Journal of Statistical Computation and Simulation
Vol. 78, No. 5, May 2008, 429–442

On the multivariate predictive distribution of multi-dimensional effective dose: a Bayesian approach

JIALIANG LI*†, CHUNMING ZHANG‡, ERIK V. NORDHEIM‡ and CHARLES E. LEHNER§

†Department of Statistics and Applied Probability, National University of Singapore, Block 16, Level 7, 6 Science Drive 2, Singapore 117546, Singapore
‡Department of Statistics and §Department of Surgical Science, University of Wisconsin, Madison, USA

(Received 16 August 2006; in final form 28 November 2006)

We propose a Bayesian procedure to sample from the distribution of the multi-dimensional effective dose. This effective dose is the set of dose levels of multiple predictive factors that produce a binary response with a fixed probability.We apply our algorithms to parametric and semiparametric logistics regression models, respectively. The graphical display of random samples obtained through Markov chain Monte Carlo can provide some insight into the predictive distribution.

Keywords: Posterior distribution; Multi-dimensional effective dose; Multivariate joint distribution; Simulating conditional distribution

1. Introduction

The methodology using the Bayesian paradigm has advanced tremendously in modern statistical analysis. Successful application of Bayesian data analysis have appeared in many scientific fields, including business, epidemiology, genetics, geography, sociology, psychometrics and economics [1]. Development of computational methods such as Markov chain Monte Carlo (MCMC) makes it feasible to evaluate complicated posterior distribution for model parameters in a relatively short time. We employ such an idea in this article to solve an underappreciated statistical problem.

Suppose in a dose response experiment we have multiple predictors, say $X_1, \ldots, X_k$, and observe the binary response Y of subjects under different combinations of values of these predictors. We are interested in finding the possible values of $X_1, \ldots, X_k$ such that they generate the outcome Y at a given probability p. We call the set of all joint values of $X_1, \ldots, X_k$ a multi-dimensional effective dose since such a parameter has been conventionally called effective dose $100p$ (ED$100p$) in a bioassay problem when $k = 1$ [2]. Whereas researchers have worked out useful solutions for the one-dimensional effective dose problem, few of

*Corresponding author. Email: stalj@nus.edu.sg

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00949650601141712

them have provided adequate operational details for conducting statistical inference about the multi-dimensional effective dose ($k > 1$).

Frequentist solutions of the estimation and confidence regions of multi-dimensional effective dose have recently been proposed by Li *et al.* [3] under a logistic regression model, motivated by analyzing the risk of decompression sickness (DCS) among deep sea divers. In this article, we treat the same problem from a Bayesian perspective and contribute an alternative approach to determine multi-dimensional effective dose. Without deriving any asymptotic approximation formula, we obtain numerical solutions straightforwardly from computation and simulation.

As is well known to statisticians, in a Bayesian statistics framework, model parameters are treated as random variables, living in suitably defined probability spaces. Therefore, we treat the multi-dimensional effective dose as a random vector with a probability distribution. We thus aim at finding its posterior distribution given the observed sample and prior information. Such a probability distribution is called a prediction distribution in Bayesian analysis and provides the basis for making inferences. We propose algorithms to simulate the conditional distribution of each univariate component as well as the joint distribution of the multivariate under the logistic regression model.

Early work on estimation of effective dose using a Bayesian approach includes Freeman [4] and Racine *et al.* [5]. Freeman [4] introduced Bayesian decision theory in the sequential design of experiments for estimating median lethal dose (ED50). The instructive account of Racine *et al.* [5] exemplified how to make inference about the ED50 from its posterior distribution and compared the Bayesian solution with the frequentist solution. Nonetheless, both articles only considered limited computation methods which might not be fully representative of current Bayesian techniques available for all kinds of scientific problems. Moreover, most of the previous work concentrated on one-dimensional effective dose and were not directly generalizable to multi-dimensional effective dose. Our article thus has a two-fold advantage: first, to update the knowledge of Bayesian treatment on one type of problems in bioassay experiments, and second, to extend the posterior analysis to multi-dimensional effective dose.

We make one further attempt to study a semiparametric model with Bayesian methods in this article as we believe that sometimes the biological mechanism between dose and response might be too complicated to follow a simple parameterized linear form. The theoretical properties of such a model have been discussed in other articles [6, 7]. We propose a method to simulate the conditional posterior distribution for multi-dimensional effective dose by using a pseudo posterior approach.

In section 2, we describe the general methods and algorithms for simulating the prediction distribution of multi-dimensional effective dose under parametric and semiparametric regression models, respectively. In section 3, we adapt our computing procedure to a real world example about the risk analysis of DCS. We conclude in section 4 with discussions of merits and limitations of our methods and indicate some future research direction.

