The Korean Communications
in Statistics Vol. 7, No. 2, 2000
pp. 371-383

Negative Exponential Disparity Based Robust Estimates of Ordered Means in Normal Models

Bhaskar Bhattacharya[1], Sahadeb Sarkar[2] and Dong-Bin Jeong[3]

Abstract

Lindsay (1994) and Basu et al (1997) show that another density-based distance called the negative exponential disparity (NED) is an excellent competitor to the Hellinger distance (HD) in generating an asymptotically fully efficient and robust estimator. Bhattacharya and Basu (1996) consider estimation of the locations of several normal populations when an order relation between them is known to be true. They empirically show that the robust HD based weighted likelihood estimators compare favorably with the M-estimators based on Huber's ψ function, the Gastworth estimator, and the trimmed mean estimator. In this paper we investigate the performance of the weighted likelihood estimator based on the NED as a robust alternative relative to that based on the HD. The NED based estimator is found to be quite competitive in the settings considered by Bhattacharya and Basu.

Keywords : Hellinger distance, negative exponential disparity, ordered normal means, weighted likelihood estimator.

1. INTRODUCTION

Let $\{X_{ij},\ j=1,2,\ldots,n_i\}$, denote a random sample from a $N(\mu_i,\sigma^2)$ population, where i = 1, 2, ..., k. Assume that the k random samples are independent, and the population variances are all equal. Bhattacharya and Basu (1996) consider the problem of estimating the vector $\mu=(\mu_1,\mu_2,\ldots,\mu_k)'$ robustly when a simple order relation among the individual population means, such as

$$\mu_1 \le \mu_2 \le \ldots \le \mu_k \tag{1.1}$$

1) Associate Professor, Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901, USA.
2) Professor, O. M. Group, Indian Institute of Management Calcutta, Post Box 16757, Calcutta 700 027, India.
3) Assistant Professor, Department of Statistics, Kangnung National University, Kangnung, 210-702, Korea.
E-mail : dj@knusun.kangnung.ac.kr

maximum likelihood estimator (MLE) of μ_i is the i-th sample mean. Under order restrictions on the μ_i's is such as (1.1), one method of finding the restricted MLEs is known as the "Pool Adjacent Violators Algorithm" or PAVA. It is described in Bhattacharya and Basu (1996, p. 165). More details on PAVA and related algorithms are given in Robertson, Wright and Dykstra (1988).

The maximum likelihood estimation method is asymptotically optimal at the true model. However, it can produce estimators that are not robust, i.e., the estimators can be highly sensitive to slight deviations from the model. Beran (1977) showed that the minimum Hellinger distance estimator is an attractive robust alternative to the MLE in parametric models since it is asymptotically fully efficient under the assumed model while being robust under small departures from the assumed model (also see Tamura and Boos, 1986; Simpson, 1987, 1989). Lindsay (1994) and Basu and Lindsay (1994) introduced the class of the minimum disparity estimators (MDEs) which contains the minimum Hellinger distance estimator as a special member. Basu, Markatou and Lindsay (1994) proposed a class of weighted likelihood estimators (WLEs). The WLEs are obtained by modifying the method of minimum disparity estimation and are considerably easier to compute. The WLEs are asymptotically fully efficient at the model and robust under data contamination like the MDEs. But, unlike the MDEs, the WLEs do not require a *transparent kernel* (Basu and Lindsay 1994) to attain full asymptotic efficiency. Bhattacharya and Basu (1996) show that the robust Hellinger distance based WLE (HDWLE) of the location parameters $\mu_1, \mu_2, \ldots, \mu_k$ perform better than the M-estimators based on Huber's ψ function, the Gastworth estimator and the trimmed mean estimator in terms of relative efficiency. In the approach of Bhattacharya and Basu, parameter estimation is done assuming the normal model to be true. Magel and Wright (1984) considered a nonparametric approach and compared different methods of estimating several location parameters under a nondecreasing order relation. However, in the parametric approach of Bhattacharya and Basu the estimation methods are more efficient than general nonparametric methods of estimation when the normal model is true.

