The Korean Communications
in Statistics Vol. 5, No. 2, 1998
pp. 517-529

Penalizing the Negative Exponential Disparity in Discrete Models[1]

Sahadeb Sarkar[2], Kijoung Song[3] and Dong Bin Jeong[4]

Abstract

When the sample size is small the robust minimum Hellinger distance (HD) estimator can have substantially poor relative efficiency at the true model. Similarly, approximating the exact null distributions of the ordinary Hellinger distance tests with the limiting chi-square distributions can be quite inappropriate in small samples. To overcome these problems Harris and Basu (1994) and Basu et al. (1996) recommended using a modified HD called penalized Hellinger distance (PHD). Lindsay (1994) and Basu et al. (1997) showed that another density based distance, namely the negative exponential disparity (NED), is a major competitor to the Hellinger distance in producing an asymptotically fully efficient and robust estimator. In this paper we investigate the small sample performance of the estimates and tests based on the NED and penalized NED (PNED). Our results indicate that, in the settings considered here, the NED, unlike the HD, produces estimators that perform very well in small samples and penalizing the NED does not help. However, in testing of hypotheses, the deviance test based on a PNED appears to achieve the best small-sample level compared to tests based on the NED, HD and PHD.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample from a discrete distribution having a probability mass function (pmf) $g \in G$, the class of all pmf's. Let $\mathcal{F}_\theta = \{ f_\theta : \theta \in \Theta \} \subset G$ be a parametric family of distributions involving the parameter θ. The model is said to be correctly specified if the true data generating density g belongs to the assumed model $\mathcal{F}_\theta$. Several authors including Beran (1977), Tamura and Boos (1986), Simpson (1987) and Lindsay (1994) showed

---

1) The research of the third author was supported in part by a grant from Office of Research Support at Kangnung National University, Korea.

2) Associate Professor, IIM Calcutta, OM Group, Joka, D.H. Road, Post Box 16757, Calcutta 700 027, India. The Author is currently on leave from the Department of Statistics, Oklahoma State University, Stillwater, OK 74078, USA, where he is Associate Professor.

3) Department of Epidemiology and Biostatistics, Case western reserve University, Cleveland, OH 44106-9920, USA.

4) Full-time Lecturer, Department of Statistics, Kangnung National University, Kangnung, 210-702, Korea.

that the minimum Hellinger distance estimator (MHDE) of θ is a very attractive robust alternative to the maximum likelihood estimator (MLE) since it is asymptotically as efficient as the maximum likelihood estimator (MLE) at the true model while achieving strong robustness properties under data contamination. However, when the sample size is small, the relative efficiency of the MHDE can be substantially poor at the true model. To overcome this problem Harris and Basu (1994) modified the Hellinger distance and defined a class of penalized Hellinger distances (PHDs). This class is indexed by a parameter, which controls the weight on the empty cells in the penalized Hellinger distance. Harris and Basu recommended a particular member of this family which is found to produce an estimator having much smaller empirical mean square error (than the MHDE) in small samples at the true model, without sacrificing robustness properties.

Simpson (1989) and Lindsay (1994) studied tests based on the Hellinger distance that are analogues of the likelihood ratio test (LRT). Basu et al. (1996) considered Hellinger distance based analogues of Rao tests and Wald tests. When the assumed model is correct, in small samples the exact null distributions of these Hellinger distance tests may be quite different from the limiting chi-square distributions. This can lead to very poor level and power values of the Hellinger distance tests in small samples compared to the likelihood based tests when the assumed model is correct. Empirical results of Basu et al. (1996) indicated that the tests based on a PHD can overcome the above problem without compromising the robustness properties. Park et al. (1995) have studied the class of combined and penalized blended weight Hellinger distances.

Lindsay (1994) introduced the minimum negative exponential disparity estimator (MNEDE) as a member of a general family of density based minimum distance estimators that contains the MHDE as a member. Results of Lindsay (1994), Basu and Lindsay (1994), Basu and Sarkar (1994) and Basu et al. (1997) showed that the MNEDE is an excellent competitor to the MHDE within the class of robust and first order efficient estimators. Moreover, the MNEDE is robust not only against outliers but also against inliers (defined as values with less data than expected), a property not shared by the MHDE (Lindsay 1994, Basu et al. 1997). Under a discrete model the unobserved values of the sample space having empty cells are common and represent extreme cases of inliers. This motivated us to investigate the small sample performance of the estimators and tests based on the negative exponential disparity (NED) and to examine the need for considering penalized negative exponential disparities (PNED).

