# Testing Exponentiality of Kullback-Leibler Information Function based on a Step Stress Accelerated Life Test

Byung Gu Park[1] · Sang Chul Yoon[2]


## Abstract

In this paper a test of fit for exponentiality and we propose the estimator of Kullback-Leibler Information functions using the data from accelerated life tests. This acceleration model is assumed to be a tampered random variable model. The procedure is applicable when the exponential parameter based on the data from accelerated life tests is or is not specified under null hypothesis. Using Simulations, the power of the proposed test based on use condition of accelerated life test under alternatives is compared with that of other standard tests in the small sample.


*Keywords* : test-of-fit, accelerated life tests, tampered random variable, step stress, Kullback -Leibler Information function.

## 1. 서 론

일반적으로 대부분의 수명검사는 정상조건에서 이루어져 왔다. 그러나 수명시험은 많은 시간과 비용을 요구하게 되므로 이를 해결하기 위해서 가속 수명시험 (accelerated life tests) 방법이 널리 이용되어 왔는데 이는 관심있는 시스템을 정상조건보다 더 높은 스트레스를 (예를들면 온도, 압력, 진동등) 준 조건에서 시험하는 방법이다. 이런 조건에서 시험한 자료를 이용하여 정상 조건에서의 시스템의 평균수명 등을 추론하는 것이다.

가속 수명시험에서 널리 이용되고 있는 방법 중의 하나는 단계 스트레스 가속수명시험 (step stress accelerated life tests)으로서 적당한 시간 간격을 두어 단계적으로 스트레스를 변화시키면서 시스템의 수명을 시험하는 방법이다. 단계 스트레스 가속수명시험에서 사용되는 모형 중 하나는 DeGroot와 Goel (1979) 이 제안한 변환확률변수 (tampered random variable ; TRV) 모형이다.

DeGroot와 Goel (1979)은 변환확률변수가 어떤 미지의 가속인자 $\alpha$에 의해 변화시간 $a$의 단위에 대한 나머지 수명의 곱으로 나타나는 낮은 스트레스에서 높은 스트레스로 변하는 효과를 제안하였다. 이 모형을 표현하기 위하여 $T$를 정상 수준에서 수명시간이라 하고 $Y$를 단계 스트

1) Professor, Department of Statistics, Kyungpook National University, Taegu, 702-701.
E-mail : bgpark@knu.ac.kr

2) Lecture, Department of Statistics, Kyungpook National University, Taegu, 702-701.
E-mail : scy@stat.knu.ac.kr

레스 가속수명시험에서의 수명시간이라 하면 단위 시험의 수명분포는 다음과 같이 표현된다.

$$Y=\begin{cases} T, & T< a \\ a+\alpha(T-a), & T> a \end{cases} \tag{1.1}$$

여기에서 $Y$는 변환확률변수 (TRV)이고, $a$는 변환점 (tampering point)이라 하며, $\alpha$는 변환계수 (tampering coefficient)라고 한다. 또한 $0<\alpha<1$이다.

Bhattacharrya 와 Soejoeti (1989)는 변환확률변수 (TRV)에서 회귀모형을 포함하는 추론을 제안하였다. Xiong (1998)은 고장 중단된 경우의 자료를 이용한 모수 추론으로 확장하였다.

한편, 많은 학자들이 쿨백-라이블러 정보 (Kullback-Leibler Information) 함수의 확률변수에 대한 엔트로피 (entropy) 추정량을 제안한 바 있다. 그 예로서, Vasicek (1976), van Es (1992), Correa (1995), Ebrahimi, Habibullahi 와 Soofi (1992), Mazzuchi, Soofi와 Soyer (2000) 등이 있다.

하지만 최근 엔트로피와 관련한 많은 연구가 다루어지고 있는 가운데 모의실험을 통한 소표본 특성에 대한 연구 또한 활발하게 진행되고 있다. 예로서, Correa (1995)는 엔트로피에 대한 추정량을 새로 제안하고, 모의실험을 통하여 제안된 방법이 Vasicek의 추정량보다 평균제곱오차 관점에서 우수하다는 것을 밝혔다. Wieczorkowski 와 Grzegorzewski (1999)는 Correa의 엔트로피 추정량을 기초한 지수분포의 대한 적합도 검정통계량을 제시하였다.

