Published in final edited form as:

Sankhya Ser B. 2011 May ; 73(1): 144–163. doi:10.1007/s13571-011-0019-7.

NONPARAMETRIC BENCHMARK ANALYSIS IN RISK ASSESSMENT: A COMPARATIVE STUDY BY SIMULATION AND DATA ANALYSIS

RABI BHATTACHARYA and LIZHEN LIN

Department of Mathematics, The University of Arizona, Tucson, AZ, 85721, USA

RABI BHATTACHARYA: rabi@math.arizona.edu; LIZHEN LIN: lizhen@math.arizona.edu

Abstract

We consider the finite sample performance of a new nonparametric method for bioassay and benchmark analysis in risk assessment, which averages isotonic MLEs based on disjoint subgroups of dosages, and whose asymptotic behavior is essentially optimal (Bhattacharya and Lin (2010)). It is compared with three other methods, including the leading kernel-based method, called *DNP*, due to Dette et al. (2005) and Dette and Scheder (2010). In simulation studies, the present method, termed *NAM*, outperforms the *DNP* in the majority of cases considered, although both methods generally do well. In small samples, *NAM* and *DNP* both outperform the MLE.

Keywords

Monotone dose-response curve estimation; effective dosage; nonparametric method; pool-adjacent-violators algorithm; mean integrated squared error; confidence interval; bootstrap

1. Introduction

Given a set of N independent observations for the purpose of estimating some parameter or functional, consider the seemingly naive procedure of (1) dividing the data into r disjoint subsets, (2) estimating the quantity based separately on each subset, and (3) averaging over the r estimates, perhaps after a bias correction. Here r is a nondecreasing function of N , $r \rightarrow \infty$ as $N \rightarrow \infty$. For parametric estimators such as the MLE based on i.i.d. observations and under standard assumptions, this naive estimator does not improve upon the MLE. But as long as $r/N \rightarrow 0$, or, better still, as $r/\sqrt{N} \rightarrow 0$ (in the absence of bias correction), it achieves the same asymptotic efficiency as the MLE. As one can check, with a little effort, that in nonparametric curve estimation based on i.i.d. data, a similar comparison prevails between the kernel based estimate and its naive counterpart.

It may then come as a surprise that, in our present context of bioassay or benchmark analysis in risk assessment, the procedure outlined above improves upon, often dramatically, the nonparametric isotonic MLE considered in Bhattacharya and Kong (2007). Indeed, with a proper choice of r , the new estimator has asymptotically optimal rates of mean integrated squared error (MISE) and asymptotic variance as shown in Bhattacharya and Lin (2010) (See Section 2). It may be pointed out that the kernel-based methods do not improve with this procedure of subgrouping and averaging, although some of them may have fine asymptotic properties (See, e.g., Müller and Schmitt (1988), Dette et al. (2005), and Dette

and Scheder (2010)). It will be shown in this paper, by extensive simulation and data analysis, that the method proposed here performs quite favorably in comparison with other leading nonparametric methods, including the new method due to Dette et al. (2005) and Dette and Scheder (2010) termed *DNP*.

It may be noted that although the problem of bioassay is an old one, effective nonparametric estimation here is relatively new. The importance of good estimation of effective dosage is still a very important problem in biology, medicine and pharmacy. In addition, concerns about the impact of pollutants in the environment on various species, including humans, have added a great sense of urgency to the development of effective means for the estimation of benchmarks in risk assessment (See, e.g., Piegorsch and Bailer (2005), Nitcheva et al. (2007), US EPA (1997)).

Consider quantal dose-response experiments in bioassay where one records 1 for the response of a subject to a drug or a chemical agent and 0 for non-response. Let $F(x)$ be the probability of response to a dose level x . The function $F(\cdot)$ is called the *dose-response curve*, and it is assumed to be *monotone increasing*. The *effective dosage* ζ_p or ED_p for a targeted response (probability) p is defined as

$$\zeta_p = ED_p = F^{-1}(p), 0 \leq p \leq 1; F^{-1}(p) := \inf\{x: F(x) \geq p\}. \quad (1.1)$$

A standard measure of effectiveness of the chemical, or drug, in bioassay is $ED_{0.5}$, often denoted as ED_{50} (See, e.g., Morgan (1992)). In environmental studies, one is generally interested in finding a level of the pollutant such that, when a subject is exposed to it, the probability of response is a small value of p such as 0.05, 0.10, etc (Piegorsch and Bailer (2005), chapter 4).