The rest of the paper is organized as follows. In Section 2 we briefly discuss the NED, and related works. In Section 3 we introduce penalized disparities in general, which include the PHDs and PNEDs as special cases. Section 4 contains small sample results for the MNEDE and minimum PNED estimators (MPNEDE) in Poisson and geometric models. In Section 5 we present small sample results for tests based on the NED and PNEDs. Section 6 contains some concluding remarks.

2. The Minimum Negative Exponential Disparity Estimator

For a random sample $X_1, X_2, \ldots, X_n$ from a discrete population let $d_n(x)$ denote the proportion of X_i's having the value x. Let the assumed parametric family $\mathfrak{F}_\theta = \{ f_\theta : \theta \in \Theta \} \subset G$ have a countable support. Let $\delta(x) = \delta(d_n, \theta, x) \equiv (d_n(x) - f_\theta(x))/ f_\theta(x)$, called the *Pearson* residual at the value x (Lindsay 1994), which depends on the data and the parameter θ. Let G be a thrice differentiable, strictly convex function with $G(0) = 0$. Then, the nonnegative disparity measure D, corresponding to G, between the data density d_n and model density f_θ is defined as

$$D = D(d_n, \theta) \equiv \sum_x G\left(\frac{d_n(x) - f_\theta(x)}{f_\theta(x)}\right) f_\theta(x). \tag{2.1}$$

A minimizer of (2.1) is called the minimum disparity estimator. The MLE of θ minimizes the *likelihood disparity* (LD)

$$\sum_x d_n(x) \ln\left(\frac{d_n(x)}{f_\theta(x)}\right) + (f_\theta(x) - d_n(x)), \tag{2.2}$$

which is obtained by letting $G_{LD}(\delta) = (\delta + 1)\, ln(\delta + 1) - \delta$ in (2.1). The MHDE minimizes the distance

$$2\sum_x [\sqrt{d_n(x)} - \sqrt{f_\theta(x)}]^2, \tag{2.3}$$

obtained by putting $G_{HD}(\delta) = 2[(\delta + 1)^{1/2} - 1]^2$ in (2.1). When $G_{NED}(\delta) = exp(-\delta) - 1 + \delta$, (2.1) defines the negative exponential disparity

$$\mathrm{NED}(g, \theta) \equiv \sum_x \left[\exp\left(-\frac{d_n(x) - f_\theta(x)}{f_\theta(x)}\right) - 1 + \frac{d_n(x) - f_\theta(x)}{f_\theta(x)} \right] f_\theta(x) dx \tag{2.4}$$

and its minimizer is the MNEDE. Lindsay (1994, Section 7.2) and also Basu et al. (1997, Section 2.2) explained why the complex looking NED produces an estimator which is asymptotically fully efficient, and robust against both outliers and inliers. The minimum disparity estimation equation, under differentiability of the model, has the form

$$-\frac{\partial D}{\partial \theta} = \sum_x A(\delta(x)) \frac{\partial f_\theta(x)}{\partial \theta} = 0,$$

where

$$A(\delta) \equiv (\delta+1)\dot{G}(\delta) - G(\delta) \tag{2.5}$$

and $\dot{G}(\delta)$ is the first derivative of $G(\delta)$. The function $G(\delta)$ can be redefined (standardized), without altering the estimating properties of the disparity D, so that its $A(\delta)$ function, defined by (2.5), satisfies $A(0) = 0$ and $\dot{A}(0) = 0$, where $\dot{A}(\delta)$ denotes the first derivative of $A(\delta)$. The function $A(\delta)$ is an increasing function on $[-1, \infty)$ and is known as the residual adjustment function (RAF) of the disparity D. The shape of the RAF controls most of the theoretical properties of the minimum disparity estimators. Note that disparities (2.2), (2.3) and (2.4) are defined in a manner that makes them asymptotically equivalent and the corresponding $G(\delta)$ functions standardized. This is why we have the $-\delta$ term in $G_{LD}(\delta)$, the multiplication factor 2 in $G_{HD}(\delta)$, and the $+\delta$ term in $G_{NED}(\delta)$.

The form of the RAFs for the likelihood disparity, HD, and NED are given by $A_{LD}(\delta) = \delta$, $A_{HD}(\delta) = 2[(\delta+1)^{1/2} - 1]$ and $A_{NED}(\delta) = 2-(2+\delta)exp(-\delta)$, respectively. A graphical display of these three RAFs can be seen in Figure 1.

