

베이지안기법에 의한 임무 신뢰도 예측[†]

전치혁* · 양희중** · 장수영* · 정의승* · 임혜란*

Mission Reliability Prediction Using Bayesian Approach[†]

C.H. Jun*, H.J. Yang**, S.Y. Chang*, E.S. Jung* and H.R. Lim*

Abstract

A Bayesian approach is proposed in estimating the mission failure rates by criticalities. A mission failure which occurs according to a Poisson process with unknown rate is assumed to be classified as one of the criticality levels with an unknown probability. We employ the Gamma prior for the mission failure rate and the Dirichlet prior for the criticality probabilities. Posterior distributions of the mission failure rate by criticalities and predictive distributions of the time to failure are derived.

1. 서 론

임무신뢰도(mission reliability)란 어떤 시스템이 이의 특정한 임무를 수행할 확률로 정의된다. 현실문제에서는 단순히 시스템의 작동 여부에만 관심을 국한시키지 않고 시스템의 임무를 몇가지로 구분하여 이러한 임무의 수행에 대한 신뢰도에 관심을 두는 경우가 많다. 예를 들어 전차(combat tank)의 경우 주요 임무로

서 기동, 사격등으로 구분하여 이러한 임무에 대한 신뢰도를 산출하고자 한다. 그러나 Kim [4]에서 언급하였듯이 임무신뢰도를 실제 산출하는데에는 여러 어려움이 있다. 특히, 임무수행 과정을 구조적분석을 통하여 신뢰도블럭선도(reliability block diagram) 또는 고장나무(fault tree) 등으로 표현하려 할 때, 실제의 많은 시스템은 이러한 표현이 어렵기 때문에 문제에 봉착하게 된다.

한편, 임무수행의 결과를 성공과 실패의 두

† 본 연구는 국방과학연구소의 장기기초연구에 의해 지원되었음.

* 포항공과대학 산업공학과

** 제주대학교 산업공학과

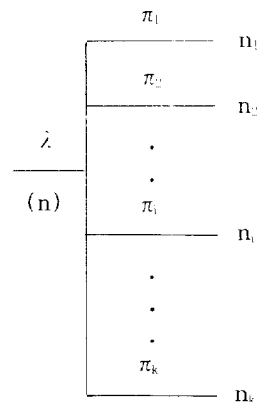
가지 만으로 분류하기는 부족한 경우가 많다. 예를 들어, 전차의 사격임무를 보면 거리측정 장치가 작동하지 않는 경우에도 사격임무가 완전히 실패하지는 않으며, 기동임무에서 속도계의 결함이 전체 임무 수행에 절대적 영향을 주지 않을 수 있다. 따라서 다소 기능이 저하된 상태의 임무 수행을 감안하는 신뢰도의 개념이 필요하다. Tillman 등[6]의 pseudo-reliability, 미군 전차에서 사용되는 임무 신뢰도 지수 (Mission Reliability Factor : MRF) [2] 등이 이와 관련된 것이다.

본 연구에서는 성능별 임무신뢰도를 예측하기 위해 임무의 고장을 치명도별로 구분하여 event tree를 작성하고 베이지안 기법을 적용하고자 한다. 즉, 여기서 산출되는 것은 임무의 치명도별 고장율의 분포 및 고장까지의 시간에 대한 예측분포(predictive distribution)이다. 베이지안 기법의 장점으로는 첫째, 자료가 불충분한 경우에도 전문가의 지식을 반영할 수 있기 때문에 효율적으로 적용할 수 있다. 둘째, 신뢰도 예측시 평균 고장률, 혹은 고장까지의 평균 시간등과 같은 점추정치(point estimate)가 아니라 이와 관련된 분포를 구함으로써 추후 고장 발생에 대한 보다 충분한 분석을 할 수 있다. 재래식 통계 방식에서는 주로 평균 고장률(λ)을 구하는데 중점을 두고 있으며 이는 평균 고장률의 역수로서 고장까지의 평균 시간($E[X]$)을 나타낸다는 것을 전제로 하고 있다. 그러나 특수한 분포 외에는 이것이 성립하지 않는다. 일반적으로 고장률의 분포는 그 빈도가 고장률이 낮은쪽에 집중되어 있고, 고장률이 높은쪽으로는 긴 꼬리(tail)를 갖는 분포를 갖기 때문에 위의 가정에 의한 예측이 들어 맞기가 어렵다. 따라서 고장률 추정은 예측의 최종 목표라 할 수 없다. 또한 고장까지의