Lindsay (1994) introduced the minimum negative exponential disparity estimator as another special member of the class of the MDEs. Basu et al (1997) show that the minimum negative exponential disparity estimator is fully efficient, and robust not only against outliers but also against inliers (defined as values with observed relative frequencies considerably smaller than expected; see Section 7.2 of Lindsay (1994)), a property the minimum Hellinger distance estimator does not share. The minimum negative exponential disparity estimator appears to be a highly promising estimator and a major competitor to the minimum Hellinger distance estimator within the class of robust and asymptotically fully efficient estimators. For more empirical evidence of full efficiency and robustness see Lindsay (1994) for discrete models, and Basu and Sarkar (1994) for the normal model. This motivated our examination of applying the NED based weighted likelihood estimation method to the settings considered by Bhattacharya and Basu (1996). In Section 2 we discuss the NED and in Section 3 we define the NED based WLE (NEDWLE). In Section 4 we present the empirical results comparing the HDWLE and NEDWLE.

2. THE NEGATIVE EXPONENTIAL DISPARITY

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed (iid) observations from a population with pdf $g \in G$, the class of all continuous densities. Let $\mathfrak{I}_\theta = \{f_\theta, \theta \in \Theta\}$ be a parametric subset of G. The assumed model is said to be correctly specified if the true data generating density $g \in \mathfrak{I}_\theta$. Let

$$\hat{g}_n(x) \equiv \int_{-\infty}^{\infty} \omega(x; y, h) dG_n(y) = \frac{1}{nh} \sum_{i=1}^{n} \omega\left(\frac{x - X_i}{h}\right)$$

be a kernel density estimate of g, where $\omega(x; y, h)$ is a family of kernel functions like the $N(y, h^2)$ density function with mean y and standard deviation h, and G_n is the empirical distribution function for the data $X_1, X_2, \ldots, X_n$. Define the smoothed model density $f_\theta^*(x)$ as

$$f_\theta^*(x) \equiv \int_{-\infty}^{\infty} \omega(x; y, h) dF_\theta(y),$$

where $F_\theta(y)$ denotes the cumulative distribution function of the model. Let

$$\delta^*(x) = \delta^*(\widehat{g_n}, \theta, x) \equiv \frac{\hat{g}_n(x) - f_\theta^*(x)}{f_\theta^*(x)},$$

called the *Pearson* residual at the value x (Lindsay 1994). Let H be a thrice differentiable, strictly convex function with $H(0) = 0$. The nonnegative disparity measure ρ corresponding to H is defined as

$$\rho = \rho(\widehat{g_n}, \theta) \equiv \int H\left(\frac{\hat{g}_n(x) - f_\theta^*(x)}{f_\theta^*(x)}\right) f_\theta^*(x) dx. \qquad (2.1)$$

A value of θ that minimizes (2.1) is called the minimum disparity estimator. If $H(\delta) = [(\delta + 1)^{1/2} - 1]^2$, (2.1) defines the Hellinger distance (Beran, 1977) given by

$$\int \left[\sqrt{\hat{g}_n(x)} - \sqrt{f_\theta^*(x)}\right]^2 dx.$$

If $H(\delta) = \exp(-\delta) - 1$ or $H(\delta) = \exp(-\delta) - 1 + \delta$, (2.1) defines the negative exponential disparity

$$NED(g,\theta) \equiv \int [\exp(-\frac{\hat{g}_n(x)-f_\theta^*(x)}{f_\theta^*(x)})-1]f_\theta^*(x)dx. \tag{2.2}$$

The disparity produced by the function $H(\delta)=(\delta+1)\ln(\delta+1)$ is called the *likelihood disparity (LD)* because its minimizer in the discrete models is the MLE (Lindsay 1994). Basu et al (1997, Section 2.2) give an explanation as to why the complex looking NED generates an estimator which is asymptotically fully efficient and at the same time robust against both outliers and inliers. Under differentiability of the model, the minimum disparity equation has the form

$$-\frac{\partial\rho}{\partial\theta} = \int A(\delta^*(x))\frac{\partial f_\theta^*(x)}{\partial\theta} = 0, \tag{2.3}$$

where

$$A(\delta) \equiv (\delta+1)\dot{H}(\delta)-H(\delta) \tag{2.4}$$

and $\dot{H}(\delta)$ denotes the first derivative of $H(\delta)$. The function $A(\delta)$ is an increasing function on $[-1,\infty)$, and without altering the estimating properties of the disparity ρ, $A(\delta)$ can be redefined (standardized) to satisfy $A(0)=0$ and $\dot{A}(0)=0$, where $\dot{A}(\delta)$ denotes the first derivative of $A(\delta)$. This standardized function is called the residual adjustment function (RAF) of the disparity and it determines various second-order measures of efficiency and robustness of the MDEs.