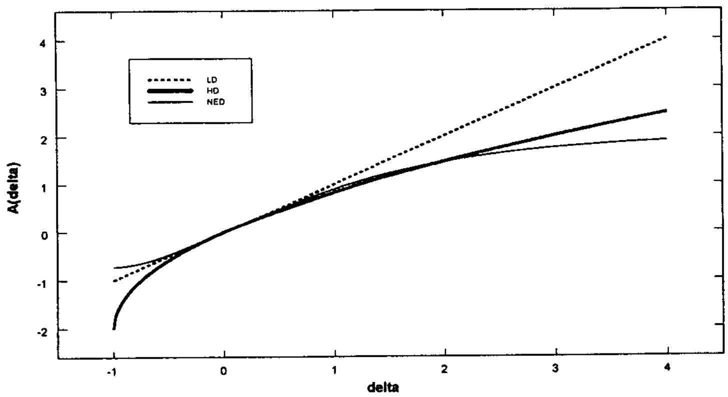

Figure 1. The Residual Adjustment Functions for LD, HD and NED

The graph shows that the RAFs for the HD and the NED heavily downweight large positive Pearson residuals (corresponding to the outliers in data), i.e., $A(\delta) << \delta$. However, the HD magnifies the effect of large negative Pearson residuals which define the inliers in data. On the other hand, the RAF for the NED has a downweighting effect on large negative Pearson residuals as well, i.e., $|A(\delta)| < |\delta|$. Thus, the shape of the RAF explains why the MNEDE is robust against both outliers and inliers while the MHDE is robust against outliers only.

3. Penalized Disparities

3.1. Estimation

Harris and Basu (1994) defined the penalized Hellinger distance (PHD) family as

$$PHD(d_n, f_\theta) = 2 \sum_{d_n(x) \neq 0} (\sqrt{d_n(x)} - \sqrt{f_\theta(x)})^2 + 2h \sum_{d_n(x) = 0} f_\theta(x) \tag{3.1}$$

where h is a real number. We can generalize the above definition of a family of penalized disparities (PDs) corresponding to any general disparity generating function $G(\delta)$ as follows:

$$PD(d_n, f_\theta) \equiv \sum_{d_n(x) \neq 0} G\left(\frac{d_n(x) - f_\theta(x)}{f_\theta(x)}\right) f_\theta(x) + [G(-1)h] \sum_{d_n(x) = 0} f_\theta(x) \tag{3.2}$$

Note that for HD, $G_{HD}(-1)$ = 2 and (3.1) is a special case of (3.2). For h = 1, (3.1) gives the ordinary Hellinger distance

$$2 \sum_{d_n(x) \neq 0} (\sqrt{d_n(x)} - \sqrt{f_\theta(x)})^2 + 2 \sum_{d_n(x) = 0} f_\theta(x),$$

of (2.3). When h = 0.5, (3.1) generates

$$2 \sum_{d_n(x) \neq 0} (\sqrt{d_n(x)} - \sqrt{f_\theta(x)})^2 + \sum_{d_n(x) = 0} f_\theta(x),$$

and this member of the PHD family was recommended by Harris and Basu (1994) for efficient and robust estimation for small sample sizes, and by Basu et al. (1996) for testing of hypotheses in small samples. Following the arguments of Harris and Basu (1994), one choice of h in (3.2) is given by h = $1/G(-1)$ since then the PD puts the same weight (equal to one) on the empty cells as the likelihood disparity of (2.2) that generates the MLE.

We now redefine the index parameter of (3.1) and (3.2) by $\lambda = G_{HD}(-1)h$. Then (3.1) and (3.2) are given by

$$PHD_\lambda(d_n, f_\theta) = 2 \sum_{d_n(x) \neq 0} (\sqrt{d_n(x)} - \sqrt{f_\theta(x)})^2 + \lambda \sum_{d_n(x) = 0} f_\theta(x), \tag{3.3}$$

and

$$PD_\lambda = PD_\lambda(d_n, f_\theta) \equiv \sum_{d_n(x) \neq 0} G\left(\frac{d_n(x) - f_\theta(x)}{f_\theta(x)}\right) f_\theta(x) + \lambda \sum_{d_n(x) = 0} f_\theta(x) \tag{3.4}$$

respectively. The $PD_{1.0}$ (for λ=1.0) in the PD family for any given $G(\delta)$ can be expected to play the same role as the $\mathrm{PHD}_{1.0}$ (for λ=1.0) in the PHD family. In the case of the NED, the family of PNEDs is given by

$$PNED_{\lambda}\equiv \sum_{d_n(x)\neq 0} G_{NED}\left(\frac{d_n(x)-f_\theta(x)}{f_\theta(x)}\right)f_\theta(x)+\lambda \sum_{d_n(x)=0} f_\theta(x). \quad (3.5)$$

