IE Interfaces
Vol. 23, No. 2, pp. 176-181, June 2010.

고객의 지연보고를 고려한 보증수리내역자료에서의 고장률 추정

박종훈[1] · 김영훈[2] · 백장현[3†] · 이창훈[1]

[1]서울대학교 산업공학과 / [2]SK C&C 소프트웨어 테크널러지팀 / [3]전북대학교 산업정보시스템공학과

Estimating Failure Rate Using Warranty Claim Data with Delayed Report of Customers

J. H. Park[1] · Y. H. Kim[2] · J. H. Baek[3] · C. H. Lie[1]

[1]Department of Industrial Engineering, Seoul National University
[2]Software Technology Team, SK C&C Co., Ltd
[3]Department of Industrial and Information Systems Engineering, Chonbuk National University

Warranty claim data analysis is a useful tool for the manufacturer because it contains many useful informations regarding reliability of the product in the real-world environments. Because of the nature of uncertainty and the incompleteness of data, some bias patterns are observed on warranty claim rate known as 'spikes'. Two types of spikes are considered. One is due to manufacturing-related failures. The other is caused by customer's behavior. This paper proposes a model by considering two types of spikes. Warranty claim data is analyzed with the proposed model. To represent spikes observed on the early warranty period, we classify failures into manufacturing-related failures and usage-related failures. Uniform distribution is assumed for the time delayed to diagnose and report by customers. By reducing maximum value of the delayed time by customers, the proposed model characterizes customer's rush in the vicinity of the warranty expiration limit. Experimental results by using the real warranty claim data show that the proposed model is better than the existing one in respect to MSE(Mean Squared Error). Moreover it is expected to estimate the failure rate more realistically with proposed model because it considers the delayed time to diagnose and report by customers.

Keyword: warranty data, spikes, delayed report, failure rate estimation

1. 서 론

보증수리내역자료(warranty claim data)란 품질보증기간(warranty limit)내에서 이루어진 보증수리요청(warranty claim) 및 각 보증수리요청에 대한 원인, 수리요청자, 수리자, 수리내역, 수리시간 등의 정보를 담고 있는 자료이다. 이러한 보증수리내역자료는, 수명시험을 통해 획득한 실험 자료에 비해, 실제 고객들의 사용특성과 환경이 반영되어있으며, 충분한 수의 데이터를 확보할 수 있다는 점에서 매력적인 자료이다. 그러나 보증수리내역자료는 실험자료에 비하여 아래와 같은 특징을 가지고 있기 때문에 분석하고 활용하는데 있어서 주의를 기울여야 한다.

첫째, 보증수리내역자료의 수리내역이 그대로 고장내역을

†연락저자 : 백장현 교수, 561-756 전북 전주시 덕진동 1가 664-14 전북대학교 산업정보시스템공학과, Tel : 063-270-2330, Fax : 063-270-2333, E-mail: jbaek@chonbuk.ac.kr
투고일(2009년 09월 18일), 심사일(1차 : 2009년 11월 24일, 2차 : 2010년 02월 11일), 게재확정일(2010년 02월 15일).

의미하진 않는다. 보증수리내역자료에는 제품의 고장 시점이 기록되는 것이 아니라 사용자가 보증 수리를 요청한 시점이 기록되기 때문이다. 고장이 치명적이고 즉시 발견 될 수 있는 하드 고장(hard failure)이라면, 고장 즉시 또는 비교적 빠른 시간에 수리를 요청하여 고장시점과 요청시점에 큰 차이가 없지만, 그렇지 않은 경우, 즉 고장발생이 시스템에 치명적인 영향을 미치지 않거나, 고장이 쉽게 파악되지 않는 소프트 고장(soft failure)의 경우에는 실제고장 시점에서 수리요청까지에 일정시간의 지연이 존재한다(Iskandar and Blischke, 2003).

