▩ 연구논문

Approximation Method for Failure Rates in a General Event Tree+

- 사건 가지상의 사고율 추정을 위한 근사적인 방법 -

Yang, Hee Joong
양 희 중*

요 지

사건 가지 상의 파라메터 추정을 위한 베이지안 접근방식이 제시된다. 먼저 일반적인 사건 가지를 따라 발생하는 사고를 예측하기 위한 모형에 대해 설명한다. 이 경우 이론적으로 베이지안 기법을 적용하는 방법에 대해 논하고 실제로 문제를 풀 경우에 발생하는 다차원 수치적분 문제를 다룬다. 감마 분포와 베타분포가 이용될 경우 위 문제를 쉽게 해결할 수 있는 근사적 방법에 대해 연구한다. 또한 사건가지상의 여러 경로가 같은 수준의 사고로 분류 될 수 있는 경우에 대해서도 위와 같은 방법에 관한 연구를 한다. 결과적으로 한 사고율이 여러 개의 파라메터의 함수로 표현되어 다차원의 수치적분이 요구되는 경우 이를 쉽게 해결 할 수 있는 근사적인 방법이 제시되어 베이지안 기법의 적용이 용이해 질 수 있다.

1. Introduction

Event Tree is one of the most frequently used tools in analyzing the occurrence of an accident and its way of propagation to more severe accident. Even though the usefulness of event trees many researchers apply classical statistical approaches on them, thereby event trees are often known as tools that only analyze the way of accident initiation and escalation after an accident occurs. NUREG report[8] can be one of the examples of the above argument. But we can get much more information from an event tree when we approach from a slightly different angle. Yang[4, 5, 6] adopted bayesian approach in updating parameters in event trees and extended the use of event trees to the adaptive parameter updating and forecasting. Sometimes many dimensional numerical integration is inevitably required when we follow bayesian approach in parameter updating. A general event tree is composed of many branches, thus as many parameters as the number of

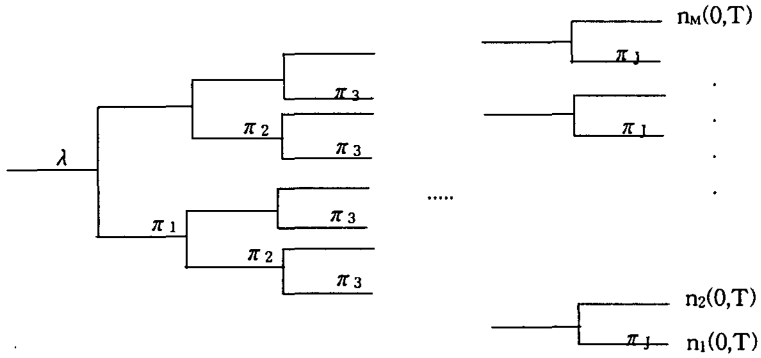
---

+ 이 논문은 99-00 학년도 청주대학교 교비 특별연구 지원비에 의하여 연구 되었슴.
* 청주대학교 산업공학과 교수

branches are involved. Therefore the probability of a specific sequence in a tree is a product of many parameters and we hardly expect such value is expressed in a nice closed form distribution. This is one of the most difficult problem in applying bayesian approach to real world problems. In this paper we propose a way of approximation method that detours such problem and makes the bayesian approach on event trees easy to perform.

2. General Event Tree Models

The failure rate that follows a specific sequence in an event tree is expressed as a function of initial failure rate, λ, and the probability of malfunction of following sub-systems, π's. Consider a general event tree in figure 1, where J sub-systems are involved and ends up with M different accident sequences. This is an event tree that shows the initiating event occurring with the rate of λ escalates to more severe accident depending on the function or malfunction of following safety backup systems. Let the rate of initiating event be λ, the probability of malfunction of jth sub-system be π_j, j=1,2,...,J.



