Article

Mathematical Formulation and Analytic Solutions for Uncertainty Analysis in Probabilistic Safety Assessment of Nuclear Power Plants

Gyun Seob Song and Man Cheol Kim *

Department of Energy Systems Engineering, Chung-Ang University, 84 Heukseok-ro Dongjak-gu, Seoul 06974, Korea; dyrnfmxm678@cau.ac.kr
* Correspondence: charleskim@cau.ac.kr; Tel.: +82-2-820-5907

Abstract: Monte Carlo simulations are widely used for uncertainty analysis in the probabilistic safety assessment of nuclear power plants. Despite many advantages, such as its general applicability, a Monte Carlo simulation has inherent limitations as a simulation-based approach. This study provides a mathematical formulation and analytic solutions for the uncertainty analysis in a probabilistic safety assessment (PSA). Starting from the definitions of variables, mathematical equations are derived for synthesizing probability density functions for logical AND, logical OR, and logical OR with rare event approximation of two independent events. The equations can be applied consecutively when there exist more than two events. For fail-to-run failures, the probability density function for the unavailability has the same probability distribution as the probability density function (PDF) for the failure rate under specified conditions. The effectiveness of the analytic solutions is demonstrated by applying them to an example system. The resultant probability density functions are in good agreement with the Monte Carlo simulation results, which are in fact approximations for those from the analytic solutions, with errors less than 12.6%. Important theoretical aspects are examined with the analytic solutions such as the validity of the use of a right-unbounded distribution to describe the uncertainty in the unavailability/probability. The analytic solutions for uncertainty analysis can serve as a basis for all other methods, providing deeper insights into uncertainty analyses in probabilistic safety assessment.

Keywords: probabilistic safety assessment; fault tree analysis; uncertainty analysis; analytic solutions; Monte Carlo simulation

1. Introduction

As many countries declare their intention to achieve carbon neutrality by 2050, the role of nuclear power is becoming more important. As nuclear safety is one of the key concerns about nuclear power, quantitative assessment of nuclear safety is essential for enhancing the safety of nuclear power plants and hence increasing the sustainability of nuclear energy.

The probabilistic safety assessment (PSA) is an analytic technique used to estimate quantitative risks and to make decisions associated with complex systems. The PSA complements deterministic safety analysis and has been widely applied to the design, operation, maintenance, and regulation of nuclear power plants. One of the major objectives of the PSA is to identify scenarios that lead to consequences of interest and to quantify the occurrence frequency of such scenarios, considering the system features [1]. The identified scenarios and their occurrence frequencies provide important insights for decision-making in the operation and safety management of nuclear power plants.

An uncertainty analysis, which is one of the major elements in PSA, involves quantifying the uncertainties of the occurrence of scenarios. Since directly quantifying the uncertainty of an entire system is practically impossible, the underlying principle of the

uncertainty analysis is "divide and conquer," i.e., decomposing a complex system into manageable parts, making the assessment separately, and then performing appropriate computations [2]. A typical method used to decompose a process is a fault tree analysis. In a fault tree analysis, the failed state of a complex system is represented by a top event and divided into logic gates and basic events. Basic events in a fault tree typically represent component failures and have their associated probabilities. These probabilities are constructed using mathematical models and associated parameters. The uncertainty associated with model parameters is called the state-of-knowledge uncertainty. For the state-of-knowledge uncertainty, subjective probability density functions (PDFs) are used as the interpretation of probability to express the degree of beliefs [3]. The subjective PDF, as a function of information, reduces the uncertainty when new information is available. The formal method to handle this information is the subjective (Bayesian) probability theory [4]. The probability of a basic event also becomes a random variable with a PDF because of the state-of-knowledge uncertainty.

As mentioned above, the uncertainty of the top event is represented by a PDF determined from the component uncertainties, which are also represented by PDFs. Various methods have been developed and applied to determine the PDF of the top event from the PDFs of basic events. In theoretical approaches, such as variable transformation [5], the PDF of the top event is synthesized by integrating with the joint PDFs of basic events [6]. In the method of moments [7], the moments of the top event are generated from the moments of basic events and then applied to an assumed distribution of the top event to approximate the PDF of the top event. In the method of discrete probability distributions [8], the continuous PDFs of basic events are discretized and combined to produce an approximate discrete PDF of the top event. A Monte Carlo simulation [9] can produce an approximate PDF of the top event by repeated sampling from the PDFs of basic events and then calculating the unavailability of the top events.

Jackson et al. [10] highlighted the difficulty in determining the exact solutions for the distribution of the top event, and this is the major motivation behind using approximate methods. Since then, the authors could find few studies that have been conducted on theoretical approaches for uncertainty analyses. Among the approximate methods, such as the method of moments, method of discrete probability distributions, and Monte Carlo simulations, the Monte Carlo simulation is most widely used in practice, owing to its ease of application to large-scale models of complex systems [11].

Although each approximate method has many advantages compared to finding analytic solutions, analytic solutions continue to serve as a basis for all approximate methods. Moreover, analytic solutions provide deeper insights into an uncertainty analysis, thus providing better explanations of the numerical results obtained from approximate methods or Monte Carlo simulations.

Starting from conventional probability theories, this paper presents a theoretical approach for uncertainty analyses in the context of PSA. Section 2 presents the mathematical formulation and analytic solutions for the uncertainty analysis in PSA. Section 3 presents the application of the analytic solutions to a simple example system and a comparison with the Monte Carlo simulation results. Finally, Section 4 presents the conclusions of this study.

2. Analytic Solutions for Uncertainty Analysis

Mathematical formulation specific to the uncertainty analysis in PSA is developed on the ground of the general probability theory. We first need to define Boolean variables in event trees and fault trees. Random variables will be assigned to the probabilities of the Boolean variables. After that, Boolean functions that are generally used in event trees and fault trees will be identified, and then the numerical functions associated with the Boolean functions will be derived and solved in their analytic forms.

