論 文 12-4-30

확률론적 안전성평가를 위한 일반 기기 신뢰도 데이타 베이스 구축 절차와 적용

Development Procedure of Generic Component Reliability Data Base in PSA and Its Application

황미정* · 김길유* · 임태진** · 정원대* · 김태운*
M.J. Hwang · K.Y. Kim · T.J. Lim · W.D. Jung · T.W. Kim
(1997년 3월 14일 접수, 1997년 12월 12일 채택)

ABSTRACT

This paper presents the development procedure and application of the generic component reliability data base considering the dependency among dependent generic compendia in NPPs (Nuclear Power Plants) PSA (Probabilistic Safety Assessment) under construction or without operating history. We use MPRDP (Multi-Purpose Reliability Data Processor) code[1~2] developed in KAERI (Korea Atomic Energy Research Institute) based on a PEB (Parametric Empirical Bayesian) procedure to estimate the reliability. The employed model in this study accounts for the relative credibility as well as the dependency among generic estimates. Numerical examples and the part of summarized reliability data table are provided as the application.

국 문 요 약

건설중이거나 기기 이력이 부족한 원자력 발전소에 대한 확률론적 안전성평가에 사용되는 일반 기기 신뢰도 데이타를 기 개발된 일반 데이타 및 발전소 데이타를 취합하여 구한다. 이를 위해 본 논문에서 사용한 계산 Code는 모수적 선험적 베이지안 방법에 근거하여 3단계 베이지안 방법[5] 으로 한국원자력연구소에서 개발한 MPRDP Code이다. 일반 자료원에서 주로 자료를 취합하였으므로 각 문헌들 사이에 존재할 수 있는 종속성을 고려하여 Code에서 처리하였다. 본 논문에서는 결과로 얻어진 기기 신뢰도 자료표의 일부분을 보여준다.

* 한국원자력연구소 종합안전평가팀
** 숭실대학교 산업공학과

1. INTRODUCTION

In NPPs PSA, the component reliability data are the most basic information. However, in assessing the safety of the plant under construction or without operating history, it is essential to develop a GDB (Generic Data Base) for component reliability since the generic distributions should be used directly to quantify the system logic models.

The component reliability data of NPPs can be classified into plant-specific data and generic data from similar plants. The generic data can further be divided into generic plant data (E_1) and generic book data (E_2). The generic data from various industry compendia are usually provided in the set of parameter estimates. We postulate that the generic estimates are statistics of raw failure data from several plants. While most generic plant data are confidential, a lot of book type data are available to get them, so maximum use of such data is desired. However, dependency may occur among the estimates when the failure data referred in two or more sources are overlapped or when the same expert joins two or more projects for developing generic data bases.

The employed estimation procedure[3] is based on the PEB framework. Each generic estimate is converted into a set of location and dispersion parameters of the PVD (Population Variability Distribution). The underlying distribution for the failure rate was assumed to be log-normal according to PSA convention. The credibility of the generic estimate is appraised by the number of plant data referred by the source, and the dependency between two sources is measured by the relative number of plant data overlapped. By this approach, we can model the inherent dependency between the sub-populations considered by the generic sources. First, we derive a simple formula for deriving the PEB estimate utilizing the whole set of generic data under independence assumption. Then we induce an efficient estimate in weighted average of functions of generic estimates, which can be used for dependent data.

The data sources were screened and the necessary data used in this study were recorded in the forms of the generic reliability data collection sheets as shown in section 5. The MPRDP code which is based on the 'three-stage' Bayesian procedure automatically calculated the resulting estimates.

2. CONSTRUCTION OF GDB

Based on the literature survey performed so far, we constructed the GDB. This data base is much more enhanced than the previous ones, in that it includes bigger population of sources especially from those of the western countries outside the U.S. In case of using the diverse data sources, however, it needs to consider the dependency among the referred sources for the dependency may occur among the estimates.