2. Methods and algorithms

2.1 *Case I: Parametric logistic regression model*

2.1.1 Posterior distribution. A binary logistic regression model for independent observations $\{(Y_i, X_{1i}, \ldots, X_{ki}) : i = 1, \ldots, n\}$ is

$$\log \frac{p_i}{1 - p_i} = \mathbf{X}_i^{*T} \beta^*, \quad i = 1, \ldots, n, \tag{1}$$

where $p_i = E(Y_i|\mathbf{X}_i)$, $\mathbf{X}_i^* = (1, \mathbf{X}_i^T)^T = (1, X_{1i}, X_{2i}, \ldots, X_{ki})^T$ and $\beta^* = (\beta_0, \beta^T)^T = (\beta_0, \beta_1, \ldots, \beta_k)^T$. The multi-dimensional effective dose Θ_p is the set of values of $\mathbf{X}$ which satisfies the following condition

$$\Theta_p = \left\{ \mathbf{x} \in \mathbb{R}^k : \beta_0 + \mathbf{x}^T\beta = \log \frac{p}{1-p}, \text{ for a given } p \right\}. \tag{2}$$

It is well known that fitting model (1) through maximum likelihood estimation results in an asymptotically consistent and efficient estimator $\hat{\beta}^*$ under regularity conditions. A frequentist solution is then to substitute $\hat{\beta}^*$ into (2) and obtain the estimator $\hat{\Theta}_p$ [3].

From a Bayesian perspective, the regression parameter β^* is stochastic. Therefore the set (2), which depends on β^*, is also a random set with a probability law.

Suppose the prior distribution of β^* is $\pi(\beta^*)$, which can be obtained from past experience, historical data or subject matter experts' opinion. The posterior distribution of β^* given the data is then

$$\begin{aligned} f(\beta^*|\text{data}) &\propto f(\text{data}|\beta^*)\pi(\beta^*) \\ &= \prod_{i=1}^{n} f(Y_i|X_{1i}, \ldots, X_{ki}, \beta^*)\pi(\beta^*) \\ &= \frac{\exp\{\sum_{i=1}^{n} \mathbf{X}_i^{*T}\beta^* Y_i\}}{\prod_{i=1}^{n}(1 + \exp\{X_i^{*T}\beta^*\})}\pi(\beta^*). \end{aligned} \tag{3}$$

We note that the distribution of the observed data is the distribution of the observed response Y given the fixed covariates $\mathbf{X}$.

When the dimension of β^* is small (*e.g.* $k \leq 3$), we can discretize the above distribution on a grid of points over the range of β^* and evaluate the density values for these grid points. Then we can draw random samples from this discrete distribution easily with a simple random variable generation routine [1, Chapter 10]. In the examples we consider in this article, the dimension of covariates is relatively small. Therefore this direct sampling method appears sufficient for our purposes.

However, the above discretizing-and-sampling method becomes expensive to implement when the dimension k is large (*e.g.* >3). In general we can choose the 'rejection sampling' approach which has been introduced in a generalized linear mixed-effects model framework by Zeger and Karim [8]. Basically we can generate a candidate $\tilde{\beta}^*$ from a known distribution $g(\beta^*)$ such as the multivariate normal distribution and accept $\tilde{\beta}^*$ if $f(\tilde{\beta}^*|\text{data})/(c \cdot g(\tilde{\beta}^*))$ is less than a random number uniformly distributed on [0, 1], where c is such that $c \cdot g(\beta^*) > f(\beta^*|\text{data})$ over the range of β^*. When the sample size is large, $f(\beta^*|\text{data})$ approximates the multivariate normal density with mean equal to the maximum likelihood estimator $\hat{\beta}^*$ and dispersion matrix equal to the inverse of the Fisher information $\mathbf{I}_{\beta^*}$ [9]. Consequently such an approach will not make too many rejections and is computationally economic. Details about this algorithm can be found in books on Bayesian statistical computing like ref. [1] and ref. [10].

The sampling technique for generating β^* is the basic construct for simulating the various distributions in this article. In the following presentation we use the discretizing-and-sampling method without re-stating the full details.

2.1.2 Algorithm for simulating the conditional distribution. Without loss of generality, we consider the $(x_2^*, \ldots, x_k^*)$−conditioning effective dose Θ_p^*, which is the value of X_1 to yield the binary outcome with a success probability p when $(X_2, \ldots, X_k)$ are given to be $(x_2^*, \ldots, x_k^*)$.

We can generate a large number, say M, of β^* independently from the posterior density $f(\beta^*|\text{data})$ and then evaluate the $(x_2^*, \ldots, x_k^*)$–conditioning effective dose by calculating

$$x_1(p)^* = \frac{\log(p/(1-p)) - \beta_0 - \sum_{j=2}^{k} \beta_j x_j^*}{\beta_1} \tag{4}$$

for each simulated β^*. We then obtain M independent realizations of $x_1(p)^*$ from its conditional distribution.

We note that a Bayesian $100(1-\alpha)\%$ credible interval for $x_1(p)^*$ can be constructed by picking the lower and upper $\alpha/2$ quantiles of the generated samples of conditional effective doses.

The Bayesian inference on the conditional distribution of X_1 given $(X_2, \ldots, X_k)$ for a specific p gives a solution similar to that of the frequentist inference on the point and interval estimations of the $(x_2^*, \ldots, x_k^*)$, conditioning effective dose $x_1(p)^*$. As we will see in the example, the Bayesian credible interval for $x_1(p)^*$ obtained by assuming a non-informative prior is numerically identical to the asymptotic pointwise confidence interval of $x_1(p)^*$.

2.1.3 Algorithm for simulating the joint distribution. It is not easy to deal with the joint multivariate distribution of Θ_p directly. By the generalized Hammersley–Clifford theorem [11] it suffices to use the conditional distribution to describe the joint distribution. The theorem, however, requires the following important assumption: for a random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$, if $P(\mathbf{X}_i = \mathbf{x}_i) > 0$ for each i, then $P(\mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n) > 0$. This so-called positivity condition is assumed throughout this article.