2.1. Mathematical Formulation and Analytic Solutions

Let $B_1, \ldots, B_n$ be the Boolean variables for the basic events in a coherent fault tree and $X_1, \ldots, X_n$ be the random variables for the probabilities that $B_1, \ldots, B_n$ are in the failed state, i.e., true state in a fault tree. Let G_{TOP} be the top event of the fault tree and X_{TOP} be the random variable for the probability that G_{TOP} is in the failed state, i.e., unavailability. X_{TOP} is given as a Boolean function h_{TOP} of B_i's and X_{TOP} is given as a numerical function g_{TOP} of X_i's, as follows:

$$G_{TOP} = h_{TOP}(B_1, \ldots, B_n) \tag{1}$$

$$X_{TOP} = g_{TOP}(X_1, \ldots, X_n) \tag{2}$$

Because X_i's and X_{TOP} are the random variables for probabilities, the images of the random variables are commonly bounded by (0, 1). The random variables for the probabilities of basic events X_i's are assumed to be independent each other due to the assumption that the basic events of a fault tree are independent of each other. The random variables X_i's have the probability density functions $f_1(x), \ldots, f_n(x)$ which represent the uncertainties. Here, x is a variable whose support is bounded by (0, 1). For any probability density function $f(x)$ for a random variable X, the cumulative distribution function $F(x)$ is given as:

$$F(x) = \Pr[0 \leq X \leq x] = \int_0^x f(x')dx' \tag{3}$$

In uncertainty analysis in PSA, right-unbounded distributions which are defined from 0 to infinity, such as lognormal distribution or gamma distribution, are also widely used to model the uncertainty. The use of such right-unbounded distributions would only be valid when the random variables are distributed in such small values, usually significantly less than 1, that the integration of the probability density function from 0 to 1, i.e., the cumulative distribution function of 1, almost approaches 1, as follows:

$$\Pr[0 \leq X \leq 1] = \int_0^1 f(x)dx = F(1) \approx 1 \tag{4}$$

The purpose of uncertainty analysis is to find the probability density function for the top event, $f_{TOP}(x)$, from the probability density functions of basic events, $f_1(x), \ldots, f_n(x)$. Because those X_i's are independent of each other, the cumulative distribution function for the top event $F_{TOP}(x)$ is given as:

$$F_{TOP}(x) = \int_0^x f_{TOP}(x')dx' = \int_{D_{TOP}} f_1(x_1) \cdots f_n(x_n)dx_1 \cdots dx_n \tag{5}$$

where D_{TOP} is the region of the combination of variables $x_1, \ldots, x_n$ that $g_{TOP}(x_1, \ldots, x_n)$ is smaller than x.

For coherent fault trees, any Boolean function can be expressed with the combination of AND and OR logics. The application of AND and OR logics to multiple (more than two) basic events can be performed by consecutively applying the logics to two events, as follows:

$$B_1 \cap B_2 \cap B_3 \cap \ldots \cap B_n = (B_1 \cap B_2) \cap B_3 \cap \ldots \cap B_n = [(B \cap B_2) \cap B_3] \cap \ldots \cap B_n \tag{6}$$

$$B_1 \cup B_2 \cup B_3 \cup \ldots \cup B_n = (B_1 \cup B_2) \cup B_3 \cup \ldots \cup B_n = [(B_1 \cup B_2) \cup B_3] \cup \ldots \cup B_n \tag{7}$$

Therefore, the probability density function for the top event, $f_{TOP}(x)$, which is calculated from the probability density functions for n independent random variables, $f_1(x), \ldots, f_n(x)$, can be found by consecutively finding probability density functions of AND and OR logics for two independent variables.

Let B_X and B_Y be the Boolean variables for two basic events, and let X and Y be two independent random variables for the probabilities of B_X and B_Y. Let Z be the Boolean

variable as the result of the Boolean function h of the two Boolean variables, B_X and B_Y, and let T be a random variable for the probability of Z, as follows:

$$Z = h(B_X, B_Y) \tag{8}$$

$$T = g(X, Y) \tag{9}$$

where $g(X, Y)$ is the numerical function associated with $h(B_X, B_Y)$.

In uncertainty analysis in PSA, Boolean variables in event trees and fault trees are basically independent of each other. When the two random variables, X and Y, are independent of each other, the cumulative distribution function of T is given by

$$F(t) = \int_{D_h} f_X(x) f_Y(y) dy dx \tag{10}$$

where D_h denotes the region of the xy plane for the Boolean function h such that the numerical function g of the two independent variables, x and y, is lower than t, as follows:

$$D_h = \{(x,y) \, | g(x,y) \le t, \, 0 \le x \le 1, \, 0 \le y \le 1\} \tag{11}$$

The PDF of T, i.e., $f(t)$, is expressed as the derivative of the cumulative distribution function of T, i.e., $F(t)$ in Equation (10), with respect to t, as follows:

$$f(t) = \frac{d}{dt} F(t) \tag{12}$$

The Leibniz integral rule shown below can be used to derive $f(t)$ [6]

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(t,x) dx = f(t, b(t)) \cdot \frac{\partial}{\partial t} b(t) - f(t, a(t)) \cdot \frac{\partial}{\partial t} a(t) + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(t,x) dx \tag{13}$$

The following three subsections present the procedures applied to derive the PDFs for the (1) AND logic, (2) OR logic and (3) OR logic with rare event approximation of two independent random variables, X and Y.

2.1.1. AND Logic

A Boolean variable Z_{AND} is defined as the result of the AND logic of two Boolean variables, B_X and B_Y, and a random variable T_{AND} is defined for the probability of Z_{AND}, as follows:

$$Z_{AND} = B_X \cap B_Y \tag{14}$$

$$T_{AND} = XY \tag{15}$$

Then, the cumulative distribution function of T_{AND}, i.e., $F_{AND}(t)$, is given by

$$F_{AND}(t) = \int_{D_{AND}} f_X(x) f_Y(y) dy dx = \int_0^t \int_0^1 f_X(x) f_Y(y) dy dx + \int_t^1 \int_0^{\frac{t}{x}} f_X(x) f_Y(y) dy dx \tag{16}$$

where D_{AND} is given by Equation (17) and shown in Figure 1 when t equals 0.5.

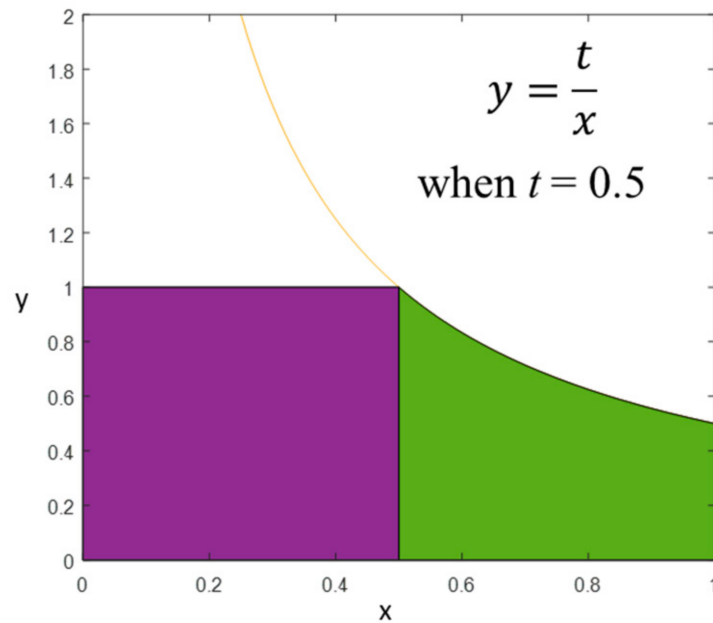$$D_{AND} = \{(x,y) \, | x \cdot y \le t, \, 0 \le x \le 1, \, 0 \le y \le 1\} \tag{17}$$

Figure 1. Region of the *xy* plane for the logical AND of two Boolean variables, B_X and B_Y.