The RAF is of the form: $A(\delta)=2[(\delta+1)^{1/2}-1]$ for the HD, $A(\delta)=2-(2+\delta)\exp(-\delta)$ for the NED, and $A(\delta)=\delta$ for the LD. These three RAFs are graphically displayed in Figure 1. It shows that the RAFs for the HD and NED satisfy $A(\delta)<<\delta$, i.e., they heavily downweight large positive Pearson residuals (corresponding to the outliers in data).

Examination of the negative side of the $\delta-$axis in Figure 1 shows that the RAF for the HD magnifies the effect of large negative Pearson residuals which define the inliers in data, and the RAF for the NED has a downweighting effect on large negative Pearson residuals. This provides a rationale for the robust behavior of the minimum negative exponential disparity estimator against both outliers and inliers and that of the minimum Hellinger distance estimator against only outliers.

3. THE NEGATIVE EXPONENTIAL DISPARITY BASED WLE

In this section we briefly discuss the computation of the WLEs, and in particular, that of

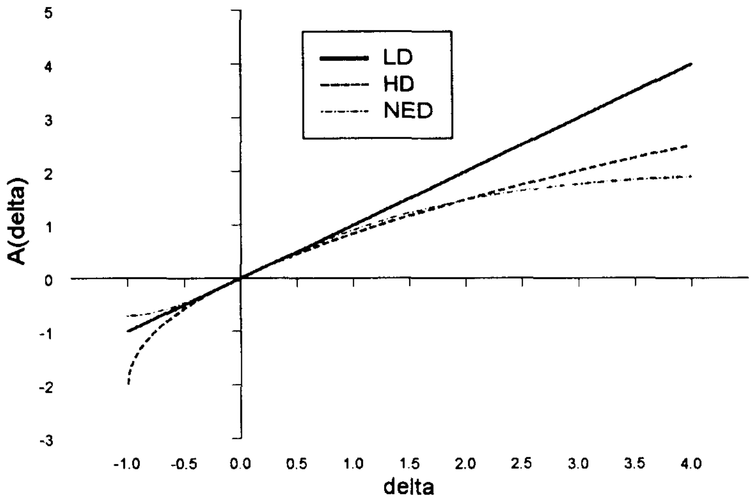

Figure 1. The Residual Adjustment Functions for LD, HD and NED

the NED based WLE. First, Basu et al (1994) write the estimating equation (2.3) as

$$\int\left[\frac{A(\delta^*(x))+1}{\delta^*(x)+1}\right](\delta^*(x)+1)\frac{\partial f_\theta^*(x)}{\partial\theta}dx=0,$$

i.e.,

$$\int w(\delta^*(x))\left[\frac{\partial f_\theta^*}{\partial\theta}/f_\theta^*(x)\right]d\widehat{G_n}(x)=0, \tag{3.1}$$

where

$$w(\delta^*(x)) = \frac{A(\delta^*(x))+1}{\delta^*(x)+1} \tag{3.2}$$

and $\hat{G}_n(x)$ the cumulative distribution function of $\hat{g}_n(x)$. By replacing the smoothed $f_\theta^*(x)$ and $\hat{G}_n(x)$ with unsmoothed $f_\theta(x)$ and $G_n(x)$, except in the weight part $w(\delta^*(x))$, one gets the following estimating equation

$$\frac{1}{n}\sum_{j=1}^{n} w(\delta^*(X_j))u(X_j,\theta)=0, \tag{3.3}$$

where $u(X_j,\theta)=\partial\ln f_\theta(x)/\partial\theta$ is the usual maximum likelihood score function. As in the case of the iteratively reweighted least squares estimator, the equation (3.3) can be solved iteratively to produce estimators called the WLEs. For a given set of initial values of θ one can construct the weights $w(\delta^*(X_j))$, then solve equation (3.3) for an improved estimate of θ, which is then used to compute new set of weights. This iterative method is continued until an appropriate convergence criterion is met. If we use $A(\delta)=2[(\delta+1)^{1/2}-1]$, and $A(\delta)=2-(2+\delta)\exp(-\delta)$ in (3.3) above we get the HDWLE and the NEDWLE, respectively.