둘째, 보증수리내역자료는 품질보증만료에 의해서 관찰이 중지된다. 보증만료 후에도 제품들은 생존하고 동작하며 고장이 발생하지만, 보증수리내역자료에서는 이러한 부분들을 고려할 수 없다. 따라서 보증수리자료는 기간만료에 의한 절단자료(truncated data)로 취급되어 왔다(Nelson, 1990; Cohen, 1991; Meeker and Escobar, 1998). 그러나 이러한 절단자료는 데이터 분석 시 추정 되는 모수의 신뢰구간이 넓어지는 등, 상대적으로 추정의 정확성이 떨어지기 때문에 Oh and Bai(2001)는 보증기간 후에 추가적으로 수집된 자료를 추가하여 현장보증자료의 분석 수행하는 연구를 통해 이러한 단점을 보완하려 하였으며, Rai and Singh(2003)은 사용량과 사용시간의 이차원 품질보증자료에 대하여 사용기간 축적률(mileage accumulation rate)의 개념을 사용하여 사용량과 사용시간의 관계를 가정하여 추정의 정확성을 높이려는 연구를 수행하였다.

셋째, 보증수리내역자료는 사용관련고장(usage-related failure), 즉 신뢰성에 관련된 고장에 대한 자료만이 존재하지 않는다. 생산과정의 품질관리에서 걸러내지 못한 불량품이나, 제품의 운송 및 판매 과정에서 발생한 고장 등에 의해서 발생하는 초기의 고장이 포함되어 있으며 이는 보증기간 초기의 보증수리요청률의 이상증가(spikes)현상을 일으키곤 한다. Majeske(2003)은 이러한 보증기간 초기의 이상증가현상을 표현하기 위하여, 생산관련고장과 사용관련고장을 구분하여 모형화한 혼합모형(mixed model)을 소개하기도 하였다.

마지막으로 보증수리내역자료에는 고객의 심리적 요인이 반영되어 있다. 품질보증기간이 만료한 즈음에는 보증수리기간이 끝나면 수리비용이 들기 때문에 고객들이 기간이 만료되기 전에 제품의 이상을 수리하려고 하기 때문이다. 특히 소프트 고장의 경우, 수리를 하지 않더라도 시스템 전체의 작동이 불가능한 것이 아니기 때문에, 평소에는 고장을 발견 못하다가 보증수리기간이 만료될 즈음에 제품에 관심을 보이면서 발견하거나, 제품점검을 실시하여 발견하는 경우 등에 의해 보증수리기간 후반에 보증수리요청률의 이상증가현상을 야기시키며, Pal and Murthy(2003)와 Rai and Singh(2004)이 이러한 이상증가현상을 설명하기 위한 연구를 수행하였다.

이상에서 언급한 바와 같이 보증수리내역자료는 실험자료와는 차별화되는 특징이 존재하며, 이 자료를 사용하여 신뢰성 분석을 실시할 경우 이러한 특징들이 반영되어야 할 것이다. 이에 본 연구는 현장 데이터로써 활용가치가 매우 높은 보증수리내역자료를 신뢰성데이터로 활용하는 과정에서 보증수리내역자료의 특성들을 반영한 고장률 추정모형을 제안하였다. 보증기간초기에는 생산관련고장과 사용관련고장이 혼재되어 보증수리요청이 발생한다고 가정하였으며, 이 과정에서 초기의 보증수리요청 이상증가현상이 표현된다. 그리고 고장에 대한 고객의 진단 및 보고에 걸리는 지연시간은 연속일양분포(uniform distribution)를 따르며, 보증만료기간에 가까워지면 지연시간이 짧아지도록 모수를 조정하여 보증기간후반의 보증수리요청의 이상증가현상 표현 할 수 있도록 모형화하였다.

본 연구의 구성은 다음과 같다. 제 2장에서는 생산관련고장과 고객의 지연보고가 고려된 고장률 추정 모형을 소개한다. 제 3장에서는 실제 보증수리내역자료를 제안된 고장률 추정모형에 적용하고, 기존에 제안된 Majeske(2003)의 혼합모형과 그 성능을 비교하였다.

2. 품질보증자료의 특성을 고려한 고장률 추정모형

2.1 기호 및 가정

본 연구에서 사용되는 기호는 다음과 같다.