The PDF of T_{AND}, i.e., $f_{AND}(t)$, can be derived as the derivative of the cumulative distribution function $F_{AND}(t)$, expressed in Equation (16), as follows:

$$
\begin{aligned}
f_{AND}(t) &= \tfrac{d}{dt} F_{AND}(t) = \tfrac{d}{dt} \int_0^t \int_0^1 f_X(x) f_Y(y) \, dy \, dx + \tfrac{d}{dt} \int_t^1 \int_0^{\frac{t}{x}} f_X(x) f_Y(y) \, dy \, dx \\
&= \tfrac{d}{dt} \int_0^t f_X(x) [F_Y(1) - F_Y(0)] \, dx + \tfrac{d}{dt} \int_t^1 f_X(x) \left[F_Y\left(\tfrac{t}{x}\right) - F_Y(0) \right] dx \\
&= f_X(t) - f_X(t) F_Y(1) + \int_t^1 \tfrac{\partial}{\partial t} f_X(x) F_Y\left(\tfrac{t}{x}\right) dx = \int_t^1 \tfrac{1}{x} f_X(x) f_Y\left(\tfrac{t}{x}\right) dx
\end{aligned}
\tag{18}
$$

It should be noted that $F_Y(1) = 1$, $F_Y(0) = 0$, and the Leibniz integral rule are applied in Equation (18).

2.1.2. OR Logic

A Boolean variable Z_{OR} is defined as the result of the OR logic of two Boolean variables, B_X and B_Y, and a random variable T_{OR} is defined for the probability of Z_{OR}, as follows:
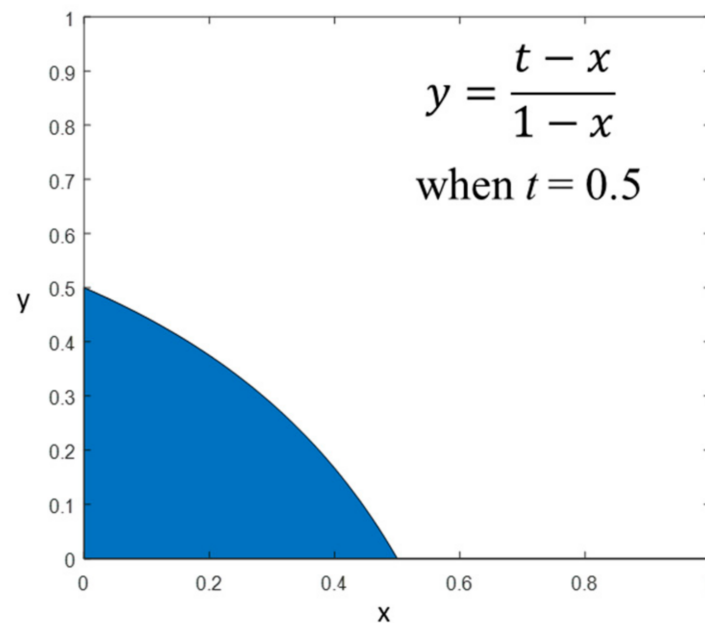
$$
Z_{OR} = B_X \cup B_Y \tag{19}
$$

$$
T_{OR} = X + Y - XY \tag{20}
$$

Then, the cumulative distribution function of T_{OR}, i.e., $F_{OR}(t)$, is given by

$$
F_{OR}(t) = \int_{D_{OR}} f_X(x) f_Y(y) \, dy \, dx = \int_0^t \int_0^{\frac{t-x}{1-x}} f_X(x) f_Y(y) \, dy \, dx \tag{21}
$$

where D_{OR} is given by Equation (22) and shown in Figure 2 when *t* equals 0.5.

$$
D_{OR} = \{(x,y) \mid x + y - x \cdot y \le t, \; 0 \le x \le 1, \; 0 \le y \le 1\} \tag{22}
$$

Figure 2. Region of the xy plane for the logical OR of two Boolean variables, B_X and B_Y.

The PDF of T_{OR}, i.e., $f_{OR}(t)$, can be derived as the derivative of the cumulative distribution function $F_{OR}(t)$, expressed in Equation (21), as follows:

$$f_{OR}(t) = \frac{d}{dt}F_{OR}(t) = \frac{d}{dt}\int_0^t \int_0^{\frac{t-x}{1-x}} f_X(x)f_Y(y)dydx = \frac{d}{dt}\int_0^t f_X(x)\left[F_Y\left(\frac{t-x}{1-x}\right) - F_Y(0)\right]dx$$

$$= f_X(t)F_Y\left(\frac{t-t}{1-x}\right) + \int_0^t \frac{\partial}{\partial t}f_X(x)F_Y\left(\frac{t-x}{1-x}\right)dx = \int_0^t \frac{1}{1-x}f_X(x)f_Y\left(\frac{t-x}{1-x}\right)dx \tag{23}$$

Similarly, $F_Y(0) = 0$, and the Leibniz integral rule are applied in Equation (23).

2.1.3. OR logic with Rare Event Approximation

In PSA, the Boolean expression for the top event is typically rearranged to the logical OR of minimal cutsets. When the probabilities of events are sufficiently lower than 1, rare event approximation is widely used in calculating the probability of the top event. Therefore, the PDF for the OR logic with rare event approximation has practical importance.

For the Boolean variable Z_{OR} for the result of the OR logic of two Boolean variables, B_X and B_Y, are given in Equation (19), and a random variable $T_{OR,REA}$ is defined for the probability of Z_{OR} with rare event approximation, as follows:

$$T_{OR,REA} = X + Y \tag{24}$$

Then, the cumulative distribution function of $T_{OR,REA}$, i.e., $F_{OR,REA}(t)$, is given by

$$F_{OR,REA}(t) = \int_{D_{OR,REA}} f_X(x)f_Y(y)dydx = \int_0^t \int_0^{t-x} f_X(x)f_Y(y)dydx \tag{25}$$

where $D_{OR,REA}$ is given by Equation (26) and shown in Figure 3 when t equals 0.5.

$$D_{OR,REA} = \{(x,y) \mid x + y \leq t, 0 \leq x \leq 1, 0 \leq y \leq 1\} \tag{26}$$