In the above we have discussed the computation of the WLEs in the one sample case. Bhattacharya and Basu (1996, Sec 2.3) described weighted likelihood estimation in multiple samples. We briefly review it here. Let $\{X_{ij}, j=1,2,\ldots,n_i\}$, i = 1, 2, ..., k, denote k independent random samples from k normal populations, i-th population having mean μ_i and variance σ^2. Let $\theta_i=(\mu_i,\sigma^2)'$. Let $\hat{g}_{i,n}(x)$ be the kernel density estimate for the i-th sample using a suitable kernel like $\omega(x;y,h)=N(y,h^2)$ density function and let $f^*_{\theta_i}(x)$ be the kernel smoothed version of the model density $f_{\theta_i}(x)$ for the i-th population. For the i-th sample, define the Pearson residual at x as

$$\delta_i^*(x)=\frac{\hat{g}_{i,n}(x)-f^*_{\theta_i}(x)}{f^*_{\theta_i}(x)},$$

and the weight

$$w_i(\delta_i^*(x))=\frac{A(\delta_i^*(x))+1}{\delta_i^*(x)+1}$$

for a chosen disparity RAF $A(\cdot)$. Then the (k+1) weighted likelihood estimating equations for $\mu_1,\mu_2,\ldots,\mu_k,\sigma^2$ are given by

$$\sum_{j=1}^{n_i} w_i(\delta_i^*(X_{ij}))u_\mu(X_{ij},\mu_i,\sigma^2)=0, \quad i=1,2,\ldots,k$$

and

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} w_i(\delta_i^*(X_{ij}))u_{\sigma^2}(X_{ij},\mu_i,\sigma^2)=0,$$

where u_μ and μ_{σ^2} are the score functions with respect to μ and σ^2 respectively in the $N(\mu_i,\sigma^2)$ model. The above equations simplify to

$$\sum_{j=1}^{n_i} w_i(\delta_i^*(X_{ij}))(X_{ij}-\mu_i)=0, \quad i=1,2,\ldots,k$$

and

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} w_i(\delta_i^*(X_{ij}))[(X_{ij}-\mu_i)^2-\sigma^2] = 0.$$

Battacharya and Basu (1996) used the "Pool Adjacent Violators Algorithm" (PAVA) for computing order restricted WLEs (based on the HD) of the parameters satisfying order restrictions defined by (1.1). In the simulation study reported in the next section, all the pooling was done from left to right (see Bhattacharya and Basu, 1996, p. 170).

4. MONTE CARLO STUDY

We conducted a Monte-Carlo study in the same settings used by Bhattacharya and Basu (1996) and computed the MLE, the WLEs based on the HD and NED of the order restricted mean parameters of normal populations under no contamination and various contamination cases. We computed the WLEs using the normal kernel $\omega(x;y,h)=N(y,h^2)$density for the bandwidth parameter h = 0.5 and h = 0.75. Let HDWLE -0.5 and HDWLE -0.75 denote the HD based WLE computed with h = 0.5 and 0.75 respectively. Similarly, NEDWLE -0.5 and NEDWLE -0.75 define the corresponding estimators based on the NED. The initial estimates of $\mu_1,\mu_2,\ldots,\mu_k,\sigma^2$ were defined as $\hat{\mu}_{i(0)}=median(X_{i1},\ldots,X_{in_i})$ and $\hat{\sigma}^2_{(0)}=$ $1.48\times M$ where $M=median\{|X_{ij}-\hat{\mu}_{i(0)}|, j=1,2,\ldots,n_i, i=1,2,\ldots,k\}$.