Bessler, Chernoff 와 Marshall (1962)은 쿨백-라이블러 정보함수를 이용하여 가속 실험에 대한 최적 축차 계획법을 제안하였으나, 그 이후 쿨백-라이블러 정보 함수를 이용한 가속수명시험 연구는 거의 이루어진 바가 없었다 (박 · 윤 · 조 (1999)를 보라). 또한 쿨백-라이블러 정보함수의 제안은 기존의 연구인 Kolmogorov-Smirnov, Cramer von Mises, Anderson-Darling등의 적합도 검정보다 우수하다는 연구결과가 있지만 가속수명시험 문제에서는 쿨백-라이블러 정보함수를 이용하여 가속수명시험에 대한 적합도 검정문제에 관심을 가지고 연구하는 것이 타당하며, 좀 더 깊이 있는 논의가 필요할 것으로 생각된다.

이런 연유로 본 연구에서는 변환확률변수에서 수명시간이 지수분포를 따르고 서로 독립일 때 단계 스트레스 가속수명시험으로부터 얻은 자료를 이용하여 쿨백-라이블러 정보함수를 사용하여 추정량을 제안하고 가속수명시험의 정상조건에서 검정력관점에서 각 추정량들의 소표본 특성을 비교하고자 한다.

따라서, 이 논문에서는 van Es(1992)가 제안한 쿨백-라이블러 정보함수에 기초한 지수모형의 단계 스트레스 가속수명모형을 제안하고, 이를 이용하여 기존의 검정통계량과 검정력을 비교하여 분석한다. 2절에서는 쿨백-라이블러 정보함수의 추정을 위하여 지수모형의 단계 스트레스 가속수명모형에서 모수의 추정량을 구하고 그 점근적 성질을 알아보며, 쿨백-라이블러 정보함수 추정에 수반되는 단계 스트레스 가속수명모형에서의 엔트로피 추정량을 제시한다. 3절에서는 2절에서 구한 추정량을 이용하여 쿨백-라이블러 정보함수의 추정량들을 제안하고 그 점근적 성질들을 함께 검토한다. 또한 제시된 통계량의 기각 영역을 조사한다. 4절에서는 제안된 지수모형의 단계 스트레스 가속수명모형에서 추정량을 이용하여 모의실험을 통하여 각 추정량들의 소표본에 대한 특성을 정상조건에서 단계 스트레스 가속수명시험과 기존의 통계량을 검정력 관점에서 비교 논의한다.

## 2. 단계 스트레스 가속수명모형에서의 모수추정

먼저, 본 절은 쿨백-라이블러 정보함수의 추정을 위한 지수모형의 단계 스트레스 가속수명모

형에서 모수의 최대우도추정량를 검토하고, 이를 이용한 van Es의 엔트로피를 이용한 검정력을 제안하고 기존의 다른 통계량들과 검정력을 통하여 비교하고자 한다.

만일, 정상조건에서의 수명시간 $T$가 모수 $\theta$인 지수분포를 따른다면 단계 스트레스 가속수명 모형인 식 (1.1)에 대한 확률밀도함수는

$$f_Y(y \mid \theta) = \begin{cases} \theta \exp\{-\theta y\} & , y < a \\ \dfrac{\theta}{\alpha} \exp\left\{-\theta\left(a + \dfrac{y-a}{\alpha}\right)\right\} & , y > a \end{cases} \tag{2.1}$$

로 나타낼 수 있다.

따라서 식 (2.1)의 우도함수는

$$\begin{aligned} L(\theta, \alpha) &= \prod_{i=1}^{n} f_Y(y_i \mid \theta) \\ &= \theta^n \alpha^{-n_2} \exp\left\{-\theta\left[V + \left(n_2 a + \frac{1}{\alpha} W\right)\right]\right\} \end{aligned} \tag{2.2}$$

이다. 여기서 $n = n_1 + n_2$, $V = \sum_{i=1}^{n_1} y_i$이고 $W = \sum_{i=1}^{n_2} (y_i - a)$이다. 즉, $n_1$은 비변환 관찰값 (untampered observation) 의 개수이고, $n_2$는 변환 관찰값 (tampered observation)의 개수이다. 그리고 $V$는 비변환 관찰치의 합이며, $W$은 변환 관찰치와 변환점의 차에 대한 합이다.

