

## Minimum Disparity Estimation for Normal Models: Small Sample Efficiency

M. J. Cho<sup>1)</sup>, C. S. Hong<sup>2)</sup>, and D. B. Jeong<sup>3)</sup>

### Abstract

The minimum disparity estimators introduced by Lindsay and Basu (1994) are studied empirically. An extensive simulation in this paper provides a location estimate of the small sample and supplies empirical evidence of the estimator performance for the univariate contaminated normal model. Empirical results show that the minimum generalized negative exponential disparity estimator (MGNEDE) obtains high efficiency for small sample sizes and dominates the maximum likelihood estimator (MLE) and the minimum blended weight Hellinger distance estimator (MBWHDE) with respect to efficiency at the contaminated model.

*Keywords:* Disparity measure, Efficiency, Minimum Hellinger distance estimator, Minimum negative exponential disparity estimator, Newton-Raphson method.

### 1. 서론

지난 오랜 동안 모수추정에 대해 통계학자들이 근본적으로 중요하게 다룬 두 가지 논제가 있었다. 하나는 모형이 적절하게 주어질 때 추정량의 효율성(efficiency)에 관한 것이고, 나머지는 참모형으로부터 벗어나 있을 때 추정량의 로버스트성(robustness)에 관한 것이다. 두 논제가 개념 상에 적지 않은 논란이 있는데 최대가능도추정량(Maximum Likelihood Estimator, MLE)은 추정량들 사이에서 점근효율(asymptotic efficiency)이 가장 높은 반면에 좋은 로버스트 성질을 갖는 추정량은 될 수 없으나 로버스트 M-추정량측은 모형에서 일차효율성을 희생하면서 로버스트성을 만족하였다(Hampel 외 1986).

효율성과 로버스트성 사이의 논란은 헬링거거리(Hellinger distance, HD)와 같은 밀도함수에 근거한 최소불일치추정량(Minimum Disparity Estimator, MDE)과 이에 관련된 연구(Beran, 1977; Stather, 1981; Tamura와 Boos, 1986; Simpson, 1987, 1989; Lindsay, 1994; Basu와 Lindsay, 1994)에 의해 적어도 부분적으로 조정되었다. 이에 관한 첫 시도로서 Beran(1977)은 최소헬링거거리추정량(Minimum Hellinger Distance Estimator, MHDE)이 이차효율성과 로버스트성을 동시에 만족할 수 있다는 사실을 발견하였다. 주어진 모형 하에서 점근효율이란 추정량이 최대가능도추정

1) Senior Researcher, Research Institute of Applied Statistics, Sungkyunkwan University, Seoul 110-745 Korea.  
E-mail : mjcho@skku.edu

2) Professor, Department of Statistics, Sungkyunkwan University, Seoul 110-745 Korea.

3) Associate Professor, Department of Information Statistics, Kangnung National University, Kangnung 210-702 Korea.

량(MLE)과 동일한 영향(력)함수(influence function)를 갖는다는 것을 의미한다. 앞에서 소개한 몇 개의 연구들은 최소헬링거리추정량(MHDE)이 이런 점에도 불구하고 로버스트 성질을 만족함을 나타낸다. Simpson(1987)은 포아송분포와 같은 몇몇의 이산분포에서 최소헬링거리추정량(MHDE)이 50% 붕괴(breakdown)가 있음을 보였다. Lindsay(1994)는 이산모형에서 소위 잔차조정 함수(Residual Adjustment Function, RAF)라 부르는 거리의 함수를 특징화시켜 최소거리추정량에 근거한 최소불일치추정량측에 대한 붕괴점(breakdown point)을 일반화시킨 결과를 도출시켰다.

Lindsay(1994) 그리고 Basu와 Sarkar(1994)는 수정된 불일치측도(Disparity Measure, DM)로써 가중헬링거리(Blended Weight Hellinger Distance, BWHD)를 제시하였고 이를 최소화시켜 얻은 추정량을 최소가중헬링거리추정량(MBWHDE)이라 하였다. 또한 Jeong과 Sarkar(2000)에 의해 제시된 일반화음지수불일치(Generalized Negative Exponential Disparity, GNED)측도를 최소화함으로써 최소일반화음지수불일치추정량(MGNEDE)을 얻게 된다.

본 논문에서는 최소일반화음지수불일치추정량(MGNEDE)이라 부르는 추정량이 소표본에서 주어진 정규모형뿐만 아니라 자료가 오염(contamination)된 상황에서도 최대가능도추정량(MLE)과 최소가중헬링거리추정량(MBWHDE)에 비해 어느 정도 효율성을 만족하는지 모의실험을 통해 알아보는데 목적이 있다. 2절에서는 일반화음지수불일치(GNED)와 가중헬링거리(BWHD)를 포함하는 불일치측도로부터 최소일반화음지수불일치추정량(MGNEDE)과 최소가중헬링거리추정량(MBWHDE)을 도출하는 방법에 대해 간략히 살펴본다. 그리고 3절에서는 정규분포의 위치모수에 대한 효율성을 세 개의 추정량(MLE, MGNEDE, MBWHDE)에 대해 살펴보기 위해서, 고려할 수 있는 여러 형태의 정규모형을 이에 적용시켰다. 끝으로 4절에서는 각 모의실험에서 얻은 결과를 구체적으로 나열하였으며, 5절은 4절에서 얻은 여러 결과에 근거한 전반적인 결론을 내린다.

## 2. 불일치측도를 이용한 모수추정

모수분포족  $\mathcal{F}_\theta = \{F_\theta, \theta \in \Theta\}$ ,  $\Theta \subset R^p$ 로부터 확률표본  $\{X_1, X_2, \dots, X_n\}$ 을 추출하자. 분포들  $F_\theta$ 는 연속이며, 이에 대응하는 확률밀도함수들을  $f_\theta$ 라고 하자. 이 때 최소불일치추정량들에 근거한 밀도함수(Lindsay 1994; Basu와 Lindsay 1994)는 비모수적 밀도추정값과 모형밀도  $f_\theta$ 사이의 양의 값을 갖는 불일치측도  $\rho_G$ 를 최소화시켜 계산한다. 그리고 불일치측도는 다음과 같이 정의된다.

$$\rho_G(f^*, f_\theta) \equiv \int G(\delta(f^*, \theta, x)) dF_\theta(x), \quad (2.1)$$

여기에서  $\delta(x) = \delta(f^*, \theta, x) = (f^*(x) - f_\theta(x)) / f_\theta(x)$ 는  $x$ 값에서 피어슨 잔차(Pearson residual)를 나타낸다. 그리고  $G$ 는 실수값을 갖고 세 번 미분가능하며, 구간  $[-1, \infty)$ 에서  $G(0) = 0$ 인 순볼록함수(strictly convex function)이고  $\delta = 0$ 일 때에만 등호가 성립하고 항상  $G(\delta) \geq 0$ 이다.  $F$ 는 경험적인 분포함수이고,  $w$ 는 평균이  $y$ 이고 표준편차가  $h$ 인 정규분포와 같은 핵함수(kernel function)의 평활족(smoothing family)이라고 할 때, 다음 정의한 비모수적 핵밀도추정량(nonparametric kernel density estimator)  $f^*(x)$ 를 사용할 수 있다.

$$f^*(x) \equiv \int w(x, y, h) dF(y) \tag{2.2}$$

참고로, 식 (2.1)은 쿨백-라이블러 발산(Kullback-Leibler divergence)을 일반화한 형태인  $f^*$ 와  $f_\theta$  사이의 Csiszar(1963)  $f$ -발산과 동일하다.

함수  $G$ 의 형태에 따라 다양한 불일치측도를 생성할 수 있다. 즉,  $G(\delta) = (\delta+1) \log(\delta+1) - \delta$ 는 가능도불일치,  $G(\delta) = \delta^2$ 은 Pearson의 카이제곱,  $G(\delta) = \{(\delta+1)^{\lambda+1} - 1\} / \lambda(\lambda+1)$ 은 멱발산족(power divergence family)(Cressie와 Read, 1984),  $G(\delta) = 2\{(\delta+1)^{1/2} - 1\}^2$ 은 두 배에 제공된 헬링거리, 그리고  $G(\delta) = e^{-\delta} - 2$ 는 음지수불일치(NED)를 생성한다.