In this study, we estimated the component reliability for about 100 of components including most of the major components considered in NPPs PSA. The total number of sources referred in this study is 31 as shown in Table 1, which include most of the reliability data bases considered by PSA studies performed so far. The 8 generic plant data from ALWR URD D/B[4] were also included as a set of book data. More plant data will be utilized whenever they are available.

Comparing with other previous GDBs, our GDB is developed by an automatic procedure, and reflects the dependency among the data sources employed. It provides the information of the updated distribution in the form of point estimates and confidence limits. The resulting distribution may not be log-normal, but they have similar shapes in most cases. The input data of system quantification are provided in the form of the mean and the error factor of the fitted log-normal distribution.

Table 1 Data source List

1. WASH-1400	17. IEEE Std 500
2. German-IAEA	18. Midland PSA
3. NUREG/CR-1205, R0	19. Oconee PRA
4. NUREG/CR-1331	20. NUREG/CR-1740
5. NUREG/CR-1362	21. Swedish NPP DB
6. Zion PSS	22. NUREG/CR-3831
7. NUREG/CR-1205, R1	23. NUREG/CR-2815
8. EPRI-NP-2433	24. NUREG/CR-4550, V3
9. Sizewell B SR	25. NUREG/CR-4550, R0
10. NREP Guide D/B	26. HWR-IAEA
11. NUREG/CR-1363, R1	27. Old PWR-IAEA
12. NUREG/CR-2886	28. NUREG/CR-4550, R1
13. NUREG/CR-2728	29. French-900 PSA
14. Shoreham PRA	30. NUREG/CR-4639
15. Millstone PSS	31. ALWR PSA KAG
16. Seabrook PSS	

3. DESCRIPTION OF MPRDP CODE

3.1 Three-stage Bayesian Procedure[5]

The Bayes' theorem is one of the most convenient ways to incorporate currently available pieces of information for updating our prior knowledge about the quantity of interest. The two-stage procedure, the first systematic procedure for analyzing failure data, was proposed by Kaplan[6] and frequently is used in many PSA studies. However, the 'two-stage' procedure can not handle generic book type data since it was mainly for processing data from generic plants' experience and plant specific experience. Later, Mosleh and Apostolakis[7,8] modified the 'two-stage' Bayesian procedure to handle generic book data as well. However, there seems to be inconsistency in their way of getting the joint likelihood function for generic plant data in the first stage, and the book data in the second stage. While the employed three-stage procedure in this study is based on the above two procedures, significant improvement were made in the proposed procedure combining generic plant experience and the generic book data.

The two-stage Bayes equation can be rewritten as follows in the three-stage procedure.

$$f(\lambda \mid E_0,E_1,E_2,E_3)=\frac{f(\lambda \mid E_0)\cdot P(E_1,E_2,E_3 \mid \lambda, E_0)}{\int_0^\infty f(\lambda \mid E_0)\cdot P(E_1,E_2,E_3 \mid \lambda, E_0)d\lambda} \quad \cdots\cdots (1)$$

where,

E_0=general engineering knowledge underlying assumptions

E_1=past information data from operating experience at similar plants

E_2=generic failure rate estimates or distributions contained in various industry compendia

E_3=plant specific performance data

The DPD (Discretized Probability Distribution) version of the above formula would be as follows.

$$P(\lambda \in C_\ell \mid E_0,E_1,E_2,E_3)\cong\frac{P(\lambda \in C_\ell \mid E_0)\cdot P(E_1, E_2, E_3 \mid \lambda, E_0)}{\Sigma_{\ell=1}^{L} P(\lambda \in C_\ell \mid E_0)\cdot P(E_1, E_2, E_3 \mid \lambda_\ell, E_0)} \quad \cdots\cdots (2)$$

In the 'three-stage' Bayesian data update, the first stage concentrates on the estimation for the parameters of the posterior distribution of λ. The set of parameter estimates takes the role of a generic compendium, E_2. If other estimates (generic book data) are not available, the estimates from the first stage can be directly used as a prior in the third stage.