Suppose that n_i subjects are given a dosage x_i ($i = 1, \dots, m$), where $x_1 < x_2 < \dots < x_{m-1} < x_m$, with the total number of observations $N = \sum_{i=1}^m n_i$. One may assume, without loss of generality, that $0 = x_1 < \dots < x_m = 1$. The number of responses observed at dosage x_i is r_i ($i = 1, \dots, m$). The likelihood function for the estimation of $F(x_i)$, $1 \leq i \leq m$, is

$$L(p_1, \dots, p_m) = \prod_{i=1}^m p_i^{r_i} (1 - p_i)^{n_i - r_i} \quad (0 \leq p_1 \leq \dots \leq p_m \leq 1); [p_i := F(x_i)]. \quad (1.2)$$

The maximum likelihood estimator (MLE) of $(p_1, \dots, p_m)$, under the monotonicity constraint, is given in Ayer et al. (1955) by the following *PAV*, or *pool-adjacent-violators algorithm* (Also see Barlow et al. (1972), p.73, and Cran (1980)):

$$\tilde{p}_i = \max_{0 \leq u < i} \min_{i \leq v < m} \frac{\sum_{j=u}^v r_j}{\sum_{j=u}^v n_j} \quad (1 \leq i \leq m). \quad (1.3)$$

One needs to estimate the entire dose-response curve in order to obtain a consistent estimator of ED_p . This was obtained by linear interpolation in the interval (x_i, x_{i+1}) in Bhattacharya and Kong (2007), where precise conditions for consistency of such an estimator and for proper confidence intervals for ED_p were derived. This method is referred as the *B-K method*, using the terminology in Dette and Scheder (2010). It is quite different from the kernel based nonparametric methods for quantal bioassay, such as those of Müller and Schmitt (1988), Park and Park (2006), Dette et al. (2005) and Dette and Scheder (2010). A description of these methods may be found in the last two articles. Simulation studies in

Dette and Scheder (2010) show that their method, which is an immediate extension of that of Dette et al. (2005), outperforms other kernel based methods, as well as the B-K method, in most cases.

For the DNP method, one first estimates $F(x)$ by a local linear estimator $\hat{F}(x)$ using a symmetric kernel K and a bandwidth λ . The p -th quantile $ED_p = F^{-1}(p)$ is then estimated as

$$\widehat{ED}_p = \int_0^1 \int_{-\infty}^p \frac{1}{\lambda_d} K_d \left(\frac{\hat{F}(x) - u}{\lambda_d} \right) du dx. \quad (1.4)$$

Here K_d is a symmetric kernel with the same properties as K (e.g, $K_d = K$). But λ_d goes to zero faster than λ .

We now introduce the PAV-based *new nonparametric adaptive method NAM*. For simplicity, let $n_i = n$ for all $i = 1, \dots, m$, and let the dosages be equidistant.

As in Bhattacharya and Kong (2007), define $\tilde{F}(x)$ of $F(x)$ as follows:

$$\tilde{F}(x) = \begin{cases} \tilde{p}_i & \text{if } x = x_i \\ \tilde{p}_i + \frac{\tilde{p}_{i+1} - \tilde{p}_i}{x_{i+1} - x_i} (x - x_i) & \text{if } x_i < x \leq x_{i+1}. \end{cases}$$

The inverse of $\tilde{F}$ is the estimate of ED_p as given by:

$$\widetilde{ED}_p = \begin{cases} x_1 & \text{if } p \leq \tilde{p}_1 \\ x_i + \frac{p - \tilde{p}_i}{\tilde{p}_{i+1} - \tilde{p}_i} (x_{i+1} - x_i) & \text{if } \tilde{p}_i < p \leq \tilde{p}_{i+1} \text{ for some } i \\ x_m & \text{if } p > \tilde{p}_m, \end{cases} \quad (1.5)$$

if $\tilde{p}_{i+1} > \tilde{p}_i$ and, more generally, by $\tilde{F}^{-1}(p) = \inf \{x : \tilde{F}(x) \geq p\}$.

From now on, we will assume, for simplicity, that there are m equidistant dosages and the same number n of i.i.d. 0 – 1 valued observations at each dosage

Let $\hat{p}_i$ denote the observed proportion of 1's at dosage x_i . Divide the observed proportions and dosages into r groups, each with approximately $s(n) = m/r$ dosages, and consider the following application of the PAV algorithm to each of the r groups of levels below:

$$\begin{aligned} & [\text{Group 1}]: (x_1, \widehat{p}_1), (x_{r+1}, \widehat{p}_{r+1}), (x_{2r+1}, \widehat{p}_{2r+1}), \dots, (x_m, \widehat{p}_m); \\ & [\text{Group 2}]: (x_1, \widehat{p}_1), (x_2, \widehat{p}_2), (x_{r+2}, \widehat{p}_{r+2}), (x_{2r+2}, \widehat{p}_{2r+2}), \dots, (x_m, \widehat{p}_m); \\ & \dots \dots \dots \\ & [\text{Group } r]: (x_1, \widehat{p}_1), (x_r, \widehat{p}_r), (x_{r+r}, \widehat{p}_{r+r}), (x_{2r+r}, \widehat{p}_{2r+r}), \dots, (x_m, \widehat{p}_m). \end{aligned} \quad (1.6)$$

Note that Group 2 through Group $r - 1$ each has $s(n) + 2$ levels, while Groups 1 and r each has $s(n) + 1$ levels. Also, except for the smallest and the largest levels (with proportions $\hat{p}_1$ and $\hat{p}_m$), the sets of levels covered by them are disjoint. Together, they comprise the set of all m dosages.

