
Burr 분포를 이용한 NHPP 소프트웨어 신뢰성장모형에 관한 연구

김희철* · 박종구**

The Study for NHPP Software Reliability Growth Model based on Burr Distribution

Hee-Cheul Kim* · Jong-Goo Park**

요 약

유한고장수를 가진 비동질적인 포아송 과정에 기초한 모형들에서 잔존 결함 1개당 고장 발생률은 일반적으로 상수, 혹은 단조증가 및 단조 감소 추세를 가지고 있다. 본 논문에서는 기존의 소프트웨어 신뢰성 모형인 Goel-Okumoto 모형과 Yamada-Ohba-Osaki 모형을 재조명하고 이 분야에 적용될 수 있는 Burr 분포를 이용한 모형을 제안하였다. 고장 간격시간으로 구성된 자료를 이용한 모수추정 방법은 최우추정법 과 일반적인 수치해석 방법인 이분법을 사용하여 모수 추정을 실시하고 효율적인 모형 선택은 편차자승합(SSE), AIC 통계량 및 콜모고로프 거리를 적용하여 모형들에 대한 효율적인 모형 선택도 시도 하였다. 소프트웨어 고장 자료 분석에서는 실제 고장자료를 통하여 분석하였다. 이 자료들에서 기존의 모형과 Burr 분포 모형의 비교를 위하여 산술적 및 라플라스 검정, 편의 검정 등을 이용하였다.

ABSTRACT

Finite failure NHPP models presented in the literature exhibit either constant, monotonic increasing or monotonic decreasing failure occurrence rates per fault. In this paper, Goel-Okumoto and Yamada-Ohba-Osaki model was reviewed, proposes the Burr distribution reliability model, which making out efficiency application for software reliability. Algorithm to estimate the parameters used to maximum likelihood estimator and bisection method, model selection based on SSE, AIC statistics and Kolmogorov distance, for the sake of efficient model, was employed. Analysis of failure using real data set for the sake of proposing shape parameter of the Burr distribution was employed. This analysis of failure data compared with the Burr distribution model and the existing model(using arithmetic and Laplace trend tests, bias tests) is presented.

키워드

Software Reliability Model, NHPP, Burr Distribution, SSE, Kolmogorov Distance

I. 서론

소프트웨어 테스트 단계에서 소프트웨어 고장수

(Number of failure)와 고장간격시간에 의해 소프트웨어 고장현상을 수리적으로 모형화하면 소프트웨어에 대한 평가를 보다 쉽게 할 수 있으며 신뢰도 모형에 의해 소프

* 남서울대학교 산업경영공학과
** 원광대학교 전기 전자 및 정보공학부

트웨어 고장 수, 소프트웨어 고장발생간격시간, 소프트웨어 신뢰도 및 고장률 등의 신뢰성 평가측도들이 추정되어 미래의 고장시간을 예측할 수도 있다.

소프트웨어 고장시간은 수명자료가 된다. 따라서 비음(Nonnegative)의 값을 가지기 때문에 이 분야에서는 주로 지수분포, 와이블분포, 감마분포 등 일반화 감마분포(혹은 어랑분포)가 많이 사용되어 왔다[1]. 그러나 이러한 감마분포 외에도 통계자료 해석시 자료를 대수변환 이 후에 정규분포로 처리하면 되는 대수정규분포(Lognormal distribution), 이 분포와 유사한 대수로지스틱 분포(Loglogistic distribution)도 이 분야에 사용이 가능하다[2]. 또, 자유도(Degree of freedom; df)에 의존하는 카이제곱(χ^2 distribution) 분포도 적용이 가능하다. 이러한 카이제곱 분포는 변화가 일어날 때까지의 대기시간을 나타내는 데 많이 사용되는 감마분포(Gamma Distribution)의 특수한 경우로도 접근할 수 있기 때문에 소프트웨어 신뢰성 수명분포로 설명할 수 있다[3].

본 논문에서는 왜도(skewness)와 첨도(kurtosis) 측면에서 보다 넓은 범위에 분포하기 때문에 어떠한 수명 단봉자료(unimodal data)에서도 적용할 수 있는 장점을 가진 분포인 2모수 Burr 분포[4, 5]를 적용하고자 한다. 이러한 Burr 분포를 수명분포로 사용한 NHPP 모형에 대한 신뢰성 척도를 추정하고 이를 바탕으로 형상모수에 따른 모형의 효율성과 그 특성을 알아보려 한다.

본 논문의 2장에서는 관련연구로서 유한 고장 NHPP 모형에 대하여 서술하였고 3장에서는 Burr 분포 신뢰성 모형 및 최우추정법을 이용한 모수추정에 대하여 설명하고 4장에서는 실제 고장자료를 이용하여 각 모형에 대한 모수추정 및 모형비교를 실시하였으며 마지막으로 5장에서는 결론을 나열 하였다.