따라서 모수 $\theta$와 $\alpha$에 대한 추정량은

$$\hat{\theta} = \frac{n - n_2}{V + n_2 a} \qquad \text{이고} \qquad \hat{\alpha} = \frac{W}{n_2} \frac{n - n_2}{V + n_2 a}$$

이다. 이때, 식 (1.1)로부터 $T_i < a$, $i = 1, 2, \cdots, n_1$인 경우는 $Y_i = T_i$가 성립하고, 이는 비변환(untampered)확률변수이다. 또 식(1.1)로부터 $T_i > a$, $i = n_1 + 1, n_1 + 2, \cdots, n$인 경우는 $T_i = (Y_i - a)/\alpha + a$의 관계가 성립하므로 식 (1.1)의 역치환 (inverse transformation)한 확률표본 $T_{n_1+1}, T_{n_1+2}, \cdots, T_n$은 모수 $\theta$인 지수분포를 따른다는 사실을 알 수 있다.

만약, $\hat{\alpha}$가 $\alpha$에 대한 추정량이라고 하면, $T_i > a$, $i = n_1 + 1, n_1 + 2, \cdots, n$인 확률변수에 대하여 $\hat{T}_i = (Y_i - a)/\hat{\alpha} + a$로 높은 스트레스 조건에서 관찰된 값을 정상조건에서의 값으로 치환 가능할 것 이다. 이렇게 치환한 관찰값들을 $\hat{T}_1$, $\hat{T}_2, \cdots$, $\hat{T}_{n_2}$로 두자.

이 때 정상조건에서의 관찰값 $T_i$는 $\hat{T}_i$, $i=1,2,\cdots,n_1$라 두고, 높은 스트레스 조건에서 관찰된 관찰값 $\hat{T}_i$, $i = 1,2,\cdots,n_2$는 $\hat{T}_i$, $i = n_1 + 1, n_1 + 2, \cdots, n$라 두자. 그러면 확률변수 $\hat{T}_1, \cdots$, $\hat{T}_n$ $\hat{T}_{n_1+1}, \cdots$, $\hat{T}_n$는 모수가 $\theta$인 지수분포를 하는 확률변수로 취급할 수 있다.

그러므로 $\hat{T}_1, \cdots$, $\hat{T}_{n_1}$, $\hat{T}_{n_1+1}, \cdots$, $\hat{T}_n$가 독립인 확률변수로서 확률밀도함수 $f(\hat{T}; \cdot)$과 분포함수 $F(\hat{T}; \cdot)$을 가진다고면 주어진 확률표본이 평균이 $\theta$을 갖는 지수분포를 단계 스트레스 가속수명을 따르는지를 검정하기 위한 관심있는 귀무가설은

$$H_0 : f(\hat{T}; \cdot) = f_0(\hat{T}; \theta), \tag{2.3}$$

이고 대립가설은

$$H_1 : f(\hat{T}; \cdot) \neq f_0(\hat{T}; \theta). \tag{2.4}$$

로 나타낼 수 있다. 여기서 $f_0(\tilde{T};\theta)$는 평균이 $\theta$을 갖는 식(1.1)에서 변환된 지수분포이다.

## 3. 단계 스트레스 가속수명모형을 이용한 쿨백-라이블러 정보함수들의 추정량

본 절은 단계 스트레스 가속수명모형의 van Es (1992) 추정량을 이용하여 쿨백-라이블러 정보함수와 엔트로피 추정량들을 제안하고 그 성질을 검토한다.