We consider the distribution of X_i conditional on $\mathbf{X}_{-i} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$ for a specific p. We note that when $\mathbf{X}_{-i}$ is given, the random behavior of X_i is completely correlated with the random variation of β^*. Hence we have

$$f_p(X_i|\mathbf{X}_{-i}) = f_p(X_i, \beta^*|\mathbf{X}_{-i}) = f_p(X_i|\beta^*, \mathbf{X}_{-i}) f_p(\beta^*|\mathbf{X}_{-i}) \tag{5}$$

Since the posterior distribution of β^*, $f(\beta^*|\text{data})$ is fixed once we observe the data, $f_p(\beta^*|\mathbf{X}_{-i}) = f(\beta^*|\text{data})$ does not depend on p or $\mathbf{X}_{-i}$. We also note that $f_p(X_i|\beta^*, \mathbf{X}_{-i})$ is simply a point mass at $((\log(p/(1-p)) - \beta_0 - \sum_{j \neq i} \beta_j X_j)/\beta_i)$. We construct a Gibbs sampler by using the full conditionals $f_p(X_i|\mathbf{X}_{-i})$ [10] such that all of the simulations may be univariate even for a multivariate distribution.

Each Markov chain is formed in the following way. At the current iteration, suppose we have $\mathbf{X}^{(t)} = (X_1^{(t)}, \ldots, X_k^{(t)})$. Given $(X_2^{(t)}, \ldots, X_k^{(t)})$, we generate a β^* from $f(\beta^*|\text{data})$ and then calculate $X_1^{(t+1)}$ by using this β^*; next, conditional on $(X_1^{(t+1)}, X_3^{(t)}, \ldots, X_k^{(t)})$, we generate another β^* from $f(\beta^*|\text{data})$ and calculate $X_2^{(t+1)}$; …; we repeat the above procedure until $X_k^{(t)}$ is updated to $X_k^{(t+1)}$. We then move to the next iteration and carry out the same updating procedure. This chain is ended at a large number, T, after convergence to a stable distribution and the last value of $\mathbf{X}^{(T)}$ is outputted as a sample from the multivariate distribution of Θ_p.

At every iteration, the updating of the k-dimension $\{\mathbf{X}^{(t)}\}$ is carried out by updating k univariate elements. Such an advantage is widely enjoyed by many types of Gibbs sampler algorithms. Not only is the sequence $\{\mathbf{X}^{(t)}\}$ a Markov chain, but also each subsequence $\{X_i^{(t)}\}$ $(i = 1, \ldots, k)$ is a Markov chain.

We replicate a large number, N, of such Markov chains (each being of length T) and obtain a random sample of size N from the joint distribution of k-dimensional effective dose

Θ_p. From this simulated sample, we can then construct the $100(1-\alpha)\%$ highest posterior density (HPD) credible region for Θ_p. The region has probability equal to the desired confidence level and has higher density for every value inside the region than every value outside of it [1].

The Bayesian inference on the joint multivariate distribution of Θ_p is closely related to the frequentist simultaneous inference about Θ_p. Usually the HPD credible region is of a different form from the simultaneous confidence region. In the example presented in this article, the HPD credible region is a bounded solution and thus more desirable compared to the unbounded band by the frequentist approach.

2.2 *Case II: Semiparametric regression model*

2.2.1 Pseudo posterior probability. Sometimes a parametric model may not be a good model to describe the biological mechanism between the predictor variables and the response variable. Such concern requests us to consider more complicated models such as the following semiparametric model

$$\begin{aligned}\log \frac{p_i}{1-p_i} &= \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \gamma(X_{k+1,i}) \\ &= \mathbf{X}_i^T \beta + \gamma(X_{k+1,i}),\end{aligned} \tag{6}$$

where the logit transformation of mean response is considered to be linearly dependent on the $\mathbf{X}_i = (X_{1i}, \ldots, X_{ki})^T$ and is dependent on $X_{k+1,i}$ through a smooth function $\gamma(\cdot)$. Both $\beta = (\beta_1, \ldots, \beta_k)^T$ and $\gamma(\cdot)$ are unknown.

This model can be fitted by using a local maximum likelihood estimation method [6], where $\gamma(\cdot)$ is locally approximated by a polynomial function. In Li's PhD dissertation [7], the author adopted this semiparametric model and subsequently estimated the multi-dimensional effective dose which is defined in this case as

$$\Theta_p = \left\{ (\mathbf{x}, x_{k+1}) \in \mathbb{R}^{k+1} : \mathbf{x}^T \beta + \gamma(x_{k+1}) = \log \frac{p}{1-p}, \text{ for a given } p \right\}. \tag{7}$$

Given the non-parametric estimator $\hat{\gamma}(\cdot)$, the likelihood of β is called a pseudo likelihood [12]. We then consider the following pseudo posterior probability

$$\begin{aligned} f(\beta|\text{data}) &\propto f(\text{data}|\beta)\pi(\beta) \\ &= \prod_{i=1}^{n} f(Y_i|X_{i1}, \ldots, X_{ik}, \beta, \hat{\gamma})\pi(\beta) \\ &= \frac{\exp[\sum_{i=1}^{n}\{\mathbf{X}_i^T\beta + \hat{\gamma}(X_{k+1,i})\}Y_i]}{\prod_{i=1}^{n}[1+\exp\{\mathbf{X}_i^T\beta + \hat{\gamma}(X_{k+1,i})\}]}\pi(\beta) \end{aligned} \tag{8}$$

In the following, we discuss algorithms to simulate the conditional distribution of X_i given $(\mathbf{X}_{-i}, X_{k+1})$ based on the pseudo posterior probability (8).