Observations from the i-th population have the form $X_{ij}=\mu_i+\varepsilon_{ij}$ where the ε_{ij} are iid random variables generated from a contaminated normal distribution of the form $(1-\varepsilon)N(0,1)$ $+\varepsilon F$, where F represents a contaminating distribution and ε is the contamination proportion. For the i-th population the target parameter to be estimated is μ_i. A pure normal distribution along with its eight different contaminated versions were considered and they are:

I.	N(0, 1)
II.	0.9 N(0, 1) + 0.1 N(0, 64)
III.	0.8 N(0, 1) + 0.2 N(0, 64)
IV.	0.9 N(0, 1) + 0.1 t(1)
V.	0.8 N(0, 1) + 0.2 t(1)
VI.	0.9 N(0, 1) + 0.1 N(0, 1)/U(0, 1)
VII.	0.8 N(0, 1) + 0.2 N(0, 1)/U(0, 1)
VIII.	0.9 N(0, 1) + 0.1 N(0, 1)/U(0, 1/3)
IX.	0.8 N(0, 1) + 0.2 N(0, 1)/U(0, 1/3).

Among the contaminating distributions, N(0, 1)/U(0, υ) denotes the distribution of the ratio of a N(0, 1) random variable and an independent uniform random variable on the interval $(0,\upsilon)$. The mean vectors with a variety of dispersions within its components were chosen as in Bhattacharya and Basu (1996).

All results are based on 5000 replications. Sample sizes considered are 10 and 20. Multiple sample cases are considered for k = 2, 3 and 4. The estimates are determined under the known order restriction relation $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$. The estimators are compared in terms of their total mean square error (TMSE), defined by $E(\sum_{i=1}^{k}(\widehat{\mu_i}-\mu_i)^2)$, which is estimated by the sum of squared errors averaged over 5000 replications. The empirical relative efficiency of an estimator is computed by taking the reciprocal of its TMSE to the smallest TMSE for all the 4 estimators. These relative efficiencies are given in Tables 1-7. We discuss the empirical results of Tables 1-7 below.

The HDWLE -0.5 is less efficient at the model and more robust under contamination than the HDWLE -.75, except one case in Table 1 (for distribution I) and one in Table 2 (for distribution VI), which can be attributed to sampling variability. On the other hand, the NEDWLE -0.5 is less efficient at the model than the NEDWLE -.75 in all cases. The NEDWLE -0.5 is more robust than the NEDWLE -.75, except generally for distributions IV and VI and n = 10 (i.e., the sample size is very small), in which cases the relative efficiencies of the two estimators are very close. The NEDWLE -.75 has considerably smaller relative efficiency for distributions III and IX for k = 2, and for k = 4 with n = 20 and the mean parameter combination (-1.1, -.7, .1, .4). On the other hand, the HDWLE -.75 has notably smaller relative efficiency values for distributions III and IX for k = 3 and k = 4 when n = 10.

For k = 2, the HDWLE -.5 outperforms the NEDWLE -.5 except for distribution V and n = 20 in Table 2. However, for k = 3, the NEDWLE -.5 is better than the HDWLE -.5 generally for distributions II, III, VII, VIII, IX and for both n = 10 and 20. For k = 4, the results are mixed for the two estimators. When $(\mu_1, \mu_2, \mu_3, \mu_4)$ = (0, 0, 0, 1) with three equal means (Table 6) the HDWLE -.5 performs very well for n = 10 and uniformly better than the NEDWLE -.5 for n = 20. This is reversed when $(\mu_1, \mu_2, \mu_3, \mu_4)$ = (-1.1, -.7, .1, .4) with four distinct means (Table 7), except for distribution IX. Therefore, the NEDWLE -.5 appears to be quite competitive to the HDWLE -.5 in the cases considered in our Monte Carlo study. Our findings support the conclusion of Basu et al (1997) that the NED provides an excellent alternative to the HD in generating a robust and asymptotically fully efficient estimator.

REFERENCES

[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., Markatou, M. and Lindsay, B. G. (1994). Robustness via maximum likelihood estimating equations. *Technical Report*, Department of Statistics, Columbia University, New York, NY.

[3] Basu, A. and Sarkar, S. (1994). 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.

[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. (1977). Minimum Hellinger distance estimates for parametric models. *Annals of Statistics*, Vol. 5, 445-463.

[6] Bhattacharya, B. and Basu, A. (1996). Robust estimates of ordered means in normal models. *Journal of Statistical Computation and Simulation*, Vol. 54, 165-175.