II. 이론적 배경에 대한 고찰

신뢰도에서 관측시간 $(0, t]$ 사이에 발견된 고장 수 $N(t)$ 를 모형화 하는데 비동질적 포아송과정(Non-homogeneous poisson process; NHPP)이 널리 사용되어 왔다. 이 과정(Process)에서 강도함수(Intensity function) 혹은 고장 발생률(Rate of occurrence of failure; ROCOF) $\lambda(t) = dE[N(t)]/dt$ 은 t 에 대한 단조(Monotonic)함수로 흔히 가정 한다[2]. 이 범주에서 지금까지 알려진 모형들

은 Goel-Okumoto 모형, Weibull 모형 그리고 Cox-Lewis 모형등이 있는데 이 모형들에 대한 강도함수는 각각 시간에 의존한 함수, 멱(Power) 함수, 대수 선형(Log-linear) 함수를 가정하였다[6].

NHPP 모형에서 평균값 함수 $m(t)$ (Mean value function)와 강도 함수 $\lambda(t)$ 는 다음과 같은 관계로 표현할 수 있다[7].

$$m(t) = \int_0^t \lambda(s) ds, \frac{dm(t)}{dt} = \lambda(t) \quad (2.1)$$

$N(t)$ 는 모수 $m(t)$ 를 가진 포아송 확률밀도함수(Probability density function)로 알려져 있다. 즉,

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!}, \quad n = 0, 1, 2, \dots, \infty \quad (2.2)$$

이처럼 시간관련 모형(Time domain models)들은 NHPP에 의해서 확률 고장 과정으로 설명이 가능하다. 이러한 모형들은 고장 강도 함수 $\lambda(t)$ 가 다르게 표현됨으로서 평균값 함수 $m(t)$ 도 역시 다르게 나타난다. 이러한 NHPP 모형들은 유한 고장 모형과 무한 고장 범주로 분류한다[8]. 유한 고장(Finite failure) NHPP 모형들은 충분한 테스트 시간이 주어지면 결함들(Faults)의 기대값이 유한 값($\lim_{t \rightarrow \infty} m(t) = \theta < \infty$)을 가지고 반면에 무한 고장(Infinite failure) NHPP 모형들은 무한 값을 가진다고 가정 된다. 유한 고장 NHPP 모형에서 충분한 테스트 시간이 주어졌을 때 탐색되어 질 수 있는 결함의 기대값을 θ 라고 표현하고 $F(t)$ 를 분포함수라고 표현하면 유한 고장 NHPP모형의 평균값 함수는 다음과 같이 표현할 수 있다[2].

$$m(t) = \theta F(t) \quad (2.3)$$

(2.3)식으로 부터 순간고장 강도함수(Instantaneous failure intensity) $\lambda(t)$ 는 다음과 같이 유도된다.

$$\lambda(t) = \theta F'(t) \quad (2.4)$$

(2.4)식을 다음과 같이 변형하여 표기할 수도 있다.

$$\lambda(t) = [\theta - m(t)] \frac{F'(t)}{1 - F(t)} = [\theta - m(t)] h(t) \quad (2.5)$$

단, $h(t) = \frac{F'(t)}{1 - F(t)}$ 는 위험함수(Hazard function, 고장률 함수)로서 소프트웨어 결함 당 고장 발생률을 의미하고 $[\theta - m(t)]$ 은 t 시점에서 소프트웨어에 남아있는 결함들의 기대값을 나타낸다. $[\theta - m(t)]$ 의 값은 시점 t 에 대한 단조 비증가 함수(Monotonically nonincreasing function)가 된다. 즉, 시간이 지남에 따라 그 고장을 찾아 제거하는 디버깅 과정을 거치면서 제거되기 때문에 감소성을 가진다.

$\lambda(t)$ 는 $h(t)$ 의 속성에 따라 달라지며 그 추세는 상수나 증가, 혹은 감소(증가)하다가 증가(감소)하는 패턴을 가질 수도 있다[2].

시간 $(0, t]$ 까지 조사하기 위한 시간 절단(Time truncated)모형은 n 번째 까지 고장시점 자료를

$$x_k = \sum_{i=1}^k t_k \quad (k = 1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \quad (2.6)$$

이라고 하면 데이터 집합 D_t 는 $\{n, x_1, x_2, \dots, x_n; t\}$ 와 같이 구성된다. n 번째까지 고장시점이 관찰된 고장 절단 모형일 경우에 데이터 집합 D_{x_n} 은 $\{x_1, x_2, \dots, x_n\}$ 으로 구성되며 이 시간 절단 모형에서의 θ 를 모수공간이라고 표시하면 우도함수는 다음과 같이 알려져 있다[7,8].

$$L_{NHPP}(\theta | D_{x_n}) = \left(\prod_{i=1}^n \lambda(x_i) \right) \exp(-m(x_n)) \quad (2.7)$$

이 분야의 기본적 모형인 Goel-Okumoto 모형[9] 은 $h(t)$ 가 정수 패턴을 가짐으로서 시점 t 에 독립이고 잘 알려진 Yamada, Ohba-Osaki 모형은 단조 비감소 패턴을 가진다[10].