X	: 사용관련고장 시간에 대한 확률변수
Y	: 고객의 지연보고시간에 대한 확률변수
p	: 출시된 제품이 생산관련고장을 포함 할 확률
w	: 품질보증기간
τ	: 고객의 최대지연보고 시간
t	: 제품의 수명
$g(x)$	: 사용관련고장시간의 확률밀도함수
$h(y\|x)$	: 고객 지연보고시간에 대한 확률밀도함수
$F(t)$	: 보증수리요청시간에 대한 누적분포함수
$f(t)$	: 보증수리요청시간에 대한 확률밀도함수
$r(t)$	: 보증수리요청률
Θ	: 사용관련고장시간에 대한 확률밀도함수의 모수
$L(T\|\Theta, p, \tau)$	: 우도함수
N	: 총 표본수
n	: 총 보증수리요청수
t_i	: 수리요청시간
T	: 수리요청시간들의 집합, 즉 $T=\{t_1, t_2, \cdots, t_n\}$

또한 보증수리내역자료의 특성인 보증기간 초기와 말기에 나타나는 보증수리요청의 이상증가 현상을 반영하기 위한 고장률 추정모형을 수립하기 위해 다음과 같이 가정하였다.

(1) 보증기간초기에는 생산관련고장과 사용관련고장이 혼재되어 보증수리요청이 발생하며, 생산관련고장이 요청

될 확률은 p이다.

(2) 사용관련고장의 발생은 모수가 Θ인 확률밀도함수를 따른다. 즉 $X \sim g(x;\Theta)$.

(3) 지연보고시간은 연속일양분포를 따른다.

(4) 고장발생 후 τ시간 내에 고장은 반드시 발견되고 수리 요청된다.

(5) 보증기간 내에 발생한 고장은 보증기간 만료 전에 모두 수리 요청된다.

가정 (1)은 보증기간초기의 수리요청 이상증가현상을 모형화하기 위한 Majeske(2003)의 연구에서 사용된 가정을 그대로 사용하였으며, 가정 (3)~(5)은 고장발생 후 수리요청까지의 지연시간과 관련된 내용으로, 평소에는 시스템에 치명적이지 않거나 파악이 쉽지 않은 고장의 경우 실제 고장시점에서 수리요청까지의 지연시간이 존재하고 그 시간의 분포는 모수가 τ인 연속일양분포를 가정하였다. 그러나 보증기간이 만료할 즈음에는 보증수리기간이 끝나면 수리비용이 든다는 사실 때문에 고장여부를 꼼꼼히 살펴보고 보증기간 만료 전에 수리요청을 하려는 고객의 심리에 의해 지연시간이 줄어들고 만료 전에는 모두 보고가 되도록 연속일양분포의 모수가 조정이 됨을 가정하였다. 따라서 지연보고시간의 분포는 고장발생시점과 보증만료기간에 따라 다르게 정의되어야 한다. 따라서 품질보증기간이 w이고 사용관련고장시점이 x일 때 고객의 지연보고시간에 대한 확률 밀도 함수 $h(y|x)$는 다음과 같이 표현된다.

$$h(y|x)=\begin{cases}\dfrac{1}{\tau} & 0\le x<w-\tau\\ \dfrac{1}{w-x} & w-\tau\le x\le w\end{cases} \quad \text{for } 0<y\le \min(\tau,\ w-x) \tag{1}$$

2.2 고장률 추정모형

보증기간동안 발생하는 수리요청은 생산관련고장이 고객의 지연보고에 의해 요청되는 경우와 사용관련고장이 고객의 지연보고에 의해 요청되는 경우로 나뉜다. 생산관련고장에 의한 보증수리요청은 이미 고장이 발생한 상태에서 지연보고만이 발생하는 상황이므로 수리요청시간은 고객의 지연보고시간의 분포에만 영향을 받는다. 반면에 사용관련고장의 경우에는 고장발생과 고객의 지연보고의 영향을 모두 받게 되며, 수리요청시간은 고장발생시간과 지연보고시간의 합의 분포를 따른다. 따라서 보증수리요청시간에 대한 누적분포함수(cummulative distribution function)는 다음과 같다.

$$F(t)=pP[Y\le t]+(1-p)P[X+Y\le t] \tag{2}$$

보증기간동안 발생하는 수리요청은 다음의 3개의 구간에서, 아래에 기술되는바와 같이, 서로 다른 형태의 고장원인과 지연보고의 특징을 가진다. 따라서 보증수리요청시간에 대한 누적분포함수 역시 다르게 표현된다.