By linear interpolation, each Group j ($j = 1, \dots, r$) provides an estimate $\tilde{F}_j$ of the dose-response curve F on $[0, 1]$, and an estimate $\tilde{\zeta}_{p,j}$ of F^{-1} . Note that while F^{-1} is defined on $[F(0), F(1)]$, $\tilde{\zeta}_{p,j}$ and $\tilde{\zeta}_p$ below are defined on $[\tilde{F}(0), \tilde{F}(1)]$. Compute

$$\tilde{F} := (1/r) \sum_{1 \leq j \leq r} \tilde{F}_j, \quad \tilde{\zeta}_p := (1/r) \sum_{1 \leq j \leq r} \widetilde{\zeta_{p,j}}, \quad (1.7)$$

and choose the values of r for which the estimated MISEs of $\tilde{F}$ and $\tilde{\zeta}$ are the smallest. These are the *NAM estimates* of F and F^{-1} . We use the bootstrap (Efron and Tibshirani (2003), Hall (1992)) for estimating MISEs. For the asymptotically optimal error rates of $\tilde{F}$ and $\tilde{\zeta}$, we refer to Bhattacharya and Lin (2010).

Remark 1.1

In benchmark analysis with environmental risk data, the majority of cases use 4, 5, or at most 6 dosages (Nitcheva et al. (2007)). Even with such small m , the NAM method performs very well, as simulation studies with $m = 5$ show (See Section 3).

In the next section, we state the asymptotic optimality results which form the theoretical basis for the finite sample comparisons carried out in Section 3 and data analysis in Section 4.

In Section 3.1, we compare the B-K, DNP, NAM methods and the MLE, using data from four parametric models, namely, Logistic, Probit, Beta and Weibull models. Both DNP and NAM outperform the MLE when the sample size n is not large. The NAM is seen to yield narrower confidence intervals than the DNP in the majority of cases. Although these intervals vary from sample to sample, the overall relative comparisons do not significantly change. Due to lack of space, we omit these calculations with alternate samples. Section 3.2 is devoted to the computation of true bias corrected MISEs in a class of models considered by Dette and Scheder (2010), with $m = 10$, $n = 5$. This provides a validation for the data adaptive NAM.

In section 4, three data examples are used for the application of the different methods to real data, and the methods are compared by the estimated MISEs.

As suggested in Dette and Scheder (2010), the bandwidth λ for the DNP is estimated using a formula of Rice (1984). Bootstrapping and bias correction for confidence intervals are carried out using the Matlab. We have also successfully implemented the computation of the MLEs for Logistic, Probit, Beta and Weibull models on the Matlab.

2. Asymptotic Optimality of NAM

As stated earlier, the new adaptive method NAM presented in this article achieves asymptotic optimality in the class of nonparametric estimators of the dose-response curve F and the effective dosage F^{-1} , under rather mild design restrictions. The basic intuition underlying this method has been described in the Introduction. We state the precise asymptotic results in this section. Their proofs are provided in Bhattacharya and Lin (2010).

One may recall the definition of the mean integrated squared error (MISE) of a curve estimator such as the NAM estimator $\tilde{F}$:

$$MISE(\tilde{F}) = \int_0^1 E(\tilde{F}(x) - F(x))^2 dx. \quad (2.1)$$

We use the notation $\asymp$ to indicate that the ratio of its two sides is bounded away from zero and infinity.

As before, let $\hat{p}_i = r_i/n$ denote the sample proportion of responses to the dose level x_i ($i = 1, \dots, m$). For the sake of simplicity, it is assumed that $n_i = n$ for all i , the dosages x_i are equidistant and lie in $[0, 1]$ (after a linear scaling). We will write $N = mn$ for the total number of observations.

Theorem 2.1

Assume that the dose-response curve F is twice differentiable with F'' bounded, and $f = F'$ is positive on $[0, 1]$. Then the following hold.

- $MISE(\tilde{F}) = O(N^{-4/5})$ as $N \rightarrow \infty$, if $m \propto n^{1/4}$, provided $r = 1$ (or, $r = O(1)$).
- If $m/n^{1/4} \rightarrow \infty$, $m = \alpha n^{3/2}/(\log n)^{5/2}$, then $MISE(\tilde{F}) = O(N^{-4/5})$, provided $r \propto (m^4/n)^{1/5}$.