$\nabla$ 을  $\theta$ 에 대한 경사도(gradient)라고 할 때, 불일치측도  $\rho_G(f^*, f_\theta)$ 를 최소화함으로써  $G$ 에 대응하는 최소불일치추정량을 얻을 수 있다. 특히,  $\rho_G$ 가  $f^*$ 와  $f_\theta$ 사이의 음지수불일치라면 이를 최소로 하여 최소음지수불일치추정량(MNEDE)을 얻어낼 수 있다.

$$\rho_{NED}(f^*, f_\theta) \equiv \int \{ \exp[-(f^*(x)/f_\theta(x)) - 1] - 2 \} dF_\theta(x)$$

또한 최소헬링거리추정량(MHDE)은 헬링거리인

$$\rho_{HD}(f^*, f_\theta) \equiv \int [(f^*(x)/f_\theta(x))^{1/2} - 1]^2 dF_\theta(x)$$

를 최소화시켜 구할 수 있다. 모형의 미분가능 하에서 최소불일치추정방정식은 다음과 같다.

$$-\nabla \rho_G = \int A(\delta(x)) \nabla F_\theta(x) = 0, \tag{2.3}$$

여기에서  $A(\delta) \equiv (\delta+1)[G'(\delta) - G(\delta)]$ 이고  $G'(\delta)$ 는  $G(\delta)$ 의 일차미분계수이다. 함수  $A(\delta)$ 는 구간  $[-1, \infty)$ 에서 증가함수이며, 불일치측도에 의해 생성된 추정값을 변화시키지 않고 표준화시킬 수 있다. 따라서 표준화된  $A(\delta)$ 에 대해  $A(0) = 0$ 과  $A'(0) = 1$ 을 얻게 된다. 이 때 표준화된 함수를 불일치의 잔차조정함수(RAF)라 부르며, 추정량의 대부분의 이론적인 성질들을 결정하게 된다. 제공된 헬링거리에 대한 잔차조정함수는  $A(\delta) = 2[(\delta+1)^{1/2} - 1]$ 이고 음지수불일치에 대한 잔차조정함수는  $A(\delta) = 2 - (2+\delta)e^{-\delta}$ 가 된다.

Basu, Sarkar 그리고 Vidyashankar(1997)는 최소음지수불일치추정량(MNEDE)이 연속형 모형에서 최소헬링거리추정량(MHDE)와 대등하게 좋은 결과를 나타낸다는 것을 보였다. 여기서 최소헬링거리추정량(MHDE)과 최소음지수불일치추정량(MNEDE)은 모두 이상점(outliers)<sup>4)</sup>에 대하여 로버스트하지만 모형 하에 기대도수보다 관찰도수가 더 작은 자료점으로 정의된 인라이어(inliers)에 대하여 최소음지수불일치추정량(MNEDE)만이 로버스트하다는 것을 증명하였다. 인라이어 또는

4) "Lindsay(1994): "Efficiency versus Robustness: The Case For Minimum Hellinger Distance And Related Method, p1095-1098"를 참고

이상점들에 대해 헬링거거리와 음지수불일치가 어떻게 반응하는지는 Lindsay(1994), Basu와 Lindsay(1994), Basu 외 2인(1997)을 참고하기 바란다. 헬링거거리와 음지수불일치에 의해 생성된 추정량들은 모두 일차 효율적이지만, 이차 효율적인 것은 음지수불일치추정량뿐이다.

최소불일치추정량은 불일치측도를 최소화함으로써 얻을 수 있고, 이는 최소불일치추정방정식의 해를 구하는 문제와 같게 됨을 알 수 있다. 이제 본 연구에 사용될 불일치측도에 대하여 알아보기로 한다. Lindsay(1994)는 보다 로버스트한 추정량으로서 수정된 불일치측도들을 제시하였다. 다음의 식 (2.4)는 가중헬링거거리측도(BWHD)이다.

$$BWHD(\lambda) = \int \frac{[f^*(x) - f_\theta(x)]^2}{2[\lambda\sqrt{f^*(x)} + \bar{\lambda}\sqrt{f_\theta(x)}]^2} dx, \tag{2.4}$$

여기서  $\lambda \in R$ 이고  $\bar{\lambda} = 1 - \lambda$ 이며, 잔차조정함수는  $A_\lambda(\delta) = \delta[w(\delta)]^{-2} + \frac{\bar{\lambda}}{2} \delta^2[w(\delta)]^{-3}$  이고, 여기에서  $w(\delta) = \lambda\sqrt{\delta+1} + \bar{\lambda}$  이다. Basu와 Sarkar(1994a)는  $\lambda = 1/9$  일 때 즉, BWHD (1/9) 를 최소화시켜 얻은 추정량(MBWHDE)이 떡발산족과 같이 다른 추정량의 비교에서 우수함을 보였다.

일반화음지수불일치측도(GNED)를 다음 식 (2.5)과 같이 정의할 수 있다.

$$GNED(\lambda) = \int \frac{\exp\left\{-\lambda\left(\frac{f^*(x)}{f_\theta(x)} - 1\right)\right\} - 1 + \lambda\left(\frac{f^*(x)}{f_\theta(x)} - 1\right)}{\lambda^2} \cdot f_\theta(x) dx \tag{2.5}$$

이때의 잔차조정함수는  $A(\delta(x)) = 2 - (2 + \delta(x))e^{-\delta(x)}$  이다. 또한 식 (2.5)의 불일치측도에서의  $G_\lambda(\delta)$  는  $\lambda$  에 따라 다음과 같이 표현된다.

$$G_\lambda(\delta) = \begin{cases} (e^{-\lambda\delta} - 1 + \lambda\delta)/\lambda^2, & \lambda \neq 0 \\ \delta^2/2, & \lambda = 0 \end{cases}$$

Jeong과 Sarkar(2000)는  $\lambda = 4/3$  일 때 즉, GNED (4/3) 을 최소로 하는 추정량(MGNEDE)이 떡발산추정량보다 우수할 수 있음을 제안하였다.

다음 절에서는 식 (2.4)와 (2.5)에서 정의된 BWHD (1/9) 와 GNED (4/3) 을 사용하여 이들을 최소화함으로써 모수를 추정하고, 각각의 최소불일치추정량인 최소일반화음지수불일치추정량(MGNEDE), 최소가중헬링거거리추정량(MBWHDE)과 최대가능도추정량(MLE)과의 상대적 효율성을 비교한다.

### 3. 모의실험 연구

본 절에서는 정규분포모형 설정 하에 GNED (4/3), BWHD (1/9) 그리고 최대가능도추정량(MLE)에 대한 이론적 결과들을 경험적으로 조사하기 위하여 몬테칼로 방법(Monte-Carlo

method)을 사용하기로 한다. 최소일반화음지수불일치추정량(MGNEDE)과 최소가중헬링거리추정량(MBWHDE)을 계산하기 위해서 필요한 밀도추정과 뉴턴-랩슨 알고리즘(Newton-Raphson algorithm)에 대해 살펴보기로 한다.

### 3.1 밀도추정

최소일반화음지수불일치추정량(MGNEDE)과 최소가중헬링거리추정량(MBWHDE)을 구하기 위해 밀도추정에서 최적의 성질을 갖는 biweight 핵을 사용하여(Simpson, 1989) 핵밀도함수  $f^*$ 를 계산한다.

$$f^*(x) = \frac{1}{nh_n} \sum_{i=1}^n \frac{15}{16} \left\{ 1 - \left( \frac{x - X_i}{h_n} \right)^2 \right\}^2 I \left| \frac{x - X_i}{h_n} \right| < 1 \left( \frac{x - X_i}{h_n} \right) \tag{3.1}$$

Parzen(1962)은 핵밀도추정값  $f^*$ 과 참밀도함수  $f_\theta$ 사이의 적분평균제곱오차(integrated mean squared error)를 최소화시키는 평활계수(smoothing efficient)  $h$ 값을 유도하였다. Devroye와 Gyorf(1985)의 biweight 핵에 관련된 평균  $L_1$  기준과 정규밀도함수  $f_\theta$ 를 고려하면 평활계수는  $h_n = (2.34 \sigma n^{-1/5})$ 이 된다. 반면에  $\sigma$ 을 모르는 경우는  $\tilde{\sigma}^{(0)} = \frac{(\text{median} | X_i - \text{median}(X) |)}{0.6745}$ 를 대신 사용할 수 있다.

### 3.2 모수추정과 효율성

<표 3-1> 모의분포의 유형

모형 1	분포	$\epsilon N(5, 1) + (1 - \epsilon)N(8, 1)$
	$N$	20, 40, 60, 100
	$\epsilon$	0, 0.05, 0.1, 0.15, 0.2
	$\sigma$	1
모형 2	분포	$\epsilon N(5, \sigma^2) + (1 - \epsilon)N(8, 1)$
	$N$	20, 40, 60, 100
	$\epsilon$	0, 0.05, 0.1, 0.15, 0.2
	$\sigma$	2, 3
모형 3	분포	$\epsilon N(5, \sigma^2) + (1 - \epsilon)N(8, 1)$
	$N$	100
	$\epsilon$	0.05
	$\sigma$	0.9, 0.8, 0.7, 0.6
모형 4	분포	$\epsilon N(8, \sigma^2) + (1 - \epsilon)N(8, 1)$
	$N$	20, 40, 60, 100
	$\epsilon$	0, 0.05, 0.1, 0.15
	$\sigma$	3

모의실험을 위한 표본추출집단의 모형은 표본크기, 오염수준(contaminated level), 오염분포(contaminated distribution)의 표준편차에 따라 그 형태가 다양하다. 이 다양성을 반영하기 위해 <표 3-1>에 주어진 네 개의 모형을 고려하기로 한다. 여기에서  $N$ 은 표본크기,  $\varepsilon$ 은 오염수준,  $\sigma$ 는 표준편차를 나타낸다.