NHPP 모형에서 테스트 시점 x_n 에서 소프트웨어 고장이 일어난다고 하는 가정 하에서 신뢰구간 $(x_n, x_n + t]$ (단, t 는 임무시간(Mission time)사이에서 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도(Reliability)

$\hat{R}(t | x_n)$ 는 다음과 같이 됨이 알려져 있다[7,9].

$$\begin{aligned} \hat{R}(t | x_n) &= e^{-\int_{x_n}^{x_n+t} \lambda(\tau) d\tau} \\ &= \exp\{-m(t + x_n) - m(x_n)\} \end{aligned} \quad (2.8)$$

2.1 Burr 분포

이 장에서 Burr 분포[4, 5]에 대하여 요약하고자 한다.

이 분포는 1942년 Burr에 의해 만들어진 분포이고 이 분포의 특징은 왜도(skewness)와 첨도(kurtosis) 측면에서 보다 넓은 범위에 분포하기 때문에 어떠한 수명 단봉 자료(unimodal data)에서도 적용할 수 있는 장점을 가진 분포라고 알려져 있다[5]. 이 분포의 확률밀도 함수와 분포함수는 다음과 같이 표현됨이 알려져 있다.

$$f(t | a, b) = abt^{a-1} / (1+t^a)^{b+1} \quad (2.9)$$

$$F(t | a, b) = 1 - (1+t^a)^{-b} \quad (2.10)$$

단, $t > 0, a > 0, b \geq 1$.

이러한 Burr 분포에 대한 연구는 많은 학자들에 의하여 연구되었는데 그 중에서 Austin[11]은 이 분포를 관리도에 적용하였으며 Dubey[12]는 고장 시간 모형으로서의 Burr 분포의 장점을 고찰 하였고 Tadikamalla[13]는 다른 분포와의 관계를 연구하였고 Upadhyay 와 Javed, Peshwani[5]은 베이지안적 접근 방법을 동원하여 4모수에 대한 모수 추정방법을 연구하기도 하였다.

위험함수는(2.9)과 (2.10)식과 관련하면 다음과 같이 유도 된다.

$$\begin{aligned} h(t | a, b) &= \frac{f(t | a, b)}{1 - F(t | a, b)} \\ &= \frac{abt^{a-1} / (1+t^a)^{b+1}}{(1+t^a)^{-b}} \\ &= \frac{abt^{a-1}}{1+t^a} \end{aligned} \quad (2.11)$$

위험함수의 형태는 그림 1 에서 보여 주듯이 b 가 1인 경우에 $a \leq 1$ 인 경우에는 감소함수이고 $a > 2$ 인 경우에는 증가하다가 감소하는 형태가 된다.

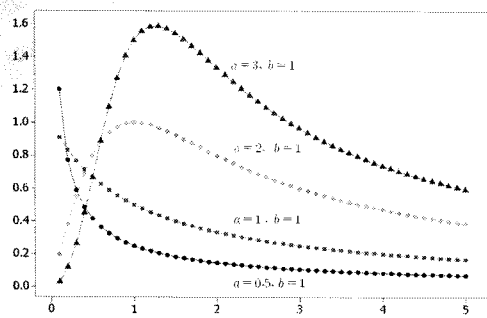


그림 1. Burr 분포의 위험함수
Fig. 1. Hazard function of Burr distribution

2.2 기존의 신뢰성 모형

기존의 모형은 주로 수명 분포가 다음과 같은 감마분포의 분포함수를 사용하고 있다[2].

$$F_{Ga}(t | k, \beta) = 1 - e^{-\beta t} \left(\sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \quad (t, k > 0, \beta > 0) \quad (2.12)$$

(2.12)식에서 감마분포의 확률밀도함수는 다음과 같이 유도 할 수 있다[14].

$$f_{Ga}(t | k, \beta) = \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \quad (k, \beta > 0, t > 0) \quad (2.13)$$

따라서 (2.3)식과 (2.4)에 관련하여 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정하거나 증가 혹은 감소하는 특징을 가지는 모형이 감마 모형이 되고 유한 고장 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta F_{Ga}(t | k, \beta) = \theta \left[1 - e^{-\beta t} \left(\sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \right] \quad (2.14)$$

$$\lambda(t) = \theta f_{Ga}(t | k, \beta) = \theta \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \quad (2.15)$$

(2.14)식의 분포함수가 폐쇄형(close form)의 형태이므로 기존의 모형은 주로 함수는 k 값이 양의 정수일 때만을 고려하고 있다.

이 분야에서 가장 기본적인 모형은 Goel-Okumoto 모형[9] 이다. 이 모형은 결함당 고장발생 시간의 분포(수명 분포)를 지수분포를 가정하였다. 따라서 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정한 형태를 가지며 (2.14)식과 (2.15)식에서 형상모수 $k=1$ 인 경우와 같이 된다. 이 모형을 개선한 모형이 잘 알려진 S 모형(Yamada, Ohba-Osaki 모형[10])이다. 따라서 이 모형은 수명 분포를 감마분포중에서 형상모수 $k=2$ 인 분포를 말한다.