Remark 2.1

It is known that, under given assumptions on F , the optimal order of the MISE in nonparametric curve estimation, is $O(N^{-4/5})$ (See, e.g., Györfi et al (2002), chapter 5, and Eubank (1999), pp. 16, 129, 172, 196, 258).

We next turn to the crucial estimation of the effective dosage curve $\tilde{\zeta}: p \rightarrow \tilde{\zeta}_p$ (See (1.7)), the main object of interest in this article.

Theorem 2.2

Under the same hypothesis on F as in Theorem 2.1, the following hold.

- If $m \propto n^{1/4}$ then with $r = 1$, one has $MISE(\tilde{\zeta}) = O(N^{-4/5})$.
- If $m/n^{1/4} \rightarrow \infty$, but $m/n^{2/3} \rightarrow \infty$, then $MISE(\tilde{\zeta}) = O(N^{-4/5})$, with the choice of r satisfying $r \propto (m^4/n)^{1/5}$.

Finally, for obtaining desired levels of confidence intervals for $\zeta_p = F^{-1}(p)$, the following result on the asymptotic distribution of $\tilde{\zeta}_p$ is important. For its statement, define $\delta^2(p) \in [1/2, 1]$ by

$$\delta^2(p) \equiv \sum_{i=1}^{m-1} (x_{i+1} - x_i)^{-2} \{(x_{i+1} - x)^2 + (x - x_i)^2\} 1_{I_i}(p) \quad (2.2)$$

where $I_i = [p_i, p_{i+1})$ for $1 \leq i \leq m-2$, $I_{m-1} = [p_{m-1}, p_m]$, $x = \tilde{\zeta}_p$. Next define,

$\bar{\delta}^2(p) = 1/r \sum_{j=1}^r \delta_j^2(p)$, where $\delta_j^2(p)$ is the quantity as in (2.2), but defined for the j -th subgroup of dosages defined in (1.6), in place of the whole group of m dosages.

Theorem 2.3

Assume the same hypothesis as in Theorem 2.1. In addition, assume $m/n^{1/4} \rightarrow \infty$. Then the following hold.

- If $m = \alpha n^{3/2}/\log^{5/2} n$, then with $r \propto (m^4/n)^{1/5}$,

$$\frac{\sqrt{rm}(\tilde{\zeta}_p - E\tilde{\zeta}_p)}{\bar{\delta}(p)} \xrightarrow{\mathcal{L}} N\left(0, \frac{p(1-p)}{f^2(p)}\right) \quad (2.3)$$

as $n \rightarrow \infty$.

- If $m = \alpha n^{2/3}/\log \log n$, then with $r \propto (m^4/n)^{1/5}/(\log \log n)^{6/5}$,

$$\frac{\sqrt{rn}(\tilde{\zeta}_p - \zeta_p)}{\bar{\delta}(p)} \xrightarrow{\mathcal{L}} N\left(0, \frac{p(1-p)}{f^2(p)}\right) \quad (2.4)$$

as $n \rightarrow \infty$.

Remark 2.2

Note that (2.3) implies that the asymptotic variance of $\tilde{\zeta}_p$ around $E\tilde{\zeta}_p$ is $O(N^{-4/5})$, the optimal order. This holds for a very broad range of m . When centered at ζ_p , the asymptotic variance of $\tilde{\zeta}_p$ is only slightly larger, namely, $O(N^{-4/5}(\log \log N)^{6/5})$. This is typically the case for optimal curve estimation. Comparing the ranges of values of m in (a) and (b), one may find it advisable to make a bias correction when m is large relative to n .

3. Comparison of Confidence Intervals

In actual bioassay experiments, it is rare that both m and n are large. In environmental risk assessment, m is almost never more than 6. It is, therefore, not clear a priori if asymptotics provide a valid guide for finite sample comparisons with small or moderate values of n and/or m . By extensive simulations, we show in this section that, indeed, the theoretical results in Section 2 are good indicators of the performance of NAM in realistic situations.

In this section, 95% confidence intervals are obtained by applying different methods to data simulated from different models as described below.

For data simulated from the Logistic model $F(x) = \frac{1}{1 + (\exp(-\alpha - \beta x))}$, we take $\alpha = -20$, $\beta = 10$. The confidence intervals for the MLEs of different quantiles are obtained by fitting the logistic model to data. Denote by $(\hat{\alpha}, \hat{\beta})$ the MLE of (α, β) . Let v_{11} , v_{22} be the asymptotic variances of $\hat{\alpha}$ and $\hat{\beta}$, respectively, and v_{12} the covariance of the two. Also write $d_p = \log(p/(1-p))$, $ED_p = (d_p - \alpha)/\beta$. Then the MLE $\widehat{ED}_p$ for ED_p is given as $(d_p - \hat{\alpha})/\hat{\beta}$, and the 95% CI for ED_p is given by using the Fieller method as $[\widehat{ED}_p - 1.96 \sqrt{\widehat{w}/\hat{\beta}}, \widehat{ED}_p + 1.96 \sqrt{\widehat{w}/\hat{\beta}}]$, where $\widehat{w} = v_{11} + 2v_{12}\widehat{ED}_p + v_{22}(\widehat{ED}_p)^2$ (See Piegorsch and Bailer (2005), pp. 30, 39-40).