[7] Hampel, F. R., Ronchetti, E., Rousseeuw, P. J. and Stahel, W. (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, New York.

[8] Huber, P. (1981). *Robust Statistics*. John Wiley & Sons, New York.

[9] Lindsay, B. G. (1994). Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *Annals of Statistics*, Vol. 22, 1081-1114.

[10] Magel, R. and Wright, F. T. (1984). Robust estimates of ordered parameters. *Journal of Statistical Computation and Simulation*, Vol. 20, 47-58.

[11] Robertson, T, Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. John Wiley & Sons, New York.

[12] Simpson, D. G. (1987). Minimum Hellinger distance estimation for the analysis of count data. *Journal of the American Statistical Association*, Vol. 82, 802-807.

[13] Simpson, D. G. (1989). Hellinger deviance test: efficiency, breakdown points, and examples. *Journal of the American Statistical Association*, Vol. 84, 107-113.

[14] 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.

APPENDIX

Table 1: Relative efficiencies of the different procedures: $(\mu_1, \mu_2) = (-0.1, 0.1)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	1.0000	.9987	.9474	.9812
II	1.0000	.9730	.9670	.9281
III	1.0000	.9324	.9339	.8266
IV	1.0000	.9746	.9514	.9593
V	1.0000	.9852	.9500	.9253
VI	1.0000	.9802	.9666	.9698
VII	1.0000	.9460	.9814	.9567
VIII	1.0000	.9699	.9768	.9498
IX	1.0000	.9109	.9543	.8601
		n = 20		
I	.9954	1.0000	.9710	.9835
II	1.0000	.9865	.9606	.9201
III	1.0000	.9286	.9325	.8260
IV	1.0000	.9970	.9886	.9831
V	1.0000	.9853	.9945	.9719
VI	1.0000	.9772	.9702	.9555
VII	1.0000	.9685	.9446	.9097
VIII	1.0000	.9736	.9496	.9052
IX	1.0000	.9264	.9658	.8693

Table 2: Relative efficiencies of the different procedures: $(\mu_1, \mu_2) = (-0.3, 0.3)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9771	1.0000	.9489	.9836
II	1.0000	.9733	.9559	.9194
III	1.0000	.9558	.9223	.8300
IV	1.0000	.9726	.9516	.9622
V	1.0000	.9704	.9444	.9279
VI	.9921	1.0000	.9611	.9663
VII	1.0000	.9779	.9670	.9399
VIII	1.0000	.9897	.9724	.9513
IX	1.0000	.9298	.9596	.8744
		n = 20		
I	.9903	1.0000	.9668	.9795
II	1.0000	.9576	.9620	.9212
III	1.0000	.9196	.9231	.8188
IV	1.0000	.9941	.9778	.9725
V	.9901	.9875	1.0000	.9751
VI	1.0000	.9788	.9785	.9639
VII	1.0000	.9526	.9457	.9128
VIII	1.0000	.9771	.9612	.9196
IX	1.0000	.9244	.9533	.8576

Table 3: Relative efficiencies of the different procedures: $(\mu_1, \mu_2, \mu_3) = (0, 0, 0)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9926	1.0000	.9767	.9998
II	.9771	.9161	1.0000	.9505
III	.9774	.8437	1.0000	.9040
IV	1.0000	.9904	.9869	.9910
V	1.0000	.9686	.9951	.9772
VI	1.0000	.9807	.9943	.9884
VII	.9872	.9470	1.0000	.9661
VIII	.9851	.9275	1.0000	.9581
IX	.9171	.8190	1.0000	.8891
		n = 20		
I	.9954	1.0000	.9895	.9999
II	.9789	.9160	1.0000	.9495
III	.9241	.8107	1.0000	.8851
IV	1.0000	.9836	.9999	.9862
V	.9948	.9564	1.0000	.9669
VI	.9938	.9668	1.0000	.9740
VII	.9869	.9383	1.0000	.9553
VIII	.9794	.9214	1.0000	.9452
IX	.9380	.8315	1.0000	.8765

