

# 사용 현장데이터를 이용한 신뢰성 분석이론의 전개와 응용

- A Note on Theoretical Development & Applications in  
Reliability Analysis using Field Data -

김 종 곁\*

Kim Jong-Gurl

박 창 규\*\*

Park Chang-Kyu

## Abstract

Field data have been recorded as the time to failure or the number of failure of systems. We consider the time to failure and covariate variables in some pre-specified follow-up or warranty period. This paper aims to investigate study on the reliability estimation when some additional field data can be collected within-warranty period or after-warranty period. A various likelihood-based methods are outlined and examined for exponential or Weibull distribution.

## 1. 서론

우리 나라의 산업은 선진국의 기술 경쟁력과 후발국의 가격 경쟁력을 극복해야 하는 큰 도전 앞에 있으며, 지금은 성장의 저해 요소를 효율적으로 제거하고 가격, 시간, 품질, 유연성 차원에서 경쟁력을 확보할 수 있는 새로운 패러다임이 필요한 시기이다.

새로운 환경에 대응해 나가기 위한 기업의 경쟁력을 확보할 수 있는 것은 선진국의 기술장벽을 효과적으로 극복하고, 제조물 책임 등 시장요구에 대응하며 제품불량 및 고장, 안전성 미비에 따른 비용을 혁신적으로 절감할 수 있는 가장 유효한 방법이 바로 신뢰성 기술이다.

제품의 품질특성을 추정하기 위한 자료는 주로 실험실에서 행해지는 수명시험, 환경 시험 등에 의해 얻어지는 수명데이터이다. 하지만, 제품의 신뢰성, 안전성, 보전성 등의 정보를 정확하게 나타내는 자료는 실제 사용현장에서 기록된 고장데이터이다. 그러므로, 실험실에서의 수명시험은 시험환경이 실제환경과 다를 수밖에 없고, 시험을 통해

\* 성균관대학교 시스템 경영공학부 교수

\*\*성균관대학교 시스템 경영공학부 박사

얻어진 데이터는 사용현장에서의 제품 수명에 대한 정보를 왜곡하여 나타낼 위험이 있다. 따라서, 이러한 위험을 줄이고 제품수명에 대한 합리적인 정보를 얻어내기 위해서는 사용현장 데이터를 수집하고 이를 분석하는 방법에 관한 연구가 무엇보다 중요하다.

본 연구에서 사용현장 데이터를 이용한 제품의 신뢰성 추정방법에 관한 이론적 개발과 응용연구의 연구동향을 분석·비교하고 새로운 연구 영역을 모색하고자 한다. 먼저, 기존의 연구내용과 결과를 설명변수, 추적조사, 보증시점과 분석시점 등의 추가적인 정보의 유·무, 다수 고장원인 등의 측면에서 제반 신뢰성 추정방법을 분류, 비교·분석하고, 이를 바탕으로 앞으로의 이론적 연구영역과 적용분야를 도출하고자 한다.

## 2. 조사 연구 내용

사용현장 데이터를 이용하여 제품의 신뢰성을 추정하는 방법은 주로 모수적 방법과 비모수적 방법으로 수명분포의 모수를 추정한다. 제품의 수명분포는 지수분포, 대수정규분포, 와이블 분포 등이 많이 이용된다. 제품 수명분포의 모수를 추정방법은 수명분포의 우도함수(likelihood function)를 이용하거나 의사우도함수(pseudo likelihood function)를 이용한 방법이다.

본 연구에서 제시되는 기호는 다음과 같다.

$N$	총 (출하)제품 수
$\theta$	제품의 수명 모수
$T_i$	제품 $i$ 의 고장시간
$x_i$	제품 $i$ 에 대한 설명변수
$T^0$	운영시간 기준으로 정해진 제품의 보증기간
$D_1$	보증기간 동안 고장난 제품의 집합
$D_2$	보증기간 동안 고장나지 않고 추적 조사된 제품의 집합
$n_u$	보증기간 동안 고장난 제품의 수
$n_c$	추적 조사된 고장난 제품의 수
$n_l$	고장나지 않은 제품 수( = $N - n_u - n_c$ )
$p^*$	추적 조사 비율
$f(t x;\theta), F(t x;\theta), \bar{F}(t x;\theta)$	설명변수 $x$ 와 수명모수 $\theta$ 가 주어진 경우 수명시간의 확률밀도함수, 분포함수, 신뢰도 함수
$r_i$	제품 $i$ 의 고장원인
$I[\cdot]$	사건이 발생하면 1이고 그렇지 않으면 0인 지시 함수

## 2.1 설명변수가 있는 경우

사용현장 데이터의 신뢰성 분석방법에는 제품에 관한 정보가 많이 포함될수록 정확한 분석을 수행할 수 있다. 서비스센터에 들어오지 않은 제품에 관한 설명변수의 정보를 파악하는 방법으로는 전화, 우편엽서 등을 이용하여 고장 시간이나 제조 특성 및 사용 환경 등의 고장 정보를 얻는 추적조사 방법이 있다. 추적조사를 할 경우 추가적인 정보 취득으로 인한 비용의 증가와 분석의 어려움은 따르지만, 데이터를 얻지 못한 제품으로부터 정보를 얻음으로써 보다 정확한 분석이 가능해 진다.

