*Journal of the Korean Data & Information Science Society*
*2013,* **24**(6), 1309–1317

http://dx.doi.org/10.7465/jkdi.2013.24.6.1309
한국데이터정보과학회지

# 일반화 지수분포를 따르는 제 1종 구간 중도절단표본에서 모수 추정

조영석[1] · 이창수[2] · 신혜정[3]

[1]부산대학교 통계학과 · [23]경운대학교 항공운항학과

접수 2013년 9월 10일, 수정 2013년 10월 1일, 게재확정 2013년 10월 11일

**요 약**

일반화 지수분포 (generalized exponential distribution)를 따르는 점진 제 1종 구간 중도절단 (progressive type-I interval censoring) 표본에서 모수 추정은 Chen과 Lio (2010)가 최대우도 추정법 (maximum likelihood estimation), 중간점 근사법 (mid-point approximation method), EM 알고리즘 (expectation maximization algorithm), 적률 추정법 (method of moments estimation; MME)으로 하였으며, 그 방법들 중 평균제곱오차 (mean square error; MSE)가 가장 작은 추정법은 중간점 근사법이다. 하지만 중간점 근사법을 바탕으로 최대우도 추정법을 이용하여 모수를 추정하려고 한다면 모수에 대한 해를 전개할 수 없기 때문에 수치 해석적인 방법을 이용하여 추정하여야 한다. 본 논문에서는 이러한 문제를 해결하기 위해서 근사 최대우도 추정법 (approximate maximum likelihood estimation)을 이용하여 두 종류의 모수를 추정하고, 모의실험을 통하여 수치해석학적인 방법을 이용한 중간점 근사법의 해 (estimate of mid-point approximation method; MP)와 제시한 두 가지 추정량을 평균제곱오차 측면에서 비교한다.

주요용어: 근사 최대우도 추정량, 일반화 지수분포, 점진 제 1종 구간 중도절단, 점진 제 2종 중도절단, 중간점 근사법.

## 1. 서론

Gupta와 Kundu (1999)가 제시한 일반화된 지수분포의 확률밀도함수 (probability density function; pdf)와 누적분포함수 (cumulative distribution function; cdf)는 다음과 같다.

$$g(t;\alpha,\lambda)=\frac{\alpha}{\lambda}e^{-t/\lambda}(1-e^{-t/\lambda})^{\alpha-1}, \tag{1.1}$$

$$G(t;\alpha,\lambda)=(1-e^{-t/\lambda})^{\alpha},\quad \alpha,\lambda>0,\ t\geq 0, \tag{1.2}$$

여기서 $\alpha=1$ 일 때, 일반화 지수분포는 지수분포를 따르게 된다. 식 (1.1) 밀도함수는 만약 $\alpha<1$일 때는 감소함수이고, $\alpha>1$일 때는 단봉함수이다. 이에 대한 연구로 Gupta와 Kundu (2003)는 수명 (life time) 자료 분석에서 일반화된 지수분포가 사용될 수 있다는 것을 언급하였고, 완전 확률표본에 기초한 일반화된 지수분포의 연구도 Gupta와 Kundu (2007)가 하였다.

[1] (609-735) 부산광역시 금정구 부산대학로 63번길 2(장전동), 부산대학교 통계학과, 교수.
[2] (730-739) 경북 구미시 산동면 강동로 730, 경운대학교 항공운항학과, 부교수.
[3] 교신저자: (730-739) 경북 구미시 산동면 강동로 730, 경운대학교 항공운항학과, 조교수.
E-mail: hjshin@ikw.ac.kr

경제, 의학, 사회과학 등의 자료에서는 중도절단 된 자료가 빈번히 발생한다. 이러한 중도 절단된 자료의 가장 일반적인 형태가 구간 중도절단 (interval censoring) 이다. 특히, Kalbfleisch와 Prentice (2002)는 사건 발생시간 (failure time) 자료의 통계적 분석방법을 연구 하였고, 수명 자료의 통계적 모형과 분석방법 연구는 Lawless (2003)가 했으며, Sun (2006)은 구간 중도절단된 자료에 대하여 다양한 방법들을 연구하였다. 중도절단 자료 중 Aggarwala (2001)는 점진 제 1종 중도절단 자료에서 지수분포에 대한 추정에 관한 연구를 했고, 지수와이블족에 대해서는 Ashour와 Afify (2007)가 연구했다. 로그정규분포는 Amin (2008)이 연구했고, 와이블분포는 Ng와 Wang (2009)이 연구했으며, 일반화 지수분포는 Chen과 Lio (2010)가 연구하였다. Shin 등 (2010)도 점진 제 1종 중도절단 자료에서 지수 분포를 바탕으로 한 추정 방법을 연구하였으며, 또한 Shin과 Lee (2012)는 부분 결측 값을 가지는 점진 제 1종 중도절단 자료에서 지수분포를 근거로 한 추정 방법을 연구하였다.