2.2.1 Goel-Okumoto모형과 Yamada-Ohba-Osaki모형

Goel-Okumoto 모형[9]에 대한 평균값 함수는 $m(t|\theta, \beta_1) = \theta(1 - e^{-\beta_1 t})$ ($\theta > 0, \beta_1 > 0$)이라고 알려져 있다. 시간 $(0, t)$ 까지 조사하기 위한 시간 절단(Time truncated)모형은 n 번째 까지 고장시점 자료를

$$x_k = \sum_{i=1}^k t_k \quad (k = 1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n)$$

이라고 하면 데이터 집합 D_t 는 $\{n, x_1, x_2, \dots, x_n; t\}$ 와 같이 구성된다.

t 을 최종 고장시점 x_n 으로 대치하고 고장 발생률 $\lambda(t|\theta, \beta_1) = \theta\beta_1 e^{-\beta_1 t}$ 을 이용하면 우도함수는 다음과 같이 표현 할 수 있다.

$$L_{GO}(\theta, \beta_1 | D_{x_n}) = \left(\prod_{k=1}^n \theta \beta_1 e^{-\beta_1 x_k} \right) \exp[-\theta(1 - e^{-\beta_1 x_n})] \quad (2.16)$$

최우추정법(MLE)을 이용하기 위한 로그우도함수를 구하면,

$$\ln L_{GO} = n \ln \theta + n \ln \beta_1 - \beta_1 \sum_{k=1}^n x_k - \theta(1 - e^{-\beta_1 x_n}) \quad (2.17)$$

으로 표현된다.

따라서 고장절단모형에서의 모수 θ 와 β_1 에 관한 편미분식은 다음과 같이 유도할 수 있다.

$$\frac{\partial \ln L_{GO}}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-\beta_1 x_n} = 0,$$

$$\frac{\partial \ln L_{GO}}{\partial \beta_1} = \frac{n}{\beta_1} - \sum_{k=1}^n x_k - \theta x_n e^{-\beta_1 x_n} = 0$$

각 모수에 대한 최우추정량 $\hat{\theta}_{MLE}$ 와 $\hat{\beta}_{1MLE}$ 은 다음 식을 만족한다.

$$\frac{n}{\theta} = 1 - \exp(-\hat{\beta}_1 x_n), \quad (2.18)$$

$$\frac{n}{\beta_1} = \sum_{k=1}^n x_k + \hat{\theta} x_n \exp(-\hat{\beta}_1 x_n) \quad (2.19)$$

(2.18)과 (2.19)식을 수치적으로 풀어 두 모수를 구할 수 있다.

Yamada-Ohba-Osaki 모형[2, 10]에 대한 평균값 함수는 $m(t|\theta, \beta_2) = \theta [1 - (1 + \beta_2 t) e^{-\beta_2 t}]$ ($\theta > 0, \beta_2 > 0$) 이라고 알려져 있고 $\lambda(t|\theta, \beta_2) = \theta \beta_2^2 t e^{-\beta_2 t}$ 가 되고 우도함수는 다음과 같다.

$$L_{YOO}(\theta, \beta_2 | D_{x_n}) = \left(\prod_{k=1}^n \theta \beta_2^2 x_k e^{-\beta_2 x_k} \right) \cdot \exp[-\theta [1 - (1 + \beta_2 x_n) e^{-\beta_2 x_n}]] \quad (2.20)$$

따라서 각 모수에 대한 최우추정량은 $\hat{\theta}_{MLE}$ 와 $\hat{\beta}_{2MLE}$ 은 다음식을 만족한다.

$$\frac{n}{\theta} = 1 - \exp(-\hat{\beta}_2 x_n) - \hat{\beta}_2 x_n \exp(-\hat{\beta}_2 x_n) \quad (2.21)$$

$$\frac{2n}{\beta_2} = \sum_{k=1}^n x_k + \hat{\theta} \hat{\beta}_2 x_n^2 \exp(-\hat{\beta}_2 x_n) \quad (2.22)$$

(2.22) 과 (2.23)식을 수치적으로 풀어 두 모수를 구할 수 있다.

III. 제안된 Burr분포를 이용한 신뢰성 모형

Burr 분포에 대한 평균값 함수는 (2.3)식과 (2.4)식을 연관하여 평균값 함수 $m_{Burr}(t|a, b)$ 와 고장발생률 $\lambda_{Burr}(t|a, b)$ 을 다음과 같이 유도 할 수 있다.

$$m(t|a, b) = \theta F_{Burr}(t|a, b) = \theta(1 - (1+t^a)^{-b}) \quad (3.1)$$

$$\lambda(t|a, b) = \theta f(t|a, b) = \theta (abt^{a-1}/(1+t^a)^{b+1}) \quad (3.2)$$