사용현장 데이터를 이용한 신뢰성 추정방법에서 추가적인 정보를 확보할 수 있는 추적조사의 구체적인 절차 및 방법에 대한 즉, 샘플링 방법에 대한 연구[4]와 단일 고장원인이 있는 제품인 경우, 제품의 수명에 영향을 주는 제품의 제조특성, 환경 특성 등과 같은 설명변수가 존재할 때, 수명 데이터가 지수분포, 와이블 분포를 따르는 모수와 설명변수가 대수선형 관계임을 가정하고 사용현장 데이터와 총 판매 제품 중 보증기간 내에 고장이 발생하지 않은 제품의 일정 비율을 추적조사 하여 얻은 데이터를 이용하여 수명 분포의 모수를 추정하였다[12].

추적조사를 하지 않는 경우 분석방법은 서비스센터를 들어오는 제품과 그렇지 않은 제품과의 수학적 모형을 가정한 후, 서비스센터를 통하여 얻어진 정보를 이용하여 제품의 수명에 관련된 특성들을 추정하는 것이다. 고장난 제품이 서비스센터에 들어오기 까지 시간지연이 있는 경우에는 사용현장데이터를 이용하여 시간지연에 관한 정보를 추정한 후 보증기간동안의 기대고장횟수와 기대보증비용을 연구했고[13], 이들의 연구를 시간지연이 제품의 고장시간에 따라 변하는 변량모형인 경우로 확장하였다[17]. 보증기간이 달력시간이고 제품의 고장은 운영시간을 따를 때, 달력시간과 운영시간 사이에 선형관계가 존재한다고 가정하고 달력시간의 사용현장 데이터를 운영시간의 데이터로 변환한 후 수명 분포의 모수를 추정하였다[18].

제품의 제조모델, 제조장소 및 제조시간 등의 제조특성과 사용조건, 사용환경 등의 환경 특성을 나타내는 설명 변수(covariate)가 있을 때의 기본가정과 사용현장 데이터를 이용한 신뢰성 분석방법에는 총 제품 수와 설명변수에 대한 사용현장 데이터의 기록형태에 따라 세 가지 경우로 나누어진다[12].

사용현장 데이터를 이용하여 제품의 신뢰성을 추정하는 방법에 쓰이는 기본가정은 다음과 같다.

- ① 제품의 수명은 서로 독립이다.
- ② 제품의 고장시간과 보증시간의 시간 척도는 같고, 보증기간 동안 고장이 발생한 모든 제품의 고장시간과 고장원인, 설명변수 값은 기록된다.
- ③ 관측 중단된 제품 중 추적 조사되는 제품은 독립적으로 선택되며, 선택된 제품의 설명변수에 대한 정보는 정확하게 얻을 수 있다.

등의 가정을 적용하여 사용현장데이터의 신뢰성분석을 연구하였다.

### 2.1.1 총 제품 수와 보증기간 동안 고장나지 않는 제품에 대한 설명변수

가 미지인 경우

보증기간 동안에 고장난 제품의 고장시간과 설명변수를 이용한 우도 함수는

$$L_T(\theta) = \prod_{i: t_i \leq T^0} \frac{f(t_i | x_i; \theta)}{F(T^0 | x_i; \theta)} \quad (1)$$

이고, 보증기간 동안 고장난 제품의 수명 데이터를 이용한 식(1)의 우도함수를 최대화하는 모수  $\theta$ 값의 최우추정량  $\hat{\theta}$ 를 구한다[12].

### 2.1.2 총 제품 수와 모든 제품에 대한 설명변수가 기지인 경우

이 경우의 우도함수는

$$L_F(\theta) = \prod_{i: t_i \leq T^0} f(t_i | x_i; \theta) \prod_{i: t_i > T^0} \bar{F}(T^0 | x_i; \theta) \quad (2)$$

이며, 식(2)는 보증기간 내에 고장난 제품의 수명 데이터와 보증기간 내에 고장나지 않은 제품의 수명 데이터를 이용하는 우도함수로 이를 최대화하는 모수  $\theta$ 값의 최우추정량  $\hat{\theta}$ 를 구한다[12].

