# A Study on CCF Parameters Using Joint Probability Distribution

Kibeom Son [a], Gyunyoung Heo [a*]
[a]*Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, Korea*
[*]*Corresponding author: gheo@khu.ac.kr*

## 1. Introduction

When performing Probabilistic Safety Assessment (PSA) of Nuclear Power Plants (NPPs), Common Cause Failure (CCF) is analysed on the basis of field test and data, and the CCF contains information about the correlation between each component. CCF modeling is necessary for cases where available data is insufficient. In this situation, simultaneous failure of the component can be explained by using the joint probability distribution [1]. The method of modeling the simultaneous failure due to a seismic event using the joint probability distribution utilizes a Beta Factor Model (BFM) [2]. However, if the above method is applied to the CCF modeling, there arises a problem that the joint probability can't be completely decomposed into basic parameter. Therefore, in this paper, we propose a new approximation method for applying the joint probability distribution to the CCF modeling. This method has the advantage of clearly separating the basic parameter from joint probability and improving the accuracy of the approximation. These points are presented in the results section as an example of two components.

## 2. Methods

In order to distinguish between the previously proposed method and the method presented in this paper, we denote 'Method 1' and 'Method 2', respectively.

### *2.1 Method 1*

This method contains the procedure to obtain the CCF basic parameters from the joint probability distributions that explain the seismic correlations, in order to reflect the simultaneous failures by the seismic events to the fault trees [2]. In this section, we describe the process of decomposing the joint probability and expressing it as a basic parameter.

If there are two components A and B, the joint probability can be expressed by a Venn diagram as shown in Figure 1.

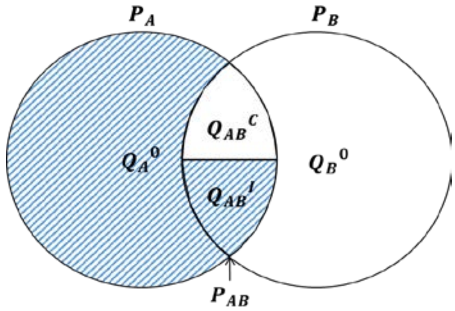


Fig. 1. The Venn diagram of the joint probability

In this case, $P_A$ is a probability that a failure occurs in the component A (irrespective of whether the component B is fail), $P_B$ is a probability that a failure occurs in the component B (irrespective of whether the component A is fail), $P_{AB}$ is the probability that simultaneous failures occur in the components A and B, $Q_A{}^0$ is the probability that a failure occurs in the component A while component B is in a normal state, $Q_B{}^0$ is the probability that a failure occurs in the component B while component A is in a normal state, $Q_{AB}{}^I$ is the probability that simultaneous failures occur in the components A and B due to independent causes, $Q_{AB}{}^C$ means the probability of simultaneous failure in the components A and B due to a common cause.

Figure 1 is expressed as the following Eq. (1), (2), and (3).

$$P_A = Q_A + Q_{AB} \quad (1)$$

$$P_B = Q_B + Q_{AB} \quad (2)$$

$$P_{AB} = Q_A Q_B + Q_{AB} \quad (3)$$

According to the definition of each term, $Q_A$, the shaded part of Figure 1, is the sum of the probabilities of component A failure and the simultaneous failure of components A and B due to the independent cause, and it is $Q_A = Q_A{}^0 + Q_{AB}{}^I$. In addition, we can recognize that $Q_{AB}$ is the probability of simultaneous failure in the components A and B due to a common cause, that is, $Q_{AB} = Q_{AB}{}^C$.

Therefore, we can express Eq. (1), (2) and (3) as Eq. (4), (5), and (6)

$$P_A = (Q_A) + (Q_{AB}) = \left(Q_A{}^0 + Q_{AB}{}^I\right) + \left(Q_{AB}{}^C\right) \quad (4)$$

$$P_B = (Q_B) + (Q_{AB}) = \left(Q_B{}^0 + Q_{AB}{}^I\right) + \left(Q_{AB}{}^C\right) \quad (5)$$

$$\begin{aligned} P_{AB} &= (Q_A Q_B) + (Q_{AB}) \\ &= \left\{\left(Q_A{}^0 + Q_{AB}{}^I\right)\left(Q_B{}^0 + Q_{AB}{}^I\right)\right\} + \left(Q_{AB}{}^C\right) \end{aligned} \quad (6)$$

The left hands of Eq. (1), (2), and (3) are constants calculated through the integration of joint probability. The $Q_A$, $Q_B$, and $Q_{AB}$ are obtained from the solutions of nonlinear simultaneous equations with three unknowns and three equations.