2.2.2 Algorithm for simulating the conditional distribution. Similar to the algorithm for simulating the conditional distribution under the parametric logistic model, we could consider the $(x_2^*, \ldots, x_k^*, x_{k+1}^*)$, conditioning effective dose $x_1(p)^*$, which is the value of

X_1 to yield outcome with a success probability p when $(X_2, \ldots, X_k, X_{k+1})$ are given to be $(x_2^*, \ldots, x_k^*, x_{k+1}^*)$.

We generate a large number, say M, of β from its posterior density $f(\beta|\text{data})$ and then evaluate the $(x_2^*, \ldots, x_k^*, x_{k+1}^*)$–conditioning effective dose by calculating

$$x_1(p)^* = \frac{\log(p/(1-p)) - \sum_{j=2}^{k} \beta_j x_j^* - \hat{\gamma}(x_{k+1}^*)}{\beta_1} \tag{9}$$

for each simulated β. We then obtain M realizations of $x_1(p)^*$ from the conditional distribution.

We note that a $100(1-\alpha)\%$ credible interval for $x_1(p)^*$ can be constructed by picking the lower and upper $\alpha/2$ quantiles of the simulated values of conditional effective doses.

3. Applications

University of Wisconsin Sea Grant-supported researchers have conducted extensive investigations to improve our knowledge of the body's susceptibility to DCS while diving. In a recent study, these researchers have conducted studies to assess the impact of two key risk factors: the pressure that a diver undergoes, usually measured by how deep he or she descends under the water, and the total exposure duration.

In this section we re-visit the data in ref. [3] to illustrate the benefits of a Bayesian analysis in finding the predictive distribution for multi-dimensional effective dose. In the Sea Grant study, sheep used as an animal model for humans, were experimental subjects to determine the effect on DCS of dive depth and duration. Each sheep underwent simulated dives (in a pressure chamber) with a designed pressure and duration and its outcomes for central nervous system DCS (CNS-DCS), limb bends (LB), respiratory DCS (RDCS) and mortality were determined thereafter. The pressure was measured in absolute atmospheres and duration at depth was measured in minutes. In the following analysis, we take log base 10 transformations for both predictor variables. All observed outcomes are coded as dichotomous variables. Our data consist of a sample of size $n = 1108$.

One of the key research questions was how to determine the values of pressure and duration that can yield the outcome(s) with a specific probability p. A frequentist analysis for multi-dimensional effective dose has been presented by Li *et al.* [3]. Parametric estimation and its associated confidence regions were given. In this article, we study the multivariate predictive distribution of this multi-dimensional effective dose by using a Bayesian approach.

As discussed in ref. [3], logistic regression model is adequate to model the dependence of mortality or RDCS outcome on covariates, but a semiparametric model is needed to fit LB or CNS-DCS appropriately. Herein, we include the results for mortality using logistic regression and the results for LB using semiparametric logistic regression.

3.1 *Case I: logistic regression model*

We analyze the mortality outcome by using the logistic regression model. The two predictors are X_1, log base 10 pressure and X_2, log base 10 duration.

We consider three types of prior distributions which correspond to none, moderate and strong prior information about the parameters.

$$\text{Prior A: } \pi(\beta^*) \propto 1,$$

$$\text{Prior B: } \frac{\pi(\beta^*) \propto \exp\{\beta_0 - \exp(\beta_0)/\eta\}}{\eta},$$

$$\text{Prior C: } \pi(\beta^*) \propto \frac{\exp\{\mathbf{a}^T\beta^*\}}{\prod_{i=1}^{n}(1+\exp\{\mathbf{X}_i^{*T}\beta^*\})^{\lambda}}.$$

Prior A is independent and locally uniform in the $k+1$ bounded parameter space. We might use a uniform prior if we have no prior knowledge about the parameters, or we want to present a simple analysis of this experiment alone. Prior B assumes an exponential prior on $\log \beta_0$ and flat priors on other regression coefficients. The parameter η is chosen to be $\exp(\hat{\beta}_0 + \gamma)$ as suggested in ref. [10], where $\gamma = 0.577216$ is the Euler's constant. Such a choice makes the prior mean of β_0 is equal to the maximum likelihood estimator $\hat{\beta}_0$. Prior C is conjugate with respect to the likelihood function. Indeed, if we denote the entire family of prior C as $P(\mathbf{a}, \lambda)$, the posterior is in the same family of distributions but with different parameters $P(\mathbf{a} + \sum_{i=1}^{n} \mathbf{X}_i^* Y_i, \lambda + 1)$. We set $\mathbf{a} = (1, \bar{X}_1, \bar{X}_2)^T$ and $\lambda = 0.5$. See Bedrick *et al.* [13] for more discussions on prior specification for logistic regression and other generalized linear models.