평균 시간을 예측하는 경우도 평균고장률을 추정할 후 이에 대한 조건부 평균시간 $E[X|\lambda]$ 을 구하고 있으므로 관측 불가능한 파라미터에 의존한 예측이 되고 있다. 이 보다는 관측 불가능한 파라미터 λ 를 적분 소거 함으로써 관측 가능한 자료에만 의존한 예측이 될수 있는 $E[X|Data]$ 를 구하는데 역점을 두어야 한다.

2. 임무 고장율에 대한 사전확률분포

어떤 특정한 임무에 대한 고장을 치명도별로 k개의 수준으로 분류할 수 있다고 하자. 무기체계 분석에서는 흔히 경고장, 중고장, 치명고장으로 나누고 있다. 그리고 시간 T에 대해 각 치명도별로 고장 자료를 수집할 수 있다고 가정하자. 이 경우 우리는 [그림 1]과 같은 Event Tree를 만들 수 있다. 여기서 n은 전체 고장수이며, $n_i(i=1,2,\dots,k)$ 는 수준 i에서의 고장수이다. (n과 n_i 는 T에 대한 함수로 볼 수 있으나 T가 고정되었다고 가정하여 T의 표기를 생략한다). 따라서 $n = \sum n_i$ 가 성립한다. 그리고 λ 는 임무의 고장율을 나타내며, π_i 는 수준 i의 고장이 발생할 확률이다.



[그림 1] 치명도별 임무 고장에 대한 Event Tree

임무의 고장은 포아손과정(Poisson process)에 의해 발생한다고 가정하자. 즉, 주어진 T 동안 발생한 고장수는 포아손분포를 따른다. 또한 n과 π_i 들이 주어진 경우 치명도별 고장수 벡터는 다항분포 (Multinomial Distribution)을 따른다고 가정한다. 따라서 고장수에 관한 확률변수와 데이터의 표기를 편의상 구분하지 않을 때, n과 (n_1, n_2, \dots, n_k) 에 관한 모델은 다음과 같다.

$$n | \lambda \sim \text{Poisson}(\lambda T)$$

$$(n_1, n_2, \dots, n_k) | n, \pi_i \text{'s} \sim \text{Multinomial}(n, \pi_1, \dots, \pi_k)$$

이때 λ 및 (π_1, \dots, π_k) 가 서로 독립이라 하면 각각의 켈레사전확률분포(Conjugate Prior Distribution)는 다음과 같다.[3]

$$\lambda \sim \text{Gamma}(\alpha, \beta)$$

$$(\pi_1, \dots, \pi_k) \sim \text{Dirichlet}(a_1, \dots, a_k)$$

(α, β) 와 (a_1, \dots, a_k) 는 사전확률분포의 모수로서 아는 것으로 가정하는데 전문가의 지식 이 이에 반영될 수 있다.

3. 임무 고장율에 대한 사후확률분포

베이지안 기법의 한 강점은 새로운 데이터가 관측될 때마다 기존의 분포를 새롭게 수정 (updating)할 수 있다는 것이다. 이와같이 수정된 분포를 사후확률분포 (posterior distribution)라 한다. 본 절에서는 임무고장율 λ 와 치명도 i의 고장 발생확률 π_i 들의 사후확률분포

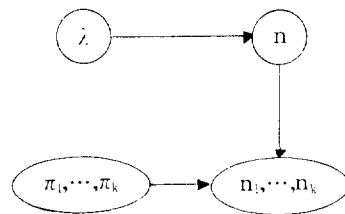
를 유도하며, 이를 바탕으로 치명도 i에서의 고장율 $\lambda_i(\lambda, \pi_i)$ 에 대한 사후확률분포를 도출하고자 한다.