The PDF of $T_{OR,REA}$, i.e., $f_{OR,REA}(t)$, can be derived as the derivative of the cumulative distribution function $F_{OR,REA}(t)$, expressed in Equation (25), as follows:

$$f_{OR,REA}(t) = \frac{d}{dt}F_{OR,REA}(t) = \frac{d}{dt}\int_0^t \int_0^{t-x} f_X(x)f_Y(y)dydx = \frac{d}{dt}\int_0^t f_X(x)[F_Y(t-x) - F_Y(0)]dx$$

$$= f_X(t)F_Y(t-t) + \int_0^t \frac{\partial}{\partial t}f_X(x)F_Y(t-x)dx = \int_0^t f_X(x)f_Y(t-x)dx \tag{27}$$
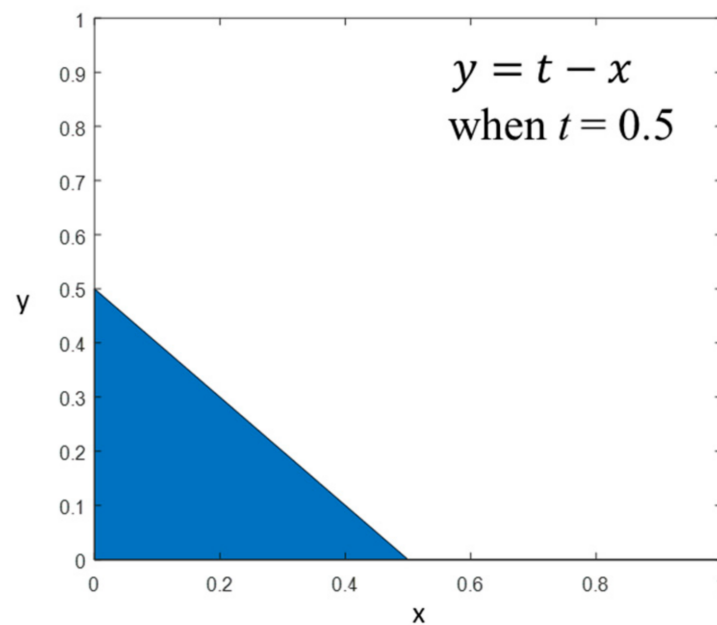
Figure 3. Region of the *xy* plane for the logical OR of two Boolean variables, B_X and B_Y, with rare event approximation.

Similarly, $F_Y(0) = 0$, and Leibniz integral rule are applied in Equation (27). It is noted that $f_{OR,REA}(t)$ is the convolution of $f_X(x)$ and $f_Y(y)$.

2.2. Implementation of Analytic Solutions

As mention previously, it is possible to obtain the PDF for any event in a coherent fault tree by repeatedly applying Equations (18), (23), or (27). The cumulative distribution function for the event can be obtained by repeatedly applying Equation (16), (21), or (25). A MATLAB$^{\text{TM}}$ code is developed to implement the analytic solutions shown above and the procedure for obtaining the PDF or the cumulative distribution function for the events of interest.

It is often difficult to determine the PDF or cumulative distribution function in a closed form, particularly when various distribution functions are associated with different basic events. Therefore, numerical integration is used when closed-form solutions for the integration in analytic solutions are not available. In numerical integration, relative and absolute tolerances are used for a trade-off between the accuracy and the computation time. In Section 3, the absolute tolerance is set to 0, and the relative tolerance is set to 0.1%. This is because the relative tolerance determines the accuracy of the numerical integration in the case of small-event probabilities.

3. Application to an Example System
3.1. Example System and Fault Tree Model

Figure 4 shows an example auxiliary feedwater system (AFWS). This system provides adequate feedwater from a condensate storage tank to a steam generator (SG) for the continuous residual heat removal from the primary system when the main feedwater system is unavailable. The AFWS comprises two pumps, a motor-driven pump, and a turbine-driven pump. Even if one of the pumps fail, the system can still perform its intended function because each pump has the capacity to provide enough feedwater to the SG. This corresponds to the logical conjunction in a fault tree because the system is in a failed state only when all the components are in the failed state.
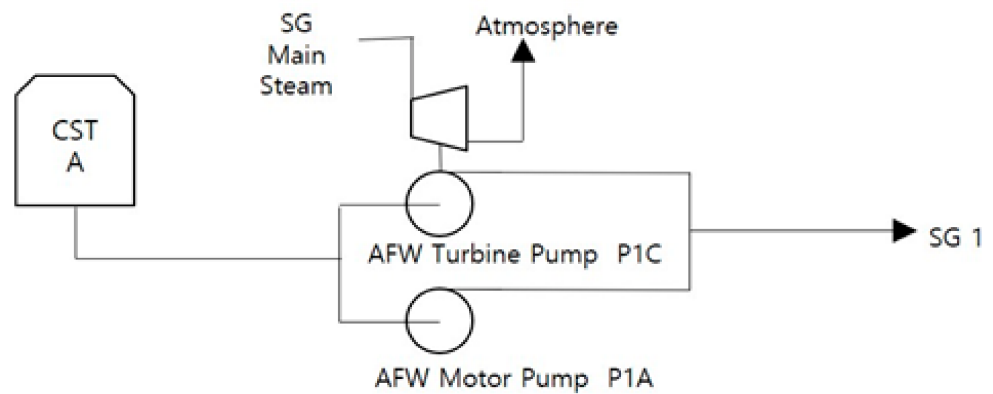
Figure 4. Example auxiliary feedwater system for a steam generator.

When a pump fails, the cause of the failure lies in one of the failure modes of the pump. Generally, failure on demand and running failure are the two most significant failure modes of a pump [12]. This corresponds to the logical disjunction in a fault tree because each failure mode prevents the pump from performing its intended function. In the case of a pump, a fail-to-start failure is one of the failures on demand and a fail-to-run failure is one of the running failures. Accordingly, the Boolean expression for the top event is given by