$\tilde{T}_1, \tilde{T}_2, \cdots, \tilde{T}_n$를 전체의 확률표본이라 하고, $f_0$를 모수 $\theta$을 가지는 지수분포의 확률밀도함수이고, $f$를 임의의 확률밀도함수라 하자. 이때 $f_0$에 대한 $f$의 쿨백-라이블러 정보함수는 아래와 같이 표현된다.

$$\begin{aligned} I(f,f_0) &= \int_{-\infty}^{\infty} \ln\left\{\frac{f(t)}{f_0}(t)\right\} f(t)\,dt \\ &= \int_{-\infty}^{\infty} \ln\{f(t)\} f(t)\,dt - \int_{-\infty}^{\infty} \ln\{f_0(t)\} f(t)\,dt \\ &= -H(f) - \int_{-\infty}^{\infty} \ln\{f_0(t)\} f(t)\,dt \end{aligned} \tag{3.1}$$

여기서

$$H(f) = \int_0^1 \ln\left\{\frac{d}{dp}F^{-1}(p)\right\} dp$$

로서 엔트로피 (entropy)라 한다. 쿨백-라이블러 정보함수는 확률밀도함수 $f_o$을 가지는 관측된 분포함수 $F_0$와 확률밀도함수 $f$를 가지는 모형함수 $F$와의 거리이다. 쿨백-라이블러 정보함수의 특징으로 $I(f:f_0) \geq 0$이고, 만약 $f=f_0$이면 $I(f:f_0)=0$이다. 따라서 $I(f:f_0)$의 값이 0에 가까워지면 관측된 분포 $F_0$는 본래의 모형분포 $F$와 같아진다.

$\tilde{T}_1, \tilde{T}_2, \cdots, \tilde{T}_n$가 확률밀도함수 식 (2.1)로부터 추출한 가속수명시험에서 추출된 확률표본일 때, van Es (1992)의 엔트로피 추정량에 기초한 $H(f)$에 대한 추정량을 $HE_{mn}$이라 둔다면 제안된 추정량은 다음과 같이 나타 낼 수 있다.

$$\begin{aligned} HE_{mn} &= \frac{1}{n-m}\sum_{i=1}^{n-m} \ln\left\{\frac{n+1}{m}\left(\tilde{T}_{(i+m)} - \tilde{T}_{(i-m)}\right)\right\} \\ &\quad + \sum_{k=m}^{n}\frac{1}{k} + \ln(m) - \ln(n+1) \end{aligned} \tag{3.2}$$

여기서 $m$은 $n/2$보다 작은 양의 정수로서 윈도우의 크기라고 한다. 그리고 $\tilde{T}_{(1)} \leq \cdots \leq \tilde{T}_{(n)}$는 표본의 크기가 $n$인 확률표본의 순서 통계량이다.

van Es의 $HE_{mn}$의 점근적 성질을 이용하여 단계 스트레스 가속수명모형에서 점근적성질을 다음과 같이 밝힐 수 있다.

**정리 3.1** 확률표본 $\tilde{T}_i$가 분포함수 $F$와 확률밀도함수 $f$를 갖고 $Var(\tilde{T}_i) < \infty$이면, $\tilde{T}_1, \cdots, \tilde{T}_n$가 $F$로부터 추출된 확률표본일 때 $n \to \infty$, $m \to \infty$에 대하여 $m/n \to 0$이면

$$HE_{mn} \xrightarrow{p} H(f) \tag{3.3}$$

이제, 지수분포의 확률표본 $Y_i$에 대하여 $\widetilde{T}_i = (Y_i - a)/\widehat{\alpha} + a$, $i = n_1+1, n_1+2, \cdots, n$로 변수 치환시켜 만든 지수분포의 단계 스트레스 가속수명모형에서의 귀무가설 $H_0$을 검정하기 위한 지수분포에 대한 쿨백-라이블러 정보함수, $I(f, f_{T_i})$는

$$\begin{aligned} I(f, f_{T_i}) &= -H(f) - \int_0^\infty \ln\{f_{T_i}(t)\} f(t)dt \\ &= -H(f) - \ln\theta + 1 \end{aligned} \tag{3.4}$$

표현 할 수 있다.

