for the SOKC, the calculated mean value of the relevant risk metric and the uncertainty about this mean value will be underestimated due to the effect of the SOKC directly. In order to account SOKC, we generally perform Monte Carlo uncertainty analysis. However, the results of Monte Carlo uncertainty analysis are not fixed and there could be a possibility of bias in the results.

This paper proposes an analytical SOKC solver to provide exact mean risk metrics accounting for the SOKC.

2. State-of-knowledge Correlation (SOKC)

The ASME/ANS standard on PRA [1] requires that both parameter and model uncertainties be addressed. For example, parameter uncertainties are addressed via the quantification process of the core damage and large early release frequencies and model uncertainties have to be identified and characterized. The ASME/ANS standard provides the two supporting requirements (QU-A3 and QU-E3) that specifically address the treatment of the state-of-knowledge correlation. The ASME/ANS standard notes the following:

- For Capability Category II and III, the mean and the distribution for the risk metric estimates are usually obtained by propagating the parameter uncertainties of the PRA inputs through the analysis using the Monte Carlo or similar sampling method.
- The difference between Capability Category II and Capability Category III is that in Capability Category II the propagation of the uncertainty is only carried out for significant contributors in the significant accident sequences and cutsets.

Table I. ASME/ANS PRA Standard Supporting Requirements Related to SOKC

	Capability Category II	Capability Category III
QU-A3	ESTIMATE the mean CDF accounting for the “state-of-knowledge” correlation between event probabilities when significant.	CALCULATE the mean CDF from internal events by propagating the uncertainty distributions, ensuring that the "state-of-knowledge" correlation between event probabilities is taken into account.
QU-E3	ESTIMATE the uncertainty interval of the CDF results. ESTIMATE the uncertainty intervals associated with parameter uncertainties (DA-D3, HR-D6, HR-G8, IEC15) taking into account the “state-of-knowledge” correlation.	PROPAGATE parameter uncertainties (DAD3, HR-D6, HR-G8, IE-C15), and those model uncertainties explicitly characterized by a probability distribution using the Monte Carlo approach or other comparable means. PROPAGATE uncertainties in such a way that the “state-of-knowledge” correlation between event probabilities is taken into account.

NUREG-1855 [2] provides on how to address the treatment of parameter uncertainty when using PSA results for risk-informed decision-making. NUREG-1855 addresses the characterization of parameter uncertainty; propagation of uncertainty; assessment of the significance of the state of-knowledge correlation; and comparison of results with acceptance criteria or guidelines. NUREG-1855 notes the following:

- In carrying out the propagation, it is important to consider the state of knowledge correlation (SOKC) between events. The SOKC arises because, for identical or similar components, the state-of-knowledge about their failure parameters is the same. In other words, the data used to obtain mean values and uncertainties of the parameters in the basic event models of these components may come from a common source and, therefore, are not independent, but are correlated.
- When the basic event mean values and uncertainty distributions are propagated in the PSA model without accounting for the SOKC, the calculated mean value of the relevant risk metric and the uncertainty about this mean value will be underestimated. The values can be underestimated due to the effect of the SOKC directly, as well as due to incorrect screening out of cutsets in truncation due to neglect of the SOKC in calculating cutset frequencies.

To account for this correlation when propagating the basic event values and their uncertainty in a Monte Carlo (or similar) sampling trial, at each pass through the process, the distribution based on the pooled data should be sampled once to obtain a failure rate, and that same failure rate should be used to generate the sample value for all the correlated basic events in the cut set equations. In general then, to account for the SOKC, the same information is used to generate the estimates of the parameters used to evaluate the probabilities of a group of basic events whose parameter values were obtained from correlated data. This means that when using a Monte Carlo (or similar) approach to propagate uncertainty, for each pass through the process the same sample value drawn from the probability distribution of the parameter should be used to calculate the basic event probability of all basic events within the group.

Let X be a random variable corresponding to a basic event that is correlated with other basic events in a minimal cut set (MCS). To illustrate SOKC, consider the simple case where two MOVs are in parallel, represented by variables X_1 and X_2 that are correlated, and system failure occurs when both fail to open. The equation when the failure probabilities of the two MOVs are identical (i.e., the distributions of the failure probabilities express the same state of knowledge) is

$$T = X^2 \tag{1}$$

where T represents system failure.

If X_1 and X_2 are considered to be independent, the equation used for system failure would be

$$T = X_1 \, X_2 \tag{2}$$

The underestimation of the mean of an MCS that contains correlated basic events is particularly significant when the expected value $E(X^n) >> E^n(X)$.

3. SOKC Effects of Log-normal Distribution

Log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then Y = ln(X) has a normal distribution. Likewise, if Y has a normal distribution, then X = exp(Y) has a log-normal distribution. A random variable which is log-normally distributed takes only positive real values.

Given a log-normally distributed random variable X and two parameters μ and σ that are, respectively, the mean and standard deviation of the variable's natural logarithm, then the logarithm of X is normally distributed. The given log-normal distribution can be notated by "Lognormal(μ, σ^2)."