$$B_{TOP} = (B_A \cup B_B) \cap (B_C \cup B_D) \tag{28}$$

where B_A, B_B, B_C, B_D, and B_{TOP} are the Boolean variables for the four basic events (AFMPS-P1A, AFMPR-P1A, AFTPS-P1C, AFTPR-P1C), and the top event (GAF-SG1) in Figure 5, respectively. Figure 5 shows the fault tree for the example AFWS with two pumps and their failure modes
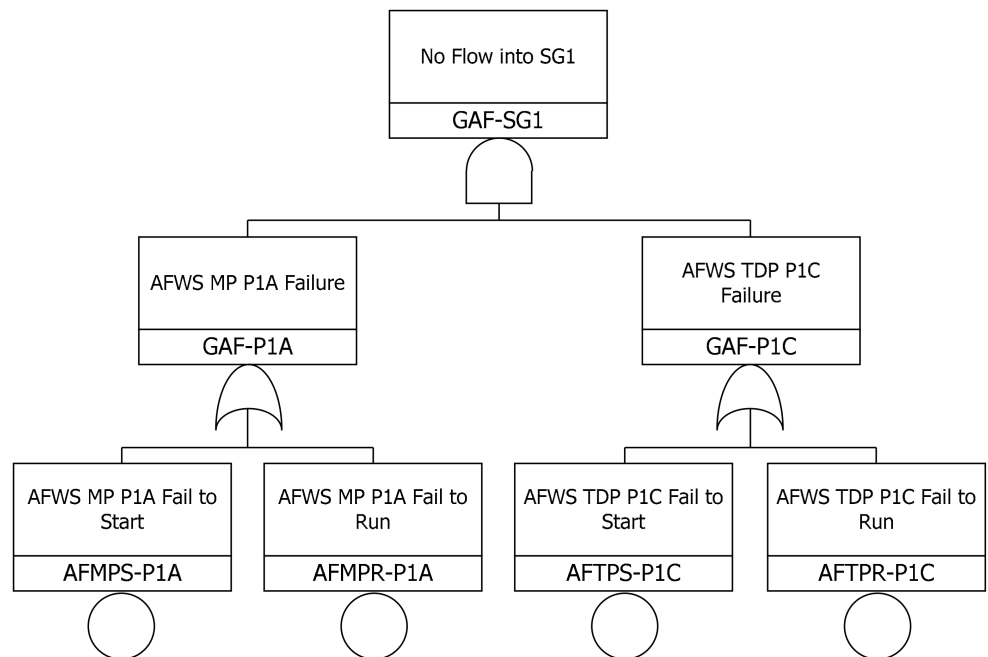

Figure 5. Fault tree for the example auxiliary feedwater system.

3.2. Uncertainty Data, Application and Comparison with Monte Carlo Simulation Results

The model parameters of the unavailability are the probability of failure on demand and the failure rate. For the PSA of nuclear power plants, various distributions have been

used to model the uncertainties in the parameters. The lognormal distribution has been widely used in many studies, such as [13]. Recent studies, such as [14], used the beta distribution for the probability of failure on demand and gamma distribution for the failure rate. Beta and gamma distributions have advantages over other distributions because they yield good results when data follow binomial and Poisson distributions, respectively, and are conjugate prior distributions in the Bayesian estimation of parameters.

Table 1 provides the industry-average data for fail-to-start failure and fail-to-run failure of motor-driven pumps and turbine-driven pumps. The source of the data is [14]. For the fail-to-start failure, a beta distribution is used to model the uncertainty in the probability of failure on demand. For the fail-to-run failure, a gamma distribution is used to model the uncertainty in the failure rate. The mission time is assumed to be 24 h in this example.

Table 1. Fail-to-start and fail-to-run data for motor-driven and turbine-driven pumps.

| Failure Mode	Distribution	Parameters	
		α	β
Motor-driven pump fail-to-start	Beta	0.909	6.175×10^2
Motor-driven pump fail-to-run	Gamma	0.500	8.619×10^4
Turbine-driven pump fail-to-start	Beta	0.414	5.976×10^1
Turbine-driven pump fail-to-run	Gamma	0.500	6.803×10^3

For fail-to-run failures, the PDF for the unavailability of the event has the same probability distribution as that for the failure rate (see Appendix A). Therefore, for fail-to-run failure events, AFMPR-P1A and AFTPR-P1C, the PDFs of the unavailability follow gamma distributions, similarly to the PDFs for the failures rates provided in 0, as shown below:

$$f_\Lambda(\lambda;\ \alpha,\ \beta)d\lambda\ =\ \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda \tag{29}$$

$$f(u)du\ =\ f_\Lambda\left(\tfrac{u}{t}\right)d\left(\tfrac{u}{t}\right) = \tfrac{\beta^\alpha}{\Gamma(\alpha)}\left(\tfrac{u}{t}\right)^{\alpha-1}e^{-\beta\frac{u}{t}}d\left(\tfrac{u}{t}\right) = \tfrac{\left(\tfrac{\beta}{t}\right)^\alpha}{\Gamma(\alpha)}u^{\alpha-1}e^{-\frac{\beta}{t}u}du\ =\ f_\Lambda\left(u;\ \alpha,\ \tfrac{\beta}{t}\right)du \tag{30}$$

Figure 6 shows the PDFs for four basic events in the example. The PDFs for the two events (GAF-P1A and GAF-P1C) and the top event (GAF-SG1) can be derived using the equations provided in Section 2 and shown in Figure 7. Figure 8 shows the probability distributions for the events in a semi-log scale, with logical relationships between the events.