Likewise, even when information type E_1 is not available, the second stage can be performed with other estimates to determine a prior generic failure rate distribution. In the third stage, the prior thus obtained is updated using equation (2). Fig. 1 shows the concept of 'Three-stage' Bayesian procedure.

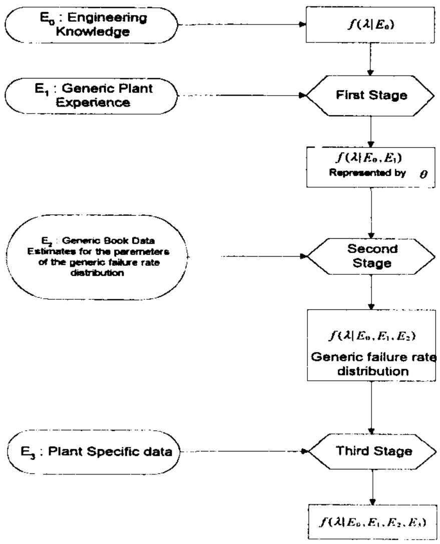

Fig. 1 The Flow Diagram of the 'Three-Stage' Bayesian Procedure

3.2 Aggregation

The generic procedure for combining generic plant data (E_1) and generic book data (E_2) can be summarized as the following two steps.

- ▶ Step 1 : Aggregate E_1 based on the engineering knowledge (E_0) with Bayesian method to provide an estimation for the PVD.
- ▶ Step 2 : Regard the estimate in step 1 as an item of generic book data, and combine it with the other book data to compute the final estimate for the PVD.

There is little problem in performing step-1. The likelihood function for the plant data can be naturally constructed, for they are mutually independent. However, performing step-2 requires much more consideration. First each set of book data is a collection of plant data, and the same plant data considered in two different books can be overlapped. This might cause significant dependency among book data. Second, the way of aggregation methods from raw data are not known, so we need some consistent assumption for the formulation of generic book data.

The dependency may stem from the overlap of experts, but the more basic sources would be duplicated use of the same raw data. When we consider the scarcity of detailed operating history, the minimum requirement for estimating the dependency would be the number of plants (N_i) referred by each book source and the number of overlapped plants (N_{ij}) between two sources.

First we reduce the problem to estimating parameters, by restricting the PVD to a member of a parametric family such as log-normal family.

The type of book estimates can also be simplified into a closed form, provided we restrict the aggregation method within the type II maximum likelihood (ML-II) method.

ML-II estimation method is a kind of PEB method. The basic idea is to select the prior π maximizing the type II likelihood function $m(x \mid \pi)$. When a prior π is parametric distribution, the ML-II approach provides a set of hyper-parameters maximizing the type II likelihood function. The determination of an ML-II prior is quite simple for many classes of priors.

To aggregate the data, the following assumptions are required.

1) The PVD estimates from the generic compendia are based on ML-II method under the log-normal model.
2) The estimation error of the individual plant estimate Ym has the same variance.($\overline{\sigma}^2_m=$

$\bar{\sigma}^2$)

Suppose that all the book data estimates are independent. Then a basic way of aggregating the book data to provide an estimate for the PVD can be derive from the lemma.

[Lemma 1] Suppose assumption 1 and 2 hold, and there is no overlap of plant data among the book data. Then the aggregated ML-II estimate $(\hat{\xi}, \hat{\sigma}^2)$ can be derived from the book data, $\{(\hat{\xi}_i, \hat{\sigma}_i^2);\ i=1, B\}$, by the following equation.

$$\hat{\xi} = \frac{1}{N}\sum_{i=1}^{B}(N_i\ \hat{\xi}_i) \quad \cdots\cdots (3)$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{B}\{N_i(\hat{\sigma}_i^2 + (\hat{\xi}_i - \hat{\xi})^2\} \quad \cdots\cdots (4)$$