근사 최대우도 추정량 (approximate maximum likelihood estimate; AMLE)은 일반적인 최대우도 추정 방법으로 풀리지 않는 모형에 대해 테일러 급수 전개 (Taylor series)를 통하여 근사적으로 모수를 추정함으로써 수치해석학적으로 접근하는 방법 대신 근사식을 사용함으로써 추정 값을 제시할 수 있다. 근사 최대우도 추정량에 대해 Balakrishnan (1989a, 1989b)이 레일리분포와 정규분포에 관한 연구를 하였고, Kang 등 (1999)은 제2종 중도절단 (type II censoring)을 바탕으로 한 레일리분포에 관한 연구를, Kang 등 (2005)은 지수분포에 관해 연구를 하였다. 그리고 Asgharzadeh (2009)이 점진 제2종 중도절단 (progressive type II censoring)을 바탕으로 한 일반화 지수분포에 관한 연구를 하였다.

본 논문에서는 점진 제 1종 구간중도 절단 자료에서 일반화 지수분포를 근거로 한 모수 추정을 근사 최대우도 추정 방법을 이용하여 모수에 대한 근사적 최대우도 추정량을 연구하고자 한다.

2절에서는 점진 제1종 구간 중도절단자료에 대해 설명하고, 근사적 최대우도 추정량을 통해 일반화 지수분포의 척도모수를 추정한다. 3절에서는 모의실험을 통해 추정량을 비교하고, 결론을 맺는다.

## 2. 일반화 지수분포의 척도모수 추정

이 절에서는 점진 제1종 구간 중도절단 자료를 간단히 설명하고, Chen과 Lio (2010)가 소개한 점진 제1종 구간 중도절단 자료에서 일반화 지수분포의 중간점 근사법과 본 논문에서 제시하는 근사 최대우도 추정법을 소개하고자 한다.

### 2.1. 점진 제 1종 구간중도 절단자료

Aggarwala (2001)에 의한 점진 제 1종 구간 중도절단자료는 미리 정해진 시간 $t_1, t_2, \cdots, t_m$ ($m$은 고정된 수)에서 각 구간에서 사건이 발생한 개체 수를 $X_1, X_2, \cdots, X_m$, 각 구간에서 임의로 제거되는 개체 수를 $R_1, R_2, \cdots, R_m$으로 둔다. 총 $n$개의 개체 중에서 첫 번째 $(0, t_1]$구간에서 발생 한 개체 수는 $X_1$이며 $X_1$을 제외한 $n - X_1$중에서 임의로 제거되는 개체 수는 $R_1$이다. 두 번째 구간 $(t_1, t_2]$에서는 $n - X_1 - R_1$중에서 발생한 개체 수가 $X_2$이며, 임의로 제거되는 개체 수는 $n - X_1 - X_2 - R_1$중에서 $R_2$이다. 이렇게 $m$번째 구간 $(t_{m-1}, t_m]$에서 발생할 개체 수는 $X_m$이고, 이 구간에서 임의로 제거하는 개체 수는 $n - \sum_{i=1}^{m} X_i - \sum_{i=1}^{m-1} R_i = R_m$으로 모두 제거한다. 따라서 확률변수 $X_1, X_2, \cdots, X_m$는 $(0, t_1], (t_1, t_2], \ldots, (t_{m-1}, t_m]$구간에서 각각 사건이 발생하는 개체 수이며 발생된 개체 수의 총합은 $\sum_{i=1}^{m} X_i$, 임의로 제거되는 개체 수의 총합은 $\sum_{i=1}^{m} R_i$이고 총 개체 수는 $n = \sum_{i=1}^{m} (X_i + R_i)$ 로 나타난다. 이러한 방법으로 얻어진 자료를 점진 제 1종 구간 중도절단 자료라고 한다.