<Figure 1> An example of a general event tree

Also let the number of accidents in time period (0,T) following the mth path from the bottom of the event tree be $n_m(0,T)$, m=1,2,...,M The maximum of M can be 2^J. Once an initiating event occurs, the probability of following a specific sequence of an event tree depends on functioning or malfunctioning of sub-systems. Let the probability vector associated with the mth accident sequence from the bottom of an event tree be θ_m. In figure 2, for example, the probability that follows the second sequence from the bottom is the probability that the first sub-system fails, the second sub-system succeeds, and the third sub-system fails. Thus θ_2 is represented by $\pi_1(1-\pi_2)\pi_3$. Therefore $\lambda\theta_m$ represents the failure rate that follows the mth sequences from the bottom of an event tree, and we let $\lambda\theta_m$ be λ_m.

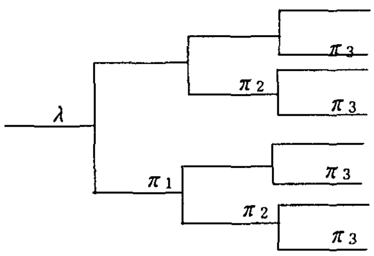
The distribution of failure rate λ is usually assumed to follow a lognormal distribution [2, 3, 7, 8]. The rational of using lognormal distribution is that a safety system provided with many redundancies tends to bunch up toward the low probability of failure. Yang[4] argued that the shape of gamma distribution also showed the tendency of positive skewness and it was a better candidate for the failure rate from the aspect of calculational efficiency. Thus, in this paper, we assume gamma distribution for λ with parameter α, β. The distributions for the failure probabilities π's are assumed to be beta distributions since beta distributions are quite flexible covering almost all forms of distribution on (0,1). The probability of jth sub-system failure π_j is assumed to be beta distribution with parameters a_j, b_j:

$$\lambda \sim \Gamma(\alpha, \beta), \qquad \pi_j \sim Be(a_j, b_j)$$

Then F_m denoting the cumulative distribution of λ_m is expressed as

$$F_m(z) = Prob\{\lambda_m = \lambda\theta_m \le z\} = \int \dots \int_{\lambda\theta_m \le z} f(\lambda)\, g(\theta_m)\, d\lambda\, d\theta_m \qquad (1)$$

where $f(\lambda)$ and $g(\theta_m)$ denotes the probability density function of λ and θ_m.



<Figure 2> An event tree that includes three sub-systems

The probability density function of λ_m, $f_m(\lambda_m)$, is obtained by differentiating equation (1) with respect to z;

$$f_m(\lambda_m) = \frac{d}{dz} F_m(z) \qquad (2)$$

The distribution resulting from the above equation can not usually be expected to be in a nice close form. Rather we are supposed to be encounter a j+1 dimensional numerical integration when we deal with the general event tree that contains j sub-systems. Such a many dimensional numerical integration is practically unsolvable or it requires considerable computer time. In order to circumvent this difficulty, we propose a method of approximation that essentially breaks down a problem involving many integrations into several repetitive steps so that each step involves only a small number of integrations.

3. An Approximation Method in a General Event Tree Model

Let the failure rate following the mth sequence from the bottom of the event tree in figure 1 be λ_m. Then λ_m is expressed as λ times θ_m, which is the probability vector associatedwith mth sequence from the bottom of the general event tree in figure 1 as defined in the previous section. Let the mean and the variance of the probability θ_m be μ_m, and σ_m, respectively. The resulting mean and the variance of λ_m can be expressed as equation (3).

$$E[\lambda_m] = \frac{\alpha}{\beta}\mu_m, \qquad Var[\lambda_m] = \frac{\alpha}{\beta^2}\mu_m^2 + \frac{\alpha^2+\alpha}{\beta^2}\mu_m^2 \tag{3}$$