3.1.1 Conditional distributions. The scatter plot of simulated samples from conditional distributions of the two components of $\Theta_{0.1}$ is shown in figure 1. At each of fifty equally spaced points over the experimental range of X_1 and X_2, we draw a sample of 1000 values of the conditional effective dose from the posterior conditional distribution. All simulated samples are plotted with light-colored '$\times$' and '$+$' for conditional distributions of X_1 and X_2, respectively. The dark-colored '$\times$' points are the simulated samples from the distribution of X_1 given X_2 truncated between upper and lower $\alpha/2$ sample quantiles. The dark-colored '$+$' points are the simulated samples from the distribution of X_2 given X_1 truncated between upper and lower $\alpha/2$ sample quantiles. The conditional modes for X_1 given X_2 are plotted in a line of Δ and those for X_2 given X_1 are plotted in a line of ∇. These two lines are close to one another but not identical in general. For the purpose of comparison, we also include the maximum likelihood estimates (solid line) and 95% pointwise confidence regions (solid curves) for $\Theta_{0.1}$ proposed in Li *et al.* [3]. Under priors A and B (upper and middle panel of figure 1), the lines of conditional modes are fairly close to the line of maximum likelihood estimates and the boundary of the 95% credible intervals for conditional distributions appear to be very close to the pointwise confidence regions. When little prior information is available, the Bayesian estimator confirms the frequentist estimator from straight-forward computer simulations. On the other hand, if we have strong evidence for the prior distribution, we might want to adjust the answer by taking into account such information. The Bayesian solution for Prior C distinguishably differ from the frequentist solution, as is observed in the lower panel of figure 1.

The conditional densities of X_1 given X_2 at four quintiles over the range of X_2 are estimated by using a kernel smoother [14] and plotted in rows 1, 3, and 5 in figure 2 for three priors. We can choose the posterior mode as the designed 'dose' value of X_1 when X_2 is given so that the mortality probability is 0.1. The conditional densities of X_2 given X_1 at quntitles of X_1 are plotted in rows 2, 4 and 6 in figure 2.

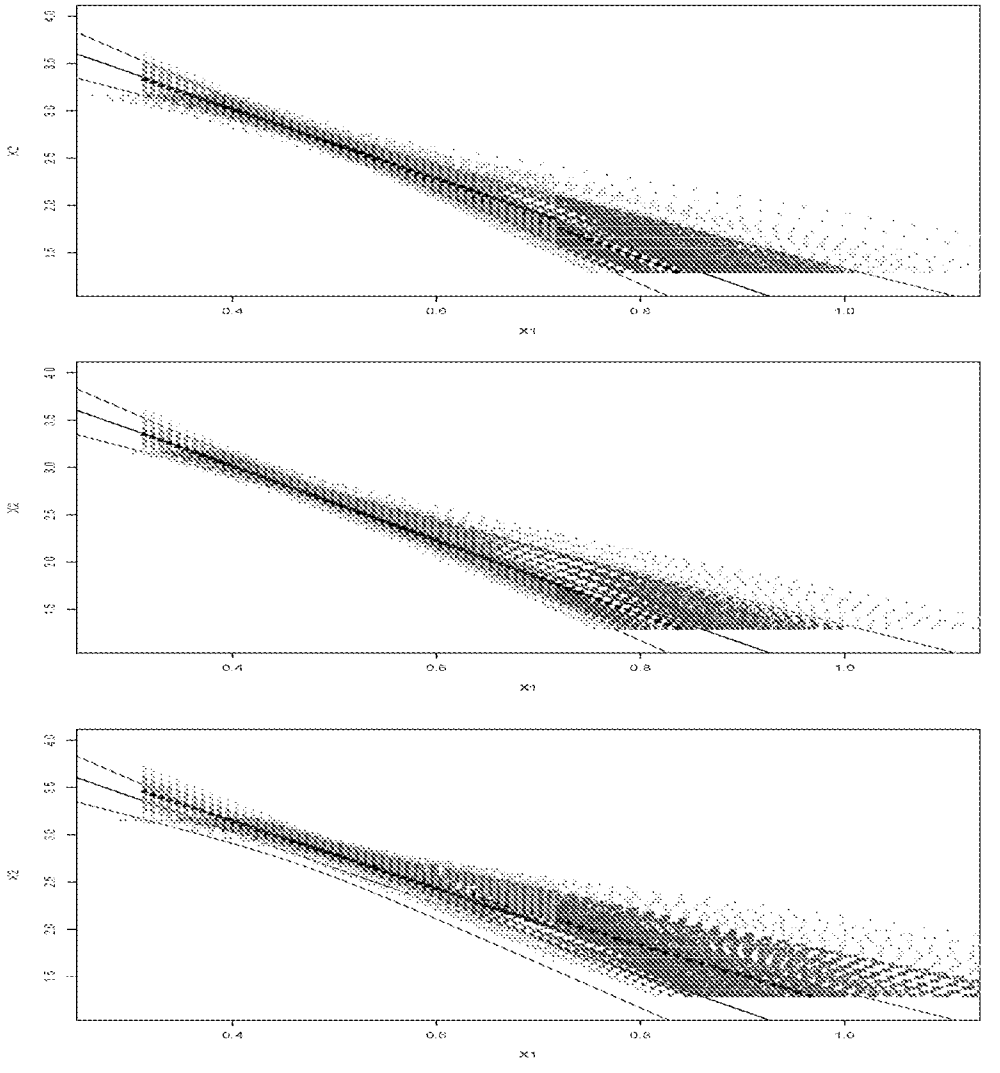

Figure 1. Simulated samples from the conditional distributions of X_1 given X_2 ('×') and that of X_2 given X_1 ('+') under Prior A (upper panel), Prior B (middle panel) and Prior C (lower panel).