### 2.2. 중간점 근사법

일반화 지수분포를 바탕으로 한 점진 제 1종 구간 중도절단표본에서 Chen과 Lio (2010)는 최대우도 추정 방법, 중간점 근사법, EM 알고리즘, 적률 추정법들을 통해 모수를 추정하였으며, 그 방법들 중 평균제곱오차가 가장 작은 중간점 근사법은 다음과 같다.

일반화 지수분포의 확률밀도함수 식 (1.1)과 누적분포함수 식 (1.2)에 의하여 Chen과 Lio (2010)의 점진 제 1종 구간 중도절단자료에서의 일반화 지수분포의 우도함수는 다음과 같다.

$$\begin{aligned} L(\lambda) &= C[G(t_1,\alpha;\lambda)]^{X_1}[1-G(t_1,\alpha;\lambda)]^{R_1} \\ &\quad \times [G(t_2,\alpha;\lambda)-G(t_1,\alpha;\lambda)]^{X_2}[1-G(t_2,\alpha;\lambda)]^{R_2} \\ &\quad \times \cdots \times [G(t_m,\alpha;\lambda)-G(t_{m-1},\alpha;\lambda)]^{X_m}[1-G(t_m,\alpha;\lambda)]^{R_m} \\ &= C\prod_{i=1}^{m}[G(t_i,\alpha;\lambda)-G(t_{i-1},\alpha;\lambda)]^{X_i}[1-G(t_i,\alpha;\lambda)]^{R_i}, \end{aligned} \tag{2.1}$$

여기서 $C = n(n-1-R_1)(n-2-R_1-R_2)\cdots(n-m+1-R_1-\cdots-R_{m-1})$, $t_0 = 0$이다.

Chen과 Lio (2010)가 소개한 중간점 근사법은 $(t_{i-1}, t_i)$ 구간의 $m_i = \frac{t_{i-1}+t_i}{2}$ 시점에서 발생한 실패 개체수를 $X_i$로 두고, $t_i$시점에서 임의로 제거되는 개체수를 $R_i$라 할 때, 식 (2.1)은 다음과 같은 간편한 모형으로 나타난다.

$$L(\lambda) \simeq C\prod_{i=1}^{m}[g(m_i,\alpha;\lambda)]^{X_i}[1-G(t_i,\alpha;\lambda)]^{R_i}. \tag{2.2}$$

따라서 로그우도 함수는

$$\begin{aligned} \ln L(\lambda) &\simeq C\sum_{i=1}^{m}[X_i\ln(g(m_i,\alpha;\lambda))+R_i\ln(1-G(t_i,\alpha;\lambda))] \\ &= \ln C + (\ln\alpha-\ln\lambda)\sum_{i=1}^{m}X_i - \frac{1}{\lambda}\sum_{i=1}^{m}X_i m_i \\ &\quad + (\alpha-1)\sum_{i=1}^{m}X_i\ln(1-e^{-m_i/\lambda}) + \sum_{i=1}^{m}R_i\ln(1-(1-e^{-t_i/\lambda})^{\alpha}) \end{aligned} \tag{2.3}$$

이고, $\lambda$를 추정하기 위하여 식 (2.3)을 $\lambda$에 대해 편미분을 하면 다음과 같다.

$$\begin{aligned} \frac{\partial \ln L(\lambda)}{\partial\lambda} &\simeq \alpha\sum_{i=1}^{m}\frac{R_i\frac{t_i}{\lambda^2}e^{-t_i/\lambda}(1-e^{-t_i/\lambda})^{\alpha-1}}{1-(1-e^{-t_i/\lambda})^{\alpha}} \\ &\quad + \frac{1}{\lambda^2}\sum_{i=1}^{m}X_i m_i - \frac{\sum_{i=1}^{m}X_i}{\lambda} - (\alpha-1)\sum_{i=1}^{m}\frac{X_i\frac{m_i}{\lambda^2}e^{-m_i/\lambda}}{1-e^{-m_i/\lambda}}. \end{aligned} \tag{2.4}$$

식 (2.4)는 일반적인 최대우도 추정방법으로는 $\lambda$에 대하여 정확한 추정값으로 전개할 수 없기 때문에 수치해석적인 방법을 이용하여 모수를 추정해야 한다. 따라서 본 논문에서는 근사 최대우도 추정방법을 이용하여 근사식을 이끌어 내어 모수의 해를 구하고자 한다.