일반적으로 사후확률분포는 다음과 같은 베이즈 법칙(Bayes Rule)에 의해 구해진다. 즉, 데이터 x에 따라 변수 θ 에 대한 사후확률분포는 다음과 같이 계산된다.

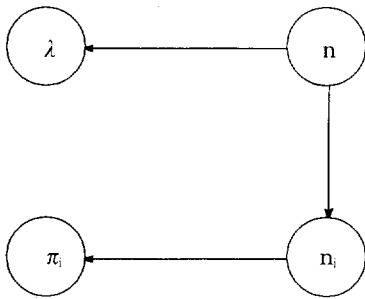
$$p(\theta | x) = p(x | \theta) p(\theta) / p(x)$$

위에서 보듯이 본 논문에서는 확률밀도함수에 대한 기호로서 편의상 모든 확률변수에 대해 p(.)를 사용하며 확률변수와 이것이 취할 수 있는 값의 표기를 구분하지 않고 혼용한다. 그러나 어떤 확률변수의 확률밀도함수인지가 불분명할 때에는 아랫첨자에 확률변수를 표기한다.

변수가 다수이고 이에 관련된 데이터의 종류가 다수있을 때 각 변수에 어떤 다른 변수와 데이터가 영향을 주는지 명확하지 않은 경우는 사후확률분포를 구하기 어려운 경우가 있다. 이러한 경우 Oliver and Yang [5]의 연구와 같이 영향도(Influence Diagram)를 사용하는 방법도 있다. [그림 2]는 [그림 1]의 Event Tree를 영향도로 표시한 것이며 Arc Reversal 과정을 통하면 [그림 3]과 같은 영향도를 얻을 수 있다. [그림 3]에서 우리는 데이터가 주어진 후 λ 와 π_i 의 사후확률분포가 서로 독립임을 알 수 있다.



[그림 2] Event Tree에 대한 영향도



[그림 3] Arc Reversal 후의 영향도

우리의 경우 T 시간동안 임무에 대한 고장이 총 n번 있고 치명도 i에 대한 고장이 n_i번 (i=1, ..., k) 있었다고 할 때, λ 및 (π₁, ..., π_k)에 대한 사후확률분포는 서로 독립이며 각각 다음과 같다.

$$\lambda | n, n_1, \dots, n_k \sim \text{Gamma}(\alpha + n, \beta + T) \quad (1)$$

$$\pi_1, \dots, \pi_k | n, n_1, \dots, n_k \sim \text{Dirichlet}(a_1 + n_1, \dots, a_k + n_k) \quad (2)$$

참고로 식(2)에서 π_i의 주변분포(marginal distribution)는 다음과 같은 베타분포(Beta Distribution)가 된다.

$$\pi_i | \text{data} \sim \text{Beta}(a_i + n_i, a_0 + n - a_i - n_i) \quad (3)$$

여기서 a₀ = ∑ a_i 이다.

다음에는 치명도별 고장율의 분포를 유도하고자 한다. 치명도 i의 고장율을 λ_i라 하면 이는 다음과 같이 표현된다.

$$\lambda_i = \lambda \pi_i \quad (i=1, \dots, k)$$

따라서 λ_i의 분포는 λ와 π_i의 결합분포(joint distribution)으로 부터 구할 수 있는데 이 결

합분포는 다음과 같이 표현된다.

$$p(\lambda, \pi_i | \text{data}) = p(\lambda | \text{data}) p(\pi_i | \text{data})$$

따라서 λ_i = x에서의 확률밀도는 다음과 같이 계산된다.

$$\begin{aligned} p_{\lambda_i}(x | \text{data}) &= \int p(x/\pi_i, \pi_i | \text{data}) (1/\pi_i) d\pi_i \\ &= \int p_{\lambda_i}(x/\pi_i | \text{data}) \cdot p(\pi_i | \text{data}) \\ &\quad (1/\pi_i) d\pi_i \end{aligned} \quad (4)$$

여기서 아랫첨자 λ_i, λ_i는 각각 i에 대한 확률밀도함수임을 의미한다.