### 2.1.3 총 제품 수는 알고 있고, 보증기간 동안 고장나지 않은 제품에 대해서는 일정비율 $p^*$ 만큼 추적 조사하여 설명변수의 값을 알 수 있을 경우

보증기간 동안 고장나지 않고 추적 조사하는 경우의 의사우도함수

$$L_{p^*}(\theta) = \prod_{i \in D_1} f(t_i | x_i; \theta) \prod_{i \in D_2} [\bar{F}(T^0 | x_i; \theta)]^{1/p^*} \quad (3)$$

이며, 여기서 우도함수의 형태는 보증기간 동안 고장난 제품의 우도함수와 보증기간 동안 고장나지 않고 추적 조사된 제품의 우도함수를 결합한 식(3)의 우도함수를 최대화하는 모수  $\theta$ 값의 최우추정량  $\hat{\theta}$ 를 구한다[12].

## 2.2 설명변수가 없는 경우

설명변수가 없는 경우의 분석방법은 제품의 사용현장 데이터에 관측중단 데이터가 포함되어 있는 경우로, 제품의 고장원인에 따른 추정방법은 주어진 데이터의 보증기간과 운영기간에 따라 다르다.

하나의 고장원인이 존재하는 제품에 대해 보증기간은 달력시간이고 고장시간은 운영시간을 따를 때, 사용현장 데이터와 총 판매 제품 중 일정 비율을 랜덤샘플링 하고, 고장이 나지 않은 제품을 추적 조사하여 얻은 관측 중단 시간데이터를 이용하여 제품의 신뢰도를 추정하는 문제를 연구하였다[25][26]. 여러 가지 수명 데이터를 분석하는 체계적인 방법들에 관한 연구[15]와 제품의 수명 데이터를 추적 조사하여 회귀분석을 통한 신뢰성 분석방법을 연구하였다[8][9]. 사용환경에서 얻어진 수명 데이터를 포함한 일반적인 수명 데이터에 대하여, 수명 데이터 분석의 중요성과 수명 데이터 분석을 수

행할 경우 발생하는 오류를 지적했다[5][6]. 실험실에서 수명 시험을 하여 얻어진 제품에 대한 수명 데이터와 보증제도를 통해 사용환경에서 얻어지는 제품에 대한 수명 데이터를 이용하여 부품에 대한 수명시간을 추정할 연구[20]와 가속수명시험(accelerated life testing)분야에서 연구되어 오고 있는 가속수명시험 설계와 이로부터 얻어진 수명 데이터의 분석을 체계적으로 정리하였다[23]. 또한, 시간지연이 제품의 고장시간에 따라 변하는 변량모형인 경우[17]와 보증기간이 달력시간이고 제품의 고장은 운영시간을 따를 때 달력시간과 운영시간 사이에 선형관계가 존재한다고 가정하고 달력시간의 사용현장 데이터를 운영시간의 데이터로 변환한 후 수명 분포의 모수를 추정하는[18] 연구를 수리 가능한 제품에 대한 경우로 확장하여 비모수적 추정방법을 제안하였고[10], 수리 가능한 제품의 수명 분포 모수가 운영시간 및 과거의 고장횟수와 선형적인 관계가 있는 경우를 연구하였다[19].

설명변수가 없는 경우의 사용현장 데이터의 분석방법에 적용되는 기본가정은 다음과 같다.

- ① 제품의 수명은 서로 독립이다.
- ② 모든 제품에 대하여, 제품의 수명  $T$ 와 보증기간  $Y$ 는 서로 독립이며 시간 척도는 운영시간(주행거리, 빈도 등)인 반면에 연구의 관측기간(제품의 보증기간)은 달력시간(달, 년 등)으로 측정된다.
- ③ 관측기간 동안 고장난 제품의 고장시간과 고장원인은 정확히 기록된다.
- ④ 추적 조사되는 제품의 비율  $p^*$ 는 0이 아니며,  $n_c \neq 0$ 이다. 이는 일치 추정량을 얻기 위함이다.
- ⑤ 추적 조사되어지는 제품은 총 제품 중 독립적으로 선택되며, 선택된 제품 중 보증기간 내에 고장나지 않은 제품의 관측중단시간은 정확하게 얻을 수 있다.