2절에서 제시된 GNED (4/3) 과 BWHD (1/9) 를 최소화하여 두 모수추정량 최소일반화음지수불일치추정량(MGNEDE)과 최소가중헬링거리추정량(MBWHDE)을 구하는 알고리즘은 다음과 같다.

### (1단계) 자료 생성

먼저 <표 3-1>에 제시된 모의분포의 네 모형에 따라 표본크기, 오염수준, 표준편차 등을 조합하여 자료  $X_i$ 를 생성한다. 이에 대해 모형 2를 예로 들면 다음과 같다.

$$X_i \sim \varepsilon N(5, \sigma^2) + (1-\varepsilon)N(8, 1), \quad i = 1, \dots, n$$

### (2단계) 모수추정(뉴턴-랩슨 알고리즘 실행)

목표모수는  $\theta = (\mu, \sigma^2)'$ 이고, 세 추정량 GNED (4/3), BWHD (1/9) 와 최대가능도추정량(MLE)에 대하여 뉴턴-랩슨 알고리즘을 이용하여 모수를 추정한다. 각 추정량들을 정리해보면 다음과 같다.

① 최대가능도추정량(MLE):

$$\hat{\mu} = \bar{X}^* = \sum_{i=1}^n X_i^* / n, \quad \hat{\sigma}^2 = \sum_{i=1}^n (X_i^* - \hat{\mu})^2 / n$$

② 일반화음지수불일치(GNED) 측도:

$$\text{GNED (4/3)} = \int \frac{9}{16} f_\theta \cdot \left[ \exp\left\{\frac{4}{3} \left(1 - \frac{f^*}{f_\theta}\right)\right\} + \frac{4}{3} \frac{f^*}{f_\theta} - \frac{7}{3} \right] dx$$

③ 가중헬링거리(BWHD) 측도:

$$\text{BWHD (1/9)} = \int \frac{81}{2} \cdot \frac{[f^* - f_\theta]^2}{[\sqrt{f^*} + 8\sqrt{f_\theta}]^2} dx$$

### (2-1단계) 초기값 계산

(1단계)에서 생성된 자료로부터 초기값을 계산한다. 추정하고자하는 모수에 대한 초기값을  $\bar{\theta}^{(0)} = (\bar{\mu}^{(0)}, \bar{\sigma}^{2(0)})'$  라 하면, 다음과 같이 구할 수 있다.

$$\bar{\mu}^{(0)} = \text{median}(X_i), \quad i = 1, \dots, n.$$

$$\bar{\sigma}^{(0)} = (\text{median} | X_i - \text{median}(X_i) |) / 0.6745$$

## (2-2단계) 뉴턴-랩슨 알고리즘 실행

뉴턴-랩슨 알고리즘에 의한 모수추정 과정은 다음의 식 (3.2)를 반복함으로써 구할 수 있다.

$$\bar{\theta}^{(j+1)} = \bar{\theta}^{(j)} - \begin{bmatrix} \frac{\partial^2 \rho}{\partial \mu^2} & \frac{\partial^2 \rho}{\partial \mu \partial \sigma^2} \\ \frac{\partial^2 \rho}{\partial \sigma^2 \partial \mu} & \frac{\partial^2 \rho}{\partial (\sigma^2)^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial \rho}{\partial \mu} \\ \frac{\partial \rho}{\partial \sigma^2} \end{bmatrix} = \bar{\theta}^{(j)} - A^{-1} \mathbf{b} \quad (3.2)$$

그리고 불일치측도  $\rho(f^*, f_\theta) = \int G(\delta) f_\theta dx$  이므로 식 (3.2)는 다음과 같이 다시 표현된다.

$$\bar{\theta}^{(j+1)} = \bar{\theta}^{(j)} - \left[ \int \frac{\partial^2}{\partial \theta \partial \theta'} \{G(\delta) f_\theta(x)\} dx \right]^{-1} \left[ \int \frac{\partial}{\partial \theta} \{G(\delta) f_\theta(x)\} dx \right] \quad (3.3)$$

먼저 불일치측도에 의한 모수추정을 위하여 핵밀도추정을 한다.

## ① 핵밀도추정

초기값  $\bar{\sigma}^{(0)}$ 을 이용하여 평활계수  $h_n = (2.3) \bar{\sigma}^{(0)} n^{-1/5}$ 를 구한다. 구해진 평활계수에 따라 핵밀도함수의 정의역과 식 (3.3)의 적분구간이 결정된다. 적분구간은 평활계수와 추출된 자료의 순서통계량  $X_{(1)}, \dots, X_{(n)}$ 을 이용하여 구할 수 있다. 적분구간은  $(X_{(1)} - h_n, X_{(n)} + h_n)$ 이고 식 (3.1)의 핵밀도함수를 계산한다.

② 불일치측도 성분인  $G(\delta) f_\theta$ 의 일·이차 편미분 계산

모수에 대한 불일치측도  $\rho_G$ 의 성분  $G(\delta) f_\theta$ 의 일차 편미분인 벡터  $\mathbf{b}$ 와 이차 편미분인 행렬  $A$ 의 각 원소들을 계산한다. 이에 관련된 구체적인 식들은 부록을 참조하기로 한다.

## ③ 적분

행렬  $A$ 와 벡터  $\mathbf{b}$ 의 각 원소들에 대하여 적분구간에 따라 적분값을 계산한다. 이때 사용된 적분방법은 사다리꼴 적분법이다.

④  $A^{-1} \mathbf{b}$  계산

적분값이 계산된 행렬  $A$ 의 역행렬  $A^{-1}$ 을 구하고 그것을 벡터  $\mathbf{b}$ 와 곱한다.

⑤  $\bar{\theta}^{(1)} = (\bar{\mu}^{(1)}, \bar{\sigma}^{2(1)})'$ 

초기값  $\bar{\theta}^{(0)} = (\bar{\mu}^{(0)}, \bar{\sigma}^{2(0)})'$ 과 ④에서 계산된 값들의 차는 1회 반복 뉴턴-랩슨 알고리즘에 의해 얻어진 모수추정값  $\bar{\theta}^{(1)} = (\bar{\mu}^{(1)}, \bar{\sigma}^{2(1)})'$ 이 된다.

## ⑥ 반복실행과 정지규칙

(2-2단계)의 ①부터 ⑤까지의 반복을 통해 최적의 모수를 구하게 되는데 다음의 정지규칙을 따른다.

$$|\bar{\theta}^{(j+1)} - \bar{\theta}^{(j)}| = \sqrt{\{\bar{\mu}^{(j+1)} - \bar{\mu}^{(j)}\}^2 + \{(\bar{\sigma}^{(j+1)})^2 - (\bar{\sigma}^{(j)})^2\}^2} < 10^{-5} \quad (3.4)$$

**(3단계) 추정된 모수의 평균 구하기**

(1단계)부터 (2단계)를 5000번 반복 실행하여 모수  $\mu$ 와  $\sigma$ 들을 구하고, 표본  $X_i$ 의 위치모수인 평균  $\mu$ 에 대한 편의(Bias)와 평균제곱오차(Mean Squared Error, MSE)를 계산하여 최대가능도추정량(MLE)에 대한 불일치측도 가중헬링거리(BWHD)와 일반화음지수불일치(GNED)의 최소불일치추정량들(MBWHDE와 MGNEDE)의 상대효율(relative efficiency, Eff)을 비교한다.

**4. 모의실험 결과**

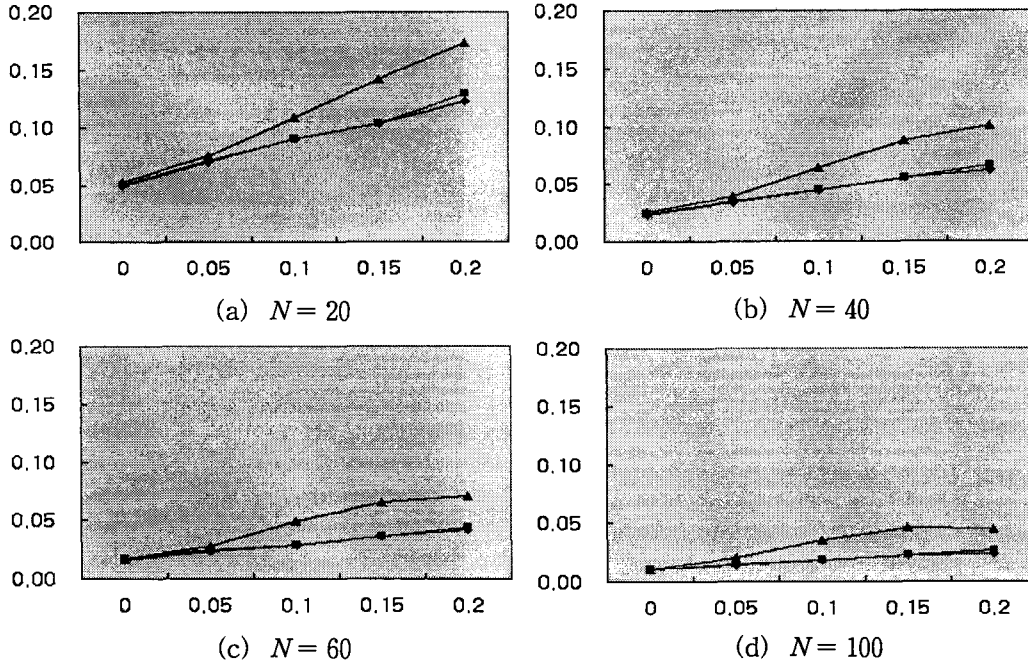
**4.1 모형 1에 대한 효율성**

모형 1은 모의분포  $\varepsilon N(\mu_1, 1) + (1 - \varepsilon)N(\mu_2, 1)$ 에서  $(\mu_1, \mu_2)$ 가 (5, 8)인 경우에 대하여 자료  $X_i$ 를 추출한 것이다. <표 4-1>은 오염수준 0.05와 0.1에 대하여 표본크기  $N = 20, 40, 60$  그리고 100에 따라 추정된 모수  $\mu$ , 표준편차, 편의, 평균제곱오차(MSE) 그리고 상대효율(Eff)을 구한 것이다.