식(1)와 (3)의 사후확률분포를 식(4)에 대입하면 λ_i의 사후확률분포는 다음과 같이 나타난다.

$$\begin{aligned} p_{\lambda_i}(x | \text{data}) &= \int \frac{(\beta + T)^{-(\alpha + n)} (x/\pi_i)^{(\alpha + n) - 1} e^{-(\beta + T)x/\pi_i}}{\Gamma(\alpha + n) B(a_i + n_i, a_0 + n - a_i - n_i) \pi_i} \\ &\quad \frac{\pi_i^{(a_i + n_i - 1)} (1 - \pi_i)^{(a_0 + n - a_i - n_i - 1)}}{d\pi_i} \\ &= K_i \int_0^1 (x/\pi_i)^{(\alpha + n) - 1} e^{-(\beta + T)x/\pi_i} \pi_i^{(a_i + n_i - 1)} \\ &\quad (1 - \pi_i)^{(a_0 + n - a_i - n_i - 1)} d\pi_i \end{aligned} \quad (5)$$

여기서 Γ(.)는 감마함수, B(...)는 베타함수이며, K_i는 다음과 같은 상수이다.

$$K_i = \frac{(\beta + T)^{-(\alpha + n)}}{\Gamma(\alpha + n) B(a_i + n_i, a_0 + n - a_i - n_i)}$$

위의 적분은 해석적으로 구해지지 않으며 수치적으로 계산하여야 한다. 수치적 적분을 위해서는 Gauss Quadrature 방법 ([7] 참조)을

이용하였다. 또한 이산적인 값 x 에 대하여 적분값을 계산하는데 x 의 간격은 λ 가 취할 수 있는 값의 구간을 설정한후 일정하게 나누어 산정할 수 있다.

식 (5)로 부터 구해진 치명도 i 의 고장율에 대한 확률밀도로 부터 백분율(percentile), 사후평균 및 사후 분산 등을 구할 수 있다. 그러나 사후평균과 사후분산은 λ 와 π_i 가 독립임을 이용하여 다음과 같이 이론적으로 구할 수 있다.

$$E[\lambda_i | data] = E[\lambda | data] \cdot E[\pi_i | data] \\ = (x+n)(a_0+n) / ((\beta+T)(a_0+n)) \quad (6)$$

$$Var[\lambda_i | data] = Var[\lambda | data] \cdot Var[\pi_i | data] \\ + (E[\lambda | data])^2 \cdot Var[\pi_i | data] \\ + Var[\lambda | data] \cdot (E[\pi_i | data])^2 \quad (7)$$

4. 임무수행시간의 예측 분포

우선 치명도별 고장을 구분하지 않고 첫 고장이 있을 때까지의 시간에 대한 예측분포(predictive distribution)를 구하고자 한다. 임무의 첫 고장 까지의 시간 y 는 고장율 λ 가 주어진 경우 이 λ 를 파라미터로 하는 지수분포를 따르므로 관측이 불가능한 λ 를 소거하면 y 에 대한 예측분포는 다음과 같다.

$$p(y | data) = \int p(y | \lambda) \cdot p(\lambda | data) \, d\lambda \\ = \int \lambda \cdot e^{-y\lambda} \cdot (\beta+T)^{x+n-1} \lambda^{x+n-1} e^{-(\beta+T)\lambda} \\ / \Gamma(x+n) \, d\lambda \\ = (x+n)(\beta+T)^{x+n} / (\beta+T+y)^{(x+n+1)} \quad (8)$$