### 2.2.1 보증기간이 달력시간인 경우

고장시간과 보증기간의 시간단위가 동일하지 않을 때 보증기간이 달력시간으로 정해져 있을 경우의 의사우도함수는

$$L_p(\theta) = \prod_{i=1}^{n_1} f(\tilde{z}_i; \theta) \left[ \prod_{j=1}^{n_2} \bar{F}(\bar{z}_j; \theta) \right]^{1 + \frac{n_1}{n_c}} \tag{4}$$

를 이용하여  $\theta$ 에 대한 최우추정량을 구할 수 있다[25][26]. 단,  $z_i = \min(T_i, Y_i)$ 이고,  $\tilde{z}_i$ 는  $T_i \leq Y_i$ 일 때의  $z_i$ 의 관측값이며,  $\bar{z}_j$ 는  $T_i > Y_i$ 이고 추적 조사된  $z_i$ 의 관측값을 나타낸다.

### 2.2.2 보증기간이 달력시간과 운영시간인 경우

보증기간이 달력시간과 운영시간의 혼합형으로 정해져 있을 경우의 의사우도함수는

$$L_p(\theta) = \prod_{i=1}^{n_{10}+n_{100}} f(\tilde{z}_i; \theta) \prod_{i=1}^{n_{11}+n_{101}} [f(\tilde{z}_i; \theta)]^{1+\frac{\widehat{n}_{100}}{n_{11}+n_{101}}} \prod_{k=1}^{n_c} [\overline{F}(z_k; \theta)]^{1+\frac{n_i}{n_c}} \quad (5)$$

와 같이된다. 단,  $\tilde{z}_i$ 는  $T_i \leq Y_i$ 일 때 운영시간의 보증기간 이후와 달력시간의 보증기간 이내에 고장이 발생한  $z_i$ 의 관측값이며,  $n_{11}$ 은 추적 조사 대상 제품 중 운영시간의 보증기간 이후 달력시간의 보증기간 이내에 고장나는 제품수,  $n_{110}$ 은 추적 조사 대상 제품 중 운영시간 및 달력시간의 보증기간 이내에 고장나는 제품수,  $n_{101}$ 은 추적 조사 대상이 아닌 제품 중 운영시간의 보증기간 이후 달력시간의 보증기간 이전에 고장이 발생하여 서비스센터를 찾는 제품수,  $n_{1010}$ 은 추적 조사 대상이 아닌 제품 중 운영시간의 보증기간 이후 달력시간의 보증기간 이전에 고장이 발생하였으나 서비스센터를 찾지 않는 제품수,  $n_{100}$ 은 추적 조사 대상이 아닌 제품 중 운영시간과 달력시간의 보증기간 이내에 고장이 발생한 제품수를 나타내며,

$$\begin{aligned} \widehat{n}_i &= \frac{N - (n_{111} + n_{110} + n_c)}{n_{111} + n_{110} + n_c} \cdot n_c, \\ \widehat{n}_{1010} &= N - n_c - \widehat{n}_i - n_{111} - n_{110} - n_{1011} - n_{100} \end{aligned} \quad (6)$$

로 주어진다.

### 2.3 다수의 고장원인이 있는 경우

다수 고장원인 모델은 생물 통계학 분야에서 연구되기 시작하여 일반 제품의 수명 데이터 분석에도 많이 응용되고 있다. 다수 고장원인이 있는 제품의 고장원인별 수명 분포가 지수 또는 와이블 분포를 따르는 경우 수명분포의 모수에 대한 최우추정치와 추정치의 분산-공분산(variance-covariance matrix)를 구했다[7]. 다수 고장원인이 있는 제품의 수명시험에 대한 연구를 체계적으로 정리한 연구[14][22]와 다수 고장원인이 있는 제품으로부터 얻어진 수명 데이터를 이용하여 신뢰성 분석을 수행하는 경우에 발생하는 문제점을 설명하고 그 해결방법을 제시하였다[5][6]. 그리고, 다수 고장원인이 있는 경우, 제품의 모든 고장시간이 기록되고, 일부 제품에 대해서는 고장원인을 조사하는 시간과 비용 때문에 정확한 고장원인을 알 수 없는 상황에서의 고장률 추정방법을 연구하였다[20]. 제품의 제조모델, 제조장소 및 제조시간 등의 제조특성과 사용조건, 사용환경 등의 환경 특성을 나타내는 설명 변수가 있을 때를 기본가정으로 한 신뢰성 추정방법[12]과 하나의 고장원인이 존재하는 제품에 대해 보증기간은 달력시간이고 고장시간은 운영시간을 따를 때, 사용현장 데이터와 총 판매 제품 중 일정 비율을 랜덤 샘플링 하고, 고장이 나지 않은 제품을 추적 조사하여 얻은 관측 중단 시간데이터로 제품의 신뢰성을 추정하는[25][26] 연구를 다수 고장원인 모델과 다수고장원인이 있고, 수리 가능한 모형으로 확장 연구하였다[1][2].