식 (2.5), 식 (3.2)와 식 (3.2)로부터 단계 스트레스 가속수명모형에서의 van Es의 쿨백-라이블러 정보함수 $I(f, f_{T_i})$의 추정량은 다음과 같이 제안한다.

$$KLHE_{mn} = \widehat{I}(f, f_{T_i}) = -HE_{mn} - \ln\widehat{\theta} + 1 \tag{3.5}$$

여기서 $\widehat{\theta}$은 식 (2.5)에 주어진 $\theta$의 최대우도추정량이다.

**정리 3.2** 변환확률변수 (TRV) $\widetilde{T}_1, \widetilde{T}_2, \cdots, \widetilde{T}_n$가 지수분포의 확률밀도함수에서 추출된 가속수명시험의 확률표본이고, $n \to \infty$, $m \to \infty$에 대하여 $m/n \to 0$이면

$$\widehat{I}(f, f_{T_i}) \xrightarrow{p} I(f, f_{T_i}) \tag{3.6}$$

를 따른다.

따라서 단계 스트레스 가속수명에서 확률밀도함수를 갖는 지수분포의 적합도를 검정하기 위한 통계량으로서 다음과 같이 제시한다.

$$\begin{aligned} IHE_{mn} = \exp(-KLHE_{mn}) &= \exp(HE_{mn} + \ln\widehat{\theta} - 1) \\ &= \frac{\exp(HE_{mn})}{\exp(-\ln\widehat{\theta}) + 1} \end{aligned} \tag{3.7}$$

여기서 지수분포와 주어진 단계 스트레스 가속수명 자료 분포와 쿨백-라이블러 정보함수에 대한 추정량들로서 이 통계량의 값들이 클수록 귀무가설이 기각될 가능성이 높아진다.

## 4. 모의실험을 통한 검정력 비교

3절에서 제안된 van Es 검정통계량과 기존의 검정통계량들 Kolmogorov-Smirnov, Cramer von Mises, Anderson-Darling, Kuiper, Watson등의 검정력을 모의실험을 통하여 비교 분석 한다.

검정통계량의 검정력 비교를 위하여 대립가설의 분포는 와이블분포, 감마분포, 균일분포, 로그 정규분포등을 고려한다.

## 참고문헌


Bhattacharrya, G. K. and Soejoeti, Z. (1989), A Tampered Failure Rate Model for Step -Stress Accelerated Life Test. *Communications in Statistics-Theory and Method*, A18, 1627-1643.

Correa, J. C. (1995), A New Estimator of Entropy, *Communications in Statistics-Theory and Method*, A24, 2439-2449.

DeGroot, M. H. and Goel, P. K. (1979), Bayesian Estimation and Optimal Design in Partially Accelerated Life Testing. *Naval Research Logistics Quarterly*, Vol. 26, 223-235.

Ebrahimi, N., Habibullah, M., and Soofi, E.S. (1992), Testing Exponentiality based ob Kullback-Leibler Information, *Journal of Royal Statistical Society*, Ser. B, Vol. 54, 739-748.

Mazzuchi, T. A., Soofi, E.S. and Soyer, R. (2000), Computation of Maximum entropy Dirichlet for modeling lifetime data, *Computational Statistics & Data Analysis*, 32, 361-378.

Miller, R. and Nelson, W. (1983), Optimum Simple Step-Stress Plans for Accelerated Life Testing, *IEEE Transactions on Reliability*, Vol. 32. 59-65.

van Es, B. (1992), Estimating Functionals Related to a Density by a Class of Statistics Based on Spacings, *Scandinavian Journal of Statistics*, Vol. 19, 61-72.

Vasicek, O. A. (1976), A Test for Normality Based on Sample Entropy, *Journal of Royal Statistical Society*, Ser. B, Vol. 38, 55-59.

Wieczorkowski, R. and Grzegorzewski, P. (1999), Entropy Estimators Improvement and Comparison, *Communications in Statistics-Simulations*, B28, No. 2, 541-567.

Xiong, C., (1998), Inferences On a Simple Step-Stress Model with Type-II Censored Exponential Data, *IEEE Transactions on Reliability*, Vol. 47, 142-146.

박병구 · 윤상철 · 조건호 (1999), 단계 스트레스 가속수명모형을 이용한 쿨백-라이블러 정보 함수의 추정, 응용통계연구, 제13권 2호, 563-573.