A : 보증기간 초기에 생산관련고장과 사용관련고장이 혼재되어 수리 요청되는 기간.

B : 모든 생산관련고장이 요청된 후, 사용관련고장만이 발생하여 일반적인 지연보고시간분포에 의해 수리 요청되는 구간.

C : 보증기간 만료에 즈음하여, 사용관련고장만이 발생하지만 고객의 심리적 요인에 의해 지연보고시간이 단축되어 수리 요청되는 구간.

A구간은 $0<t\le\tau$인 구간으로, A구간에서 발생하는 수리요청은 가정 (1)에 의하여 p의 확률로 생산관련고장이며 $1-p$의 확률로 사용관련고장이다. 생산관련고장은 $t=0$에서 이미 고장이 발생한 상황이므로 A구간동안 지연보고에 의해 나뉘어져서 수리요청이 발생하며, 가정 (4)에 의하여 A구간에서 모든 생산관련고장은 수리요청 된다. 사용관련고장에 대한 수리요청은 이전의 임의의 시점에서 고장이 발생한 후 모수가 τ인 연속일양분포를 따르는 지연보고에 의해 수리요청이 발생하므로 일부는 t시점 이전에 수리요청이 발생하고 일부는 t시점 이후에 수리요청이 발생한다. 따라서 A구간에서 발생하는 보증수리요청의 시간에 대한 누적분포함수는 다음과 같다.

$$\begin{aligned}F(t)&=pP[Y\le t]+(1-p)P[X+Y\le t]\\ &=p\int_0^t\frac{1}{\tau}dy+(1-p)\int_0^t\frac{t-x}{\tau}g(x)dx\end{aligned} \tag{3}$$

B구간은 $\tau<t\le w-\tau$인 구간으로, 생산관련고장은 A구간에서 이미 모두 수리요청이 완료되었기 때문에, B구간에서는 사용관련고장만이 모수가 τ인 연속일양분포를 따라 지연되어 수리요청이 발생한다. 이때 $t-\tau$ 이전에 발생한 사용관련고장은 t시점에서 모두 수리가 요청되지만, $t-\tau$ 이후에 발생한 사용관련고장은 일부만이 t시점 이전에 수리가 요청 되고 나머지는 그 이후에 요청된다. 따라서 B구간에서 발생하는 보증수리요청의 시간에 대한 누적분포함수는 다음과 같다.

$$\begin{aligned}F(t)&=pP[Y\le t]+(1-p)P[X+Y\le t]\\ &=p+(1-p)\left(\int_0^{t-\tau}g(x)dx+\int_{t-\tau}^{t}\frac{t-x}{\tau}g(x)dx\right)\end{aligned} \tag{4}$$

C구간은 $w-\tau<t<w$인 구간으로, 생산관련고장은 A구간에서 이미 모두 수리요청 되었으며, $t-\tau$ 이전에 발생한 사용관련고장 역시 t시점에서 모두 수리가 요청된다. 그러나 $t-\tau$부터 $w-\tau$사이에 발생하는 사용관련고장은 모수가 τ인 연속일양분포를 따라 지연되어 수리요청이 발생하므로 일부만이 t시점 이전에 수리가 요청 되고 나머지는 그 이후에 요청된다.

마지막으로 $w-\tau$ 이후에 발생하는 사용관련고장은 고객의 심리적요인에 의해 연속일양분포의 모수가 $\tau-x$인 일양분포로 조정되어 수리요청이 발생한다. 따라서 C구간에서 발생하는 보증수리요청의 시간에 대한 누적분포함수는 다음과 같다.

$$F(t)=pP[Y\le t]+(1-p)P[X+Y\le t] \quad (5)$$
$$=p+(1-p)\left(\begin{array}{l}\int_0^{t-\tau} g(x)dx \\ +\int_{t-\tau}^{w-\tau}\frac{t-x}{\tau}g(x)dx \\ +\int_{w-\tau}^{w}\frac{t-x}{\tau-x}g(x)dx\end{array}\right)$$