### 2.3. 근사 최대우도 추정량

확률변수 $y = t/\lambda$로 두면, 일반화 지수분포에서 $y$의 확률밀도함수와 누적분포함수는 다음과 같다.

$$f(y) = \alpha e^{-y}(1-e^{-y})^{\alpha-1},\ \alpha>0,\ y>0, \tag{2.5}$$

$$F(y) = (1-e^{-y})^{\alpha},\ \alpha>0,\ y>0. \tag{2.6}$$

식 (2.2)에서 $y_i = t_i/\lambda$, $z_i = m_i/\lambda$로 두면 우도함수는 다음과 같이 나타낼 수 있다.

$$L(\lambda) \simeq C \prod_{i=1}^{m} \left[\frac{1}{\lambda} f(z_i)\right]^{X_i} [1 - F(y_i)]^{R_i} . \tag{2.7}$$

식 (2.7)의 로그 우도함수를 $\lambda$에 대하여 편미분을 하면 다음과 같다.

$$\begin{aligned} \frac{\partial \ln L(\lambda)}{\partial \lambda} &\simeq -\frac{1}{\lambda} \sum_{i=1}^{m} X_i \frac{f'(z_i) z_i}{f(z_i)} - \frac{1}{\lambda} \sum_{i=1}^{m} X_i + \frac{1}{\lambda} \sum_{i=1}^{m} R_i y_i \frac{f(y_i)}{1 - F(y_i)} \\ &= -\frac{1}{\lambda} \left[\sum_{i=1}^{m} X_i + \sum_{i=1}^{m} X_i \frac{f'(z_i) z_i}{f(z_i)} - \sum_{i=1}^{m} R_i y_i \frac{f(y_i)}{1 - F(y_i)}\right] = 0. \end{aligned} \tag{2.8}$$

식 (2.8)에서 $\frac{f'(z_i)}{f(z_i)}$와 $\frac{f(y_i)}{1-F(y_i)}$때문에 $\lambda$에 관한 정확한 형태의 해를 전개 할 수 없다. 이러한 이유로 $\lambda$에 관한 해를 구하기 위해 근사 최대우도 추정방법을 이용한다.

근사 최대우도 추정은 Balakrishnan (1989a, 1989b)이 제시한 방법으로, 본 논문에서는 $\frac{f'(z_i)}{f(z_i)}$와 $\frac{f(y_i)}{1-F(y_i)}$를 테일러 급수전개를 통해 1차식으로 근사식을 유도하고, 다른 함수식 $\frac{f'(z_i)z_i}{f(z_i)}$와 $\frac{f(y_i)y_i}{1-F(y_i)}$도 테일러 급수전개를 통하여 같은 방법으로 근사식을 유도하고 두 가지 방법을 이용하여 추정된 모수를 비교하고자 한다.

먼저, $\frac{f'(z_i)}{f(z_i)}$와 $\frac{f(y_i)}{1-F(y_i)}$의 함수식을 $\xi_i = F^{-1}(p_{i:m:n})$ 지점에서 테일러 급수 전개를 할 때, Balakrishnan과 Aggarwala (2000)에 의하여

$$\begin{aligned} p_{i:m:n} &= E(U_{i:m:n}) \\ &= 1 - \prod_{j=m-i+1}^{m} \frac{j + R_{m-i+1} + \cdots + R_m}{j + 1 + R_{m-i+1} + \cdots + R_m} \end{aligned}$$

이고, 여기서 $U_{i:m:n}$는 균일분포 $U(0,1)$에서 $i$번째 점진 제 2종 중도절단 순서 통계량이다.

$\frac{f'(z_i)}{f(z_i)}$와 $\frac{f(y_i)}{1-F(y_i)}$를 테일러 급수 전개를 통해 1차식으로 근사시키면 다음과 같다.

$$\frac{f'(z_i)}{f(z_i)} \simeq (\beta_i + \gamma_i z_i), \tag{2.9}$$

$$\frac{f(y_i)}{1 - F(y_i)} \simeq \delta_i + \eta_i y_i, \tag{2.10}$$

여기서

$$\begin{aligned} \beta_i &= (\alpha - 1)\left[\frac{e^{-\xi_i}}{1 - e^{-\xi_i}} + \frac{\xi_i e^{-\xi_i}}{(1 - e^{-\xi_i})^2}\right] - 1, \\ \gamma_i &= -(\alpha - 1)\frac{e^{-\xi_i}}{(1 - e^{-\xi_i})^2}, \\ \delta_i &= \frac{1}{1 - F(\xi_i)}\left[f(\xi_i) - f'(\xi_i)\xi_i - \frac{f^2(\xi_i)\xi_i}{1 - F(\xi_i)}\right], \\ \eta_i &= \frac{1}{1 - F(\xi_i)}\left[f'(\xi_i) + \frac{f^2(\xi_i)}{1 - F(\xi_i)}\right]. \end{aligned}$$