Wikipedia, the free encyclopedia, tells us the following characteristics of Log-normal distribution:

1) If X ~ Lognormal(μ, σ^2) is distributed log-normally, then

$$X^a \sim \text{Lognormal}(a\mu,\ a^2\sigma^2). \tag{3}$$

2) If X_j ~ Lognormal(μ_j, σ_j^2) are n independent log-normally distributed variables, and $Y = \prod_{j=1}^{n} X_j$, then Y is also distributed log-normally:

$$Y \sim \text{Lognormal}\ \left(\textstyle\sum_{j=1}^{n} \mu_j,\ \sum_{j=1}^{n} \sigma_j^2\right). \tag{4}$$

3) Let X_j ~ Lognormal(μ_j, σ_j^2) be independent log-normally distributed variables with possibly varying μ and σ parameters, and $Y = \sum_{j=1}^{n} X_j$. The distribution of Y has no closed-form expression, but can be reasonably approximated by another log-normal distribution Z at the right tail. Its probability density function at the neighborhood of 0 has been characterized and it does not resemble any log-normal distribution. A commonly used approximation due to L.F. Fenton (but previously stated by R.I. Wilkinson and mathematical justified by Marlow) is obtained by matching the mean and variance of another lognormal distribution:

$$\sigma_Z^2 = ln\left[\frac{\sum e^{2\mu_j+\sigma_j^2}(e^{\sigma_j^2}-1)}{(\sum e^{\mu_j+\sigma_j^2/2})^2} + 1\right],$$
$$\mu_Z = ln\left[\sum e^{\mu_j+\sigma_j^2/2}\right] - \frac{\sigma_j^2}{2}. \tag{5}$$

4. Analytical SOKC Solver for PSA

In probabilistic safety assessment (PSA) for a nuclear power plant, the risk measures (e.g., core damage frequency and large early release frequency) are computed from cut set equations as

$$R = \textstyle\sum_{i=1}^{all} E(MCS_i) \tag{6}$$

Each MCS contains independent basic events X's and correlated sets T's as

$$MCS = T_1 \ \ T_2 \ \ldots\ X_1\, X_2 \ \ldots \tag{7}$$

where X's and T's are also independent of each other. If all basic event is distributed log-normally, then mean and distribution of the expected value of a MCS can be exactly calculated by Eq.(3) and Eq.(4).

The mean of a risk metric accounting for the SOKC can also be exactly calculated by Eq.(6), but the distribution of the risk metric accounting for the SOKC could not be analytically calculated. However, the distribution can be approximated by Eq.(5).

The distribution of the risk metric accounting for the SOKC can be calculated by Monte Carlo uncertainty analysis approach [3, 4]. The newly proposed approach to account for the SOKC is a combined one which calculates mean values by the proposed analytical solver and error factors (as distribution parameters) by Monte Carlo simulation.

5. Application to Example PSA model

In order to assess the adequacy of the proposed approach to NPP PSA models, an example PSA model is selected as follows:

- Level 1 internal event PSA model of a plant
- CDF : 1.093E-6/years (by rare event approximation)
- # of MCSs : 24,083
  - SOKC sets : 3,641

− Non-correlated sets : 20,442

Table II shows the Monte Carlo simulation results of Example PSA model.

Table II: Monte Carlo simulation results (from 100 runs of sample size 1E5)

	E(CDF)	EF of CDF
Mean (A)	1.103E-6	4.496
standard deviation (B)	1.335E-8	2.044E-2
(B) / (A)	1.210 %	0.455 %
minimum	1.071E-6	4.434
maximum	1.142E-6	4.544

Using the proposed analytical SOKC solver based on Eq. 3, 4, and 6, the CDF for Example PSA model accounting SOKC is calculated as

- 1.10262285E-6/years (by the analytical solver).

This value is exact for Example problem.

Using Eq. 5, we can approximate the mean and distribution of Example problem as follows:

- μ = -15.52384, σ^2 = 3.612046
- E(CDF) = 1.1026224E-6/years
- EF = 22.7851

Comparing with Table II, it is shown that the approximation equation (Eq. 5) provides a very close expected value, but overestimates EF.

5. Conclusions

An analytical SOKC solver is developed in this study. This SOKC solver provides exact risk metrics accounting for the SOKC without Monte Carlo uncertainty analysis. The distribution of risk metrics accounting for the SOKC has no exact solver, but it can be reasonably obtained by Monte Carlo technique.

REFERENCES

[1] Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications, ASME/ANS RA-Sa-2009, ASME/ANS, 2009.

[2] Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision making, NUREG-1855 Rev.1, USNRC, 2013.

[3] J. Choi, Improved Monte Carlo Method for PSA Uncertainty Analysis, KNS Spring Meeting, Gyeongju, Oct. 27-28, 2016.

[4] J. Choi, How to Treat SOKC Effects in PSA Quantification, KNS Autumn Meeting, Gyeongju, Oct. 26-27, 2016.