We take samples from the Probit model $F(x) = \Phi((x - \mu)/\sigma)$ with $\mu = 0.5$ and $\sigma = 0.3$.

Simulations from the Beta distribution $F(x) = \int_0^x ((B(\alpha, \beta))^{-1} t^{\alpha-1} (1-t)^{\beta-1}) dt$ ($0 \leq x \leq 1$) are carried out with $\alpha = 2$ and $\beta = 3$.

The last set of simulations are from the Weibull model $F(x) = 1 - \exp(-(x/\alpha)^\beta)$ with $\alpha = 2$ and $\beta = 1.5$.

The computations of the MLEs of the parameters of the models are carried out on the Matlab, starting with some properly chosen easily computable consistent estimators as initial values.

Bootstrapping is used for constructing bias-corrected confidence intervals for the DNP and the NAM estimates and the MLEs of the last three models. For the NAM, the value of r in each case is the one among $r = 1, 2, 3$ for which the bootstrap estimate of $MISE(\tilde{\zeta}_p)$ is the smallest.

Simulation studies were carried out for computing confidence bands for the cases $m = 5, 10$, and $n = 5, 10, 25, 50$, with samples from each of the four models-Logistic, Probit, Beta and

Weibull, as mentioned above. However, to save space, graphs of confidence bands are provided for the DNP, NAM and MLE, only for data obtained from the Probit and Weibull models, and omitting $n = 50$. Graphs and tables for all the four models, and for all the four sample sizes, may be found in our website www.math.arizona.edu/~lizhen/about.html. In the graphs, the blue line represents the true curve of quantiles, the green lines represent 95% confidence intervals for the MLE obtained by bootstrapping. The red lines with circles represent the confidence intervals obtained by NAM with $r = 2$ for $m = 5$ and $r = 3$ for $m = 10$, and the dark dashed lines represent the confidence bounds of DNP estimates. When the bias is not corrected, the confidence bands of the DNP and NAM still include the true curve in almost all cases. Here we include only the bias corrected figures for the DNP and NAM, and, of course, Tables 1-12 are the same whether the bias is corrected or not.

A number of significant features are revealed by the graphs and tables. First, both the DNP and the NAM tend to perform better than the MLE for small values of n such as $n = 5$. The MLE also seems to be more bias prone in small samples than the two nonparametric methods. Secondly, for $m = 10$ and $n = 5$, DNP and NAM are more or less on par in performance. This makes sense because the DNP was originally devised for large m and small n (such as $n = 1$) (See Dette et al. (2005)). For larger n ($n = 10$ or more), the NAM outperforms the DNP in almost all cases. A succinct comparison of the two nonparametric methods, by the lengths of confidence intervals, for all four models (Logistic, Probit, Beta and Weibull) is provided in Table 13. Here an entry (i, j) means that for i dose levels the NAM has shorter intervals than the DNP, and for j dosages the DNP has shorter intervals, for the particular model and particular values of (m, n) as indicated.

Although we do not show the confidence intervals for the B-K method ($r = 1$) in order to avoid cluttering up of the graphs, our computations show that they are wider than those of the DNP and NAM in most cases.

An entry (i, j) in Table 13 indicates that for the estimation of i effective dosages (for the particular model, and particular combination of (m, n) for the cell), the NAM provides shorter confidence intervals than the DNP, and for j of them the DNP has shorter confidence intervals.

3.1. True Bias-Corrected MISEs in A Class of Models

In this short subsection, we compute bias-corrected MISEs or integrated variance, for the DNP and NAM estimates for a class of models considered by Dette and Scheder (2010) with $m = 10$, $n = 5$. These computations in Table 14 show that NAM corresponds to $r = 3$ in the population, as it does in the samples. The models are listed as follows:

1. Linear model:
$$F(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 0.3 \\ 0.4x + 0.48 & \text{if } 0.3 \leq x \leq 0.8 \\ x & \text{if } 0.8 \leq x \leq 1 \end{cases}$$
2. Normal Model 1 : $F(x) = \Phi((x - \mu)/\sigma)$, with $\mu = 0.5$, $\sigma = 0.5$.
3. Normal Model 2: $F(x) = \Phi((x - \mu)/\sigma)$, with $\mu = 0.5$, $\sigma = 0.1$.
4. Weibull Model: $F(x) = 1 - \exp(-x^\gamma)$, with $\gamma = 0.52876$.
5. Cauchy Model: $F(x) = 1/2 + 1/\pi \arctan((x - \mu)/\sigma)$, with $\mu = 0.15$, $\sigma = 0.05$.
6. Beta Model: $F'(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-x)^{\beta-1}x^{\alpha-1}$, with $\alpha = 2$, $\beta = 3$.
7. Logistic Model: $F(x) = 1/(1 + \exp(5 - 15x))$.