Monte Carlo simulations are performed with 1,000,000 samples. Simulations are performed on the same set of minimal cutsets obtained for the mean values of basic event probabilities [15]. Their results are compared with the analytic solutions, as shown in Figures 6–8. The analytic solutions and Monte Carlo simulation results are in good agreement.
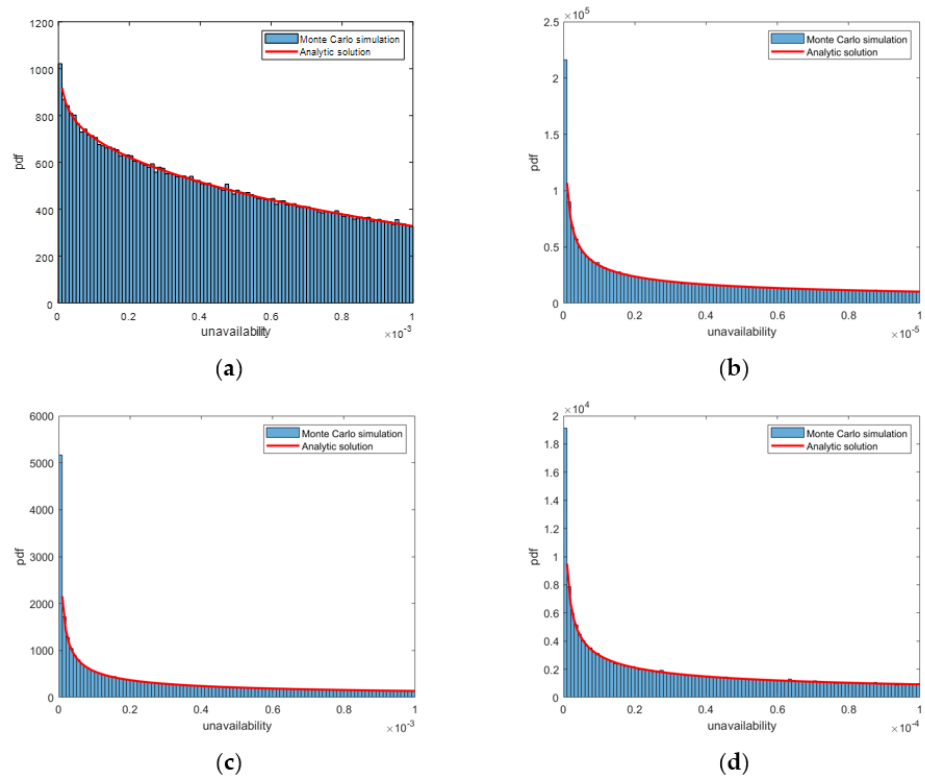
Figure 6. Probability density functions (PDFs) for basic events along with Monte Carlo simulation result: (**a**) AFMPS-P1A (beta distribution); (**b**) AFMPR-P1A (gamma distribution); (**c**) AFTPS-P1C (beta distribution); (**d**) AFTPR-P1C (gamma distribution).
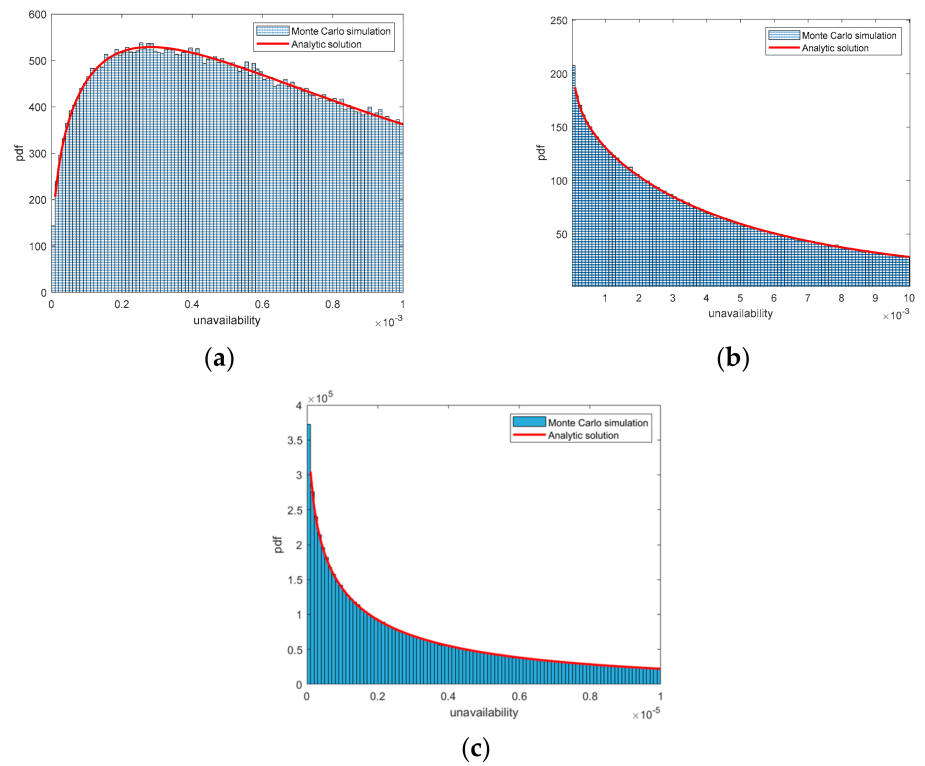
Figure 7. Analytic solutions for the PDFs of (**a**) GAF-P1A; (**b**) GAF-P1C; (**c**) GAF-SG1 (**top event**), along with Monte Carlo simulation results.
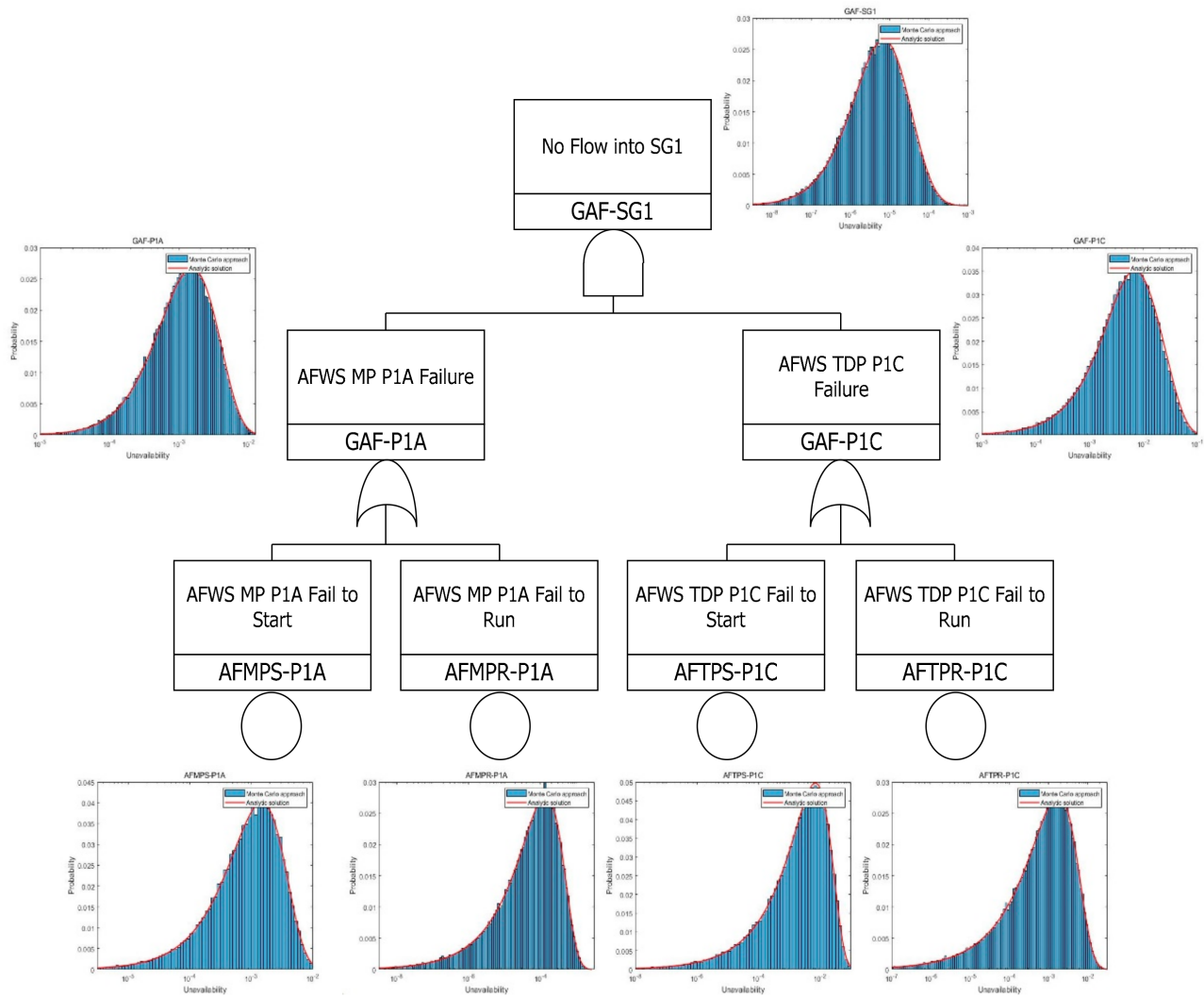
Figure 8. Probability distribution in a semi-log scale with logical relationships between events and comparison with Monte Carlo simulation results.