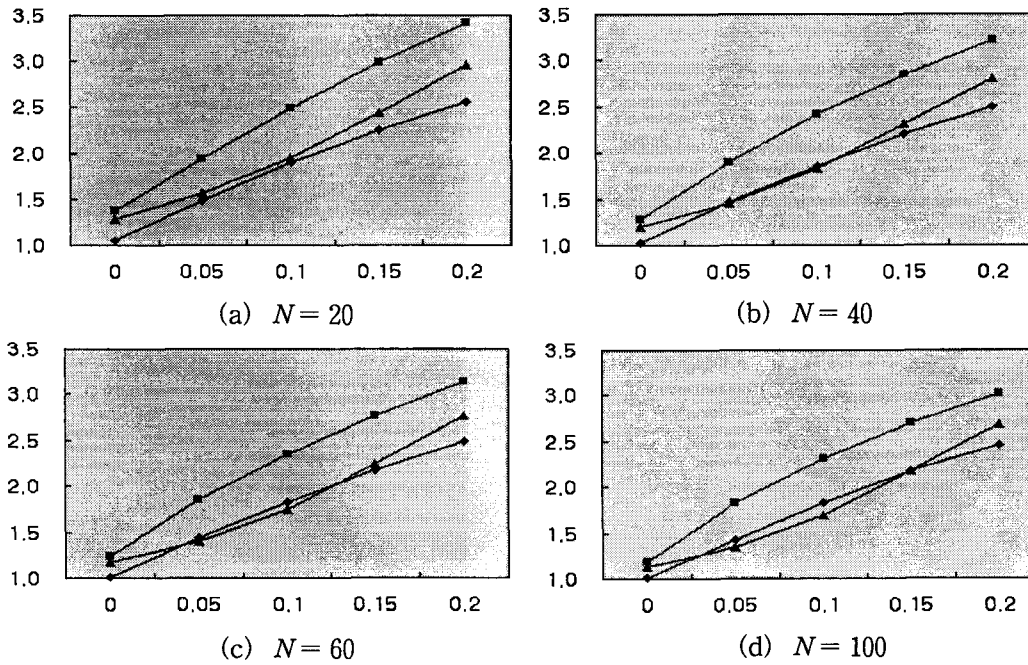
<표 4-1> 모형 1에 대한 MLE, MBWHDE, MGNEDE의 편의, MSE, Eff

N	추정량	$\varepsilon = 0.05$					$\varepsilon = 0.1$				
		$\mu$	$\sigma$	편의	MSE	Eff	$\mu$	$\sigma$	편의	MSE	Eff
20	MLE	7.856	1.167	0.070	1.489	—	7.699	1.319	0.090	1.898	—
	MBWHDE	7.877	1.335	0.071	1.935	0.770	7.728	1.516	0.090	2.488	0.763
	MGNEDE	7.932	1.190	0.076	1.561	0.954	7.818	1.319	0.108	1.959	0.969
40	MLE	7.853	1.186	0.035	1.470	—	7.696	1.334	0.045	1.859	—
	MBWHDE	7.861	1.347	0.035	1.895	0.775	7.721	1.522	0.045	2.412	0.771
	MGNEDE	7.932	1.175	0.040	1.453	1.011	7.820	1.305	0.064	1.826	1.018
60	MLE	7.848	1.183	0.024	1.443	—	7.703	1.335	0.029	1.834	—
	MBWHDE	7.850	1.341	0.024	1.855	0.778	7.720	1.514	0.029	2.354	0.779
	MGNEDE	7.925	1.162	0.028	1.399	1.032	7.825	1.289	0.049	1.748	1.049
100	MLE	7.853	1.187	0.014	1.434	—	7.698	1.341	0.018	1.831	—
	MBWHDE	7.845	1.337	0.015	1.823	0.786	7.710	1.507	0.018	2.311	0.792
	MGNEDE	7.929	1.147	0.020	1.348	1.063	7.820	1.278	0.035	1.691	1.082





<그림 4-1> 모형 1의 오염수준과 편의 (—◆— MLE —■— MBWHDE —▲— MGNEDE)



<그림 4-2> 모형 1의 오염수준과 MSE (—◆— MLE —■— MBWHDE —▲— MGNEDE)

먼저 <표 4-1>에서 편의에 대해 살펴보면 표본크기와 오염수준에 관계없이 MGNEDE가 MLE와 MBWHDE에 비해 다소 크다. 반면에 평균제곱오차(MSE)는 표본크기가 20인 경우를 제외하고

40 이상인 경우에 대하여 MGNEDE가 표본크기가 증가할수록 MLE와 MBWHDE보다 작아진다. 따라서 MLE에 대한 상대효율 즉,

$$Eff = \frac{MSE(ML\epsilon)}{MSE(MBWHDE)} \quad \text{또는} \quad \frac{MSE(ML\epsilon)}{MSE(MGNEDE)}$$

를 고려하면, MBWHDE는 MLE보다 효율적인 추정량이 되지 못하는 반면에 표본크기 40 이상인 경우에 MGNEDE가 MLE보다 더 효율적인 추정량이라 할 수 있다. 이 사실은 MGNEDE의 분산이 상대적으로 다른 추정량에 비해 작다는 것을 의미한다.

<그림 4-1>과 <그림 4-2>는 각 표본크기에 대하여 오염수준이 증가할 때, 편의와 평균제곱오차(MSE)를 그래프로 나타낸 것이다. 표본크기에 대하여 MLE와 MBWHDE는 오염수준이 증가하더라도 편의는 거의 차이가 없는 반면에, MGNEDE는 오염수준이 0.05 보다 클 때 다른 두 추정량에 비해 다소 편의가 커지기는 하지만 추정량간의 차이는 미미하다. 더구나 <그림 4-2>에서 보는 바와 같이, 편의의 값이 평균제곱오차(MSE)에 거의 영향을 미치지 못하고 있다. 결과적으로 MGNEDE는 다른 두 추정량 MLE와 MBWHDE에 비해 편의는 다소 크지만, 추정된 분산이 작기 때문에 평균제곱오차(MSE)가 작게 된다. 즉 제한된 조건하에 MGNEDE가 다른 두 추정량에 비해 효율적인 추정량이라 할 수 있다.

#### 4.2 모형 2에 대한 효율성

<표 4-2>  $\sigma = 2$ , MLE, MBWHDE, MGNEDE의 편의, MSE, Eff

N	추정량	$\epsilon = 0.05$					$\epsilon = 0.1$				
		$\mu$	$\sigma$	편의	MSE	Eff	$\mu$	$\sigma$	편의	MSE	Eff
20	MLE	7.851	1.215	0.075	1.658	—	7.701	1.401	0.109	2.219	—
	MBWHDE	7.887	1.335	0.074	1.985	0.835	7.760	1.554	0.106	2.727	0.814
	MGNEDE	7.954	1.157	0.073	1.472	1.126	7.887	1.254	0.121	1.799	1.234
40	MLE	7.850	1.230	0.039	1.607	—	7.696	1.430	0.053	2.172	—
	MBWHDE	7.864	1.377	0.042	2.038	0.789	7.727	1.645	0.053	2.901	0.749
	MGNEDE	7.954	1.139	0.042	1.367	1.175	7.894	1.214	0.078	1.597	1.360
60	MLE	7.851	1.239	0.027	1.604	—	7.699	1.437	0.036	2.152	—
	MBWHDE	7.852	1.407	0.031	2.107	0.761	7.717	1.693	0.035	3.018	0.713
	MGNEDE	7.957	1.122	0.033	1.310	1.225	7.898	1.199	0.064	1.529	1.408
100	MLE	7.851	1.242	0.016	1.584	—	7.699	1.438	0.021	2.119	—
	MBWHDE	7.831	1.457	0.020	2.219	0.714	7.699	1.734	0.022	3.108	0.682
	MGNEDE	7.956	1.110	0.023	1.265	1.252	7.897	1.175	0.054	1.450	1.461

모형 2는 모의분포  $\epsilon N(\mu_1, \sigma^2) + (1-\epsilon)N(\mu_2, 1)$  에서  $(\mu_1, \mu_2)$ 가 (5, 8) 이고  $\sigma$ 가 2와 3인 경우에 대하여 표본  $X_i$ 를 추출한 것이다. <표 4-2>는  $\sigma=2$ 인 경우로, 두 오염수준 (0.05, 0.1)이 주어져 있을 때 표본크기에 따라 위치모수  $\mu$ 에 대한 추정량들의 효율을 비교한 것이다.