4. Data Examples

4.1. Data Example I: Remission Cancer Data

In this subsection, in a real data example, estimates of $ED_{0.5}$ (also denoted as ED_{50}) are obtained by fitting parametric models and also using nonparametric methods such as the B-K, NAM and DNP. The source of the data is Lee (1974). This remission cancer data set has 27 observations (binary 0-1 variable with 1 for remission of cancer). There are 14 'dose levels' between 8 and 38 for the explanatory variable labeling index (LI). This labeling index measures proliferative activity of cells after a patient receives an injection of tritiated thymidine, representing the percentage of cells that are labeled. The number of observations are 1 for 6 levels, 2 for 3 levels and 3 for 5 levels. We are interested in estimating the quantile curve and, in particular, the effective dosage $ED_{0.5}$.

Lee (1974) fitted the logistic model to the data, Dette and Scheder (2010) fitted Cauchy and Weibull models. The maximum likelihood estimators for $ED_{0.5}$ are 26.05, 23.65 and 26.09 respectively under the three models. The estimated curves are given by

$$\hat{F}(x) = \frac{1}{1 + \exp(-3.777 + 0.145x)} \text{ for the Logistic model, } \hat{F}(x) = 1 - \exp(-(0.00028x)^{2.3954}) \text{ for the Weibull model } \hat{F}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x - 23.6474}{6.2391}\right) \text{ for the Cauchy model.}$$

For the nonparametric estimates of $ED_{0.5}$, Dette and Scheder (2010) computed the DNP estimate as 20.35229, and the B-K estimate as 8. After checking repeatedly, our calculations show that the B-K estimate is actually 17.8462. Also, for an estimate of $ED_{0.5}$ using the NAM method proposed in this paper, we obtained 21.2 with $r=2$ and 25.1667 with $r=3$. Since the number of dosages is much larger here than the numbers of responses at the individual level, we would prefer $r=3$ for the NAM. Note that the NAM estimate is then closer to all the MLE's than the DNP estimate. The estimates of $ED_{0.5}$ are given in Table 15 for the different methods. The estimates of ED_p with $p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$ for both parametric and nonparametric methods are illustrated in Figure 13. Asymptotic comparisons are not feasible here using bootstrapping, since the numbers of responses n_j are very small. This data example is a follow up of the calculations carried out in Dette and Scheder (2010), and it is only meant to illustrate that the NAM may be applied even with such small sample sizes.

4.2. Data Example II: Data on Insect Mortality

The Logistic model together with the nonparametric methods B-K, NAM ($r=2, 3$) and DNP, are used to fit the data in this example which records kills of *Tribolium confusum* following 5-hour exposures to known concentrations of Carbon Disulphide. The data are due to Strand and may be found in Bliss (1935). This data set has 8 dosages from 49.06 mg/litre to 76.04 mg/litre, with around 30 insects for each dosage.

The estimates of ED_p for $p = 0.1, 0.2, \dots, 0.9$ are obtained and the estimated quantile curves are illustrated in Figure 14.

The MISE is calculated for the Logistic model, Probit model, Weibull model and for each of the nonparametric methods, through bootstrapping. The NAM with $r=3$ yields the smallest estimated MISE 0.9130 compared to 1.0413 of the Probit model and 1.9374 of DNP as shown in Table 16.

4.3. Data Example III: the Toxicity of Rotenone

In this example using data from Martin (1942), we are investigating the relation between the dose of rotenone (mg/l) and the response of exposed insects. The percentage of insects dead or seriously affected out of about 50 insects are recorded for each of the 5 dosages from 2.6mg/l to 10.2mg/l. The parametric estimates are obtained by fitting the Logistic, Probit and Weibull to data together with the nonparametric estimates of the B-K, NAM and DNP. The results are shown in Figure 15.

Comparisons are carried out among the different methods by their estimated MISEs through bootstrapping as shown in Table 17. The Weibull models yields the smallest estimated MISE 0.1626 followed by that of the NAM estimates which is 0.1888.

5. Concluding Remark

Simulation studies and data analysis carried out for this article show that the new method NAM for risk assessment in bioassay and environmental studies perform remarkably well for small, moderate and large sample sizes that are generally used in practice. This also demonstrates that the desirable asymptotic properties of the method, derived in Bhattacharya and Lin (2010), are good guides for its finite sample behavior.

Acknowledgments

The authors wish to thank Professors Pranab K. Sen and Walt Piegorsch for helpful suggestions.