3.3. Discussions

Although a gamma distribution is a right-unbounded distribution, the domain of the random variables for the unavailability is bounded by (0, 1). As a result, the integration of the PDF from 1 to infinity is inevitably truncated. Table 2 lists the truncated values, i.e., $1 - F(1)$, of the events. It is first noted that the truncated values are zero when using beta distributions, AFTPS-P1C and AFMPS-P1A, because the beta distributions are bounded by (0, 1). When gamma distributions are used, the truncated values are low enough to justify the use of the right-unbounded distribution to describe the uncertainty in the unavailability/probability. Similarly, the analytic solutions enable the analysts to quantify the truncation errors due to the use of a right-unbounded distribution for unavailability/probability in uncertainty analysis, which is hardly quantified with Monte Carlo simulation.

Table 2. Truncated values of events in the fault tree.

Events	Truncated values
GAF-SG1	1.76×10^{-6}
GAF-P1A	0
GAF-P1C	1.88×10^{-6}
AFMPS-P1A	0
AFMPR-P1A	9.00×10^{-16}
AFTPS-P1C	0
AFTPR-P1C	7.00×10^{-15}

When the mean unavailability of the events is significantly lower than 1, the application of rare event approximation is valid for the logical OR of the events. Figure 9 compares the PDFs for the top event unavailability between the exact results calculated using Equation (23) and the rare event approximation made using Equation (27). The PDFs are in good agreement. In 0, the maximum error is 1.39% when the unavailability is 1×10^{-3}, the rightmost point in the figure. Thus, the rare event approximation is found to provide a reasonable approximation to the exact result, not only for the mean unavailability but also for the PDF, that characterizes the uncertainty in the unavailability of the event. The use of rare event approximation in uncertainty analysis with analytic solutions is expected to significantly reduce the computational burden in calculating the PDFs of events in a fault tree with limited errors.
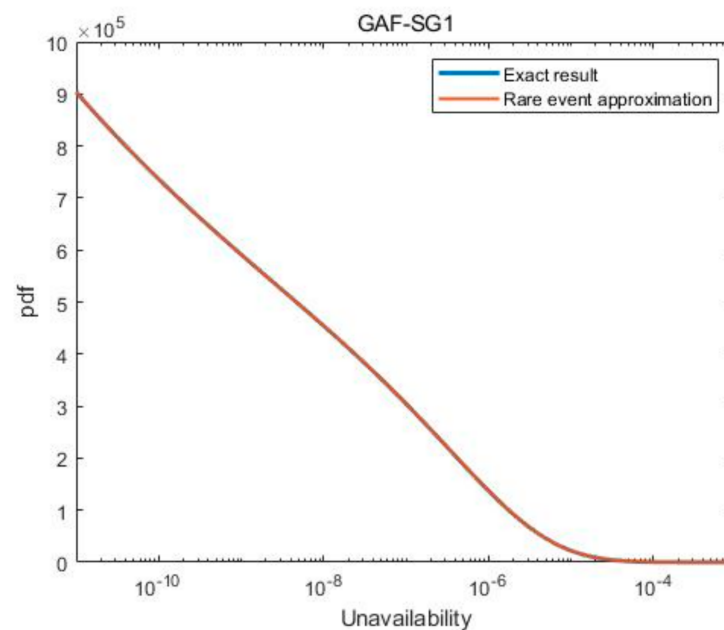
Figure 9. Comparison of PDFs for the top event between exact results and rare event approximation.

In the case of lognormal distributions, there are many previous studies such as Fenton-Wilkinson [16], Schwartz-Yeh [17], and El-Shanawany et al. [18], on finding the approximate PDF of the unavailability of an event. However, there have been few discussions on theoretical methods or approximations on PDFs when using probability distributions other than lognormal distributions, or when using combinations of different probability distributions. For example, there are few theoretical methods or approximation methods to determine the PDFs when beta and gamma distributions are logically combined, as in GAF-P1A and GAF-P1C in the example above, as well as in the case of their combination (GAF-SG1). The importance of uncertainty analysis with beta and gamma distributions can be emphasized,

with the recent trend in employing them for modeling the uncertainty of events owing to the advantages mentioned above. The mathematical formulation and analytic solutions, presented in Section 2, provide a theoretical foundation for the uncertainty analysis with probability distributions other than lognormal distributions or a combination of probability distributions. As beta and gamma distributions are widely adopted in uncertainty analysis of PSA, further research on the combination of those distributions and their approximations are expected on the theoretical foundations described in this paper.

4. Conclusions

Monte Carlo simulations are widely used for uncertainty analysis in PSA. Although approximation approaches have been developed for specific probability distributions, such as a lognormal distribution, few theoretical approaches have been proposed for general probability distributions.

This study provides a mathematical formulation and analytic solutions for the uncertainty analysis in PSA. Starting from the definitions of variables, mathematical equations are derived to synthesize the PDFs for the logical AND, logical OR, and logical OR with rare event approximation of two independent events. When there are more than two events, the derived equations can be applied consecutively.

The effectiveness of the analytic solutions is demonstrated by applying them to an example system. The PDFs synthesized with the analytic solutions are found to be in good agreement with the Monte Carlo simulation results. Hence, the analytic solutions can be used to validate Monte Carlo simulation results and explain those aspects that cannot be done using other methods. For example, the analytic solutions can provide truncated values when right-unbounded distributions, such as lognormal and gamma distributions, are used to model the uncertainty in the unavailability of an event.

Thus far, there have been few discussions on theoretical methods or approximation methods to determine the PDFs of such events when using probability distributions other than lognormal distributions or when using combinations of various probability distributions. Monte Carlo simulations seem to be the most widely used for such conditions. The analytic solutions proposed herein provide a new method for performing uncertainty analysis when using a combination of various probability distributions, e.g., a combination of beta and gamma distributions, that has been widely used in PSA in recent times.

The theoretical approach for uncertainty analyses remains a basis for all other approaches, such as approximation methods and Monte Carlo simulations. Moreover, a theoretical approach provides deeper insights into the uncertainty analysis that cannot be provided by approximation methods or Monte Carlo simulations. Although the theoretical approach is computationally burdensome, we found that it has significant potential and should be further investigated. The mathematical formulation and analytic solutions provided in this study are expected to serve as a basis for future studies on theoretical approaches for the uncertainty analysis in PSA.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In the practice of PSA, two types of unavailability are widely considered for hardware components: the probability of failure on demand for fail-to-start failures and the probability of failure in performing intended functions under a specified mission time for fail-to-run failures. For fail-to-start failures, the probability of failure on demand equals the unavailability; hence, the PDF for the unavailability is the same as that for the probability of failure on demand. For fail-to-run failures, it is shown below that the PDF for the unavailability has the same probability distribution as that for the failure rate under specified conditions.