이상의 내용을 정리하면 보증수리요청의 시간에 대한 누적분포함수는 다음과 같다.

$$F(t)=\begin{cases} p\int_0^t\frac{1}{\tau}dy+(1-p)\int_0^t\frac{t-x}{\tau}g(x)dx, & \text{for } 0<t\le\tau \\ p+(1-p)\left(\begin{array}{l}\int_0^{t-\tau}g(x)dx \\ +\int_{t-\tau}^{t}\frac{t-x}{\tau}g(x)dx\end{array}\right), & \text{for } \tau<t\le w-\tau \\ p+(1-p)\left(\begin{array}{l}\int_0^{t-\tau}g(x)dx \\ +\int_{t-\tau}^{w-\tau}\frac{t-x}{\tau}g(x)dx \\ +\int_{w-\tau}^{w}\frac{t-x}{\tau-x}g(x)dx\end{array}\right), & \text{for } w-\tau<t\le w\end{cases} \quad (6)$$

식 (6)의 결과와 보증수리내역자료를 사용하여 최우추정법(Maximum Likelihood Estimate)으로 $g(x)$의 모수(Θ)와 생산관련 고장 확률(p) 그리고 최대지연보고시간(τ)를 추정한다.

제품의 수명자료는 보증수리내역자료에서 보증수리 요청일과 해당제품의 판매일의 차이를 계산하여 수리요청시간을 구할 수 있으며, 해당 제품의 총판매수량자료와 보증수리 요청이 발생한 제품을 비교함으로써 보증수리기간동안 수리요청이 발생한 제품과 발생하지 않은 제품의 수를 확인 할 수 있다. 따라서 만약 총 판매수량이 N이고, 이중 n개의 제품이 수리요청 되어 수리요청시간 $\{t_1, t_2, \cdots, t_n\}\in T$을 확보하게 되었다면, 우도함수는 다음과 같다.

$$L(\Theta, p, \tau|T)=\prod_{i=1}^{n} f(t_i|\Theta, p, \tau)(\overline{F}(w|\Theta, p, \tau))^{N-n} \quad (7)$$

이때 보증수리요청의 확률밀도함수 $f(t)$는 누적분포함수 $F(x)$를 미분하여 구할 수 있으며, 식 (7)를 최대화하는 Θ, p, τ를 추정한다.

3. 모형적용

3.1 보증수리자료의 정리

본 연구의 내용을 적용할 실제 보증수리내역자료는 자동차의 보증수리내역자료를 사용하였다. 이 자료에는 생산일(Production Date), 판매일(Sales Date), 보증수리요청일(Warranty Claim Date), 자동차의 차대번호(VIN : Vehicle Identification Number)와 원인부품 등이 표시되어있다. 보증수리내역자료 중, 판매일 기준으로 품질보증만료까지의 기간에 해당하는 자료를 모두 가지지 못한 불충분 자료는 제거하였으며, 추정을 위해서 판매일과 보증수리요청일의 차이로부터 수리요청시간을 도출하였다. 차대번호를 바탕으로 동일한 자동차가 동일한 원인부품의 보증수리요청이 두번이상 요청된 경우, 첫 번째 보증수리요청만을 수리요청시간으로 인정하였다. 이와 같은 방법을 통해 최종적으로 확보된 총 표본수는 62,478개였다.

대상단위는 시스템(자동차) 전체가 아니라 부품군을 그 대상으로 하였다. 시스템 전체가 아닌 부품군을 대상으로 삼은 이유는 고장 유형에 따라서 고객의 지연보고시간이 다르기 때문에, 만약 차량 전체의 고장을 대상으로 삼을 경우 각 부품별로 가지는 다양한 고객의 지연보고시간이 중첩되어, 모형의 특성이 잘 반영되지 않을 것이라 판단되었기 때문이다. 또한 개개의 부품이 아닌 부품군으로 그 대상을 설정한 이유는 자료 입력에 의한 불명확성을 제거하기 위함이다. 예를 들어 나사의 채결불량에 의한 고장이라고 하더라도, 보증수리내역자료를 수집하는 정비소에서 경우에 따라 볼트를 고장원인으로 지적하거나, 또는 너트로 고장 원인을 다르게 지적하는 경우가 존재하고 이러한 경우 자료입력의 불명확성에 의해 분석결과의 신뢰성이 저하될 수도 있다. 그러나 부품군을 대상으로 할 경우에는 이와 같은 문제점을 어느 정도 해결할 수 있다. 실제로 부품군이라고 표현은 하지만, 하나의 부품이 여러 개의 단위부품 코드로 나뉘어 기록되어져 있거나, 현장에서는 마치 하나의 부품처럼 취급하는 경우가 많기 때문에, 대상단위를 부품군으로 하는 것이 분석의 효과성을 높일 것이라 판단되었다.