그러면 식 (2.8)을 테일러급수전개를 통해 구한 근사식은 다음과 같이 나타낼 수 있다.

$$\begin{aligned}\frac{\partial \ln L(\lambda)}{\partial \lambda} &\simeq -\frac{1}{\lambda}\left[\sum_{i=1}^{m} X_i + \sum_{i=1}^{m} X_i(\beta_i+\gamma_i z_i)z_i - \sum_{i=1}^{m} R_i(\delta_i+\eta_i y_i)y_i\right] \\ &= -\frac{1}{\lambda}\left[\sum_{i=1}^{m} X_i\lambda^2 + \left[\sum_{i=1}^{m} X_i\beta_i m_i - \sum_{i=1}^{m} R_i\delta_i t_i\right]\lambda + \sum_{i=1}^{m} X_i\gamma_i {m_i}^2 - \sum_{i=1}^{m} R_i\eta_i {t_i}^2\right] = 0.\end{aligned} \tag{2.11}$$

이를 이용하여 $\lambda$에 관한 이차식을 근의 공식을 이용하여 양의 근을 구하게 되면 다음과 같다.

$$\hat{\lambda}_{AMLE1} = \frac{-B+\sqrt{B^2-4AC}}{2A}, \tag{2.12}$$

여기서

$$\begin{aligned}A &= \sum_{i=1}^{m} X_i,\ B = \sum_{i=1}^{m} X_i\beta_i m_i - \sum_{i=1}^{m} R_i\delta_i t_i, \\ C &= \sum_{i=1}^{m} X_i\gamma_i {m_i}^2 - \sum_{i=1}^{m} R_i\eta_i {t_i}^2.\end{aligned}$$

또한, $\frac{f'(z_i)z_i}{f(z_i)}$와 $\frac{f(y_i)y_i}{1-F(y_i)}$의 함수식을 $\xi_i = F^{-1}(p_{i:m:n})$지점에서 테일러 급수 전개를 통해 1차식으로 근사시키면 다음과 같다.

$$\frac{f'(z_i)z_i}{f(z_i)} \simeq {\beta'}_i + {\gamma'}_i z_i, \tag{2.13}$$

$$\frac{f(y_i)y_i}{1-F(y_i)} \simeq {\delta'}_i + {\eta'}_i y_i, \tag{2.14}$$

여기서

$$\begin{aligned}{\beta'}_i &= (\alpha-1)\left[\frac{e^{-\xi_i}\xi_i}{1-e^{-\xi_i}} - \frac{e^{-\xi_i}(1-e^{-\xi_i}-\xi_i)\xi_i}{(1-e^{-\xi_i})^2}\right], \\ {\gamma'}_i &= (\alpha-1)\frac{e^{-\xi_i}(1-e^{-\xi_i}-\xi_i)}{(1-e^{-\xi_i})^2} - 1, \\ {\delta'}_i &= \frac{1}{1-F(\xi_i)}\left[-f'(\xi_i){\xi_i}^2 - \frac{f^2(\xi_i){\xi_i}^2}{1-F(\xi_i)}\right], \\ {\eta'}_i &= \frac{1}{1-F(\xi_i)}\left[\frac{f'(\xi_i)\xi_i+f(\xi_i)}{1-F(\xi_i)} + \frac{f^2(\xi_i)\xi_i}{1-F(\xi_i)}\right].\end{aligned}$$

그러면 식 (2.8)을 테일러 급수 전개를 통해 구한 근사식은 다음과 같이 나타낼 수 있다.

$$\begin{aligned}\frac{\partial \ln L(\lambda)}{\partial \lambda} &\simeq -\frac{1}{\lambda}\left[\sum_{i=1}^{m} X_i + \sum_{i=1}^{m} X_i({\beta'}_i+{\gamma'}_i z_i) - \sum_{i=1}^{m} R_i({\delta'}_i+{\eta'}_i y_i)\right] \\ &= -\frac{1}{\lambda}\left[\left[\sum_{i=1}^{m} X_i + \sum_{i=1}^{m} X_i{\beta'}_i - \sum_{i=1}^{m} R_i{\delta'}_i\right]\lambda + \sum_{i=1}^{m} X_i{\gamma'}_i m_i - \sum_{i=1}^{m} R_i{\eta'}_i t_i\right] = 0.\end{aligned} \tag{2.15}$$