The result is very useful, for equation (4) does not require the estimate of variance $\bar{\sigma}^2$ of individual plant data. Equation (3) and (4) also imply that $\hat{\xi}$ and $\hat{\sigma}^2$ are weighted sums of $\hat{\xi}_i$ *s*, and $\{\hat{\sigma}_i^2 + (\hat{\xi}_i - \hat{\xi})^2\}$ *' s*, respectively. The coefficient for the ith book data is proportional to the sample size *Ni*. This weighting scheme minimizes the variance of the estimators. Here we propose a minimum variance linear unbiased estimator for the case when the book estimates are dependent.

[Lemma 2] Suppose assumption 1 and 2 hold, and some plant data referred in some generic compendia overlap. Then the minimum variance linear unbiased estimator for the location parameter is given by estimation (5), and it's variance reduces to equation (6).

$$\hat{\xi}_L = (1^T \Sigma_0^{-1}\hat{\xi})/(1^T \Sigma_0^{-1}1), \quad \cdots\cdots (5)$$

where $1 \equiv (1,\ldots,1)^T$, $\hat{\xi} = (\hat{\xi}, \hat{\xi}, \ldots \hat{\xi}_B)^T$,

and

$$\Sigma_0 = \begin{bmatrix} 1/N_1 N_{12}/(N_1N_2)\ldots N_{1B}/(N_1N_B) \\ N_{12}/(N_1N_2) 1/N_2 \ldots N_{2B}/(N_2N_B) \\ \cdot\ \cdot\ \cdot \\ N_{1B}/(N_1N_B) N_{2B}/(N_2N_B)\ldots 1/N_B \end{bmatrix}$$

$$Var(\hat{\xi}_L) = (\sigma^2 + \sigma^{-2})/(1^T \Sigma_0^{-1}1) \quad \cdots\cdots (6)$$

When there exist book I and j such that $N_i = N_j = N_{ij}$, the matrix Σ_0 turns out to be singular. Since $N_i = N_j = N_{ij}$ implies perfect overlap between two sources, one of them should be removed from the data set.

[Remark] Equation (5) of the above lemma is identical with the generalized least square estimate of the following linear model.

$\hat{\xi} = \xi 1 + u$, $u \sim (0, \Sigma_0)$

Hence the Gauss-Markov theorem also implies that the estimator given in equation (5) is the best linear unbiased estimator.

In order to find similar type of estimator for the scale parameter, let's consider the type of the estimator $\hat{\sigma}^2$ to be linear combination of $\{\hat{\sigma}_i^2 + (\bar{y}_i - \bar{\bar{y}})^2\}$ ' s. Then we have the following theorem.

[Theorem 1] Suppose assumption 1 and 2 hold, and some plant data referred in some generic compendia overlap. Suppose we restrict the estimator of the scale parameter to be linear combinations of $w_i = \hat{\sigma}_i^2 + (\bar{y}_i - \bar{\bar{y}})^2$ ' s. Then the minimum variance estimator among them is given by equation (7), and it's variance reduces to equation (8).

$$\hat{\sigma}_L^2 = (1^T \Sigma_0^{-1} w) / (1^T \Sigma_0^{-1} 1), \quad \cdots\cdots (7)$$

where $1 \equiv (1,\cdots,1)^T$, $w = (w_1, w_2, \cdots w_B)^T$.

$$Var(\hat{\sigma}_L^2) = 2(\sigma^2 + \bar{\sigma}^2)^2 \times \{(1-2/N)/(1^T \Sigma_0^{-1}1) + 1/N^2\} \quad \cdots\cdots (8)$$

3.3 The Structure of MPRDP Code

MPRDP code was developed to provide wide range of options, and so it is composed of various subroutines. According to the input options, only a part of these subroutines are to be executed. The flow diagram of the MPRDP is in Fig. 2. In Fig. 2, MLE is maximum likelihood estimates and BE is best esimates.