Following the arguments of Harris and Basu (1994), it can be shown that minimizing a general disparity D is equivalent to maximizing a penalized Kullback-Leibler divergence (Kullback and Leibler, 1951). To see this, for x such that $d_n(x)\neq 0$, define

$$f_\theta^*(x)=d_n(x)exp\left(-G\left(\frac{d_n(x)-f_\theta(x)}{f_\theta(x)}\right)/\left(\frac{d_n(x)}{f_\theta(x)}\right)\right), \quad (3.6)$$

a data driven modification of the model density $f_\theta(x)$. Let $KL(d_n, f_\theta^*)$ denote the Kullback-Leibler divergence between d_n and f_θ^*, defined by

$$KL(d_n, f_\theta^*)=\sum_{d_n(x)\neq 0} d_n(x)ln\left(\frac{d_n(x)}{f_\theta^*(x)}\right).$$

Then, the disparity D can be expressed as

$$D=\sum G\left(\frac{d_n(x)-f_\theta(x)}{f_\theta(x)}\right)f_\theta(x)=KL(d_n, f_\theta^*)+\sum_{d_n(x)=0} f_\theta(x).$$

Thus the disparity D can be thought of as a penalized Kullback-Leibler divergence between d_n and f_θ^*, where the penalty is $\sum_{d_n(x)=0} f_\theta(x)$. Note, however, that the modified function f_θ^* is not a proper density unless $d_n(x)=f_\theta(x)$ for all x.

3.2. Testing

Consider the problem of testing

$$H_o: \theta=\theta_o \text{ against } H_a: \theta\neq\theta_o. \quad (3.7)$$

Basu et al. (1996) considered three classes of tests corresponding to the likelihood ratio test (LRT), the Rao test and the Wald test based on the PHD. One can define these tests based on any penalized disparity (3.4). In particular, we do this for the PNED_λ. Let $\hat{\theta}_{G_{NED},\lambda}$ denote a value of θ that minimizes (3.5). Let I_θ denote the Fisher information about θ in $f_\theta(x)$. The deviance test statistic for the PNED$_\lambda$ has the form $2n[NED_\lambda(d_n, f_{\theta_o}) - NED_\lambda(d_n, f_{\widehat{\theta}_{G_{NED}}}, \lambda)]$, and the Rao test statistic is defined by $n a^T_{G_{NED},\lambda,n,\theta_o}(I_{\theta_o})^{-1} a_{G_{NED},\lambda,n,\theta_o}$, where $a_{G_{NED},\lambda,n,\theta_o} = \partial[NED_\lambda(d_n, f_{\theta_o})]/\partial\theta$. Finally, the Wald test statistic based on the PNED$_\lambda$ is given by $n(\widehat{\theta}_{G_{NED},\lambda} - \theta_o)^T I_{\theta_o}(\widehat{\theta}_{G_{NED},\lambda} - \theta_o)$.

The Hellinger deviance test was proposed by Simpson (1989). By arguments similar to those of Basu et al. (1996) and using results of Lindsay (1994), the limiting distribution of the deviance, Rao and Wald tests based on any general penalized disparity is χ^2_p where p is the dimension of the parameter θ.

4. Simulation Results for Estimators

The data were generated from the $(1-\varepsilon)Poi(\theta=2) + \varepsilon Poi(\theta^* = 15)$ distribution, where ε = 0, 0.1. The target parameter is the mean of the *Poisson*(2) component, and a contaminating distribution is defined by the second component *Poisson*(15) when ε = 0.1. Four different sample sizes, n = 10, 20, 50 and 100, were considered. All results were based on 5000 replications. The simulations were performed using the MICROSOFT FORTRAN POWER STATION on WINDOWS 95.

Harris and Basu (1994) showed that the MPHDE$_{1.0}$ (for λ=1.0) in small samples has the best performance under both contamination and no contamination cases. In Table 1 we present the empirical means and mean square errors (MSEs) of the MPNEDEs for λ = 0.431, 0.718, 1.0 and 1.293 (obtained by λ = G(−1)h for h $\approx$ 0.6, 1.0, 1.4, 1.8) compared to the MPHDE$_{1.0}$ (for λ=1). The MPHDE$_{0.718}$ (for λ=0.718) represents the usual MNEDE. From Table 1, we see that at the model ($\varepsilon=0$ case) the empirical mean values of the MNEDE are closer to the target value than the MPHDE$_{1.0}$ and the MPNEDE$_{1.0}$, while the MSE values of the MNEDE are comparable to those of the MPHDE$_{1.0}$. Under the contamination case ($\varepsilon=0.1$), the MNEDE has the smallest MSE values among all estimators, which is most noticeable for the sample size 10.