대상이 되는 제품의 품질보증기간은 3년(157주)이며, 일일평균 보증수리요청 건수가 분석에 무리가 될 정도로 적으며, 요일효과에 의해서 자료의 균일성이 부족하기 때문에 고장자료의 시간단위를 주(week)단위로 하여 분석을 수행하였다.

3.2 최우추정

<그림 1>은 부품군 A의 사용시간에 따른 수리요청 빈도를 나타내고 있다. 막대그래프는 4주씩 묶어준 기간의 수리요청 빈도이며, 여기서 볼 수 있듯이 수리요청의 건수가 품질보증기간 시작과 종료에 즈음하여 많이 발생하는 이상증가현상이 존재함을 확인 할 수 있다.

부품군 A의 수리요청시간 자료를 사용하여 모수를 추정하

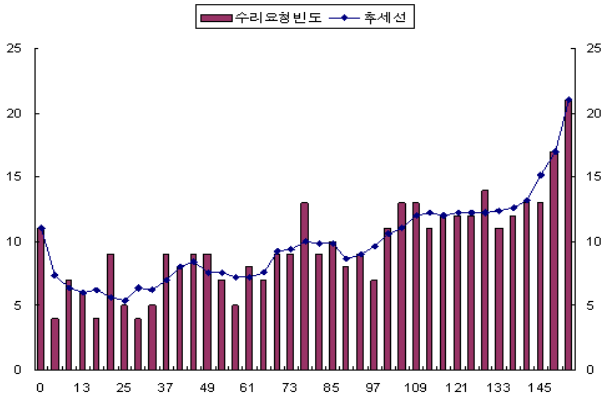

그림 1. 부품군 A의 수리요청빈도

표 1. τ에 따른 추정결과 및 MSE

τ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$	MSE
3	1.79×10^{-4}	1.43×10^{0}	1.65×10^{-4}	2.76×10^{0}
10	1.68×10^{-4}	1.42×10^{0}	2.92×10^{-4}	2.68×10^{0}
17	1.65×10^{-4}	1.41×10^{0}	4.08×10^{-4}	2.73×10^{0}
23	1.44×10^{-4}	1.37×10^{0}	4.96×10^{-4}	2.79×10^{0}

는 과정에서 사용관련 고장시간의 확률밀도함수를 식 (8)과 같이 척도모수(scale parameter)가 α이고 형상모수(shape parameter)가 β인 와이블분포(weibull distribution)로 가정하였다.

$$g(x) = \alpha^{\beta}\beta x^{\beta-1} e^{-(\alpha x)^{\beta}} \tag{8}$$

최우추정을 위해, 업체의 기존정보를 참고하여 각 모수들의 가능한 범위를 결정하고 해당범위에 대하여 순차적검색(sequential search)을 통하여 우도함수를 최대화하는 α, β, p, τ를 추정하였다. <표 1>은 고객의 최대지연보고시간에 따른 최우추정의 결과를 나타낸 것이다. 여기서 살펴 볼 수 있듯이 고객의 최대지연보고시간, 즉 τ가 커짐에 따라 $\hat{\alpha}$와 $\hat{\beta}$의 값이 작아짐을 알 수 있는데, 이는 τ가 커짐으로 상대적으로 많은 수의 보증수리요청이 이상증가현상으로 인한 것으로 간주되어, 실제 사용에 대한 고장률을 낮게 추정하기 때문이다. 실제 자료와의 평균제곱오차(MSE : Mean Squared Error)를 살펴보면 τ가 10주일 때 가장 좋은 결과를 보이고 있다. 이는 <그림 1>에서 초기의 이상증가현상이 10주 이전에 이상증가현상이 발생한 후 다시 요청률이 줄어들고 있는 사실과 일치한다.