먼저 편의에 대해 살펴보면 모형 1의 경우와 유사한 결과가 나타난다. 즉 오염수준과 표본크기에 관계없이 MGNEDE의 편의는 MLE와 MBWHDE의 것보다 큰 차이가 없다. 그러나 모형 2에서의 평균제곱오차(MSE)는 표본크기가 20인 소표본인 경우나 100인 대표본인 경우에 관계없이 MGNEDE의 평균제곱오차(MSE)가 MLE와 MBWHDE의 것보다 상대적으로 훨씬 작다. 따라서 효율성을 비교해보면 MGNEDE가 MLE보다 효율적인 추정량이 된다.

MGNEDE의 MLE에 대한 상대효율이 오염수준과 표본크기에 관계없이 높게 나타나고, 특히 표본수가 20인 소표본의 경우에 그 상대효율이 1.126으로 MGNEDE가 보다 효율적인 추정량임을 알 수 있다. 이러한 현상은 오염수준이 0.05에서 0.1로 높아질수록 상대효율이 더 높아진다. 또한  $\sigma=3$ 인 경우에는(표 생략) 이러한 현상이 더욱 두드러진다. 소표본인 경우 MGNEDE의 분산과 편의 모두가 MLE와 MBWHDE의 것들보다 작으므로, 모형 1의 경우보다 MGNEDE의 평균제곱오차(MSE)가 다른 추정량보다 더 작게 되며, 상대적 효율성이 더욱 높아지게 된다. 따라서 MGNEDE는 다른 두 추정량보다 효율적인 추정량이라 할 수 있다.

<표 4-3>  $\sigma < 1$ , MLE, MBWHDE, MGNEDE 의 편의, MSE, Eff

분포의 $\sigma$	추정량	$\epsilon = 0.05$					$\epsilon = 0.1$				
		$\mu$	$\sigma$	편의	MSE	Eff	$\mu$	$\sigma$	편의	MSE	Eff
0.9	MLE	7.850	1.186	0.014	1.431	—	7.697	1.334	0.018	1.811	—
	MBWHDE	7.844	1.326	0.014	1.792	0.799	7.709	1.488	0.018	2.250	0.805
	MGNEDE	7.922	1.156	0.019	1.369	1.046	7.810	1.290	0.034	1.723	1.051
0.8	MLE	7.850	1.184	0.014	1.427	—	7.700	1.328	0.018	1.794	—
	MBWHDE	7.847	1.317	0.014	1.764	0.809	7.713	1.474	0.018	2.207	0.813
	MGNEDE	7.921	1.162	0.019	1.382	1.032	7.805	1.301	0.034	1.752	1.024
0.7	MLE	7.851	1.178	0.014	1.413	—	7.701	1.321	0.018	1.773	—
	MBWHDE	7.850	1.303	0.014	1.726	0.818	7.713	1.459	0.018	2.160	0.821
	MGNEDE	7.918	1.166	0.019	1.392	1.015	7.797	1.313	0.032	1.781	0.996
0.6	MLE	7.851	1.178	0.013	1.410	—	7.698	1.318	0.017	1.765	—
	MBWHDE	7.851	1.297	0.014	1.709	0.825	7.709	1.453	0.017	2.143	0.824
	MGNEDE	7.917	1.171	0.019	1.406	1.003	7.788	1.323	0.031	1.807	0.977

### 4.3 모형 3에 대한 효율성

모형 3은 모의분포  $\epsilon N(\mu_1, \sigma^2) + (1-\epsilon)N(\mu_2, 1)$  에서  $(\mu_1, \mu_2)$ 는 (5, 8) 이고  $\sigma$ 는 1보다

작은 값 (0.9, 0.8, 0.7, 0.6)을 갖는 경우에 대하여 표본  $X_i$ 가 추출된 것이다. <표 4-3>는 표본크기  $N=100$ 에 대하여 오염수준이 각각 0.05와 0.1인 경우, 표준편차가 변함에 따라 추정된 모수  $\mu$ 의 편의, 평균제곱오차(MSE) 그리고 상대효율(Eff)을 구한 것이다.

오염수준이 0.05인 경우를 살펴보면  $\sigma$ 가 0.6에서 0.9로 갈수록 MGNEDE의 표준편차가 MLE와 MBWHDE보다 점점 더 작아짐을 알 수 있다. 그리고 모형 1과 2의 경우에서와 마찬가지로 MGNEDE의 편의가 MLE와 MBWHDE의 것보다 다소 크지만 큰 차이가 없다. 따라서  $\sigma$ 가 1에 가까워질수록 MGNEDE의 평균제곱오차(MSE)가 다른 두 추정량 MLE와 MBWHDE보다 더 작아지므로 MLE에 대한 MGNEDE의 상대효율이 높아짐을 알 수 있다.

오염수준 0.1인 경우에 MLE와 MBWHDE의 편의가 MGNEDE보다 오염수준 0.05일 때에 비해 차이가 더 크게 나타나며, 그 영향으로  $\sigma$ 가 0.6과 0.7인 경우에 MGNEDE의 평균제곱오차(MSE)가 MLE 보다 크게 되어 상대효율이 떨어졌다. 그러나  $\sigma$ 가 0.8과 0.9로써 1에 가까워질수록 MGNEDE의 표준편차가 다른 추정량보다 크게 작아짐으로 MLE에 대한 상대효율이 다시 높아지게 된다.

<표 4-4> 모형 4에 대한 MLE, MBWHDE, MGNEDE의 편의, MSE, Eff

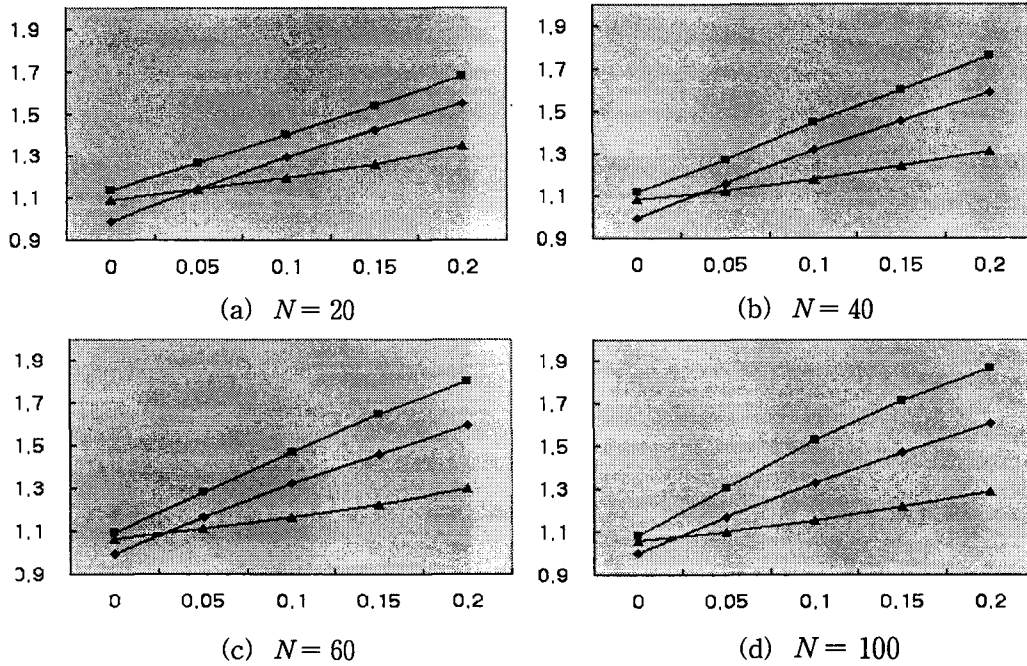
N	추정량	$\epsilon=0.05$					$\epsilon=0.1$				
		$\mu$	$\sigma$	편의	MSE	Eff	$\mu$	$\sigma$	편의	MSE	Eff
20	MLE	8.000	1.145	0.070	1.468	—	8.001	1.294	0.094	1.902	—
	MBWHDE	8.001	1.267	0.068	1.766	0.832	8.000	1.402	0.087	2.207	0.862
	MGNEDE	8.002	1.140	0.060	1.413	1.039	7.999	1.196	0.070	1.573	1.210
40	MLE	7.997	1.160	0.034	1.430	—	8.001	1.317	0.046	1.852	—
	MBWHDE	7.995	1.269	0.039	1.716	0.833	8.002	1.447	0.056	2.259	0.820
	MGNEDE	7.996	1.121	0.028	1.309	1.092	8.001	1.179	0.033	1.454	1.274
60	MLE	7.999	1.163	0.023	1.410	—	8.003	1.322	0.031	1.829	—
	MBWHDE	8.000	1.282	0.030	1.733	0.814	8.003	1.473	0.045	2.312	0.791
	MGNEDE	8.000	1.113	0.019	1.273	1.108	8.002	1.166	0.021	1.402	1.305
100	MLE	7.997	1.173	0.014	1.413	—	8.003	1.330	0.018	1.818	—
	MBWHDE	7.996	1.307	0.023	1.787	0.790	7.999	1.528	0.034	2.452	0.741
	MGNEDE	7.997	1.102	0.012	1.235	1.144	8.001	1.153	0.013	1.353	1.343