To examine the performance of the estimators for other parameter values, we also computed the empirical mean and MSE of the estimators under $(1-\varepsilon)Poi(\theta)+\varepsilon Poi(\theta^*)$, ε = 0, 0.1, for (θ, θ^*) = (2,10), (1,8), and (5,18). The results had a similar pattern as in the case of $(1-\varepsilon)Poi(2)+\varepsilon Poi(15)$, and they are not presented here for brevity.

As a second example of a discrete model, the geometric distribution was considered, in which case the data were generated from $(1-\varepsilon)Geo(\theta=0.5)+ \varepsilon Geo(\theta^*=0.1)$, ε = 0, 0.1. The target parameter was the inverse of the mean of *Geo*(0.5), and sample sizes of 10, 20, 50, and 100 were used. The results are presented in Table 2. We observe that the MNEDE (i.e., $\text{MPNEDE}_{0.718}$) performs better than the $\text{MPHDE}_{1.0}$ and the $\text{MPNEDE}_{1.0}$ in terms of both empirical bias and MSE at the model, while showing the smallest MSE under contamination. Although the $\text{MPNEDE}_{0.431}$ has the best performance among robust estimators at the model ($\varepsilon = 0$), it is generally outperformed by the MNEDE under contamination.

Thus, under the *Poisson* and the *Geometric* distributions settings considered here the $\text{MPNEDE}_{1.0}$ does not show improvements over the MNEDE in small datasets, unlike the $\text{MPHDE}_{1.0}$ over the MHDE. In fact, the MNEDE performs as well as or better than the $\text{MPHDE}_{1.0}$ under both contamination and no contamination cases.

5. Simulation Results for Test Statistics

As in Section 4, all results presented in this section were based on 5000 replications with four different sample sizes 10, 20, 30 and 100 for the deviance tests. We also computed the Rao tests and Wald tests but did not present the numbers for brevity. For the Poisson model, the data were generated from the mixture $(1-\varepsilon)Poi(\theta=2)+\varepsilon Poi(\theta^*=15)$ distributions for ε = 0, 0.1. The empirical levels and powers were computed for $H_o: \theta=2$ and $H_a: \theta=3$ respectively for the nominal level 0.05. The empirical level and power were calculated as the proportion of test statistic values exceeding the $\chi^2(1)$ critical value. These are presented in Tables 3 and 4.

We now discuss the results of Table 3. Under both no contamination and contamination cases, improvements in the level values of the negative exponential deviance test (NEDT) are provided by the penalized negative exponential deviance test for λ=1.0 ($\text{PNEDT}_{1.0}$). At the model, the levels of $\text{PNEDT}_{1.0}$ are the closest to the nominal level 0.05 among all non-LRT tests for sample sizes 10 and 20, while performing well under contamination. Our computations on the empirical levels of the Rao tests and Wald tests (not presented here) showed the following. The levels of the Rao test based on the ordinary NED is overly conservative and the level values of the Rao tests based on the PNEDs increase as the penalty parameter λ increases. On the other hand, the empirical levels of the Wald tests based on the NED and PNED are higher than the nominal level, as for the HD and PHD. The Wald test based on the ordinary NED appeared to have the best overall level performance among all other Wald tests under contaminated and uncontaminated cases.

The empirical powers of the deviance tests based on the hypothesis $H_a: \theta=3$ are presented in Table 4. At the model the power of the $\text{PNED}_{1.0}$ test is the closest to that of the LRT. Under contamination the likelihood ratio test incur a significant loss in power whereas the robust tests maintain their powers better. The empirical powers of the $\text{PNED}_{1.0}$ tests are generally higher than the $\text{PHD}_{1.0}$ and NED tests both under contamination and no contamination. Note that the ordinary Hellinger deviance test has much higher empirical powers than other tests, which is expected since it has much higher empirical level values.

We also did simulations under the mixture $(1-\varepsilon)Poi(\theta)+\varepsilon Poi(\theta^*=8)$ distributions for $H_o: \theta=1$ vs. $H_a: \theta=2$, and under the mixture $(1-\varepsilon)Poi(\theta)+\varepsilon Poi(\theta^*=18)$ distributions for $H_o: \theta=5$ vs. $H_a: \theta=8$. The results were similar to the case discussed above and we do not present the simulation results for brevity.