The distribution of θ_m is relatively sharp compared to the shape of the distribution of λ especially when $\frac{\sigma_m}{\mu_m}$ is small. Also the posterior variance of θ_m decreases to zero as failures are accumulated. In this case θ_m can be approximated to be a constant. The distribution resulting from the multiplication of constant to a gamma distribution is also another gamma distribution;

$$\lambda_m \sim \Gamma(\alpha_m, \beta_m)$$

where α_m, β_m are obtained as solutions of following equations :

$$\frac{\alpha_m}{\beta_m} = \frac{\alpha}{\beta}\mu_m, \qquad \frac{\alpha_m}{\beta_m^2} = \frac{\alpha}{\beta^2}\mu_m^2 + \frac{\alpha^2+\alpha}{\beta^2}\mu_m^2$$

Then we the following parameters of λ_m :

$$\alpha_m = \alpha\frac{\mu_m^2}{(\alpha+1)\sigma_m^2+\mu_m^2}, \qquad \beta_m = \frac{\beta\mu_m}{(\alpha+1)\sigma_m^2+\mu_m^2}$$

Since the distribution of λ_m is approximated to be a gamma distribution, the predictive distribution for the time to next accident following the same sequence can be obtained as a closed form distribution which is a shifted pareto distribution as following;

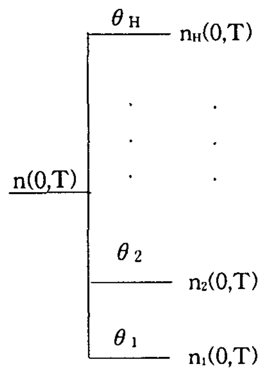
$$p(x_m) = \int p(x_m \mid \lambda_m)\, p(\lambda_m)\, d\lambda_m = \left(\frac{\beta_m}{\beta_m + x_m}\right)^{\alpha_m}\left(\frac{\alpha_m}{\beta_m + x_m}\right)$$

The above approximation method help us detour the difficulty of many dimensional numerical integration that we often encounters when to make a bayesian forecasting. This method becomes more accurate when $\frac{\sigma_m}{\mu_m}$ is small. As we acquire more and more data, the distribution of θ_m becomes sharper and sharper, resulting in smaller values of $\frac{\sigma_m}{\mu_m}$. Therefore we can expect more accurate approximations as we observe more data.

4. Approximation Method When Several Sequences Ends Up to Same Levels of Accidents

There are often cases when the different sequences end up with same level of accidents. In this case it has more meaning to analyze the specific level of accidents rather than to analyze the specific sequence of accidents. The event tree in figure 1 is simplified to as figure 3, assuming M different accident sequences are combined to H different levels of accidents. The number of accident on the left branch n(0,T) follows the poisson distribution with parameter λ. The number of accidents on right hand side branches, $n_h(0,T)$, is the number that is obtained by randomly partitioning the total number of accident n(0,T) with the probability θ_h. Therefore $n_h(0,T)$, h= 1,2, ...,H, also follows poisson distribution with parameter $\lambda\theta_h$. Then the time to next accident that follows each path, X_h, follows exponential distribution with parameter $\lambda\theta_h$. Let's assume, for example, that the accidents that follow the 2th, 6th, and 9th paths are classified as the same level of accidents. The time to such level of accident can occur following any of these paths so that the time to such accident, X, is the minimum of time to accident that follows any of these paths;

$$X = \min[X_2, X_6, X_9]$$



$$n(0,T) = \sum_{h=1}^{H} n_h(0,T)$$

Figure 3. An event tree that combines same levels of accidents.

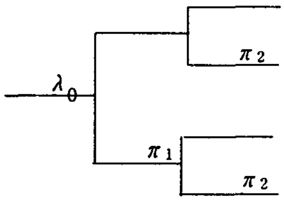
Therefore X follows exponential distribution with parameter obtained by summing all parameters of exponential distributions for different sequences;

$$X \mid \lambda, \theta_1, \ldots \theta_M \sim \mathrm{Exp}(\lambda_2 + \lambda_6 + \lambda_9)$$

$\lambda_2+\lambda_6+\lambda_9$ is expressed as $\lambda(\theta_2+\theta_6+\theta_9)$, and the distribution of $\theta_2+\theta_6+\theta_9$ is considerably sharp compared to that of λ. As we explained before the distribution of $\lambda_2+\lambda_6+\lambda_9$ is approximated as a gamma distribution that makes easy to obtain the distribution of time to the next accident of a specific level.