단, $t > 0, a > 0, b \geq 1$. 시간 (0,t] 까지 조사하기 위한 시간 절단(Time truncated)모형은 n 번째 까지 고장시점 자료를

$$x_k = \sum_{i=1}^k t_k \quad (k=1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n)$$

이라고 하면 데이터 집합 D_t 는 $\{n, x_1, x_2 \dots x_n, t\}$ 와 같이 구성된다.

t 을 최종 고장시점 x_n 으로 대치하고 평균값 함수 $m(t|a, b)$ 와 고장발생률 $\lambda(t|a, b)$ 을 이용하면 우도함수 $L_{Burr}(\theta, a, b | D_{x_n})$ 는 다음과 같이 표현 할 수 있다.

$$L_{Burr}(\theta, a, b | D_{x_n}) = \prod_{k=1}^n \theta (abx_i^{a-1}/(1+x_i^a)^{b+1}) \cdot \exp(-\theta(1 - (1+x_n^a)^{-b})) \quad (3.3)$$

따라서 최우추정법을 이용하기 위한 Burr 분포 모형 로그우도함수는 다음과 같이 유도된다.

$$\ln L_{Burr}(\theta, a, b | D_{x_n}) = n \ln \theta + n \ln a + n \ln b + (a-1) \sum_{k=1}^n \ln x_i - (b+1) \sum_{k=1}^n \ln(1+x_i^a) - \theta + \theta(1+x_n^a)^{-b} \quad (3.4)$$

형상 모수 b 값은 상수(사전에 알고 있는 경우)라고 가정 했을 때 최우추정법을 이용하기 위하여 (3.4)식을 θ 와 a 에 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{\partial \ln L_{Burr}(\theta, a, b)}{\partial \theta} = \frac{n}{\theta} - 1 + (1+x_n^a)^{-b} = 0 \quad (3.5)$$

$$\frac{\partial \ln L_{Burr}(\theta, a, b)}{\partial a} = \frac{n}{a} + \sum_{k=1}^n \ln x_i + \quad (3.6)$$

$$(b+1) \sum_{k=1}^n \frac{x_i^a \ln x_i}{1+x_i^a} - \theta b(1+x_n^a)^{-b-1} x_n^a \ln x_n = 0$$

따라서 고장절단 모형에서의 각 모수에 대한 최우추정량은 $\hat{\theta}$ 와 \hat{a} 은 다음 식을 만족한다.

$$\frac{n}{\hat{\theta}} = 1 - (1 + x_n^{\hat{a}})^{-b} \quad (3.7)$$

$$\begin{aligned} \frac{n}{\hat{a}} = & - \sum_{k=1}^n \ln x_k - (b+1) \sum_{k=1}^n \frac{x_k^{\hat{a}} \ln x_k}{1 + x_k^{\hat{a}}} \\ & + \theta b (1 + x_n^{\hat{a}})^{-b-1} x_n^{\hat{a}} \ln x_n \end{aligned} \quad (3.8)$$

(3.7) 식과 (3.8) 식을 비선형 연립방정식(수치 해석적 방법)을 이용하여 풀면 최우추정치 \hat{a}_{MLE} 와 $\hat{\theta}_{MLE}$ 의 값을 구할 수 있다.

본 논문에서는 Burr 분포의 형상 모수 b 값은 1, 2, 3 과 같은 상수인 경우를 본 논문에서는 제시하고자 한다. 예를 들어 $b = 1$ 인 경우에는 다음과 같은 연립방정식을 유도할 수 있으며 다른 경우도 유사한 방법을 이용할 수 있다.

$$\frac{n}{\hat{\theta}} = 1 - (1 + x_n^{\hat{a}})^{-1} \quad (3.9)$$

$$\begin{aligned} \frac{n}{\hat{a}} = & - \sum_{k=1}^n \ln x_k - 2 \sum_{k=1}^n \frac{x_k^{\hat{a}} \ln x_k}{1 + x_k^{\hat{a}}} \\ & + \theta (1 + x_n^{\hat{a}})^{-2} x_n^{\hat{a}} \ln x_n \end{aligned} \quad (3.10)$$

IV. 소프트웨어 고장 자료 분석

이 장에서 실제 고장자료를 이용하여 형상모수에 따른 Burr 모형을 분석하고자 한다. 이 고장자료는 NTDS (Naval Tactical Data System)에 의해 발생된 소프트웨어 고장자료로서 Goel과Okumoto[9], Mazzuchi 와 Soyer[15], Kuo와 Yang[8] 등이 이 고장자료를 이용하여 소프트웨어 모형을 제안 한 바 있다. 본 연구도 이 자료를 이용하고자 한다. 이 자료는 표 1 에 나열 되어 있고 제시하는 신뢰모형들을 분석하기 위하여 우선 자료에 대한 추세 검정이 선행 되어야 한다[2,16].