In the Eq. (4), (5), and (6), the nonlinear simultaneous equations with four unknowns and three equations are established, and ${Q_A}^0 + {Q_{AB}}^I$, ${Q_B}^0 + {Q_{AB}}^I$, ${Q_{AB}}^C$ can be obtained as a solution. In other words, to separate the ${Q_A}^0 + {Q_{AB}}^I$, it is necessary to have additional equations.

### *2.2 Method 2*

In order to apply the joint probability distribution to the CCF, it is necessary to separate the cause of the simultaneous failure of the components A and B into whether the cause is an independent cause or a common cause. In the case of $\rho = 0$, there exists only the independent cause. In the case of $\rho = 1$, only the common cause exists. However, since the exact ratio of these two causes can't be estimated in the range of $0 < \rho < 1$, we assume that the terms expressing the simultaneous failure of components A and B are variables that vary according to the correlation coefficient rather than the fixed value. Therefore, when a simultaneous failure occurs in the components A and B, the variables representing the ratio of the causes are approximation value.

In the case of Method 2, we added a formula expressing the simultaneous failure of the component due to the independent cause in view of the above situation. This equation utilizes the correlation coefficient, which is the input value of the integral calculation, as a kind of weight, and it can express a situation where the simultaneous failure rate by the independent cause decreases as the correlation coefficient increases.

In Figure 1, $Q_A$ is decomposed into ${Q_A}^0 + {Q_{AB}}^I$, and ${Q_{AB}}^I$ is expressed as $P_A P_B(1-\rho)$, and these mean Eq. (7), (8), (9), (10).

$$P_A = {Q_A}^0 + {Q_{AB}}^I + {Q_{AB}}^C \quad (7)$$

$$P_B = {Q_B}^0 + {Q_{AB}}^I + {Q_{AB}}^C \quad (8)$$

$$P_{AB} = {Q_{AB}}^I + {Q_{AB}}^C \quad (9)$$

$$P_A P_B = \frac{{Q_{AB}}^I}{1-\rho} \quad \because {Q_{AB}}^I = P_A P_B(1-\rho) \quad (10)$$

The $\rho$ is a correlation coefficient, and Eq. (1) corresponds to Eq. (7), and Eq. (2) corresponds to Eq. (8).

There are four unknowns, four equations, and four solutions are drawn denoted ${Q_A}^0$, ${Q_B}^0$, ${Q_{AB}}^I$, ${Q_{AB}}^C$, since $Q_A$ and $Q_B$ of the Method 1 can be decomposed to ${Q_A}^0 + {Q_{AB}}^I$ and ${Q_B}^0 + {Q_{AB}}^I$. Eq. (10) is an approximate expression of the situation where there is only an independent cause when $\rho = 0$ and only common cause when $\rho = 1$.

### *2.3 Discussions on Equation (10)*

In Section 2.3, we explain in more detail the reason why we define it as ${Q_{AB}}^I = P_A P_B(1-\rho)$.

i) According to the definition of correlation coefficient, if $\rho = 0$, there is no ${Q_{AB}}^C$, and if $\rho = 1$, ${Q_A}^0$ and ${Q_{AB}}^I$ do not exist. This is shown in Figure 2.

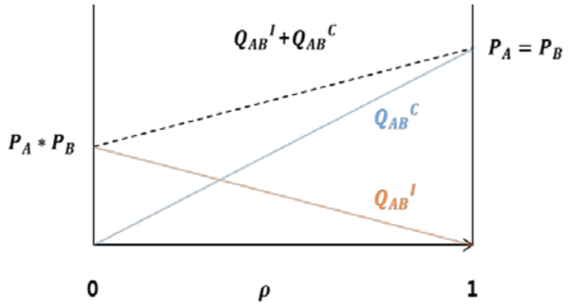


Fig. 2. The hypothetic ratio of ${Q_{AB}}^I/{Q_{AB}}^C$

Assuming that $P_A = P_B$, the components A and B are independent events, $P_{AB} = P_A * P_B = {Q_{AB}}^I$ if $\rho = 0$ according to Eq. (9). In addition, $P_A = {Q_{AB}}^C$ if $\rho = 1$ according to Eq. (7). In the boundary of such a correlation coefficient, independent cause and common cause can be calculated. However, it is not possible to know what nonlinearity ratio of ${Q_{AB}}^I/{Q_{AB}}^C$ has in the range between $\rho = 0$ and $\rho = 1$.

ii) If we assume that ${Q_{AB}}^I = P_A P_B$ and there is no weight, then irrespective of the correlation coefficient, ${Q_{AB}}^I$ will be a constant and only ${Q_{AB}}^C$ will change according to the integral of Eq. (9). This phenomenon is in violation of the definition for the correlation coefficient, so that in the range of $0 < \rho < 1$, the ratio of ${Q_{AB}}^I/{Q_{AB}}^C$ is unknown without an exact solution, so proper approximation is necessary. Thus, by the condition that the $\rho$ increases, ${Q_{AB}}^I$ decrease, and $\rho_{max} = 1$, we assume that the first-order linear relationship, such as Eq. (10).