Computation of each density is based on the 1000 simulated sample points. The densities at different quintiles are similar to each other, with centered peaks and roughly symmetric shapes. The non-informative Prior A (rows 1 and 2) gives more unimodal densities. The exponential Prior B (rows 3 and 4) gives more local modes and bumps than other two priors.

3.1.2 Joint distributions. We next consider simulating the joint distribution of X_1 and X_2 at $p = 0.1$ by using the algorithm we described in section 2.1.3. We noticed that for a chain of length $T = 10,000$, the simulated values of X_1 and X_2 can deviate too far from the experimental range.We thus restricted our final analysis to generated observations that are reasonably close to the experimental range. In 1000 simulations for Prior A, 593 observations of X_1 are within $[-10, 10]$ and 603 observations of X_2 are within $[-40, 40]$. The rest of the simulated observations can go way up to $[-10^6, 10^7]$ for X_1 and $[-10^9, 10^6]$ for X_2. Those

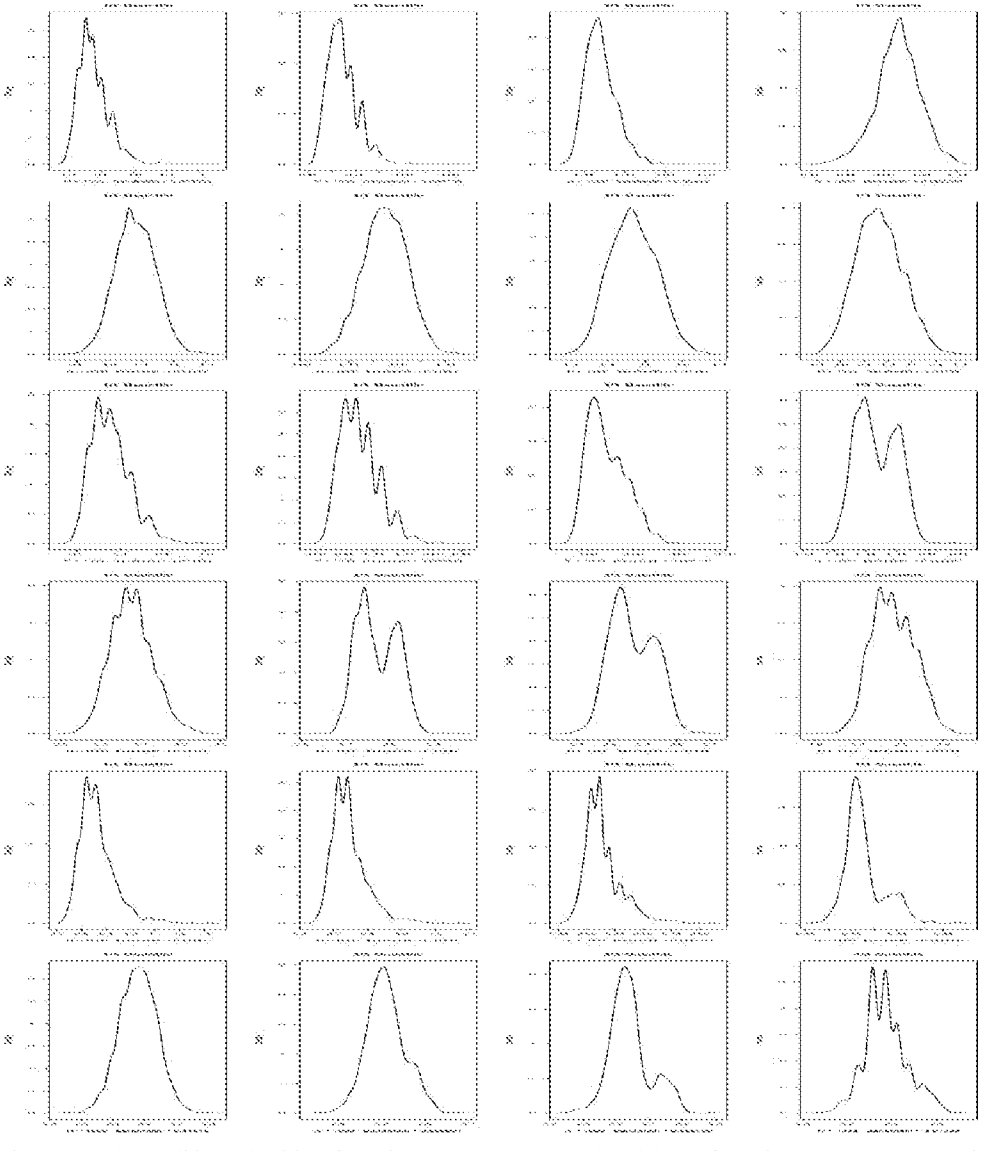

Figure 2. The conditional densities of X_1 given X_2 (rows 1, 3 and 5) and those of X_2 given X_1 (rows 2, 4 and 6) for causing mortality with $p = 0.1$ under Priors A (rows 1 and 2), B (rows 3 and 4) and C (rows 5 and 6).