추세 분석에는 산술평균 검정(Arithmetic mean test)과 라플라스추세 검정(Laplace trend test)등이 있다. 이 검정을 실시한 결과 그림 2 에서 산술평균 검정결과 고장시

간이 증가함에 따라 산술 평균이 거의 증가 추세를 보이고 있으므로 신뢰성장(Reliability growth) 속성을 가지고 있고 그림 3 에 나타난 라플라스 추세 검정의 결과도 라플라스 요인(Factor)이 증가하다가 고장시간이 증가함에 따라 음수로서 감소하기 하는 신뢰성장(Reliability growth) 속성을 나타내고 있다. 신뢰성장 모형을 이 자료에 적용시키는 것이 가능하다.[2, 16,17]. 본 논문에서는 모수추정을 용이하게 하기 위하여 실제자료에서 수치 변환된 자료(failure times $\times 10^{-2}$)를 이용하였다.

소프트웨어 신뢰성 모형의 모수 추정은 최우추정법을 이용하였고 비선형 방정식의 계산방법은 수치 해석적 기본 방법인 이분법(Bisection method)을 사용하였다. 이러한 계산은 초기값을 10^{-2} 와 10 을, 허용 한계(Tolerance for width of interval)는 10^{-10} 을 주고 수렴성을 확인 하면서 충분한 반복 횟수인 100번을 C-언어를 이용하여 모수 추정을 수행하였다. 그러나 $b=3$ 이상인 경우는 비효율적이면서 근이 수렴되지 않아 본 논문에서는 고려대상에서 제외 시켰다.

기존의 모형과 Burr 분포신뢰도 모형에 대한 모수의 추정값들의 결과는 표 2 에 요약되었다.

이 분야에서는 일반적으로 모형 선택의 하나의 방법으로 편차자승합(SSE [7])을 이용할 수 있다. 본 논문은 SSE 기준 뿐 아니라 이러한 SSE 기준을 보완하는 AIC(Akaike Information Criterion) 통계량을 이용하여 비교도 포함 시켰다. 이 통계량은 적용모형에 대하여 자유도와 관련된 우도함수를 최대화 시켜주는 척도이다. 이 척도는 다음과 같이 정의된다[7,9].

$$AIC = -2 \log(\max \text{ of likelihood function}) + 2N$$

단, N 은 적용모형에 대한 모수의 수를 의미하고 이러한 AIC 척도는 SSE와 마찬가지로 작은 값을 가지는 모형이 상대적으로 효율적인 모형으로 간주된다. 이러한 결과는 표 3 에 요약되었다. 이 표에서 형상모수($a=1,2$)에 따른 Burr 신뢰도 모형들이 이 분야에서 기존에 알려진 모형인 Yamada-Ohba-Osaki 모형이나 Goel-Okumoto 모형에 비해 상대적으로 효율적 모형으로 나타나고 있다.

예측 오류(prediction error)의 비정상성(nonstationarity)에 대한 척도는 Kolmogorov 거리(distance)[2.18]로 측정되는데 이 거리가 클수록 상대적으로 비정상성을 내포

하고 있다. 그림 4는 S-Plus 소프트웨어[18]를 이용하여 Kolmogorov 검정에 대한 그림을 보여주고 있고 이 그림에서도 Burr 분포 신뢰도 모형이 상대적으로 비정상성 속성이 덜 내포하고 있음을 알 수 있다. 표 4는 Kolmogorov 거리를 요약한 표로서 모형에 대한 치우침(bias)을 나타내고 있다. 이 표에서도 Burr 분포 신뢰도 모형($\alpha=1, 2$)들이 상대적으로 비정상성 속성이 덜 내포하고 있음을 알 수 있다.

표 1. 고장 간격 자료
Table. 1 Failure Interval Time Data

Failure number	Failure Interval Time	Failure Time	Failure Time $\times 10^{-2}$
1	9	9	0.09
2	12	21	0.21
3	11	32	0.32
4	4	36	0.36
5	7	43	0.43
6	2	45	0.45
7	5	50	0.5
8	8	58	0.58
9	5	63	0.63
10	7	70	0.7
11	1	71	0.71
12	6	77	0.77
13	1	78	0.78
14	9	87	0.87
15	4	91	0.91
16	1	92	0.92
17	3	95	0.95
18	3	98	0.98
19	6	104	1.04
20	1	105	1.05
21	11	116	1.16
22	33	149	1.49
23	7	156	1.56
24	91	247	2.47
25	2	249	2.49
26	1	250	2.5
27	87	337	3.37
28	47	384	3.84
29	12	396	3.96
30	9	405	4.05
31	135	540	5.4