When the random variable for the time to failure T is modeled using an exponential distribution with a constant failure rate λ, the unavailability u due to the fail-to-run failures as a function of the mission time t is given by

$$u = \Pr[T \leq t] = F(t) = \int_0^t \lambda e^{-\lambda\tau}d\tau = 1 - e^{-\lambda t} \tag{A1}$$

When λt is significantly lower than 1, Equation (28) can be approximated as

$$u = 1 - e^{-\lambda t} \approx \lambda t \tag{A2}$$

The random variables Λ and U are defined for the failure rate λ and unavailability u, as follows:

$$U = 1 - \Lambda e^{-\Lambda t} \approx \Lambda t \tag{A3}$$

The cumulative distribution function of U, i.e., $F_U(u)$, is given by

$$F_U(u) = \int_0^u f_U(u')du' = \Pr[U \leq u] \approx \Pr[\Lambda t \leq u] = \Pr\left[\Lambda \leq \frac{u}{t}\right] = \int_0^{\frac{u}{t}} f_\Lambda(\lambda)d\lambda = F_\Lambda\left(\frac{u}{t}\right) \tag{A4}$$

where $f_U(u)$ and $f_\Lambda(\lambda)$ are the PDFs for U and Λ, respectively, and $F_\Lambda(\lambda)$ is the cumulative distribution function of Λ. The PDF of U, i.e., $f_U(u)$, can be derived as the derivative of the cumulative distribution function $F_U(u)$, expressed in Equation (A1), as follows:

$$f_U(u)du = \frac{dF_U(u)}{du}du = \frac{dF_\Lambda\left(\frac{u}{t}\right)}{d\left(\frac{u}{t}\right)}\frac{d\left(\frac{u}{t}\right)}{du}du = f_\Lambda\left(\frac{u}{t}\right)\frac{1}{t}du = f_\Lambda\left(\frac{u}{t}\right)d\left(\frac{u}{t}\right) = f_\Lambda(\lambda)d\lambda \tag{A5}$$

Therefore, the PDF for the unavailability due to fail-to-run failures has the same probability distribution as that for the failure rate $f_\Lambda(\lambda)$ when the time to failure is exponentially distributed and the product of the failure rate and mission time λt is significantly lower than 1.

References

1. Adamec, P.; Ahmed, K.; Babar, A.K.; Berg, H.P.; Boneham, P.; Burgazzi, L.; Campbell, F.; Carlsson, L.; Caruso, G.; Čepin, M. *Applications of Probabilistic Safety Assessment (PSA) for Nuclear Power Plants*; International Atomic Energy Agency: Vienna, Austria, 2001.
2. Winkler, R.L. Uncertainty in probabilistic risk assessment. *Reliab. Eng. Syst. Saf.* **1996**, *54*, 127–132. [CrossRef]
3. Apostolakis, G. The Concept of Probability in Safety Assessments of Technological Systems. *Science* **1990**, *250*, 1359–1364. [CrossRef] [PubMed]
4. Apostolakis, G.E. Uncertainty in Probabilistic Safety Assessment. *Nucl. Eng. Des.* **1989**, *115*, 173–179. [CrossRef]
5. Cook, A.S.; Downs, T. Estimating Manufacturing Yield by Means of Jacobian of Transformation. *IEE Proc. G (Electron. Circuits Syst.)* **1982**, *129*, 127–133. [CrossRef]
6. Papoulis, A.; Saunders, H. Probability, Random Variables and Stochastic Processes (2nd Edition). *J. Vib. Acoust.* **1989**, *111*, 123–125. [CrossRef]
7. Ulmeanu, A.P. Analytical Method to Determine Uncertainty Propagation in Fault Trees by Means of Binary Decision Diagrams. *IEEE Trans. Reliab.* **2012**, *61*, 84–94. [CrossRef]
8. Kaplan, S. On The Method of Discrete Probability Distributions in Risk and Reliability Calculations–Application to Seismic Risk Assessment. *Risk Anal.* **1981**, *1*, 189–196. [CrossRef]

9. Levy, L.L.; Moore, A.H. A Monte Carlo Technique for Obtaining System Reliability Confidence Limits from Component Test Data. *IEEE Trans. Reliab.* **1967**, *R-16*, 69–72. [CrossRef]

10. Jackson, P.S.; Hockenbury, R.W.; Yeater, M.L. Uncertainty Analysis of System Reliability and Availability Assessment. *Nucl. Eng. Des.* **1982**, *68*, 5–29. [CrossRef]

11. Durga Rao, K.; Kushwaha, H.S.; Verma, A.K.; Srividya, A. Quantification of Epistemic and Aleatory Uncertainties in Level-1 Probabilistic Safety Assessment Studies. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 947–956. [CrossRef]

12. Jung, W.D.; Lee, Y.H.; Hwang, M.J. *Procedure for Conducting Probabilistic Safety Assessment: Level 1 Full Power Internal Event Analysis*; IAEA: Vienna, Austria, 2003.

13. Eide, S.A.; Chmielewski, S.V.; Swantz, T.D. *Generic, Component Failure Data Base for Light Water and Liquid Sodium Reactor Pras*; Idaho National Laboratory (INL): Idaho Falls, ID, USA, 1990.

14. Eide, S.A.; Wierman, T.E.; Gentillon, C.D.; Rasmuson, D.M.; Atwood, C.L. *Industry-Average Performance for Components and Initiating Events at US Commercial Nuclear Power Plants*; NUREG/CR-6928; Nuclear Regulatory Commission: Washington, DC, USA, 2007.

15. Rauzy, A. Notes on Computational Uncertainties in Probabilistic Risk/Safety Assessment. *Entropy* **2018**, *20*, 162. [CrossRef] [PubMed]

16. Fenton, L. The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems. *IRE Trans. Commun. Syst.* **1960**, *8*, 57–67. [CrossRef]

17. Schwartz, S.C.; Yeh, Y.S. On the Distribution Function and Moments of Power Sums with Log-Normal Components. *Bell Syst. Tech. J.* **1982**, *61*, 1441–1462. [CrossRef]

18. El-Shanawany, A.B.; Ardron, K.H.; Walker, S.P. Lognormal Approximations of Fault Tree Uncertainty Distributions. *Risk Anal.* **2018**, *38*, 1576–1584. [CrossRef] [PubMed]