We studied the performance of the tests under the geometric distribution also. We generated data fromfrom $(1-\varepsilon)Geo(\theta)+\varepsilon Geo(\theta^*=0.1)$ distributions for $H_o: \theta=0.5$ vs. $H_a: \theta=0.3$. The nominal level of 0.10 was considered in this case for sample sizes 10, 20, 50, and 100. The results presented in Tables 5 and 6 seem to be similar to those in the distribution case. The empirical levels of the deviance tests based on the LD, HD, $PHD_{1.0}$, NED, and $PNED_{1.0}$ were plotted in Figure 2 for testing and for sample sizes between 10 and 200 at intervals of 5. Note that in the graph the labels PHD and PNED represent the $PHD_{1.0}$ and $PNED_{1.0}$ tests respectively. The level of the deviance test based on the $PNED_{1.0}$ is seen to be closer to the nominal level than other (including $PHD_{1.0}$) robust tests for small sample sizes up to about 125.

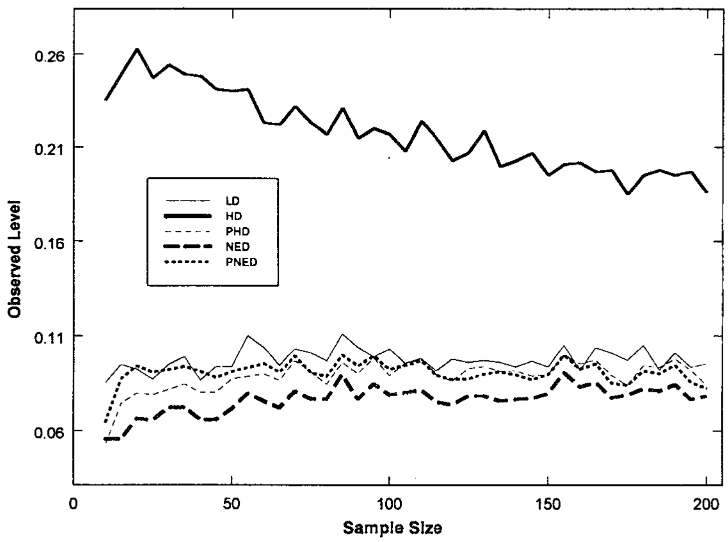

Figure 2. Observed Levels for the Geometric Example

6. Concluding Remarks

We have studied the small sample performance of estimators and tests based on the PNED under Poisson and geometric models. At the model, penalizing the NED did not lead to improvements over the ordinary NED in small samples, unlike in the case of the $PHD_{1.0}$ over the HD. In fact, the NED appears to perform as well as or better than the $PHD_{1.0}$. This can be explained by the fact that the HD is very sensitive to inliers whereas the NED is robust against them, and usually a considerable number of inliers are encountered in small datasets under discrete models in the form of empty cells. However, in testing of hypotheses, the deviance test based on the $PNED_{1.0}$ appears to improve upon that based on the ordinary NED and thus it can be used as a good robust alternative to the tests of hypothesis based on the HD, the PHD, and the NED.

References

[1] Basu, A., Harris, I. R. and Basu, S. (1996). Tests of hypotheses in discrete models based on the penalized Hellinger distance. *Statist. Prob. Letters*, 27, 367-373.

[2] Basu, A. and Lindsay, B. G. (1994). Minimum disparity estimation for continuous models: Efficiency, distributions and robustness. *Ann. Inst. Statist. Math.* 46, 683-705.

[3] Basu, A. and Sarkar, S. (1994). The trade-off between robustness and efficiency and the effect of model smoothing in minimum disparity inference. *J. Statist. Comput. Simul.* 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*, 58, 349-370.

[5] Beran, R. (1977). Minimum Hellinger distance estimates for parametric models. *Ann. Statist.* 5, 445-463.

[6] Harris, I. R. and Basu, A. (1994). Hellinger distance as a penalized log likelihood. *Commun. Statist. Simula.* 23(4), 1097-1113.

[7] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statist.* 22, 79-86.

[8] Lindsay, B. G. (1994). Efficiency versus robustness: the case of minimum Hellinger distance and related methods. *Ann. Statist.* 22, 1081-1114.

[9] Park, C., Basu, A. and Basu, S. (1995). Robust minimum distance inference based on combined distances. *Commun. Statist. Simula.* 24, 653-673.

[10] Simpson, D. G. (1987). Minimum Hellinger distance estimation for the analysis of count data. *J. Amer. Statist. Assoc.* 82, 802-807.

[11] Simpson, D. G. (1989). Hellinger deviance test: efficiency, breakdown points, and examples. *J. Amer. Statist. Assoc.* 84, 107-113.