5. Numerical example

Consider the event tree in figure 4 that shows the escalation process up to more severe accident from initiating events.



<Figure 4.> An event tree that includes two sub-systems

The counts at the end of each sequence can be considered as randomly partioned numbers from the total number of initiating events with probabilities of $(1-\pi_1)(1-\pi_2)$, $(1-\pi_1)\pi_2$, $\pi_1(1-\pi_2)$, and $\pi_1\pi_2$ from the top to the bottom sequence, respectively. Then the number of accidents that pass through each sequence follows a Poisson distribution with a rate of λ_0 times the associated probabilities. Let's assume that the accident following the first sequence from the top of the event tree remains as the same severity accident as the initiating event due to the successful operation of the first and the second sub-systems, and the accidents following remaining three sequences escalate to more severe accidents due to the failure of operation of one or both of the sub-systems. Let X_i be the time to the next accident following the ith sequence from the top of the event tree. Then time to the next accident is exponentially distributed with the appropriate rate:

$$X_2 \sim Exp(\lambda_0(1-\pi_1)\pi_2)$$
$$X_3 \sim Exp(\lambda_0\pi_1(1-\pi_2))$$
$$X_4 \sim Exp(\lambda_0\pi_1\pi_2)$$

Since the severe accident may happen following any sequence except the first sequence in the event tree, the time to next severe accident, X, is minimum of X_2, X_3, and X_4;

$$(X \mid \lambda_0, \pi_1, \pi_2) = Min.[X_2, X_3, X_4 \mid \lambda_0, \pi_1, \pi_2] \sim Exp(\lambda_0(\pi_1+\pi_2-\pi_1\pi_2))$$

The above is because the minimum of exponentially distributed independent random variables is also exponential with rate of the sum of all individual rates. Let the resulting rate of more severe accident $\lambda_0(\pi_1+\pi_2-\pi_1\pi_2)$ be λ_1.

To find the distribution of λ_1, let λ_0, π_1, π_2 follow distributions of F_1, F_0, G_1, and G_2, respectively. Then the cumulative distribution of λ_1 is

$$F_1(z) = Prob.\{\lambda_1 = \lambda_0(\pi_1+\pi_2-\pi_1\pi_2) \le z\}$$

$$= \int\int\int_{\lambda_0(\pi_1+\pi_2-\pi_1\pi_2) \le z} f_0(\lambda_0) g_1(\pi_1)\, g_2(\pi_2)\, d\lambda_0\, d\pi_1\, d\pi_2$$

$$= \int_0^1 \int_z^{\frac{z}{\pi_2}} \int_0^{\frac{z-\lambda_0\pi_2}{\lambda_0(1-\pi_2)}} f_0(\lambda_0) g_1(\pi_1) g_2(\pi_2)\, d\pi_1\, d\lambda_0\, d\pi_2$$

$$= \int_0^1 \int_z^{\frac{z}{\pi_2}} f_0(\lambda_0)\, G_1\!\left(\frac{z-\lambda_0\pi_2}{\lambda_0(1-\pi_2)}\right) g_2(\pi_2)\; d\lambda_0\, d\pi_2 \tag{4}$$

By differentiating equation (4), we obtain

$$f_1(z) = \int_0^1 \int_z^{\frac{z}{\pi_2}} f_0(\lambda_0)\, g_1\!\left(\frac{z-\lambda_0\pi_2}{\lambda_0(1-\pi_2)}\right) g_2(\pi_2) \frac{1}{\lambda_0(1-\pi_2)}\, d\lambda_0 d\pi_2 \tag{5}$$

Here we want to find a closed form distribution that approximates the distribution obtained by equation (5). Let $\lambda_0 \sim \Gamma(\alpha_0, \beta_0)$ and $E[\pi_1+\pi_2-\pi_1\pi_2]=\mu_1$, and $Var[\pi_1+\pi_2-\pi_1\pi_2]=\sigma_1^2$. The mean and variance of the more severe accident are calculated from

$$E[\lambda_1] = \frac{\alpha_0}{\beta_0}\mu_1, \qquad Var[\lambda_1] = \frac{\alpha_0}{\beta_0^2}\mu_1^2 + \frac{\alpha_0^2+\alpha_0}{\beta_0^2}\mu_1^2 \tag{6}$$