다수 고장원인 모델의 신뢰성 분석방법은 설명변수가 있는 경우와 설명변수가 없는 경우로 분류할 수 있으며, 기본가정은 2.1과 2.2절에 제시한 첫 번째 가정에서 제품의 고장원인은  $k$ 개가 있으며, 그 고장원인별로 수명은 독립이라는 것으로 바뀌어질 뿐,

나머지는 동일하게 적용한다.

설명변수가 없고 다수의 고장원인이 있는 제품의 사용현장 데이터에 대해서 고장원인별 제품의 수명분포에 관한 의사우도함수는

$$L_p(\theta) = \prod_{i=1}^{n_u} \prod_{m=1}^k [f_m(t_i; \theta) \prod_{l \neq m}^k \overline{F}_l(t_i; \theta)]^{I(r_i=m)} \prod_{j=1}^{n_c} \left[ \prod_{m=1}^k \overline{F}_m(t'_j; \theta) \right]^{1 + \frac{n_j}{n_c}} \quad (7)$$

이다[1]. 이 우도함수는 보증기간 중 고장원인  $m$ 에 의해 고장이 발생하여 고장시간  $t_i$ 를 이용한 우도함수에 대한 기여와 보증기간 내에 고장이 나지 않은 제품 중에서 추적조사를 통해 관측중단시간  $t'_j$ 를 이용한 우도함수 기여 그리고, 보증기간 내에 고장도 발생하지 않고 추적조사도 되지 않은 제품의 우도함수 기여를 이용한 이 우도함수를 최대화하는 모수  $\theta$ 값의 최우추정량  $\hat{\theta}$ 를 구한다.

설명변수가 있고 다수의 고장원인이 있는 제품의 사용현장 데이터에 대해서 고장원인별 제품의 수명분포에 관한 의사우도함수는

$$L_p(\theta) = \prod_{i \in D_1} \prod_{m=1}^k [f_m(t_i | x_i; \theta) \prod_{l \neq m}^k \overline{F}_l(t_i | x_i; \theta)]^{I(r_i=m)} \cdot \prod_{i \in D_2} \left[ \prod_{m=1}^k \overline{F}_m(T^0 | x_i; \theta) \right]^{1/p^*} \quad (8)$$

이다[1]. 이 우도함수는 2.1.3절의 총 시험 제품수는 알고 있고, 보증기간 동안 고장나지 않은 제품에 대해서는 일정비율  $p^*$ 만큼 추적 조사하여 설명변수의 값을 알 수 있을 경우에서 다수고장원인이 포함된 의사우도함수이며, 식(7)과 식(8)을 비교해보면  $1 + n_j/n_c = 1/p^*$ 로 표현되었다는 점을 제외하고는 같은 형태를 갖는다.

## 2.4 보증기간 후의 사용현장 데이터를 추가한 신뢰성 분석

제품의 제조모델, 제조장소 및 제조시간 등의 제조특성과 사용조건, 사용환경 등의 환경 특성을 나타내는 설명 변수가 있을 때를 기본가정으로 한 신뢰성 추정방법[12]과 하나의 고장원인이 존재하는 제품에 대해 보증기간은 달력시간이고 고장시간은 운영시간을 따를 때, 사용현장 데이터와 총 판매 제품 중 일정 비율을 랜덤샘플링 하고, 고장이 나지 않은 제품을 추적 조사하여 얻은 관측 중단 시간데이터로 제품의 신뢰성을 추정하는[25][26]의 연구를 확장하여 보증기간 후의 데이터가 추가 될 경우에 대한 신뢰성 분석을 연구하였다[3][24]. 이 분석에서 쓰고 있는 기본 가정은 다음과 같다.

- ① 제품의 보증기간과 고장시간은 모두 달력시간이다.
- ② 제품의 고장이 발생하여 서비스센터에 들어올 때까지의 시간은 무시할 수 있다.
- ③ 서비스센터에 들어온 제품의 고장시간은 정확히 기록된다.
- ④ 보고확률  $p_1$ 은 기지이며, 모든 제품에 대하여 동일하다.
- ⑤ 고장이 발생한 제품이 서비스센터에 들어올 확률은 제품의 고장시간이 보증기간 이내인 경우에는 1이고, 보증시점과 분석시점 사이인 경우에는 보고확률  $p_1 (< 1)$ 로

시점에 관계없이 일정하다.