#### 4.4 모형 4에 대한 효율성

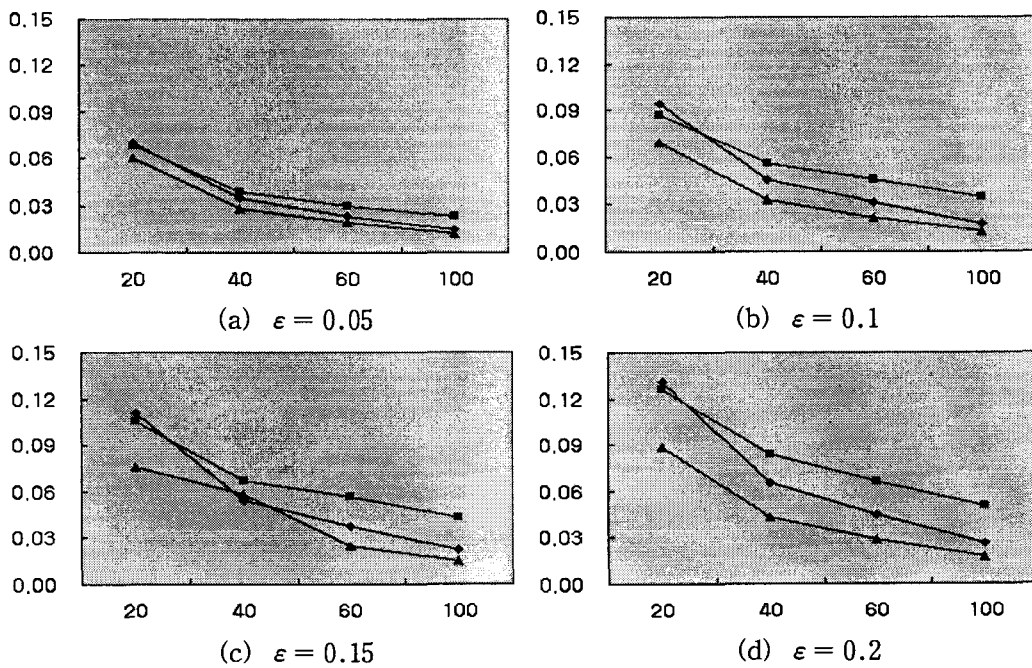
마지막으로 모의분포  $\epsilon N(8, 3^2) + (1-\epsilon)N(8, 1)$ 에서 오염수준 0, 0.05, 0.1, 0.15, 0.2와 표본크기 20, 40, 60, 100에 대하여 표본을 추출한다.

<표 4-4>는 두 오염수준 0.05와 0.1이 주어져 있을 때 표본크기 20, 40, 60, 100에 대하여 세 가지 추정량의 편의, 평균제곱오차(MSE) 그리고 상대효율(Eff)을 계산한 것이다. 모형 4의 경우는 모형 1과 모형 2와는 달리, 표본크기 또는 오염수준에 관계없이, MGNEDE의 편의가 MLE와

MBWHDE의 것보다 작고, 또 추정된 표준편차가 다른 어떤 추정량에 비해 작기 때문에 평균제곱 오차(MSE)가 가장 작게 되는 것은 당연한 결과이다.



<그림 4-3> 모형 4의 오염수준과 표준편차 (—◆— MLE—■— MBWHDE—▲— MGNEDE)

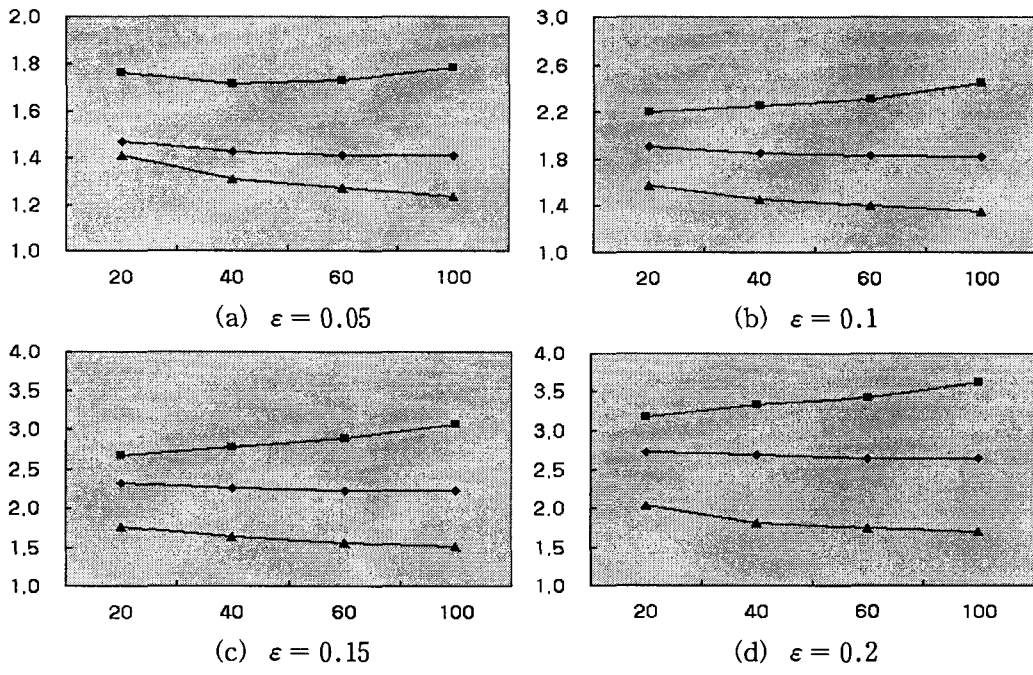


<그림 4-4> 모형 4의 표본크기와 편의 (—◆— MLE—■— MBWHDE—▲— MGNEDE)

<그림 4-3>은 표본크기 20, 40, 60, 100에 대하여 각 오염수준의 증가에 따른 표준편차의 변화를 그래프로 나타낸 것이다. 오염수준이 0인 경우를 제외한 모든 경우에서 MGNEDE의 표준편차가 MLE와 MBWHDE의 것 보다 작고, 특히 표본크기에 관계없이 오염수준이 커질수록 추정량들 간의 표준편차의 차이가 점점 커짐을 알 수 있다. 따라서 모형 2와 모형 3의 경우와 마찬가지로 MGNEDE의 상대적 효율이 더 높아지게 된다.

<그림 4-4>는 오염수준별로 각 표본크기를 변화시켰을 때 편의를 그래프로 나타낸 것이다. 오염수준이 높아질수록 추정량들 간의 편의의 차이가 커짐을 알 수 있고, 오염수준과 표본크기에 관계없이 MGNEDE의 편의가 MLE와 MBWHDE의 것보다 작게 나타난다.

<그림 4-5>는 오염수준별로 각 표본크기를 변화시켰을 때 평균제곱오차(MSE)를 나타낸 그래프이다. 결과적으로 오염수준이나 표본크기에 관계없이, MGNEDE에 대한 분산과 편의 모두가 작기 때문에 MLE와 MBWHDE의 것 보다 당연히 작게 된다. 따라서 MLE에 대한 MGNEDE의 상대적 효율이 높아진다.



<그림 4-5> 모형 4의 표본크기와 MSE (◆ MLE ■ MBWHDE ▲ MGNEDE)

### 5. 결론

본 논문에서는 정규모형의 위치모수에 대하여 불일치측도 일반화음지수불일치(GNED)와 가중헬링거거리(BWHD)를 사용하여 최소불일치추정량인 최소일반화음지수불일치추정량(MGNEDE)과 최소가중헬링거거리추정량(MBWHDE)을 구하고, 소표본인 경우의 모의실험을 통해 이 두 추정량과 최대가능도추정량(MLE)간의 효율성을 비교하였다. 각 모형에 따라 그 실험결과가 조금씩 차이가 있었으나 모형에 관계없이 최소일반화음지수불일치추정량(MGNEDE)의 분산이 다른 두 추정량 최대가능도추정량(MLE)과 최소가중헬링거거리추정량(MBWHDE) 보다 작게 나타남을 알 수 있었

다. 이는 오염수준과 표본크기에 상관없는 결과이다.