3.3 기존 모형과의 성능비교

본 연구의 결과와 Majeske(2003)가 제안한 혼합모형(Mixture Model)의 성능을 비교하여 보았다. Majeske(2003)의 혼합모형 역시 보증기간초기에 생산관련고장과 사용관련고장이 혼재되어 보고됨을 가정하였고, 데이터를 절단하지 않고 분석하였다는 점에서 두 모형의 비교는 의미가 있다고 판단된다.

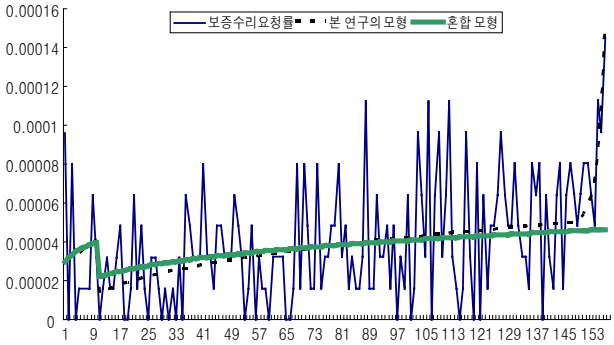

그림 3. 부품군 A에 대한 제안된 모형과 Majeske(2003)의 혼합모형 비교

표 3. 본 연구의 모형과 Majeske(2003)의 혼합모형 비교

		본 연구의 모형	혼합 모형
부품군 A	$\hat{\alpha}$	1.68×10^{-4}	1.12×10^{-4}
	$\hat{\beta}$	1.42×10^{0}	1.28×10^{0}
	$\hat{p}$	2.92×10^{-4}	1.83×10^{-4}
	MSE	2.68×10^{0}	3.76×10^{1}
부품군 B	$\hat{\alpha}$	9.41×10^{-7}	1.93×10^{-6}
	$\hat{\beta}$	6.10×10^{-1}	6.75×10^{-1}
	$\hat{p}$	1.59×10^{-4}	2.08×10^{-4}
	MSE	1.75×10^{0}	2.14×10^{0}
부품군 C	$\hat{\alpha}$	1.64×10^{-4}	4.19×10^{-4}
	$\hat{\beta}$	1.58×10^{0}	2.14×10^{0}
	$\hat{p}$	3.60×10^{-4}	3.55×10^{-4}
	MSE	1.54×10^{0}	1.85×10^{0}

<그림 2>는 부품군 A의 자료를 본 연구의 제안모형과 Majeske (2003)의 혼합모형으로 추정하여 보증수리요청률함수로 도시한 것이다. 이때 보증수리요청률 함수는 식 (9)을 사용하여 도출하였다.

$$r(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{\overline{F}(t)} \tag{9}$$

부품군 A 이외에 부품군 B와 C에 대해서도 추가적으로 추정하였고, 그 결과를 <표 2>에 나타내었다. 본 연구에서 제안한 모형과 Majeske(2003)의 혼합모형의 성능을 비교하기 위하여 본 연구의 제안모형과 Majeske(2003)의 혼합모형 각각에서 계산한 보증수리요청률함수, $r(t)$를 해당 구간에서 적분하여 기대보증수리요청빈도를 계산하여 수집된 자료와의 평균제곱오차를 식 (10)와 같이 계산하였는데, A, B, C 부품군 모두에 대하여 본 연구의 모형이 더 나은 성능을 보여주고 있음을 알 수 있다.

$$MSE = \frac{\sum_{k=1}^{157}\left(k^{th}\text{주의 수리요청빈도} - \int_{k-1}^{k} r(t)dt\right)^2}{157} \tag{10}$$