[12] Tamura, R. N. and D. D. Boos (1986). Minimum Hellinger distance estimation for multivariate location and covariance. *J. Amer. Statist. Assoc.* 81, 223-229.

APPENDIX

Table 1. Empirical Means and MSEs of the estimators for different disparities and sample sizes n = 10, 20, 50 and 100 under model $(1\text{-}\varepsilon)Poisson(2)+\varepsilon Poisson(15)$

	$\varepsilon=0.0$							
	n=10		n=20		n=50		n=100	
Disparity	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
LD	2.002	0.202	1.999	0.100	2.002	0.039	2.002	0.020
HD	1.800	0.280	1.857	0.127	1.921	0.046	1.952	0.023
$PHD_{1.0}$	1.943	0.216	1.962	0.105	1.976	0.041	1.984	0.021
$PNED_{.431}$	2.023	0.227	2.031	0.113	2.018	0.043	2.010	0.021
$PNED_{.718}$(NED)	1.956	0.227	1.983	0.108	1.996	0.041	1.999	0.021
$PNED_{1.0}$	1.900	0.240	1.943	0.110	1.976	0.041	1.988	0.021
$PNED_{1.293}$	1.849	0.264	1.906	0.117	1.958	0.042	1.978	0.021
	$\varepsilon=0.1$							
LD	3.274	3.295	3.293	2.577	3.295	2.053	3.303	1.888
HD	1.821	0.550	1.864	0.148	1.938	0.053	1.979	0.026
$PHD_{1.0}$	1.974	0.389	1.992	0.127	2.004	0.050	2.015	0.026
$PNED_{.431}$	2.054	0.287	2.055	0.134	2.038	0.051	2.028	0.026
$PNED_{.718}$(NED)	1.971	0.283	1.996	0.125	2.011	0.047	2.015	0.025
$PNED_{1.0}$	1.907	0.295	1.950	0.127	1.987	0.048	2.003	0.026
$PNED_{1.293}$	1.850	0.319	1.907	0.134	1.965	0.048	1.991	0.026

Table 2. Empirical Means and MSEs of the estimators for different disparities and sample sizes n = 10, 20, 50 and 100 under model $(1\text{-}\varepsilon)Geometric(0.5)+\varepsilon Geometric(0.1)$

	$\varepsilon=0.0$							
	n=10		n=20		n=50		n=100	
Disparity	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
LD	0.524	0.013	0.514	0.007	0.504	0.003	0.501	0.001
HD	0.615	0.026	0.584	0.014	0.546	0.005	0.528	0.002
$PHD_{1.0}$	0.554	0.016	0.539	0.009	0.520	0.003	0.512	0.001
$PNED_{.431}$	0.511	0.014	0.504	0.008	0.498	0.003	0.498	0.001
$PNED_{.718}$(NED)	0.546	0.015	0.529	0.009	0.511	0.003	0.505	0.001
$PNED_{1.0}$	0.573	0.018	0.548	0.010	0.521	0.003	0.511	0.002
$PNED_{1.293}$	0.597	0.023	0.565	0.012	0.530	0.004	0.516	0.002
	$\varepsilon=0.1$							
LD	0.417	0.028	0.393	0.024	0.371	0.022	0.364	0.021
HD	0.591	0.024	0.557	0.012	0.506	0.003	0.478	0.002
$PHD_{1.0}$	0.518	0.016	0.501	0.009	0.474	0.004	0.460	0.003
$PNED_{.431}$	0.481	0.016	0.474	0.010	0.461	0.005	0.457	0.004
$PNED_{.718}$(NED)	0.521	0.015	0.503	0.009	0.476	0.004	0.464	0.003
$PNED_{1.0}$	0.551	0.017	0.526	0.010	0.488	0.004	0.470	0.003
$PNED_{1.293}$	0.576	0.021	0.545	0.011	0.499	0.004	0.477	0.002

Table 3. Empirical levels of deviance tests for different disparities and sample sizes n = 10, 20, 50 and 100 at nominal level 0.05 under the Poisson model