On the contrary, we can approximate the case of the $\rho = 1$ such as ${Q_{AB}}^C = \rho P_A = \rho\left(\frac{P_A+P_B}{2}\right)$, where the ${Q_{AB}}^I/{Q_{AB}}^C$ ratio of the two ways is different. Therefore, when both of ${Q_{AB}}^I = P_A P_B(1-\rho)$ and ${Q_{AB}}^C = \rho\left(\frac{P_A+P_B}{2}\right)$ are applied, we can't obtain the solution. If a weight is applied to ${Q_{AB}}^C$, the value of ${Q_{AB}}^C$ is overestimated in the range of high correlation coefficient. So, we choose the assumption ${Q_{AB}}^I = P_A P_B(1-\rho)$.

## 3. Results

In this paper, the failure is judged based on the normal distribution where the mean is 0 and the standard deviation is 1. In addition, Monte Carlo integration and nonlinear simultaneous equations were solved using MASS and rootSolve packages of R programming code.

$$f_{A,B}(a,b) = \frac{1}{2\pi\sigma_A\sigma_B\sqrt{1-\rho_{AB}{}^2}}\exp\left[\frac{-1}{2(1-\rho_{AB}{}^2)}\left\{\left(\frac{A-\mu_A}{\sigma_A}\right)^2 - 2\rho_{AB}\left(\frac{A-\mu_A}{\sigma_A}\right)\left(\frac{B-\mu_B}{\sigma_B}\right) + \left(\frac{B-\mu_B}{\sigma_B}\right)^2\right\}\right] \quad (11)$$

The $\rho_{AB}$ means Pearson's product moment correlation between A and B. The joint probability density function of the bivariate normal distribution can be expressed by a general formula such as Eq. (8). According to the definitions of $P_A$, $P_B$, $P_{AB}$, the integration interval of the joint probability applied in this paper is as shown in Table I, and the corresponding interval means a component failure.

Table I. The integration interval of joint probability

| Interval | Component A | | Component B | |
|---|---|---|---|---|
| | Min. | Max. | Min. | Max. |
| $P_A$ | $-\infty$ | 0 | $-\infty$ | $+\infty$ |
| $P_B$ | $-\infty$ | $+\infty$ | $-\infty$ | 0 |
| $P_{AB}$ | $-\infty$ | 0 | $-\infty$ | 0 |

*3.1 The results of Method 2*

Figure 3 is a graph of the Method 2, which calculates the values of the $Q_A{}^0$, $Q_{AB}{}^I$, $Q_{AB}{}^C$ according to the correlation coefficient. In this paper, we can easily calculate as (12), (13), and (14) for the two symmetrical components.

$$Q_{AB}{}^I = P_A{}^2(1-\rho) \quad (12)$$
$$Q_{AB}{}^C = P_{AB} - P_A{}^2(1-\rho) \quad (13)$$
$$Q_A{}^0 = P_A - P_{AB} \quad (14)$$

If $\rho = 0$ and 1, it can be regarded as an exact solution, but it can not be obtained at $0 < \rho < 1$. Since weights are given by using correlation coefficient, $Q_{AB}{}^I$ and $Q_{AB}{}^C$ show a linear trend.

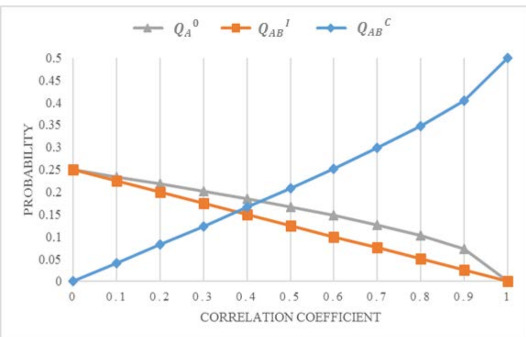


Fig. 3. Calculated basic parameter by Method 2

*3.2 The comparison of Method 1 and Method 2*

In this paper, we compared the approximate results of Methods 1 and 2 through Eq. (15) expressing a mutually exclusive Venn diagrams. The left hand of Eq. (15) is calculated differently according to the joint probability decomposition way of Method 1 and Method 2, and the right hand is a value calculated through integration.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (15)$$

In the above equation, $Q_A = Q_B$ is defined by symmetry between the components, and $P(A) = P_A$, $P(A \cap B) = P_{AB}$ are applied.