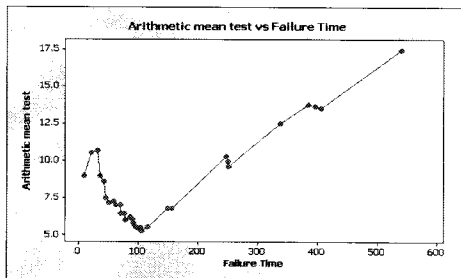


그림 2. 산술평균검정
Fig. 2 Arithmetic mean test

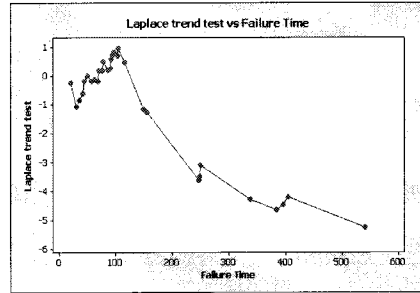


그림 3. 라플라스 추세 검정
Fig. 3 Laplace trend test

표 2. 각 모형의 모수 추정값
Table. 2 Parameter estimation of each model

Model	MLE
Goel-Okumoto Model	$\hat{\beta}_{1, MLE} = 0.39812$
	$\hat{\theta}_{MLE} = 35.073$
Yamada-Ohba-Osaki Model	$\hat{\beta}_{2, MLE} = 0.68351$
	$\hat{\theta}_{MLE} = 34.862$
Burr Model ($b=1$)	$\hat{a}_{MLE} = 1.71259$
	$\hat{\theta}_{MLE} = 32.726$
Burr Model ($b=2$)	$\hat{a}_{MLE} = 1.90508$
	$\hat{\theta}_{MLE} = 32.247$
Burr Model ($b=2.5$)	$\hat{a}_{MLE} = 1.91675$
	$\hat{\theta}_{MLE} = 31.016$

표 3. 각 모형의 SSE 및 AIC의 값
Table. 3 SSE and AIC of each model

Model	SSE	AIC
Goel-Okumoto Model	9013.172	109.1432
Yamada-Ohba-Osaki Model	5049.463	104.0325
Burr Model ($b=1$)	1022.814	79.2567
Burr Model ($b=2$)	3890.297	98.9871
Burr Model ($b=2.5$)	6876.1376	113.1589

V. 결 론

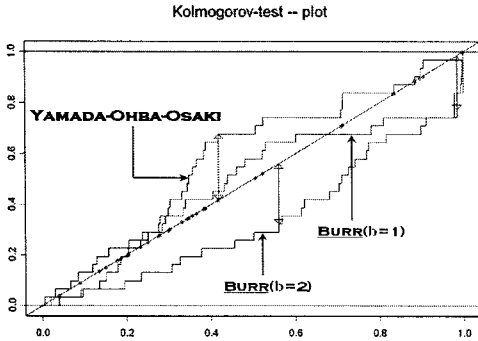


그림 4. 모형에 대한 Kolmogorov 검정 그림
Fig. 4 Kolmogorov test plot

표 4. 모형에 대한 Kolmogorov 거리
Table. 4 Kolmogorov distance

Model	Kolmogorov distance
Goel-Okumoto	0.464987
Yamada-Ohba-Osaki Model	0.258784
Burr Model(b=1)	0.196432
Burr Model(b=2)	0.237895

t 는 임무시간(Mission time)에 대한 즉, 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도(Reliability) $\hat{R}(t | x_n)$ 는 그림 5 와 같다. 이 그림에서도 Burr 분포 신뢰도 모형($a=1, 2$)들이 상대적으로 효율적인 신뢰도를 나타내고 있다.

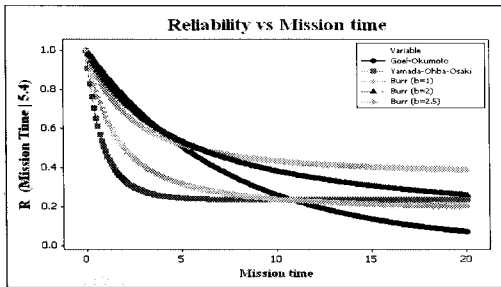


그림 5. 각 모형에 대한 신뢰도
Fig. 5 Reliability of each model

소프트웨어 신뢰성은 개발의 최종단계에 있는 테스트 공정이나 실제 사용단계에 있어서 소프트웨어 내에 존재하는 고장 수나 고장 발생 시간에 의해서 효과적으로 평가할 수 있는 상황으로 그 평가 기술이 중요하게 된다. 따라서 소프트웨어 개발의 테스트 공정이나 실제 사용단계에 있어서 고장 발생 환경이나 고장 발생 현상을 수리적으로 모형화가 가능하면 평가를 할 수 있다. 테스트 시간이나 혹은 실행 시간, 발생된 고장 수와 고장 발생 시간과의 관계를 효율적으로 관리함으로써 소프트웨어 신뢰도를 성장 시킬 수 있다. 이러한 과정을 소프트웨어 성장 과정이라고 볼 수 있다.