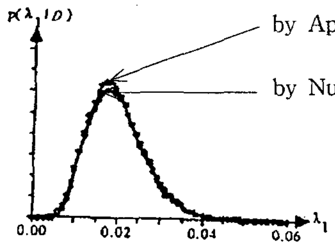
As justified in the previous section, λ_1 can be approximated by $\Gamma(\alpha_1, \beta_1)$. We obtain parameters α_1, β_1 from equation (6):

$$\alpha_1 = \alpha_0 \frac{\mu_1^2}{(\alpha_0+1)\sigma_1^2+\mu_1^2}, \qquad \beta_1 = \frac{\beta_0\mu_1}{(\alpha_0+1)\sigma_1^2+\mu_1^2} \tag{7}$$

We arbitrarily assume λ_0 follows gamma distribution with parameters (100, 200), π_1, and π_2 follow beta distribution with parameters (10, 500), and (5, 150), respectively. Then the parameters of the arrival rate of more severe accident is approximately obtained by equation (7);

$\lambda_1 \sim \Gamma(99.8, 3890)$

The gamma distribution obtained by the approximation method suggested in this paper is compared with the distribution obtained by the numerical integration based on equation (4), that is in figure 5. Figure 5 shows that the approximation is fairly close. Furthermore, as we acquire more and more data the distributions of π_1, and π_2 become sharper and sharper, resulting in smaller values of σ_1/μ_1. Therefore we can expect more accurate approximations as we observe more data.



<Figure 5> Distributions obtained by numerical integration and obtained by approximation

6. Summary

A bayesian approach is adopted with the use of general event tree models. We find the way of detouring the difficulty that often encounters when to apply bayesian forecasting. Gamma distribution for the accident initiating rate, and beta distribution for failure probabilities are adopted. The failure rate following specific accident sequence in an event tree is approximated to be another gamma distribution that makes the forecasting time to next accident be easy. In our case the predictive distribution is ended up with shifted pareto distribution. We also analyze the case where several different accident sequences result to same levels of accidents. In such case, similar approach is applied and also provides with closed form distribution for the time to next accident. The simple and easy results derived in this paper can be applied in predicting time to next accidents or number of accidents in a given period of time in a general event tree.

References

[1] Aitchison, I.R., and Dunsmore, "Statistical Prediction Analysis", Cambridge University Press, 1975.

[2] Apostolakis, G., and Moleh, A., "Expert Opinion and Statistical Evidence: An Application to Reactor Core Melt Frequency", *Nuclear Science and Engineering*, Vol.70. pp.135-149, 1979.

[3] Chow, T.C., and Oliver, R.M., "Predicting Nuclear Incidents", *Journal of Forecasting*, Vol.7, pp.49-61, 1988.

[4] Heejoong Yang, "The Influence of Human Errors on Nuclear Incidents", Doctoral dissertation, University of California, Berkeley, 1989.

[5] Heejoong Yang, et.el., "A Bayesian Approach to Prediction of System Failure Rates by Criticalities Under Event Trees", International Journal of Production Economics, Vol.60, pp.623-628, 1999.

[6] Oliver, R.M., and Heejoong Yag, "Bayesian Updating of Event Tree Parameters to Predict High Risk Incidents", Influence Diagrams, Belief Nets and Decision Analysis, John Wiley and Sons, pp.277-296, 1990.

[7] Rasmussen, N.C., "Method of Hazard Analysis and Nuclear Safety engineering", Annals New Your Academy of Science , pp.20-36, 1981.

[8] U.S. Nuclear Regulatory Commission, "Reactor Risk Reference Document", Office of Nuclear Regulatory Research, NUREG-1150, Vol.1-3, 1987.