In the case of Method 1, the approximate value of the left hand is obtained by Eq. (16), because it is impossible to distinguish whether the cause of the simultaneous failure is due to the independent cause or common cause.

$$P^{M1}(A \cup B) = Q_A + Q_B + Q_{AB} \quad (16)$$

In the case of Method 2, the cause of the simultaneous failure is clearly distinguished. So, Eq. (17) is used.

$$P^{M2}(A \cup B) = Q_A{}^0 + Q_B{}^0 + Q_{AB}{}^I + Q_{AB}{}^C \quad (17)$$

Here, $P^{M1}$ means Method 1 and $P^{M2}$ means Method 2.

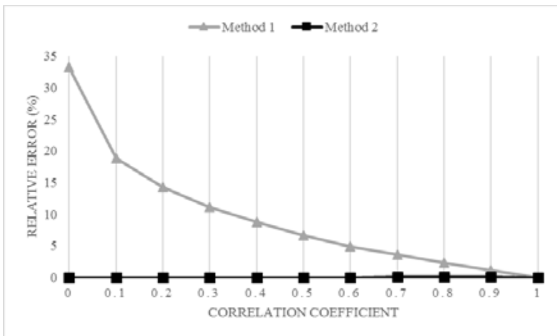


Fig. 4. The relative error of each method

Figure 4 is a graph showing the relative error by comparing the values of Eq. (16) and (17) on the left hand of Eq. (15) with the integral value of the right hand. In Method 1, the difference between the left hand and the right hand is large. This is because the left hand value is overestimated due to the limitation that $Q_A$ can not be divided into $Q_A{}^0 + Q_{AB}{}^I$. In other words, the left hand of Method 1 is the result of the duplicate calculation of the probability of $Q_{AB}{}^I$ as in Eq. (18).

$$\begin{aligned} P^{M1}(A \cup B) &= (Q_A) + (Q_B) + (Q_{AB}) \\ &= \left(Q_A{}^0 + Q_{AB}{}^I\right) + \left(Q_B{}^0 + Q_{AB}{}^I\right) + \left(Q_{AB}{}^C\right) \\ &= Q_A{}^0 + Q_B{}^0 + 2 * Q_{AB}{}^I + Q_{AB}{}^C \end{aligned} \quad (18)$$

The models described in equations (1), (2), and (3) are referred to in the report on CCF modeling guidelines as 'Approximate Formulate-Basic Parameter Model' [3].

The reason for expressing 'Approximation' is estimated that the simultaneous failure is classified as an independent cause or a common cause under the expert judgment, such as impact vector. The method presented in this paper is an attempt to distinguish between independent causes and common causes using $\rho$ itself, and emphasizes that $\rho$ can explain information about simultaneous failures. However, it is still necessary to study whether the independent cause and the common cause can be separated through some function of $\rho$.

### 4. Conclusion

In this paper, we have proposed a joint probability decomposition method which can be utilized for the CCF modeling by using joint probability distribution, and the points of this method are explained. This method can completely decompose joint probability into basic parameters and can be used for CCF modeling when field data is insufficient. In order to apply this method to three or more components, $\rho_{ABC}$, which can express the correlation between at least three variables in a comprehensive manner, is required. To generalize this approximation method, it is necessary to study how to make the simultaneous equations with high order.

This topic is an ongoing research, and methodology and related source code will be released through future papers.

### Acknowledgments

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KOFONS), using the financial resource granted by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea (No. 1705001).

### REFERENCES

[1] K.H. Jin, G.Y. Heo, Copula Study for Common Cause Failures under Asymmetric Conditions, Transactions of the Korean Nuclear Society Spring Meeting, Jeju, Korea, May 32-24, 2019

[2] M.F. Pellissetti, U. Klapp. Integration of correlation models for seismic failures into fault tree based seismic PSA, Structural Mechanics in Reactor Technology SMiRT 21, New Delhi, India, November 2011

[3] International Atomic Energy Agency, Guidelines on Modeling Common-Cause Failures in Probabilistic Risk Assessment, NUREG/CR-5485, 1998