4. APPLICATION

4.1 Motor Operated Valve (MOV) Demand on Failure

The generic data for the demand on failure of the MOV from 21 compendium and 6 plants are summarized in Table 2. The number of plant data used by each source and the list of dependent sources estimated are given in the table.

We provide the coefficient vector $p_i(I)$ under assumption of independence as well as $p_i(D)$ under dependence. For this case, the parameter estimates do not change significantly by considering dependency, because the generic data are relatively homogeneous.

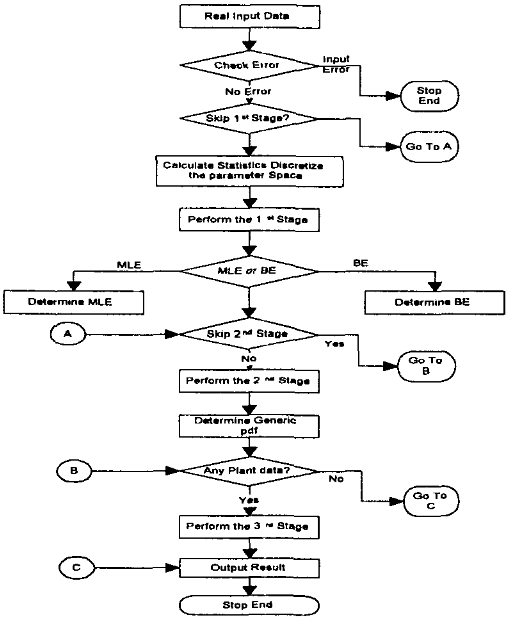

Fig. 2 The Flow Diagram of MPRDP

4.2 Solenoid Operated Valve (SOV) Demand on Failure

The generic data for the demand on failure of the SOV from 14 compendium are summarized in Table 3.

Some of the coefficient $p_i(D)$ became negative, which implies that the plant data in the ith book are referred too much in other sources, so the impact should be removed. For this case, by considering dependency, the parameter estimates for the EF change significantly.

Table 2 MOV Demand on Failure Data

No	date	mean (10^{-3}/ry)	EF	$\hat{\xi}_i$	$\hat{\sigma}_i^2$	N_i	Dependent Sources	P_i (I)	P_i (D)
1	75.10	1.25	3.0	-6.91	0.45	20	2,6,9,10,16, 25,28,30	0.02	0.00
2	79	6.1	3.0	-5.32	0.45	22	1	0.02	0.03
6	81.12	1.6	4.7	-6.88	0.88	30	1,16,25,28, 30,31	0.03	0.03
9	82.06	0.67	6.0	-7.90	1.19	25	1	0.02	0.04
10	82.06	1.0	10.0	-7.89	1.96	60	1,13	0.06	-0.04
11	82.10	4.2	1.1	-5.47	0.00	66	1,15,16,17, 25,28,31	0.06	0.08
13	83.01	3.0	10.0	-6.79	1.96	60	10,23,25,28	0.06	0.25
15	83.08	2.6	8.7	-6.82	1.74	50	1,17,30,31	0.05	0.06
16	83.12	4.3	3.7	-5.77	0.64	50	1,6,11,17,31	0.05	0.06
17	84.01	4.0	10.0	-6.50	1.96	70	11,25,28,31	0.07	0.09
18	84.05	7.0	3.0	-5.18	0.45	25	6,25,28	0.02	0.03
19	84.06	4.0	9.5	-6.46	1.88	20	1,6,25,28, 30,31	0.02	0.02
21	85	7.2	5.8	-5.50	1.14	8	None	0.01	0.01
23	85.08	3.6	15.8	-7.03	2.82	65	13,31	0.06	-0.15
25	87.09	3.0	10.0	-6.79	1.96	80	28,31	0.07	-0.05
26	88	0.79	1.2	-7.15	0.01	30	N/A	0.03	0.04
27	88	2.9	1.5	-5.87	0.06	30	N/A	0.03	0.04
28	90.01	3.0	10.0	-6.79	1.96	90	Most	0.08	0.13
29	90.04	3.0	10.0	-9.09	1.96	44	None	0.04	0.06
30	90.05	6.15	1.6	-5.13	0.08	120	Most	0.11	0.16
31	91.11	4.0	4.7	-5.96	0.88	100	6,11,16,19, 25,28	0.09	0.08
32		3.76	1.9	-5.66	0.15	6		0.01	0.01
Independence		4.50	11.1	-6.48	2.14	1071			
Dependence		4.84	9.0	-6.22	1.78				