위의 가정들 중 ①~③ 가정은 사용현장 데이터 분석에 사용되는 가정들이다. N개 제품에 대해 보증기간 후의 사용현장 데이터를 이용한 신뢰성 분석방법의 우도함수는 보증시점  $T_1$  이전의 시간  $t_i$ 에 고장이 발생하여 서비스센터에 들어온 제품의 우도함수 기여 그리고, 보증 시점  $T_1$ 과 분석시점  $T_2$  사이의 시간에  $t_i$ 에 고장이 발생하여 서비스센터에 들어온 제품의 우도함수 기여와 분석시점  $T_2$ 까지 서비스센터에 들어오지 않은 제품의 우도함수에 대한 기여는, 보증시점과 분석시점 사이에서 고장이 발생하였지만 서비스센터에 들어오지 않은 경우와 실제로 분석시점까지 고장이 나지 않은 경우를 모두 고려하면 다음과 같다.

$$L_F(\theta) = \prod_{i=1}^N [f(t_i; \theta)]^{I_{r \in S_1}} \cdot [p_1 f(t_i; \theta)]^{I_{r \in S_2}} \cdot [p_1 \bar{F}(T_2; \theta) + (1 - p_1) \bar{F}(T_1; \theta)]^{I_{r \in S_3}} \quad (9)$$

여기서,  $S_1$ 은 N개의 제품 중 보증시점  $T_1$  이전에 서비스센터에 들어온 제품의 집합이고,  $S_2$ 는 보증시점  $T_1$ 과 분석시점  $T_2$  사이에서 서비스센터에 들어온 제품의 집합이며,  $S_3$ 는 분석시점  $T_2$ 까지 서비스센터에 들어오지 않은 제품의 집합을 나타낸다. 식 (9)의 우도함수를 이용하여 최우 추정량과 추정량의 점근 성질을 규명하였다. 특히, 보고확률을 알고 있는 경우와 모르는 경우를 나누어 분석을 하였는데, 보고확률을 모르는 경우에는 보고확률을 추정하는 방법으로 EM 알고리즘과 확률질량함수를 이용한 추정방법을 연구하였다.

지금까지 사용현장 데이터를 이용하여 제품의 신뢰성 분석을 수행한 연구들을 정리하면 <표 1>과 같다.

### 3. 결론

지금까지 사용현장 데이터의 신뢰성 분석에 관한 연구들은 제품의 추가적인 정보의 유·무와 고장원인의 수에 따른 최우 추정량과 추정치의 점근 성질을 규명하는 연구내용들이다. 이를 정리하면 다음과 같다.

#### ① 제품의 추가적인 정보의 유·무에 관한 연구

제품의 수명 데이터, 설명변수, 더 많은 제품의 정보를 파악하기 위한 추적조사 그리고, 설명변수와 수명 데이터의 모수와의 관계가 선형식으로 주어질 때 제품의 최우 추정량을 추정하는 신뢰성 분석방법이다.

#### ② 고장원인의 수에 따른 연구

다수 고장원인을 갖는 제품인 경우, 제품의 고장원인별로 분리하여 최우 추정치를 구하는 것보다 다수고장원인 모델을 이용하여 추정치를 구하는 것이 더 정확함을 평균제곱합의 값을 통해 효율적임을 보이고 있다[12][25][26]. 따라서, 사용현장 데이터를



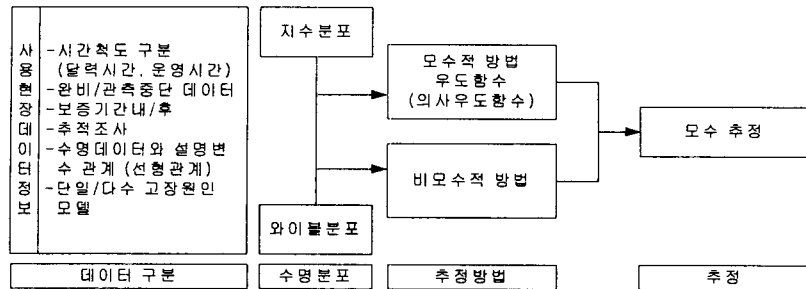
이용한 제품의 신뢰성을 분석할 경우 수명 데이터를 구성하는 정보는 수명 데이터, 설명변수, 추적조사를 통해 확보하고, 고장원인이 다수 일 경우 다수 고장원인 모델을 이용하는 것이 더 정확한 신뢰성 분석방법이다.