Contaminating Proportion	Disparity	Model: $(1-\varepsilon)Poisson(2)+\varepsilon Poisson(15)$ n=10	n=20	n=50	n=100
	LD	0.045	0.054	0.050	0.048
	HD	0.118	0.108	0.084	0.077
	$PHD_{1.0}$	0.034	0.044	0.049	0.052
ε=0.0	$PNED_{.431}$	0.017	0.028	0.036	0.043
	$PNED_{.718}$(NED)	0.031	0.036	0.040	0.046
	$PNED_{1.0}$	0.044	0.046	0.043	0.047
	$PNED_{1.293}$	0.065	0.056	0.049	0.051
	LD	0.546	0.713	0.941	0.995
	HD	0.129	0.118	0.090	0.079
	$PHD_{1.0}$	0.042	0.046	0.056	0.061
ε=0.1	$PNED_{.431}$	0.018	0.026	0.039	0.048
	$PNED_{.718}$(NED)	0.033	0.033	0.040	0.048
	$PNED_{1.0}$	0.051	0.046	0.042	0.049
	$PNED_{1.293}$	0.071	0.060	0.049	0.053

Table 4. Empirical powers of deviance tests for different disparities and sample sizes n = 10, 20, 50 and 100 at nominal level 0.05 under the Poisson model

Contaminating Proportion	Disparity	Model: $(1-\varepsilon)Poisson(3)+\varepsilon Poisson(15)$ n=10	n=20	n=50	n=100
	LD	0.475	0.810	0.993	1.000
	HD	0.694	0.910	0.998	1.000
	$PHD_{1.0}$	0.462	0.789	0.994	1.000
ε=0.0	$PNED_{.431}$	0.279	0.646	0.986	1.000
	$PNED_{.718}$(NED)	0.397	0.742	0.992	1.000
	$PNED_{1.0}$	0.491	0.806	0.994	1.000
	$PNED_{1.293}$	0.566	0.849	0.997	1.000
	LD	0.404	0.418	0.455	0.512
	HD	0.659	0.869	0.993	1.000
	$PHD_{1.0}$	0.401	0.703	0.976	1.000
ε=0.1	$PNED_{.431}$	0.242	0.564	0.964	1.000
	$PNED_{.718}$(NED)	0.358	0.674	0.977	1.000
	$PNED_{1.0}$	0.455	0.750	0.984	1.000
	$PNED_{1.293}$	0.530	0.803	0.987	1.000

Table 5. Empirical levels of deviance tests for different disparities and sample sizes n = 10, 20, 50 and 100 at nominal level 0.10 under the geometric model

Contaminating Proportion	Disparity	Model: $(1-\varepsilon)Geometric(0.5)+\varepsilon Geometric(0.1)$ n=10	n=20	n=50	n=100
	LD	0.087	0.092	0.094	0.103
	HD	0.233	0.263	0.241	0.217
	$PHD_{1.0}$	0.055	0.080	0.087	0.089
ε=0.0	$PNED_{.431}$	0.026	0.050	0.059	0.074
	$PNED_{.718}$(NED)	0.058	0.066	0.071	0.079
	$PNED_{1.0}$	0.090	0.094	0.091	0.093
	$PNED_{1.293}$	0.111	0.137	0.122	0.109
	LD	0.335	0.455	0.679	0.855
	HD	0.191	0.189	0.146	0.219
	$PHD_{1.0}$	0.059	0.073	0.136	0.277
ε=0.1	$PNED_{.431}$	0.038	0.074	0.155	0.274
	$PNED_{.718}$(NED)	0.044	0.066	0.126	0.240
	$PNED_{1.0}$	0.057	0.080	0.109	0.209
	$PNED_{1.293}$	0.105	0.102	0.095	0.175

Table 6. Empirical powers of deviance tests for different disparities and sample sizes n = 10, 20, 50 and 100 at nominal level 0.10 under the geometric model

Contaminating Proportion	Disparity	Model: $(1-\varepsilon)Geometric(0.3)+\varepsilon Geometric(0.1)$ n=10	n=20	n=50	n=100
	LD	0.676	0.906	0.999	1.000
	HD	0.949	0.997	1.000	1.000
	$PHD_{1.0}$	0.666	0.904	0.999	1.000
ε=0.0	$PNED_{.431}$	0.398	0.715	0.986	1.000
	$PNED_{.718}$(NED)	0.573	0.864	0.997	1.000
	$PNED_{1.0}$	0.742	0.933	0.999	1.000
	$PNED_{1.293}$	0.819	0.967	0.999	1.000
	LD	0.407	0.447	0.551	0.708
	HD	0.865	0.965	0.996	1.000
	$PHD_{1.0}$	0.492	0.704	0.933	0.994
ε=0.1	$PNED_{.431}$	0.274	0.498	0.871	0.991
	$PNED_{.718}$(NED)	0.448	0.705	0.946	0.996
	$PNED_{1.0}$	0.612	0.821	0.974	0.998
	$PNED_{1.293}$	0.713	0.894	0.985	0.999