Research supported by NIH grant R21-ES016791 and NSF grant DMS 0806011

References

1. Ayer M, Brunk HD, Ewing GM, Reid WT, Silverman E. An empirical distribution function for sampling with incomplete information. *Ann Math Statist.* 1955; 26:641–647.
2. Barlow, RE.; Bartholomew, DJ.; Bremner, JM.; Brunk, HD. *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression.* Wiley; London: 1972.
3. Bhattacharya R, Kong M. Consistency and asymptotic normality of the estimated effective dose in bioassay. *Journal of Statistical Planning and Inference.* 2007; 137:643–658.
4. Bhattacharya R, Lin L. An adaptive nonparametric method in benchmark analysis for bioassay and environmental studies. *Statistics & Probability Letters.* 2010 In Press.
5. Bliss CI. The calculations of dose-mortality curve (with an appendix by Fisher, R.A.). *Annals of Applied Biology.* 1935; 22:134–167. Table IV.
6. Cran GW. AS149 amalgamation of means in the case of simple ordering. *Appl Statist.* 1980; 29(2): 209211.
7. Dette H, Scheder R. A finite sample comparison of nonparametric estimates of the effective dose in quantal bioassay. *Journal of Statistical Computation and Simulation.* 2010; 80(5):527–544.
8. Dette H, Neumeyer N, Pliz KF. A note on nonparametric estimation of the effective dose in quantal bioassay. *J Amer Statist Assoc.* 2005; 100:503–510.
9. Efron, B.; Tibshirani, RJ. *An Introduction to the Bootstrap.* Chapman & Hall; London: 1993.
10. Eubank, RL. *Nonparametric Regression and Spline Smoothing.* 2. Marcel Dekker; New York: 1999.
11. Györfi, L.; Kohler, M.; Krzyżak, A.; Walk, H. *A Distribution-Free Theory of Nonparametric Regression.* Springer-Verlag; New York: 2002.
12. Hall, PG. *The Bootstrap and Edgeworth Expansion.* Springer; New York: 1992.
13. Lee ET. A computer program for linear logistic regression analysis. *Comput Programs Biomed.* 1974; 4:80–92. [PubMed: 4434745]

14. Martin JT. The problem of the evaluation of rotenone-containing plants. VI. The toxicity of l-elliptone and of poisons applied jointly, with further observations on the rotenone equivalent method of assessing the toxicity of derris root. *Ann Appl Biol.* 30:293–300.
15. Morgan BJT. Analysis of Quantal Response Data. Monographs on Statistics and Applied Probability. 1992; 46
16. Müller HG, Schmitt T. Kernel and probit estimation in quantal bioassay. *J Amer Statist Assoc.* 1988; 83(403):750–759.
17. Nitcheva DK, Piegorsch WW, West RW. On use of the multistage dose-response model for assessing laboratory animal carcinogenicity. *Regulatory Toxicology and Pharmacology.* 2007; 48:135–147. [PubMed: 17490794]
18. Park D, Park S. Parametric and nonparametric estimators of $ED_{100\alpha}$. *Journal of Statistical Computation and Simulation.* 2006; 76(8):661–672.
19. Piegorsch, WW.; Bailer, AJ. *Analyzing Environmental Data.* John Wiley & Sons; 2005.
20. U.S.EPA. Exposure Factors Handbook (Final Report). U.S. Environmental Protection Agency; Washington, DC: 1997. EPA/600/P-95/002F a-c

Figure 1. [Probit]

Figure 2. [Probit]

Figure 3. [Probit]

Figure 4. [Probit]

Figure 5. [Probit]

Figure 6. [Probit]

Figure 7. [Weibull]

Figure 8. [Weibull]

Figure 9. [Weibull]

Figure 10. [Weibull]

Figure 11. [Weibull]

Figure 12. [Weibull]

Figure 13.

Figure 14.

Figure 15.

Table 1

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=5$

DNP:	0.1487	0.2320	0.2779	0.2871	0.2944	0.3224	0.3601	0.4471	0.4750
NAM:	0.2331	0.2479	0.2861	0.3009	0.3627	0.3681	0.3795	0.3597	0.2624

Table 2

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=10$

DNP:	0.2159	0.3341	0.4126	0.4264	0.3729	0.3423	0.3422	0.4062	0.2959
NAM:	0.2144	0.2835	0.3498	0.3716	0.3611	0.3404	0.3265	0.4057	0.2604

Table 3

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=25$

DNP:	0.2382	0.2785	0.2533	0.2119	0.2286	0.2034	0.2014	0.2269	0.2443
NAM:	0.1575	0.2329	0.2091	0.1857	0.1884	0.1817	0.1797	0.1939	0.1954

Table 4

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=5$

DNP:	0.1238	0.1962	0.2452	0.3028	0.3492	0.3638	0.3565	0.3171	0.2571
NAM:	0.2030	0.2497	0.3244	0.3479	0.3625	0.3627	0.3691	0.3409	0.2962