points can still be regarded as the appropriate samples from the predictive distribution of $\Theta_{0.1}$ if we assume an unnecessarily large parameter space. However, in order to make the results more interpretable while still maintaining a sufficiently big sample size, we shall restrict our attentions on those $X_1 \in [-10,\ 10]$ and $X_2 \in [-40,\ 40]$. The kernel density estimators of the marginal distributions of X_1 and X_2 are displayed in figure 3. The peak for the density of X_1 is around 0.472 and the peak for the density of X_2 is around 2.85 for Prior A. Similar selections are conducted for priors B and C, too.

The points simulated from the posteriors are shown in scatter plots in figure 4. We estimate the bivariate density by using a 2D kernel smoother and draw contour lines corresponding to 80%, 90% and 95% of the highest probability in this distribution. These contours provide the

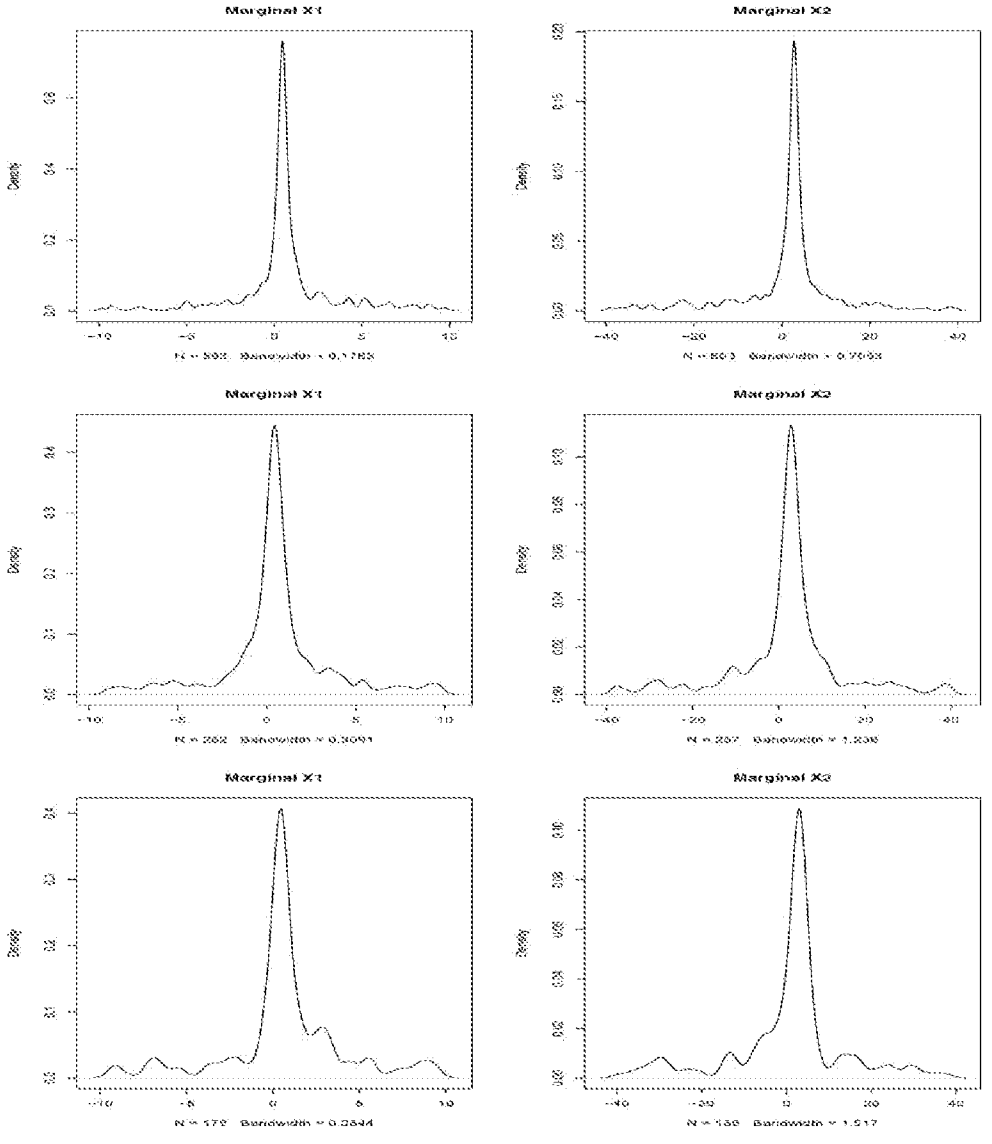

Figure 3. The marginal distribution of X_1 (left) and X_2 (right) simulated from their joint distribution for causing mortality with $p = 0.1$ under Priors A (row 1), B (row 2) and C (row 3).