즉, 위의 분포는 Inverse-Beta($1, x+n, \beta+T$) 분포에 해당한다. 여기서 Inverse-Beta(k, g, h) 분포의 밀도함수는 다음과 같다[1].

$$p(y) = \frac{h^k y^{k-1}}{B(k, g) (h+y)^{k+g}} \quad y > 0$$

참고로 $k=1$ 일 때 Inverse-Beta 분포는 파레토분포 (Pareto distribution)가 되므로 식 (8)은 파레토분포의 밀도함수로 간주할 수 있다.

치명도별 임무수행시간의 예측분포 역시 위와 유사한 방법으로 구할 수 있다. 치명도 i 의 첫 고장까지의 시간 y_i 역시 파라미터 λ_i 가 주어진 경우 지수분포를 따르므로 예측분포밀도함수는 다음과 같다. 이 때 λ_i 의 사후분포 $p(\lambda_i | data)$ 는 식(5)와 같다.

$$p(y_i | data) = \int \lambda_i \exp(-\lambda_i y_i) p(\lambda_i | data) d\lambda_i \quad (9)$$

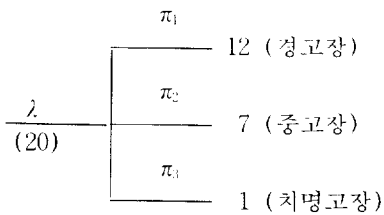
여기서 $p(\lambda_i | data)$ 는 3절에서 기술한 바와 같이 이산적 값 λ_i 에 대해 수치적으로 산출되므로 위의 적분식 (9) 또한 다음과 같이 근사적으로 계산한다. 즉, $y_i=y$ 에서의 확률밀도값은 다음과 같다.

$$p_{y_i}(y | data) = \sum_x x \exp(-xy) p_{\lambda_i}(x | data) \Delta\lambda_i \quad (10)$$

여기서 $\Delta\lambda_i$ 는 식 (5)의 평가시 λ_i 값간의 간격이다.

5. 예제

어떤 시스템의 특정 임부에 대한 시험기간 중 ($T=1$ 가정) 고장을 경고장, 중고장, 치명 고장의 세가지로 구분하고 있는데 다음과 같은 자료를 얻었다고 하자.



그리고 λ 에 대한 사전확률분포는 Gamma(2, 0, 0.1), π_1, π_2, π_3 에 대한 사전확률분포는 Dirichlet(7,2,1) 분포를 따르는 것으로 알려져 있다고 하자. 따라서 λ 에 대한 사전 기대치 및 분산은 각각 20, 200이며, 경고장, 중고장, 치명 고장 비율(확률)에 대한 사전기대치는 각각 0.7, 0.2, 0.1이다.

위와 같은 시험기간 중 자료를 얻은 후 λ 및 (π_1, π_2, π_3) 의 사후확률분포는 식(1), (2)로 부터 얻을 수 있다. 즉, λ 에 대한 사후 기대치 및 분산은 각각 20, 18.18로 바뀌게 된다. 여기서 분산이 현격히 줄어들었음을 볼 수 있다. 또한 (π_1, π_2, π_3) 의 사후분포는 Dirichlet(19,9,2)가 되므로 π_1, π_2, π_3 각각의 사후기대치는 각각 0.63, 0.30, 0.07로 바뀌게 된다.

식(5)에 따라 $\lambda_1, \lambda_2, \lambda_3$ 에 대한 사후확률분포를 구해 [그림 4]에 동시에 도시하였다. 이 들의 25%, 50%, 75%, 90%, 95%에서의 percentile을 구하면 <표 1>과 같다.