모형 1인 경우에 추정된 모수는 최대가능도추정량(MLE)에 비해 다른 두 추정량의 편의가 조금 높게 나타나지만 근소한 차이를 보이고 있다. 그러나 최소일반화음지수불일치추정량(MGNEDE)의 분산이 다른 추정량들에 비하여 작으며 추정량간의 분산의 차이가 편의에 의한 차이보다 크게 되어 평균제곱오차(MSE)가 최대가능도추정량(MLE)과 최소가중헬링거거리추정량(MBWHDE) 보다 작게 나타남으로써 상대효율이 높아졌다. 특히 오염수준 0.05와 0.1인 경우에 표본크기가 40이상의 자료집단에서의 추정된 모수는 최대가능도추정량(MLE)과 최소가중헬링거거리추정량(MBWHDE) 보다 최소일반화음지수불일치추정량(MGNEDE)의 평균제곱오차(MSE)가 작게 나타났고 따라서 최소일반화음지수불일치추정량(MGNEDE)이 더 효율적인 추정량이라 할 수 있다. 모형 2와 3의 경우에도 같은 결과가 나타났다. 모형 4의 모의분포에서는 최소일반화음지수불일치추정량(MGNEDE)의 분산과 편의가 모두 다른 두 추정량보다 작게 나타남으로 MSE는 더욱 상대적으로 작아지게 된다.

모의실험 결과, 최소일반화음지수불일치추정량(MGNEDE)이 다른 두 추정량 최대가능도추정량(MLE)과 최소가중헬링거거리추정량(MBWHDE) 보다 더 효율적이었다. 이러한 양상은 오염수준, 표본크기, 척도모수가 커짐에 따라 더 두드러진다. 따라서 최대가능도추정량(MLE)에 대한 최소일반화음지수불일치추정량(MGNEDE)의 상대효율 역시 표본크기, 오염정도, 척도모수가 커짐에 따라 최대가능도추정량(MLE)에 대한 최소가중헬링거거리추정량(MBWHDE)의 상대효율 보다 크게 된다. 특히 척도모수만이 다른 두 정규분포가 혼합되어 있는 경우(예, 모형 4) 두 상대효율의 차이는 더욱 커지게 된다. 이상값이 포함되지 않은 경우(즉, 오염수준=0)에 추정량들의 평균제곱오차(MSE)는 거의 차이가 없는 반면에, 이상값이 포함되어 있는 경우 대부분 최소일반화음지수불일치추정량(MGNEDE)의 평균제곱오차(MSE)가 최대가능도추정량(MLE)의 것보다 작다. 이와 같이 효율성의 측면에서 자료에 이상값이 포함되어 있는지 여부에 관계없이 최소일반화음지수불일치추정량(MGNEDE)은 최대가능도추정량(MLE)의 대안으로 사용할 수 있다.

## 참고문헌

- [1] Basu, A. and Lindsay, B. G. (1994). Minimum disparity estimation for continuous models: efficiency, distributions and robustness, *Annals of the Institute of Statistical Mathematics*, Vol. 46, 683-705.
- [2] Basu, A. and Sarkar, S. (1994a). The Trade-Off Between Robustness and Efficiency and The Effect of Model Smoothing in Minimum Disparity inference, *Journal of Statistical Computation and Simulation*, Vol. 50, 173-185.
- [3] Basu, A. and Sarkar, S. (1994b). Minimum disparity estimation in the errors-in-variables model, *Statistics & Probability Letter*, Vol. 20, 69-73.
- [4] Basu, A., Sarkar, S. and Vidyashankar, A. N. (1997). Minimum negative exponential disparity estimation in parametric models, *Journal of Statistical Planning and Inference*, Vol. 58, 349-370.
- [5] Beran, R. J. (1977). Minimum Hellinger distance estimates for parametric models, *Annals of Statistics*, Vol. 5, 445-463.
- [6] Cressie, N. and Read, T. (1984). Multinomial Goodness-of-fit Tests, *Journal of the Royal*

- Statistical Society B*, Vol. 46, No. 3, 440-464.
- [7] Csiszar, I. (1963). Eine informations theoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizitat von Markoffschen Ketten. *Publ. Math. Inst. Hungar. Acad. Sci.*, Vol. 3, 85-107.
- [8] Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation: The  $L_1$  View*, John Wiley, New York.
- [9] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*, John Wiley, New York.
- [10] Jeong, D. B. and Sarkar, S. (2001). Negative Exponential Disparity Based Deviance and Goodness-of-fit Tests for Continuous Models: Distribution, Efficiency and Robustness, *Journal of the Korean Statistical Society*, Vol. 30: 1, 41-61.
- [11] Jeong, D. B. and Sarkar, S. (2000). Negative exponential disparity family based goodness-of-fit tests for multinomial models. *Journal of Statistical Computation and Simulation*, Vol. 65, 43-61.
- [12] Lindsay, B. G. (1994). Efficiency versus robustness: The case for minimum Hellinger distance and related methods, *Annals of Statistics*, Vol. 22, 1081-1114.
- [13] Parzen, E. (1962), On estimation of a probability density function and its mode, *Annals of Mathematical Statistics*, Vol. 33, 1065-1076.
- [14] Sarkar, S., Basu, A. and Shin, D. W. (1996). Small sample comparisons for the blended weight chi-square goodness-of-fit test statistics. *Communications in Statistics, Theory and Method*, Vol. 25, 211-226.
- [15] Simpson, D. G. (1987). Minimum Hellinger distance estimation for the analysis of count data, *Journal of the American Statistical Association*, Vol. 82, 802-807.
- [16] Simpson, D. G. (1989). Hellinger Deviance Tests: Efficiency, Breakdown Points, and Examples, *Journal of the American Statistical Association*, Vol. 84, 107-113.
- [17] Stather, C. R. (1981). *Robust Statistical Inference using Hellinger Distance Methods*, Unpublished Ph. D. Dissertation, La Trobe University, Melbourne, Australia.
- [18] Tamura, R. N. and Boos, D. D. (1986). Minimum Hellinger Distance Estimation for Multivariate Location and Covariance, *Journal of the American Statistical Association*, Vol. 81, 223-229.



## 부록

뉴턴-랩슨 알고리즘에서 식 (3.2)의 행렬  $A$ 와 벡터  $\mathbf{b}$ 의 원소들은 각각 불일치측도  $\rho(f^*, f_\theta)$ 를 모수  $\mu$ 와  $\sigma^2$ 에 대하여 일·이차 편미분함으로써 얻을 수 있다. 여기에서 핵함수  $f^*$ 는 모수가 포함되어 있지 않으므로 형식적으로 편미분되지만 실제 계산에서는 제외된다.  $\rho(f^*, f_\theta)$ 의 편미분 결과는 다음과 같다.

(1)  $\rho(f^*, f_\theta)$ 의 모수  $\mu, \sigma^2$ 에 관한 편미분

$$\textcircled{1} \quad \frac{\partial \rho}{\partial \mu} = \frac{\partial \rho}{\partial f^*} \cdot \frac{\partial f^*}{\partial \mu} + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial f_\theta}{\partial \mu} = \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial f_\theta}{\partial \mu}$$

$$\textcircled{2} \quad \frac{\partial \rho}{\partial \sigma^2} = \frac{\partial \rho}{\partial f^*} \cdot \frac{\partial f^*}{\partial \sigma^2} + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial f_\theta}{\partial \sigma^2} = \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial f_\theta}{\partial \sigma^2}$$

$$\begin{aligned} \textcircled{3} \quad \frac{\partial^2 \rho}{\partial \mu^2} &= \frac{\partial^2 \rho}{\partial f^{*2}} \left( \frac{\partial f^*}{\partial \mu} \right)^2 + 2 \frac{\partial^2 \rho}{\partial f^* \partial f_\theta} \cdot \frac{\partial f^*}{\partial \mu} \cdot \frac{\partial f_\theta}{\partial \mu} + \frac{\partial^2 \rho}{\partial f_\theta^2} \left( \frac{\partial f_\theta}{\partial \mu} \right)^2 + \frac{\partial \rho}{\partial f^*} \cdot \frac{\partial^2 f^*}{\partial \mu^2} + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial^2 f_\theta}{\partial \mu^2} \\ &= \frac{\partial^2 \rho}{\partial f_\theta^2} \left( \frac{\partial f_\theta}{\partial \mu} \right)^2 + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial^2 f_\theta}{\partial \mu^2} \end{aligned}$$

$$\begin{aligned} \textcircled{4} \quad \frac{\partial^2 \rho}{\partial (\sigma^2)^2} &= \frac{\partial^2 \rho}{\partial f^{*2}} \left( \frac{\partial f^*}{\partial \sigma^2} \right)^2 + 2 \frac{\partial^2 \rho}{\partial f^* \partial f_\theta} \cdot \frac{\partial f^*}{\partial \sigma^2} \cdot \frac{\partial f_\theta}{\partial \sigma^2} + \frac{\partial^2 \rho}{\partial f_\theta^2} \left( \frac{\partial f_\theta}{\partial \sigma^2} \right)^2 + \frac{\partial \rho}{\partial f^*} \cdot \frac{\partial^2 f^*}{\partial (\sigma^2)^2} + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial^2 f_\theta}{\partial (\sigma^2)^2} \\ &= \frac{\partial^2 \rho}{\partial f_\theta^2} \left( \frac{\partial f_\theta}{\partial \sigma^2} \right)^2 + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial^2 f_\theta}{\partial (\sigma^2)^2} \end{aligned}$$