본 논문에서는 Burr 분포를 이용한 신뢰도 모형을 제안하였다. 고장 간격시간으로 구성된 NTDS 자료를 이용하여 기존의 모형과 Burr 분포 신뢰도 모형에 대하여 최우 추정법을 이용하여 모수 추정을 실시하였다. 효율적인 모형 비교를 위한 편차자승합의 결과는 Burr 분포 신뢰도 모형($a=1,2$)들이 기존에 잘 알려진 Yamada-Ohba-Osaki 모형이나 Goel-Okumoto보다 우수함을 보이고 있고 콜모고로프 거리의 결과도 Burr 분포 신뢰도 모형($a=1,2$)이 상대적으로 정상성에 가까운 모형으로 간주할 수 있다. 분석된 자료에 대한 평가에서도 산술평균 검정과 라플라스 추세 검정을 실시한 결과도 신뢰성장이 되고 있음을 나타내고 있다. 신뢰도에 대한 평가도 역시 Burr 모형이 신뢰도가 높음을 알 수 있다. 따라서 Burr 분포를 이용한 신뢰도 모형도 이 분야에서 효율적으로 이용할 수 있는 모형이 됨을 알 수 있었다. 향후 이러한 Burr 분포를 이용한 베이지안 접근 방법이나 수리적인 추정과 검정 부분에 대한 수리적인 접근이 기대 된다.

참고문헌

[1] 김 희철, 최 유순, 박 종구, “어랑분포를 이용한 NHPP 소프트웨어 신뢰성장 모형에 관한 연구”, 한국해양정보통신학회논문지, 10권1호. pp.7-14, 2006

[2] S. S. Gokhale and K. S. Trivedi. “A time/structure based software reliability model”. Annals of Software Engineering, 8, pp. 85-121. 1999

[3] 김 희철, “카이제곱 NHPP에 의한 소프트웨어 신뢰성 모형에 관한 연구”, 한국컴퓨터정보학회논문지, 제 11권 1호, pp. 45-53, 2006.

[4] I. W. Burr, “Cumulative frequency functions”. Ann. Math. Statistic. Vol 13, pp 215-232, 1942.

[5] S. K. Upadhyay and I. A. Javed and M. Peshwani, “Bayesian analysis of generalized four-parameter Burr distribution via Gibbs sampler”, METRON-International Journal of statistics, Vol, LXII, n.1, pp.115-135, 2004.

[6] J. F. Lawless, “Statistical Models and Methods for Lifetime Data”, John Wiley & Sons, New York, 1981.

[7] H. Pham and L. Nordmann and X. Zhang, “A General Imperfect-Software -Debugging Model with S-Shaped Fault-Detection Rate”, IEEE Trans. on reliability, Vol, 48, No 2, pp, 169-175, 1999.

[8] L. Kuo and T. Y. Yang, “Bayesian Computation of Software Reliability”, Journal of the American Statistical Association, Vol.91, pp. 763-773, 1996.

[9] A. L. Goel and K. Okumoto, “Time-Dependent Error-Detection Rate Models for Software Reliability and Other Performance Measures”. IEEE Trans. on Reliability, R-28(3):pp. 206-211, Aug. 1979.

[10] S. Yamada, M. Ohba and S. Osaki. “S-Shaped Reliability Growth Modeling for Software Error Detection”. IEEE Trans. on Reliability. R-32(5): pp. 475-485, Dec. 1983.

[11] J. A. Austin, “Control chart constants for largest and smallest in sampling from a normal distribution using the generalized Burr estimation”, Technometrics, Vol. 15, pp. 931-933, 1971.

[12] S. D. Dubey, “Statistical treatment of certain life testing and reliability problems”, ARL TR pp. 73-0155, AD 774537, 1973.

[13] P. K. Tadikamalla, “A look at the Burr and related distributions”, Inter. Statist. Rev., 48, pp. 337-344, 1980

[14] V. K. Rohatgi, “Statistical inference”, pp. 398-416. JOHN WILEY & SONS, INC, New York, 1984.

[15] T. A., Mazzuchi and R., Soyer, “A Bayes Empirical Bayes Model for Software Reliability”. IEEE Transactions on Reliability, 37, pp 248-254, 1988

[16] K. Kanoun and J. C. Laprie. Handbook of Software Reliability Engineering, M.R.Lyu, Editor, chapter Trend Analysis, pp.401-437.

[17] S. Brocklehurst and B. Littlewood, “Handbook of Software Reliability Engineering”, M.R.Lyu, Editor, chapter Techniques for Prediction Analysis and Recalibration, pp.119-166, 1996

[18] S. Selvin, “Modern Applied Biostatistical Methods Using S-Plus”, pages 141-184. Oxford University Press, New York, 1998.

저자소개

김 희 철(Hee-Cheul Kim)



1992년 동국대학교 통계학과 석사
1998년 동국대학교 통계학과 박사
2000년 3월 ~ 2004년 2월 송호대학
정보산업계열 조교수

2005년 3월 ~ 현재 남서울대학교 산업경영공학과 전임
강사

※관심분야: 소프트웨어 신뢰성 공학, 웹 프로그래밍,
전산통계, 인터넷비즈니스

박 종 구(Jong-Goo Park)



1975년 동국대학교 전자정보처리학과
석사
1999년 동국대학교 통계학과 박사

1981년 ~ 현재 원광대학교 전기전자 및 정보공학부 교수
※관심분야: 전문가시스템, 소프트웨어공학, 소프트웨어
신뢰성 공학, 모바일 프로그래밍