<표 1> 사용 현장데이터를 이용한 신뢰성 분석에 관한 연구

기존 연구	수명 분포	설명변수	추적 조사	고장 원인	내 용
Holt et al.(1980)	지수	유	유	단일 고장 원인	수명 데이터의 회귀분석방법 적용
Kalbfleisch & Prentice(1980)	지수, 와이블	유, 무	유, 무	단일 고장 원인	수명 데이터 분석방법 분류
Holt와 Scott(1981)	지수	유	유	단일 고장 원인	수명 데이터의 다수 요인들이 있을 경우 회귀분석 적용
Hahr. & Meeker(1982a)	일반적인 경우 (와이블, 대수정규)	무	무	다수 고장 원인	고장데이터 신뢰성분석
Ams'er, Brush & Saperstein(1982)		무	유		사용현장 자료의 추적조사방법
Miyamura(1982)	일반적인 경우	무	무	단일 고장 원인	보증제도와 고장데이터를 이용한 부품에 대한 수명시간 추정
Nelson(1982)	일반적인 경우	무	무	다수 고장 원인	여러 가지 고장데이터의 체계적인 분석
Lawless(1983)	지수, 대수정규	무	유		사용환경에서의 데이터 수집방법
Kalbfleisch & Lawless(1988)	지수, 와이블	유	유	단일 고장 원인	설명변수 데이터를 이용한 수명분포의 모수 추정량과 점근 분산
Suzuki(1985a,b)	지수, 와이블	무	유	단일 고장 원인	시험 제품 일부를 추적조사, 보증기간은 달력시간인 경우
Herman & Patell(1971)	지수, 와이블	무	무	다수 고장 원인	수명분포의 모수 추정
Miyakawa(1984)	일반적인 경우	무	무	다수 고장 원인	일부 고장원인이 알 수 없는 경우 수명시간 추정
Nelson(1990)	지수, 대수정규, 와이블	무	무	다수 고장 원인	가속수명시험 하에서의 수명분포 모수 추정
Bai et al.(1995a,b)	지수, 와이블	유, 무	유	다수 고장 원인	다수 고장원인 모형과 다수 고장인 경우 수리 가능한 경우의 신뢰성 분석
Kalbfleisch, J.D., Lawless, J.F. and Robinson, J.A. (1991)	지수, 와이블	유	무	단일 고장 원인	시간지연 정보를 추정 후 기대고장횟수와 기대보증비용 추정
Lawless(1994)	지수	유	무	단일 고장 원인	Kalbfleisch et al.(1991)의 연구를 제품 고장 시간에 따라 변하는 변량모형로 확장
Lawless, J.F., Huj. & Cao, J.(1995)	지수	유	무	단일 고장 원인	달력시간 사용현장데이터를 운영시간으로 변환하여 수명모수 추정
Hu, X. J. & Lawless, J.F.(1996)		유	무	단일 고장 원인	수리 가능한 제품에 대해 비모수적 추정방법 제안
Lawless, J.F. & Thiagarajah, K.(1996)	지수	유	무	단일 고장 원인	수리 가능한 제품에 대해 수명분포 모수가 운영시간 및 과거의 고장횟수와 선형적인 관계 적용
Oh and Bai.(2001)	와이블	무	무	다수 고장 원인	보증기간과 보증기간 후의 정보를 이용한 신뢰성 분석

③ 사용현장 데이터를 이용한 모수 추정방법

사용현장 데이터를 이용한 수명분포의 모수 추정방법은 <그림 1>과 같이 나타낼 수 있다.



< 그림 1 > 사용현장 데이터를 이용한 모수 추정방법

따라서, 사용현장 데이터의 신뢰성 분석방법에서 새롭게 모색될 수 있는 이론적 연구영역과 적용분야는 다음과 같다.

- ① 제품의 수명 데이터와 설명변수의 수학적 관계에 따른 신뢰성분석
- ② 추적 조사에서의 정확한 샘플링 방법
- ③ 추적 조사비용을 고려한 신뢰성 분석
- ④ 고장원인에 따른 적절한 모델의 선정방법
- ⑤ 제품의 고장원인을 정확히 모르는 마스크(masked)되어 있는 경우의 신뢰성 분석
- ⑥ 실험실 데이터와 사용현장데이터의 정보를 결합한 베이지안(Bayesian) 신뢰성 분석
- ⑦ 사용현장 데이터의 신뢰성 분석을 통한 제품안전시스템, 신뢰성경영시스템 그리고, 리스크 경영시스템의 적용성을 고려하여야 한다.

#### 4. 참고문헌

- [1] 배도선, 최인수, 황용근, “고장 원인이 여럿인 제품의 사용현장데이터 분석,” 응용통계연구, 제8권 1호, pp89-104, 1995.
- [2] 배도선, 윤형제, 최인수, “수리 가능한 제품의 사용현장데이터 분석,” 응용통계연구, 제8권 2호, pp133-145, 1995.
- [3] 배도선, 오영석, “보증기간 후의 사용현장데이터를 추가한 신뢰성 분석”, KAIST 석사논문, 1997.
- [4] Amster, S. J., Brush, G. G., and Saperstein, B., "Planning and Conducting Field-Tracking Studies." *Bell System Technical Journal*, 61, pp2333-2364, 1982.
- [5] Hahn, G. J. and Meeker, W. Q., "Pitfalls and Practical Considerations in Product Life Analysis, Part I ; Basic Concepts and Dangers of Extrapolation," *Journal of Quality Technology*, 14, pp144-152, 1982a.
- [6] Hahn, G. J. and Meeker, W. Q., "Pitfalls and Practical Considerations in Product Life Analysis, Part II; Mixtures of Product Populations and More General Models," *Journal of Quality Technology*, 14, pp177-185, 1982b.
- [7] Herman, R. J. and Patell K. N., "Maximum Likelihood for Multi-Risk Model",