이를 이용하여 $\lambda$에 관한 일차식 근을 구하게 되면 다음과 같다.

$$\hat{\lambda}_{AMLE2} = \frac{-\sum_{i=1}^{m} X_i{\gamma'}_i m_i + \sum_{i=1}^{m} R_i{\eta'}_i t_i}{\sum_{i=1}^{m} X_i + \sum_{i=1}^{m} X_i{\beta'}_i - \sum_{i=1}^{m} R_i{\delta'}_i}. \tag{2.16}$$

## 3. 모의실험

Aggarwala (2001)에 의한 점진 제 1종 구간 중도절단 표본의 사건이 발생한 개체수 $X_1, X_2, \cdots, X_m$와 $R_i$를 생성하는 알고리즘은 다음과 같다.

$X_1 \sim BIN[n, G(t_1)]$이고, $i = 2, 3, \cdots, m$에 대하여

$$\begin{aligned} & X_i | X_{i-1}, \cdots, X_1, R_{i-1}, \cdots, R_1 \\ & \sim BIN\left[n - \sum_{j=1}^{i-1}(X_j + R_j), \frac{G(t_i) - G(t_{i-1})}{1 - \sum_{j=1}^{i-1}[G(t_j) - G(t_{j-1})]}\right] \\ & = BIN\left[n - \sum_{j=1}^{i-1}(X_j + R_j), \frac{G(t_i) - G(t_{i-1})}{1 - G(t_{i-1})}\right]. \end{aligned}$$

다음과 같이 m개의 이항 확률변수를 발생시키는 알고리즘을 수행한다.

- 단계 1. 각각의 변수를 $i = 0$, $xsum = 0$, $rsum = 0$로 초기화 한다.
- 단계 2. 다음 $i$로,
- 단계 3. 만약, $i = m + 1$이면 알고리즘을 끝낸다.
- 단계 4. $BIN[n - xsum - rsum, G(t_i) - G(t_{i-1})/[1 - G(t_{i-1})]]$을 따르는 $X_i$를 발생시킨다.
- 단계 5. $R_i = floor[p_i(n - xsum - rsum - X_i)]$을 계산한다.
- 단계 6. $xsum = xsum + X_i$, $rsum = rsum + R_i$을 계산한다.
- 단계 7. 단계 2로 돌아간다.

여기서 $t_0$는 개체 $n$개의 관찰 시작시점이고 끝나는 시점은 $t_m$이며, $p_1, p_2, \cdots, p_{m-1}$, $(p_m = 1)$는 정해진 시간 $t_1, t_2, \cdots, t_m$으로 나누어진 $m$개의 각 구간에서 제거 할 개체들의 비율이고, 이것과 함께 $R_m = n - \sum_{i=1}^{m} X_i - \sum_{i=1}^{m} R_i$은 고정된 수이다.

일반화 지수분포를 바탕으로 점진 제 1종 구간 중도절단표본에서 식 (2.8)의 해 $MP$를 수치 해석학적인 방법인 이분법 (bisection method)을 이용하여 구한 평균제곱오차와 제시한 두 가지 추정량 $\hat{\lambda}_{AMLE1}$와 $\hat{\lambda}_{AMLE2}$의 평균제곱오차를 비교하기 위하여 본 논문에서는 개체 $n = 30, 50$, 구간 $m = 9$, 간격 $t = 0.5$, 모수 $\alpha = 0.9, 2$, $\lambda = 0.5, 1.5$로 두고, 각 구간에서 개체들의 제거할 비율은 $P_1 = (0.2, 0.2, 0.2, 0.2, 0, 0, 0, 0, 0, 1)$, $P_2 = (0.5, 0.5, 0.5, 0.5, 0, 0, 0, 0, 0, 1)$로 했을 때, 1000번 반복하여 모의실험을 실시한 결과는 다음 Table 3.1과 같다.