Table 4: Relative efficiencies of the different procedures: $(\mu_1, \mu_2, \mu_3) = (0, 0, 1)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9880	1.0000	.9671	.9974
II	.9798	.9314	1.0000	.9671
III	.9723	.8623	1.0000	.9204
IV	1.0000	.9951	.9825	.9945
V	1.0000	.9753	.9877	.9823
VI	1.0000	.9833	.9882	.9952
VII	.9965	.9590	1.0000	.9818
VIII	.9902	.9409	1.0000	.9743
IX	.9340	.8402	1.0000	.9113
		n = 20		
I	.9935	1.0000	.9860	.9997
II	.9514	.9269	1.0000	.9577
III	.9321	.8241	1.0000	.8877
IV	1.0000	.9864	.9969	.9906
V	.9980	.9628	1.0000	.9757
VI	1.0000	.9765	.9994	.9833
VII	.9909	.9442	1.0000	.9644
VIII	.9802	.9252	1.0000	.9507
IX	.9471	.8456	1.0000	.8890

Table 5: Relative efficiencies of the different procedures: $(\mu_1, \mu_2, \mu_3) = (-1.5, 0, 1.5)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9851	1.0000	.9564	.9962
II	.9863	.9378	1.0000	.9738
III	.9767	.8579	1.0000	.9254
IV	1.0000	.9971	.9766	.9955
V	1.0000	.9772	.9871	.9839
VI	1.0000	.9859	.9815	.9946
VII	1.0000	.9594	.9986	.9819
VIII	1.0000	.9531	.9992	.9899
IX	.9371	.8465	1.0000	.9272
		n = 20		
I	.9918	1.0000	.9807	.9990
II	.9831	.9251	1.0000	.9581
III	.9266	.8208	1.0000	.8896
IV	1.0000	.9879	.9957	.9913
V	.9990	.9649	1.0000	.9778
VI	1.0000	.9780	.9968	.9849
VII	.9923	.9475	1.0000	.9662
VIII	.9823	.9296	1.0000	.9573
IX	.9460	.8448	1.0000	.8951

Table 6: Relative efficiencies of the different procedures: $(\mu_1, \mu_2, \mu_3, \mu_4) = (0, 0, 0, 1)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9891	1.0000	.9596	.9874
II	.9981	.9451	1.0000	.9863
III	.9713	.8563	1.0000	.9433
IV	1.0000	.9912	.9842	.9945
V	1.0000	.9730	.9930	.9801
VI	1.0000	.9846	.9822	.9905
VII	1.0000	.9613	.9976	.9793
VIII	.9921	.9490	1.0000	.9714
IX	.9408	.8466	1.0000	.9142
		n = 20		
I	.9912	.9988	.9752	1.0000
II	1.0000	.9415	.8736	.7721
III	1.0000	.8850	.8780	.7277
IV	1.0000	.9894	.9098	.9171
V	1.0000	.9656	.8746	.8168
VI	1.0000	.9765	.8875	.8521
VII	1.0000	.9601	.8635	.7950
VIII	1.0000	.9514	.8543	.7690
IX	1.0000	.8802	.8981	.7315

Table 7: Relative efficiencies of the different procedures: $(\mu_1, \mu_2, \mu_3, \mu_4) = (-1.1, -.7, .1, .4)$

	HDWLE		NEDWLE	
Distribution	h = 0.5	h = 0.75	h = 0.5	h = 0.75
		n = 10		
I	.9867	1.0000	.9571	.9951
II	1.0000	.9517	.9991	.9868
III	.9733	.8671	1.0000	.9437
IV	1.0000	.9935	.9841	.9960
V	1.0000	.9761	.9984	.9847
VI	1.0000	.9883	.9814	.9927
VII	1.0000	.9673	.9980	.9840
VIII	.9947	.9539	1.0000	.9798
IX	.9497	.8609	1.0000	.9294
		n = 20		
I	.9926	1.0000	.9836	.9988
II	.9836	.9301	1.0000	.9611
III	.9491	.8420	1.0000	.9099
IV	1.0000	.9876	.9959	.9897
V	.9933	.9589	1.0000	.9704
VI	.9987	.9776	1.0000	.9841
VII	.9879	.9465	1.0000	.9599
VIII	.9913	.9468	1.0000	.9699
IX	1.0000	.9313	.7018	.7015