5. RESULT

The output form of the MPRDP code is shown in Fig. 3. We evaluated mainly the safety-related components in this study. As the result we get the mean value and the error factor for about 100 the components. Table 4 is the part of the result of this study. The full set of the result will be published forward. In this table, the reliability and the error factor is compared with that not considering the dependency among the referred sources.

Table 3 SOV Demand Failure Data

No	date	mean (10^{-3}/ry)	EF	$\hat{\xi}_i$	$\hat{\sigma}_i^2$	N_i	Dependent Sources	P_i (I)	P_i (D)
1	75.10	1.25	3.0	-6.91	0.45	20	2,10,16,20, 25	0.04	0.02
2	79	1.3	20.0	-6.00	3.32	22	1	0.04	0.05
10	82.06	1.0	3.0	-7.13	0.45	60	1,110,23,253	0.11	-0.07
13	83.01	1.0	3.0	-7.13	0.45	60	1,6,11,17	0.11	0.45
16	83.12	2.4	9.5	-6.97	1.87	50	6,25	0.09	0.13
18	84.05	2.4	9.5	-6.97	1.87	25	1,6,25	0.04	0.06
19	84.04	0.013	12.7	-12.4	2.39	20	1,6,25	0.04	0.05
20	84.07	1.7	19.1	-7.99	3.22	70	1,25	0.12	0.17
21	85	0.25	5.2	-8.80	1.01	8	None	0.01	0.02
23	85.08	0.72	3.5	-7.53	0.58	65	13,24	0.11	-0.29
24	86.11	1.0	3.0	-7.13	0.45	15	13,20,23,25	0.03	0.02
25	87.09	2.0	3.0	-6.44	0.45	80	Most	0.14	0.18
27	88	1.7	2.3	-6.51	0.26	30	N/A	0.05	0.08
29	9.04	0.23	2.7	-8.56	0.36	44	None	0.08	0.12
Independence		2.17	13.7	-7.40	2.53	569			
Dependence		3.28	19.5	-7.35	3.27				

Component Name : MOTOR OPERATED VALVE
Code : MV

Boundary : Include the valve body, all its internal ...

Failure Mode : Failure to operate on demand
Code : O (Fail to Open), C (Fail to Close)
Type (*) : 0

Failure Data

No	Mean	EF	SIG	L-Mean	SIG-TR	WEIGHT	SOURCE
1	1.25e-3	3.0	.67	-6.91	.92	.00	WASH-1400
2	6.10e-3	3.0	.67	-5.32	1.25	.03	German-IAEA
6	1.60e-3	4.7	4.7	-6.88	1.32	.03	Zion PSS
9	6.70e-4	6.0	6.0	-7.90	4.02	.04	Sizewell B SR
...	...	...	...	...	...	...	...

Value Calculated : Mean=4.84e-3; EF=9.0; Correl. No.=

Rationale :

Fig. 3 MPRDP Code Output Form

6. CONCLUSIONS

The GDB developed in this study was constructed by updating the data provided by sources. The data updating work was performed by the MPRDP code.