Table 5

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=10$

DNP:	0.1916	0.2217	0.2194	0.2188	0.2293	0.2444	0.2735	0.2962	0.3149
NAM:	0.1534	0.1904	0.2139	0.2072	0.2229	0.2519	0.2580	0.2700	0.2671

Table 6

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=25$

DNP:	0.1380	0.2559	0.2370	0.2489	0.2447	0.1917	0.1216	0.1506	0.2003
NAM:	0.1227	0.1853	0.2108	0.2197	0.1695	0.1396	0.1215	0.1169	0.1274

Table 7

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=5$

DNP:	0.6969	1.0458	1.1129	1.1841	1.2247	1.1968	1.1445	1.0868	0.9663
NAM:	0.9469	1.0721	1.1421	1.1569	1.1716	1.1863	1.1274	0.9948	0.6963

Table 8

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=10$

DNP:	0.2949	0.5102	0.6378	0.7579	0.8538	1.0105	1.0884	1.2061	1.2235
NAM:	0.1982	0.4228	0.5281	0.6325	0.7405	0.9112	1.0303	1.1790	1.2932

Table 9

The Length of Confidence Intervals for DNP and NAM($r=2$) for $m=5, n=25$

DNP:	0.1768	0.4657	0.5843	0.6694	0.8776	0.8382	0.7878	0.8514	1.3161
NAM:	0.1219	0.3443	0.4422	0.5409	0.6491	0.6968	0.7236	0.8510	0.8382

Table 10

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=5$

DNP:	0.3918	0.7013	0.8197	0.8952	0.9395	0.9487	0.9247	0.8603	0.8180
NAM:	0.6207	0.6573	0.8481	0.8931	0.9406	0.9233	0.8597	0.8192	0.7286

Table 11

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=10$

DNP:	0.2442	0.5456	0.7215	0.9547	1.1352	1.2921	1.3359	1.2411	1.1389
NAM:	0.2163	0.4406	0.5700	0.7428	0.9899	1.1956	1.2320	1.3621	1.9545

Table 12

The Length of Confidence Intervals for DNP and NAM($r=3$) for $m=10, n=25$

DNP:	0.1969	0.5101	0.7398	0.7254	0.6308	0.6309	0.8897	1.0056	0.8166
NAM:	0.1608	0.3385	0.4566	0.5475	0.5700	0.5768	0.6072	0.7201	0.8915

Table 13

Comparison by Number of Shorter Confidence Intervals

Model				
<i>m</i>	<i>n</i>	Logistic	Probit	Beta Weibull
5	5	(5.4)	(2.7)	(5.4) (6.3)
5	10	(6.3)	(9.0)	(9.0) (8.1)
5	25	(9.0)	(9.0)	(9.0) (9.0)
5	50	(9.0)	(9.0)	(9.0) (9.0)
10	5	(4.5)	(2.7)	(5.4) (6.3)
10	10	(7.2)	(8.1)	(7.2) (7.2)
10	25	(8.1)	(9.0)	(7.2) (8.1)
10	50	(9.0)	(8.1)	(9.0) (9.0)

Table 14

Integrated variance with $m=10$, $n=5$.

	B-K	r=2	r=3	DNP	NAM
Linear Model:	0.0140	0.0109	0.0051	0.0088	0.0051($r=3$)
Normal1 Model:	0.0192	0.0150	0.0077	0.0130	0.0077($r=3$)
Normal2 Model	0.0028	0.0021	0.0010	0.0014	0.0010($r=3$)
Weibull Model:	0.0263	0.0210	0.0127	0.0210	0.0127($r=3$)
Cauchy Model:	0.0051	0.0045	0.0027	0.0030	0.0027($r=3$)
Beta Model:	0.0076	0.0055	0.0028	0.0039	0.0028($r=3$)
Logistic Model:	0.0036	0.0025	0.0015	0.0017	0.0015($r=3$)

Table 15

The estimates of ED_{50} using different methods.

Methods:	B-K	r=2	r=3	DNP	Logistic	Cauchy	Weibull
$\widehat{ED}_{50}$:	17.8462	21.2	25.1667	20.35229	26.5	23.65	26.09

Table 16

Estimated MISEs of the estimates of the B-K, DNP, NAM and MLEs.

Methods:	B-K	r=2	r=3	DNP	NAM (r=3)	Logistic	Probit	Weibull
MISE:	2.2090	1.3258	0.9130	1.9374	0.9130	1.2355	1.0413	1.2077

Table 17

Estimated MISEs of the estimates of the B-K, DNP, NAM and MLEs.

Methods:	B-K	r=2	DNP	NAM (r=2)	Logistic	Probit	Weibull
MISE:	0.4070	0.1888	0.3512	0.1888	0.2066	0.1908	0.1626