4. 결론 및 추후연구과제

본 연구에서는 생산관련고장에 의한 품질보증기간 초기의 보증수리요청의 이상증가와 고객의 지연보고로 품질보증기간 말기에 일어나는 보증수리요청의 이상증가 현상을 표현할 수 있는 모형을 제안하였다. 이를 위해서 보증기간 초기에는 생산관련고장과 사용관련고장이 혼재되어 요청되는 상황에서 확률적으로 요청되도록 모형화 하였다. 또한 고객의 지연보고의 시간에 대한 분포를 도입하여, 연속일양분포를 가정하였으며, 품질보증기간 말기에는 이를 수정해줌으로 수리요청의 이상증가 현상을 설명할 수 있는 모형을 수립하였다. 또한 본 연구에서 제안된 모형을 이용하여 실제 자료에서 그 모수를 추정하여 적용가능성을 보였으며, 더 나아가 기존연구에서 제안된 모형과 비교를 실시하여, 기존 모형보다 더 나은 결과를 가질 수 있음을 보여주었다.

그러나 자동차의 수명은 사용기간 뿐 아니라 주행거리를 기초로 추정하는 경우도 가능하며, 주행거리를 시간 단위로 하여 추정하였을 경우, 본 연구에서 제안된 방법의 적용가능성을 확인할 필요가 있다고 판단된다. 따라서 추후 연구로 사용기간이 아닌 주행거리를 기준으로 하는 고객의 지연보고의 특성을 고려한 모형개발이 가능할 것이라 판단된다. 또한 보증수리내역자료에서의 고객의 지연보고 특성에 대한 연구가 크게 이루어지지 않은 상황에서, 본 연구는 지연보고의 분포를 Majeske(2003)가 그의 연구에서 가정한 일양분포를 참고하여 확장하여 가정하였다. 그러나 고객의 지연보고에 대한 현실적인 연구가 추가되어 그 분포를 현실화 할 수 있다면 모형의 정확성과 활용가치를 높일 수 있을 것이다.

참고문헌

Cohen, A. C. (1991), Truncated and censored samples: theory and applications, Marcel Dekker, New York.

Iskandar, B. P. and Blischke, W. R. (2003), Reliability and warranty analysis of a motorcycle based on claims data, Case Studies in Reliability and Maintainenance, 623-656, New Jersey, John Wiley and Sons, Inc.

Majeske, K. D. (2003), A mixture model for automobile warranty data, *Reliability Engineering and System Safety*, **81**(1), 71-77.

Meeker, W .Q. and Escobar, L. A. (1998), Statistical methods for reliability data, John Wiley and Sons, New York.

Nelson, W. (1990), Hazard plotting of left truncated life data, Journal of Quality Technology, **22**(3), 230-232.

Oh, Y. S. and Bai, D. S. (2001), Field data analyses with additional after warranty failure data, *Reliability Engineering and System Safety*, **72**(1), 1-8.

Pal, S. and Murthy, G. S. R. (2003), An application of Gumbel's bivariate exponential distribution in estimation of warranty cost of motor cycle, *The International Journal of Quality and Reliability Management*, **20**(4), 488-502.

Rai, B. and Singh N. (2003), Hazard rate estimation from incomplete and unclean warranty data, *Reliability Engineering and System Safety*, **81**(1), 79-92.

Rai, B. and Singh, N. (2004), Modeling and analysis of automobile warranty data in presence of bias due to customer-rush near warranty expiration, *Reliability Engineering and System Safety*, **86**(1), 83-94.

박 종 훈
동국대학교 산업공학과 학사
서울대학교 산업공학과 석사
서울대학교 산업공학과 박사
현재 서울대학교 산업공학과 연구원
관심분야 : 신뢰성공학, 품질공학, 응용통계, 확률모형

김 영 훈
서울대학교 산업공학과 학사
서울대학교 산업공학과 석사
TMAX AP컨설팅
현재 SK C&C 소프트웨어 테크널러지팀
관심분야 : 신뢰성공학, 응용통계

백 장 현
서울대학교 산업공학과 학사
서울대학교 산업공학과 석사
서울대학교 산업공학과 박사
현재 전북대학교 산업정보시스템공학과 교수
관심분야 : 정보통신, 경영과학, 응용통계

이 창 훈
서울대학교 공과대학대학 졸업
캔사스 주립대학 산업공학 석사
캔사스 주립대학 산업공학 박사
현재 서울대학교 산업공학과 교수
관심분야 : 신뢰성공학, 품질공학, 이동통신망 성능분석 등