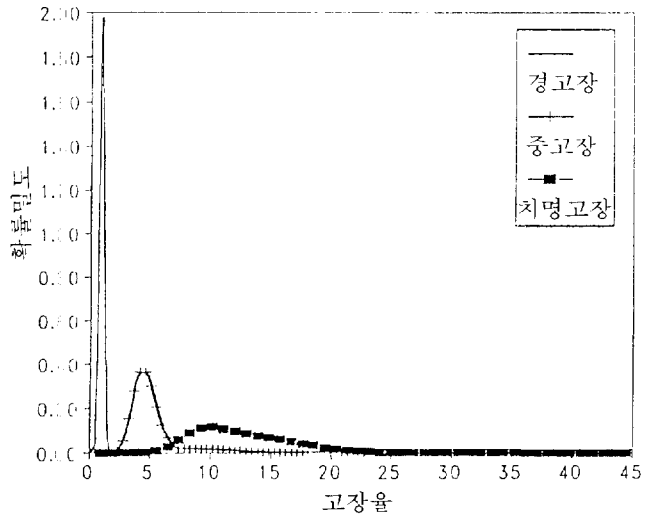


그림 4 | 치명도별 고장율의 사후분포

<표 1> 치명도별 고장율의 백분율

고장종류	λ _i	0.25	0.50	0.75	0.90	0.95
경고장	λ ₁	0.25	9.26	12.04	14.82	17.60
		0.50	12.04	14.82	17.60	19.45
		0.75	14.82	17.60	19.45	21.30
		0.90	17.60	19.45	21.30	23.15
		0.95	19.45	21.30	23.15	25.00
		0.99	21.30	23.15	25.00	26.85
중고장	λ ₂	0.25	4.17	4.63	5.56	6.95
		0.50	4.63	5.56	6.95	9.73
		0.75	5.56	6.95	9.73	14.18
		0.90	6.95	9.73	14.18	21.30
		0.95	9.73	14.18	21.30	30.00
치명고장	λ ₃	0.25	0.80	0.90	1.10	1.21
		0.50	0.90	1.10	1.21	1.41
		0.75	1.10	1.21	1.41	1.61
		0.90	1.21	1.41	1.61	1.81
		0.95	1.41	1.61	1.81	2.01

참고로 치명도별 고장율에 대한 사전 및 사후 평균과 분산을 식(6), (7)을 이용하여 구하면 <표 2>와 같다.

<표 2> 치명도별 고장율의 평균과 분산

	사 전		사 후	
	평균	분산	평균	분산
경고장율	14.0	109.5	12.6	10.4
중고장율	4.0	16.7	6.0	4.5
치명고장율	2.0	6.9	1.4	0.9

식(8)에 의한 임무 고장시간의 예측분포는 Inverse-Beta(1, 22, 1.1)로 단조감소하는 형태를 띠고 있다. [그림 5]는 경고장, 중고장, 치명고장까지의 시간에 대한 예측분포(확률밀도함수)를 나타낸 것이다. 그리고 <표3>에는 치명도별 고장까지의 시간에 대한 백분율을 나타내었다.

<표 3> 치명도별 고장시간 예측분포에 대한 백분율

경고장시간	0.25	0.023
	0.50	0.058
	0.75	0.120
	0.90	0.205
	0.95	0.274
중고장시간	0.25	0.060
	0.50	0.144
	0.75	0.296
	0.90	0.508
	0.95	0.676
치명고장시간	0.25	0.312
	0.50	0.744
	0.75	1.512
	0.90	2.568
	0.95	3.384