**Table 3.1** The MSE of $MP$, $\hat{\lambda}_{AMLE1}$ and $\hat{\lambda}_{AMLE2}$

| | | | $P_1$ | | | $P_2$ | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $\alpha$ | $\lambda$ | $MP$ | $\hat{\lambda}_{AMLE1}$ | $\hat{\lambda}_{AMLE2}$ | $MP$ | $\hat{\lambda}_{AMLE1}$ | $\hat{\lambda}_{AMLE2}$ |
| 30 | 0.9 | 0.5 | 0.0949424 | 0.0009865 | 0.0011313 | 0.0828587 | 0.0010460 | 0.0013419 |
| | | 1.5 | 0.2415405 | 0.0109000 | 0.0119900 | 0.3786083 | 0.0178100 | 0.0202500 |
| | 2 | 0.5 | 0.0180952 | 0.0009734 | 0.0005158 | 0.0155885 | 0.0025178 | 0.0006957 |
| | | 1.5 | 0.1630433 | 0.0129900 | 0.0070500 | 0.2380221 | 0.0393700 | 0.0126700 |
| 50 | 0.9 | 0.5 | 0.0901574 | 0.0006712 | 0.0008168 | 0.0745206 | 0.0006455 | 0.0008748 |
| | | 1.5 | 0.1542370 | 0.0064900 | 0.0072100 | 0.2359320 | 0.0103800 | 0.0132100 |
| | 2 | 0.5 | 0.0142492 | 0.0008777 | 0.0003210 | 0.0097684 | 0.0026315 | 0.0003927 |
| | | 1.5 | 0.1175097 | 0.0121600 | 0.0046400 | 0.2142455 | 0.0267800 | 0.0083500 |

Table 3.1의 결과를 보면 제시한 추정량 $\hat{\lambda}_{AMLE1}$와 $\hat{\lambda}_{AMLE2}$의 평균제곱오차가 $MP$의 평균제곱오차 보다 더 작게 나타나 제시한 추정량이 평균제곱오차 측면에서 일치성을 가지며 더 좋은 추정량임을 알 수 있다. $\alpha = 0.9$인 경우는 $\hat{\lambda}_{AMLE1}$의 평균제곱오차가 $\hat{\lambda}_{AMLE2}$의 평균제곱오차 보다 더 작게 나타나 $\hat{\lambda}_{AMLE1}$가 평균제곱오차 측면에서 일치성을 가지며 더 좋은 추정량임을 알 수 있고, $\alpha = 2$인 경우는 $\hat{\lambda}_{AMLE2}$의 평균제곱오차가 $\hat{\lambda}_{AMLE1}$의 평균제곱오차 보다 더 작게 나타나 $\hat{\lambda}_{AMLE2}$가 평균제곱오차 측면에서 일치성을 가지며 더 좋은 추정량임을 알 수 있다.

본 논문에서는 일반화 지수분포를 바탕으로 한 점진 제 1종 중도절단 표본에서 모수를 추정하기 위하여 평균제곱오차가 가장 작은 중간점 근사법에서 근사 최대우도 추정법을 이용하여 테일러 급수 전개를 달리하여 모수의 추정량을 두 가지로 추정하였다. 모의실험을 통하여 두 가지 추정량 $\hat{\lambda}_{AMLE1}$와 $\hat{\lambda}_{AMLE2}$의 평균제곱오차와 $MP$의 평균제곱오차를 비교한 결과, 본 논문에서 제시한 추정량이 평균제곱오차 측면에서 일치성을 가지며 더 좋은 추정량으로 나타났다. $\alpha = 0.9$인 경우는 $\hat{\lambda}_{AMLE1}$가 평균제곱오차 측면에서 일치성을 가지며 $\hat{\lambda}_{AMLE2}$보다 더 좋은 추정량으로 나타났고, $\alpha = 2$인 경우는 $\hat{\lambda}_{AMLE2}$가 평균제곱오차 측면에서 일치성을 가지며 $\hat{\lambda}_{AMLE1}$보다 더 좋은 추정량으로 나타났다. 또한, 중간점 근사법을 이용하여 모수를 추정하려면 모수에 관한 정확한 형태의 해를 전개하기 어렵지만 본 논문에서 제시한 추정방법을 이용하면 모수에 대한 해를 전개 할 수 있다.

## References


Aggarwala, R. (2001). Progressive interval censoring: Some mathematical results with applications to inference. *Communications in Statistics-Theory and Methods*, **30**, 1921-1935.

Ashour, S. K. and Afify, W. M. (2007). Statistical analysis exponentiated weilbull family under type I progressive interval censoring with random removal. *Journal of Applied Sciences Resarch*, **3**, 1851-1863.