$$\begin{aligned} \textcircled{5} \quad \frac{\partial^2 \rho}{\partial \mu \partial \sigma^2} &= \frac{\partial^2 \rho}{\partial f^{*2}} \cdot \frac{\partial f^*}{\partial \mu} \cdot \frac{\partial f^*}{\partial \sigma^2} + \frac{\partial^2 \rho}{\partial f^* \partial f_\theta} \left( \frac{\partial f^*}{\partial \mu} \cdot \frac{\partial f_\theta}{\partial \sigma^2} + \frac{\partial f^*}{\partial \sigma^2} \cdot \frac{\partial f_\theta}{\partial \mu} \right) + \frac{\partial^2 \rho}{\partial f_\theta^2} \cdot \frac{\partial f_\theta}{\partial \mu} \cdot \frac{\partial f_\theta}{\partial \sigma^2} \\ &\quad + \frac{\partial \rho}{\partial f^*} \frac{\partial^2 f^*}{\partial \mu \partial \sigma^2} + \frac{\partial \rho}{\partial f_\theta} \frac{\partial^2 f_\theta}{\partial \mu \partial \sigma^2} = \frac{\partial^2 \rho}{\partial f_\theta^2} \cdot \frac{\partial f_\theta}{\partial \mu} \cdot \frac{\partial f_\theta}{\partial \sigma^2} + \frac{\partial \rho}{\partial f_\theta} \cdot \frac{\partial^2 f_\theta}{\partial \mu \partial \sigma^2} = \frac{\partial^2 \rho}{\partial \sigma^2 \partial \mu} \end{aligned}$$

불일치측도  $\rho(f^*, f_\theta)$ 을 모수  $\mu$ 와  $\sigma^2$ 에 관하여 편미분한 식의 각 항들을 살펴보면 불일치측도를 밀도함수  $f_\theta$ 으로의 일·이차 편미분이 필요하다. 그 식들은 다음과 같다.

(2)  $\rho(f^*, f_\theta)$ 의  $f_\theta$ 에 관한 일·이차 편미분

(Case 1) 일반화음지수불일치(GNED) 측도

$$\text{GNED}(\lambda) = \sum \frac{1}{\lambda^2} f_\theta \cdot \left[ \exp \left\{ -\lambda \left( \frac{f^*}{f_\theta} - 1 \right) \right\} - 1 + \lambda \left( \frac{f^*}{f_\theta} - 1 \right) \right]$$

을  $f_\theta$ 에 관하여 일·이차 편미분하면 다음과 같다.

$$\textcircled{1} \quad \frac{\partial \rho}{\partial f_\theta} = \frac{1}{\lambda^2} \left[ \left( 1 + \lambda \frac{f^*}{f_\theta} \right) \exp \left\{ -\lambda \left( \frac{f^*}{f_\theta} - 1 \right) \right\} - (1 + \lambda) \right]$$

$$\textcircled{2} \quad \frac{\partial^2 \rho}{\partial f_\theta^2} = \frac{f^{*2}}{f_\theta^3} \cdot \exp \left\{ -\lambda \left( \frac{f^*}{f_\theta} - 1 \right) \right\}$$

특히  $\lambda = 4/3$  일 때,

$$\text{GNED}(4/3) = \sum \frac{9}{16} f_\theta \cdot \left[ \exp \left\{ \frac{4}{3} \left( 1 - \frac{f^*}{f_\theta} \right) \right\} + \frac{4}{3} \frac{f^*}{f_\theta} - \frac{7}{3} \right]$$

이고 편미분 결과는 다음과 같다.

$$\textcircled{3} \quad \frac{\partial \rho}{\partial f_\theta} = \frac{9}{16} \left[ \left( 1 + \frac{4}{3} \frac{f^*}{f_\theta} \right) \exp \left\{ \frac{4}{3} \left( 1 - \frac{f^*}{f_\theta} \right) \right\} - \frac{7}{3} \right]$$

$$\textcircled{4} \quad \frac{\partial^2 \rho}{\partial f_\theta^2} = \frac{f^{*2}}{f_\theta^3} \cdot \exp \left\{ \frac{4}{3} \left( 1 - \frac{f^*}{f_\theta} \right) \right\}$$

(Case 2) 가중헬링거리(BWHD) 측도

$$\text{BWHD}(\lambda) = \sum \frac{[f^* - f_\theta]^2}{2[\lambda\sqrt{f^*} + \bar{\lambda}\sqrt{f_\theta}]^2}$$

을  $f_\theta$ 에 관하여 일·이차 편미분하면 다음과 같게 된다.

$$\textcircled{5} \quad \frac{\partial \rho}{\partial f_\theta} = \frac{1}{2} \left\{ (f_\theta - f^*) \left( 2\lambda\sqrt{f^*} + \bar{\lambda}\sqrt{f_\theta} + \bar{\lambda} \frac{f^*}{\sqrt{f_\theta}} \right) \right\} \times \frac{1}{(\lambda\sqrt{f^*} + \bar{\lambda}\sqrt{f_\theta})^3}$$

$$\textcircled{6} \quad \frac{\partial^2 \rho}{\partial f_\theta^2} = \frac{1}{4} \left( 4\lambda^2 f^* + \lambda \bar{\lambda} \sqrt{f^*} \sqrt{f_\theta} + 4 \bar{\lambda}^2 \frac{f^{*2}}{f_\theta} + 6\lambda \bar{\lambda} \frac{f^* \sqrt{f^*}}{\sqrt{f_\theta}} + \lambda \bar{\lambda} \frac{f^{*2} \sqrt{f^*}}{f_\theta \sqrt{f_\theta}} \right) \\ \times \frac{1}{(\lambda\sqrt{f^*} + \bar{\lambda}\sqrt{f_\theta})^4}$$

특히  $\lambda = 1/9$  일 때,

$$\text{BWHD}(1/9) = \sum \frac{81}{2} \cdot \frac{[f^* - f_\theta]^2}{[\sqrt{f^*} + 8\sqrt{f_\theta}]^2}$$

이고 편미분 결과는 다음과 같다.

$$\textcircled{7} \quad \frac{\partial \rho}{\partial f_\theta} = -81 \left\{ (f^* - f_\theta) \left( \sqrt{f^*} + 4\sqrt{f_\theta} + 4 \frac{f^*}{\sqrt{f_\theta}} \right) \right\} / (\sqrt{f^*} + 8\sqrt{f_\theta})^3$$

$$\textcircled{8} \quad \frac{\partial^2 \rho}{\partial f_\theta^2} = 81 \left( f^* + 2\sqrt{f^*}\sqrt{f_\theta} + 64 \frac{f^{*2}}{f_\theta} + 12 \frac{f^*\sqrt{f^*}}{\sqrt{f_\theta}} + 2 \frac{f^{*2}\sqrt{f^*}}{f_\theta\sqrt{f_\theta}} \right) \times \frac{1}{(\sqrt{f^*} + 8\sqrt{f_\theta})^4}$$

마지막으로 밀도함수  $f_\theta$ 을 모수  $\mu$ 와  $\sigma^2$ 에 관하여 편미분함으로써 식 (3.2)의 우변 두 번째 항이 모두 완성된다.

(3)  $f_\theta$ 의  $\mu, \sigma^2$ 에 관한 일·이차 편미분

정규밀도함수  $f_\theta$ 을 모수  $\mu$ 와  $\sigma^2$ 에 대하여 편미분하여 정리하면 다음과 같다.

$$\textcircled{1} \quad \frac{\partial f_\theta}{\partial \mu} = \frac{(x-\mu)}{\sigma^2} f_\theta$$

$$\textcircled{2} \quad \frac{\partial f_\theta}{\partial \sigma} = \left\{ \frac{(x-\mu)^2}{\sigma^3} - \frac{1}{\sigma} \right\} f_\theta$$

$$\textcircled{3} \quad \frac{\partial^2 f_\theta}{\partial \mu^2} = \left\{ \frac{(x-\mu)^2}{\sigma^4} - \frac{1}{\sigma^2} \right\} f_\theta$$

$$\textcircled{4} \quad \frac{\partial^2 f_\theta}{\partial (\sigma^2)^2} = \left\{ \frac{(x-\mu)^4}{\sigma^6} - 5 \frac{(x-\mu)^2}{\sigma^4} + \frac{2}{\sigma^2} \right\} f_\theta$$

$$\textcircled{5} \quad \frac{\partial^2 f_\theta}{\partial \mu \partial \sigma^2} = \left\{ \frac{(x-\mu)^3}{\sigma^5} - 3 \frac{(x-\mu)}{\sigma^3} \right\} f_\theta$$