- Technometrics*, 13, pp385-396, 1971.
- [8] Holt, D., and Scott, A. J., "Regression Analysis Using Survey Data," *The Statistician*, 30, 169-178, 1981.
- [9] Holt, D., Smith, T. M. F., and Winter, P. D., "Regression Analysis of Data From Complex Surveys," *Journal of the Royal Statistical Society*, Sep. A, 143, pp474-487, 1980.
- [10] Hu, X. J. and Lawless, J. F., "Estimation of Rate and Mean Functions From Truncated Recurrent Event Data," *Journal of American Statistical Association*, 91, pp300-310, 1996.
- [11] Kalbfleisch, J. D., and Lawless, J. F., "Likelihood Analysis of Multi-state Models for Disease Incidence and Mortality," *Statistics in Medicine*, 7, 149-169, 1988.
- [12] Kalbfleisch, J. D. and Lawless, J. F., "Estimation of Reliability in Field Performance Studies," *Technometrics*, 30, pp365-388, 1988.
- [13] Kalbfleisch, J. D., Lawless, J. F. and Robinson, J. A., "Methods for the Analysis and Prediction of Warranty Claims," *Technometrics*, 33, pp273-285, 1991.
- [14] Kalbfleisch, J. D. and Prentice, R. L., *The Statistical Analysis of Failure Time Data*, New York ; John Wiley, 1980.
- [15] Lawless, J. F., *Statistical Models and Methods for Lifetime Data*, New York ; John Wiley, 1982.
- [16] Lawless, J. F., "Statistical Methods in Reliability," *Technometrics*, 25, pp305-335, 1983.
- [17] Lawless, J. F., "Adjustments for reporting delays and the prediction of occurred but not reported events," *The Canadian Journal of Statistics*, 22, pp15-31, 1994.
- [18] Lawless, J. F., Hu, J. and Cao, J., "Methods for the Estimation of Failure Distributions and Rates from Automobile Warranty Data." *Lifetime Data Analysis*, 1, pp227-240, 1995.
- [19] Lawless, J. F. and Thiagarajah, K., "A Point-Process Model Incorporating Renewals and Time Trends, With Application to Repairable System," *Technometrics*, 38, pp131-138, 1996.
- [20] Miyakawa, M., "Analysis of Incomplete Data in Competing Risks Model", *IEEE Transactions on Reliability*, Vol. R-33, No. 4, pp293-296, 1984.
- [21] Miyamura, T., "Estimating Component Failure Rates From Combined Component and Systems Data : Exponentially Distributed Component Lifetimes," *Technometrics*, 24, pp313-318, 1982.
- [22] Nelson, W., *Applied Life Data Analysis*, New York ; John Wiley, 1982.

- [23] Nelson, W., *Accelerated Testing - Statistical Models, Test Plans, and Data Analyses*, New York ; John Wiley, 1990.
- [24] Oh Y., Bai D., "Field data analyses with additional after-warranty failure data", *Reliability Engineering & System Safety*, 72, pp1-8, 2001.
- [25] Suzuki, K., "Estimation of Lifetime parameters From Incomplete Field Data," *Technometrics*, 27, pp263-272, 1985.
- [26] Suzuki, K., "Nonparametric Estimation of Lifetime Distributions From a Record of Failures and Follow - Ups," *Journal of the American Statistical Association*, 80, pp68-72, 1985.

## 저 자 소 개

### 김 종 걸 :

- 서울대 학사, 석사(컴퓨터 과학 및 통계학), KAIST 박사(산업공학)
- 호주 퀸즈랜드 공과대학기계, 제조, 의공학부 초빙교수
- 산업자원부 신뢰성위원, 품질전문위원(현)
- IEC/TC56 전문위원, 한국 품질보증/PL 연구회장(현)
- 한국 신뢰성학회 부회장(현)
- 성균관대학교 시스템경영공학부 교수(품질, 신뢰성)

### 박 창 규 :

- 성균관대학교 학사, 석사, 박사(산업공학)
- 한국 품질보증/PL 연구회원
- 경기공업대학 겸임교수