Amin, Z. H. (2008). A note on the parameter estimation for the lognormal distribution distribution based on progressively type I interval censored samples. *Model Assisted Statistics Application*, **3**, 169-176.

Asgharzadeh, A. (2009). Approximate MLE the scaled generalized exponential distribution under progressive type II censoring. *Journal of the Korean Statistical Society*, **38**, 223-229.

Balakrishnan, N. (1989a). Approximate maximum likelihood estimation of the mean and standard deviation of the normal distribution based on type II censored samples. *Journal of Statistical Computation and Simulation*, **32**, 137-148.

Balakrishnan, N. (1989b). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring. *IEEE Transactions on Reliability*, **38**, 355-357.

Balakrishnan, N. and Aggarwala, R. (2000). *Progressive censoring: Theory, methods and applications*, Birkhauser, Boston.

Chen, D. G. and Lio, Y. L. (2010). Parameter estimation for generalized exponential distribution under progressive type I interval censoring. *Computational Statistics and Data Analysis*, **54**, 1581-1591.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, **41**, 179-196.

Gupta, R. D. and Kundu, D. (2003). Discriminating between weibull and generalized exponential distributions. *Computational Statistics and Data Analysis*, **43**, 117-130.

Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, **137**, 3537-3547.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The statistical analysis of failure time data*, Second edition, John Wiley, New York.

Kang, S. B., Cho, Y. S. and Hwang, K. M. (1999). AMLE for the Rayleigh distribution with type II censoring. *Journal of the Korean Data & Information Science Society*, **10**, 405-413.

Kang, S. B., Lee, S. K. and Choi, H. T. (2005). Reliability estimation for the exponential distribution under multiply type II censoring. *Proceeding of Autumn Conference of the Korean Data & Information Science Society*, 13-26.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data*, John Wiley, New York.

Ng, T. H. K and Wang, Z. (2009). Statistical estimation for the parameters of weibull distribution based on progressively type I interval censored sample. *Journal of Statistical Computation and Simulation*,

**79**, 145-159.

Shin, H. J., Lee, K. H. and Cho, Y. S. (2010). Parameter estimation for exponential distribution under progressive type I interval censoring. *Journal of the Korean Data & Information Science Society*, **21**, 927-934.

Shin, H. J. and Lee, K. H. (2012). Estimation in the exponential distribution under progressive type I interval censoring with semi-missing data. *Journal of the Korean Data & Information Science Society*, **23**, 1271-1277.

Sun, J. (2006). *The statistical analysis of interval-censored failure time data*, Springer Verlag, New York.

*Journal of the Korean Data & Information Science Society*
*2013,* **24**(6), 1309–1317

http://dx.doi.org/10.7465/jkdi.2013.24.6.1309
한국데이터정보과학회지

# Estimation for the generalized exponential distribution under progressive type I interval censoring

Youngseuk Cho[1] · Changsoo Lee[2] · Hyejung Shin[3]

[1]Department of Statistics, Pusan National University
[23]Department of Flight Operation, Kyungwoon University

Received 10 September 2013, revised 1 October 2013, accepted 11 October 2013

**Abstract**

There are various parameter estimation methods for the generalized exponential distribution under progressive type I interval censoring. Chen and Lio (2010) studied the parameter estimation method by the maximum likelihood estimation method, mid-point approximation method, expectation maximization algorithm and methods of moments. Among those, mid-point approximation method has the smallest mean square error in the generalized exponential distribution under progressive type I interval censoring. However, this method is difficult to derive closed form of solution for the parameter estimation using by maximum likelihood estimation method. In this paper, we propose two type of approximate maximum likelihood estimate to solve that problem. The simulation results show the obtained estimators have good performance in the sense of the mean square error. And proposed method derive closed form of solution for the parameter estimation from the generalized exponential distribution under progressive type I interval censoring.

*Keywords*: Approximate maximum likelihood estimation, generalized exponential distribution, mid-point approximation method, progressive type I interval censoring, progressive type II interval censoring

[1] Professor, Department of Statistics, Pusan National University, Busan 609-735, Korea.
[2] Associate professor, Department of Flight Operation, Kyungwoon University, Geongbuk 730-739, Korea.
[3] Corresponding author: Assistant professor, Department of Flight Operation, Kyungwoon University, Geongbuk 730-739, Korea. E-mail: hjshin@ikw.ac.kr