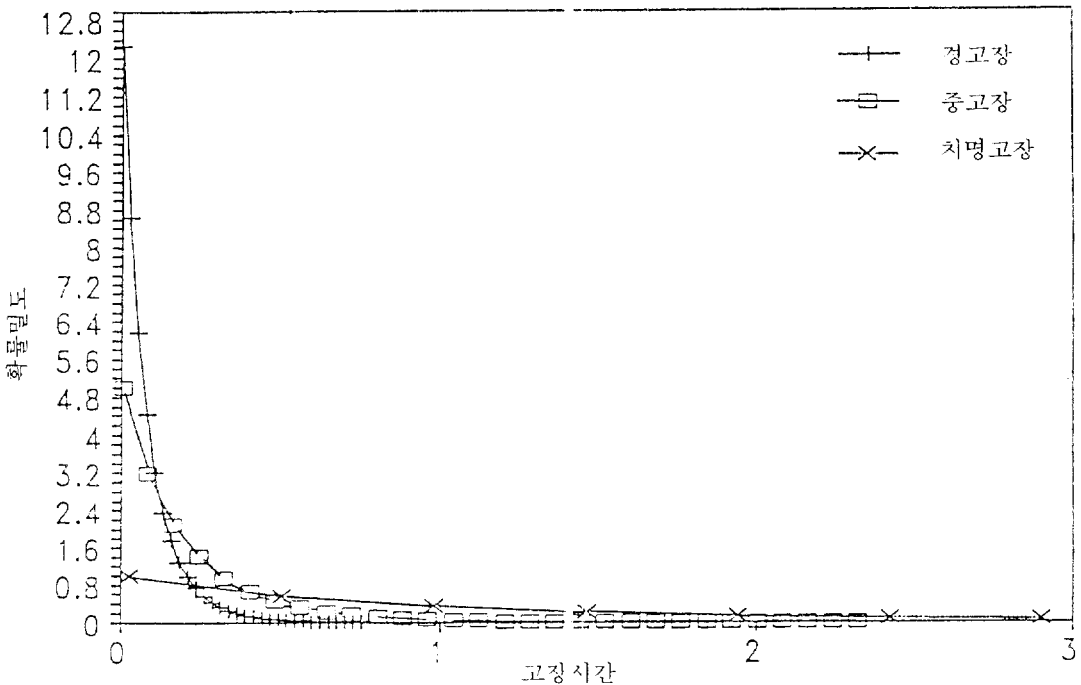


그림 5] 치명도별 고장시간의 예측분포

6. 결 론

본 논문에서는 임부의 수행 또는 고장에 대한 구조적 분석이 어려운 경우 전문가의 사전 지식과 실제로 얻게되는 고장자료를 베이지안 분석을 통하여 결합시킴으로써 치명도별 고장율, 고장까지의 시간에 대한 분포 등을 얻고자 하였다. 이러한 분포로부터 우리는 평균치에 대한 점추정치만을 추구하는 기존의 신뢰도분석과는 달리 보다 다양한 고장에 대한 정보를 얻을 수 있다. 수치적 적분을 하는 계산부담은 약간 있으나 최근의 컴퓨터 능력을 감안할 때 그다지 어려움은 아닐 것으로 사료된다.

Diagrams, Belief Nets and Decision Analysis. Chapter 12, Wiley, Chichester, 1990.

- [6] Tillman, F.A., C.H. Lie and C.L. Hwang, "Analysis of Pseudo-Reliability of a Combat Tank System and Its Optimal Design," *IEEE Transactions on Reliability*. Vol.R-25, No.4 (1976), pp.239-242.
- [7] Yakowitz, S. and F. Szidarovszky, *An Introduction to Numerical Computations*. 2nd Ed., Maxwell Macmillan, 1990.

참 고 문 헌

- [1] Aitchison, J. and I.R. Dunsmore, *Statistical Prediction Analysis*. Cambridge University Press, 1975.
- [2] Chrysler Corporation, M48A5/M60A1 Baseline Armor Reliability Test, Final RAM Report, March 1977.
- [3] DeGroot, M.H., *Optimal Statistical Decisions*. McGraw-Hill, New York, 1970.
- [4] Kim, Chul, "Analysis for Mission Reliability of a Combat Tank," *IEEE Transactions on Reliability*. Vol.38, No.2 (1989), pp. 242-245.
- [5] Oliver, R.M. and H.J. Yang, "Bayesian Updating of Event Tree Parameters to Predict High Risk Incidents," in R.M. Oliver and J.Q. Smith (ed.), *Influence*