Table 4 Summary of Component Reliability Data Base

Component	Failure Mode	Code	Dependent		Independent	
			Mean	E.F	Mean	E.F.
Motor operated valve	fail to open	MVO	4.84e-3	9.0	2.74e-3	3.7
	fail to close	MVC	4.84e-3	9.0	2.74e-3	3.7
	fail remain open	MVT, MVP	1.70e-7	10.9	1.05e-7	3.2
	catastrophic internal leakage	MVL	5.36e-7	39.6	8.51e-7	66.3
Solenoid operated Valve	fail to open	LVO	2.11e-3	9.4	9.37e-4	5.0
	fail to close	LVC	2.11e-3	9.4	9.37e-4	5.0
	transfer closed	LVT	5.94e-7	10.9	3.54e-7	4.6
Air operated Valve	fail to open	AVO	2.17e-3	5.0	1.11e-3	2.4
	fail to close	AVC	2.17e-3	5.0	1.11e-3	2.4
	transfer closed	AVT	4.11e-7	10.5	3.01e-7	6.0
Check valve (other than stop check)	fail to open	CVO	6.50e-4	12.5	4.57e-4	8.4
	fail to close	CVC	6.50e-4	12.5	4.57e-4	8.4
	transfer closed	CVT	3.80e-8	8.4	9.44e-8	8.9
	reverse leakage	CVL	1.66e-6	19.4	2.00e-6	21.5
...	...	...	...	...	...	...

The advantage of this database is that it incorporates huge population of data sources and it is based on systematic and consistent procedure. Specially, we considered the dependency among the data sources employed.

By applying the method to some set of data, we can indicate when the generic estimates are relatively homogeneous, dependency does not affect the location or dispersion parameter estimates very much. However, when there are heterogeneous generic estimates, dependency affects the results very much especially on the dispersion parameter.

As the directions for further study, generic data collection program for both foreign and domestic data should be constructed. The development of database system with appropriate data collection sheet would be major part of this project.

REFERENCES

1) T.J. Lim, C.K. Park, J.H. Park, and T.W. Kim, Survey and comparison of generic component reliability data and establishment of preliminary generic data base for us in PSA, KAERI/TR-364, May 1993.
2) Park, Chang K. and T.J. Lim, Introduction to MPRDP, a reliability data processing code, Proceeding. 9th Pacific Basin Nuclear Conference, pp. 685~692, May 1994.
3) T.J., Lim, Estimating the population variability distribution using dependent estimates from generic sources, The Korean OR/MS Society, pp. 43~59, December 1995.
4) ALWR KAG, Advanced light water reactor utility requirements documents, Chapter 1, Appendix A. PRA key assumptions and groundrules, EPRI, December 1992.
5) T.J. Lim, and C.K. Park, "Enhancement of the Bayesian Procedure for Developing a Reliability Data Base," Proc. of PSA Int'l Topical Meeting, Florida, 1993.
6) S. Kaplan, "On a Two-Stage Bayesian Procedure for Determining Failure Rates from Experimental Data," IEEE Trans. on Power Apparatus and Systems, Vol. PAS-102, No. 1, pp. 195~202, 1983.
7) A. Mosleh, and G. Apostolakis, "The Development of a Generic Data Base for Failure Rates," ANS/ENS Int'l Topical Meeting on Probabilistic Safety Methods and Applications, 1985.
8) A. Mosleh, "Bayesian Modeling of Expert-to-expert Variability and Dependence in Estimating Rare Event Frequencies," Reliability Eng. and System Safety, Vol. 38, pp. 47~57, 1992.
9) J.O. Berger, Statistical Decision Theory and Bayesian Analysis, 2nd ed., p. 608, Springer-Verlag, New York, 1985.
10) WASH-1400, Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants, USNRC, 1975.
11) Hubble and Miller, NUREG/CR-1363, Data Summaries of Licensee Event Reports of Valves at U.S. Commercial Nuclear Power Plants, rev. 1, EG & G, 1982.
12) NUREG/CR-4550, Analysis of Core Damage Frequency from Internal Events: Methodology Guidelines, SNL, 1987.
13) Carlson, NUREG/CR-2728, Interlim Reliability Evaluation Program Procedure Guide, SNL, 1983.