Bayesian HPD credible regions for $\Theta_{0.1}$. Such regions can cover a future predicted joint value of X_1 and X_2 corresponding to the mortality probability $p = 0.1$ with a fixed confidence level. It is noted that some credible regions for priors B and C could be disconnected, reflecting the multimodality of the posterior distributions. For the purpose of comparison, we also include the simultaneous confidence bands for $\Theta_{0.1}$ by using Scheffe's method described in ref. [3]. Under the non-informative Prior A, all simulated points fall inside the simultaneous confidence band, agreeing with the frequentist region. It is hypothesized that the 95% contour plot for simulated points from the infinite $\mathbb{R}^2$ space is the same as the frequentist band. Currently it is difficult, if not infeasible, to check this since the program can only create contour graphs for a limited range of values.

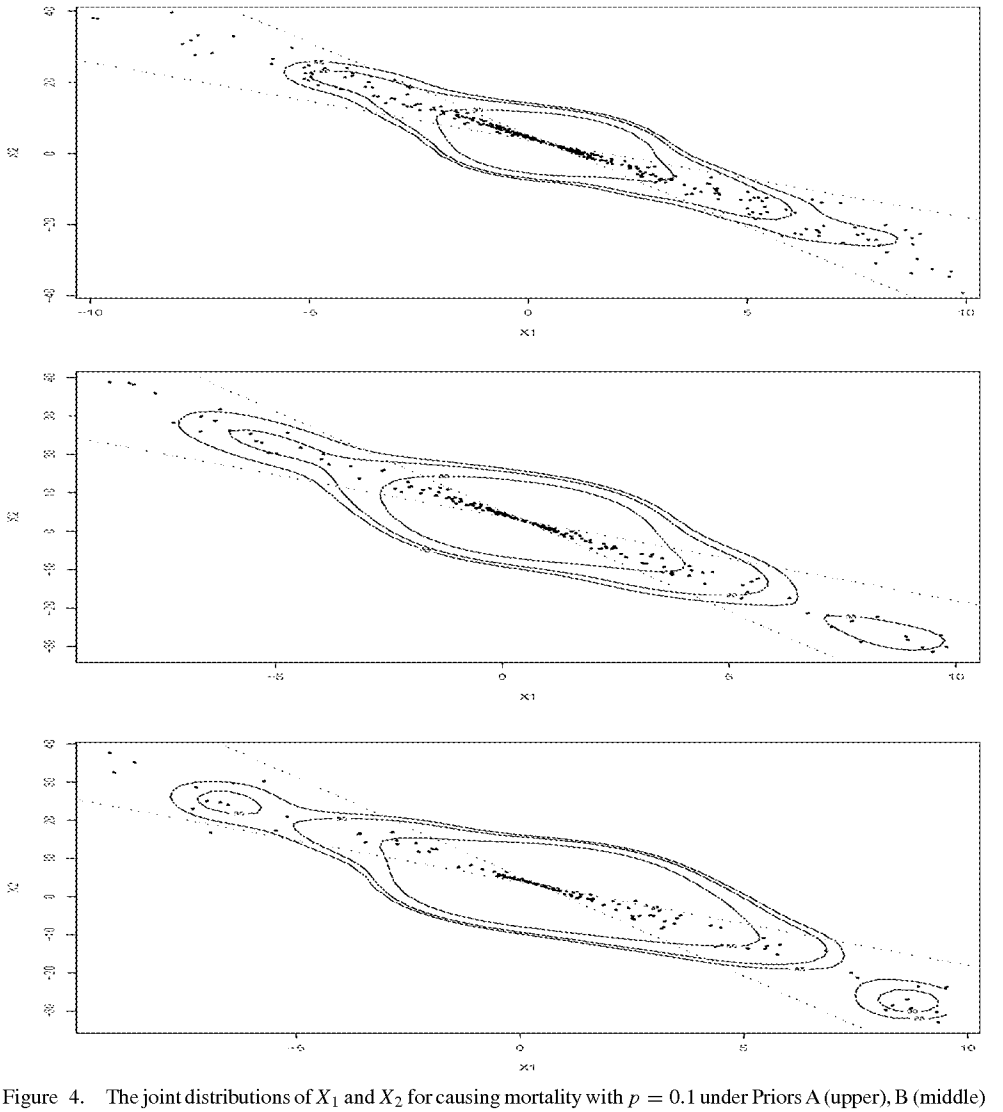

Figure 4. The joint distributions of X_1 and X_2 for causing mortality with $p = 0.1$ under Priors A (upper), B (middle) and C (lower). The solid lines are 80%, 90% and 95% HPD credible regions for $\pi_{0.1}$, respectively. The dashed lines are the 95% simultaneous confidence bands for $\Theta_{0.1}$.

3.2 *Case II: Semiparametric regression model*

We analyze the LB outcome by using the semiparametric regression model. The variable X_2, log base 10 duration, is fitted as a parametric component and X_1, log base 10 pressure, is fitted as a non-parametric component. We consider a non-informative prior for $\pi(\beta)$ in the computation.

The two-dimensional effective dose for $p = 0.3$ and the 95% pointwise confidence region for X_2 at a given X_1 are displayed as solid lines in both panels of figure 4. The conditional distribution of X_2 given X_1 is sampled by using the algorithm described in section 2.2 and depicted by gray circles in the upper panel of figure 5. At each X_1, an independent chain of length 10,000 is generated to compute X_2. The densities of X_2 given X_1 at four quintiles are estimated by a kernel smoother from the simulated sample and are displayed in figure 6.

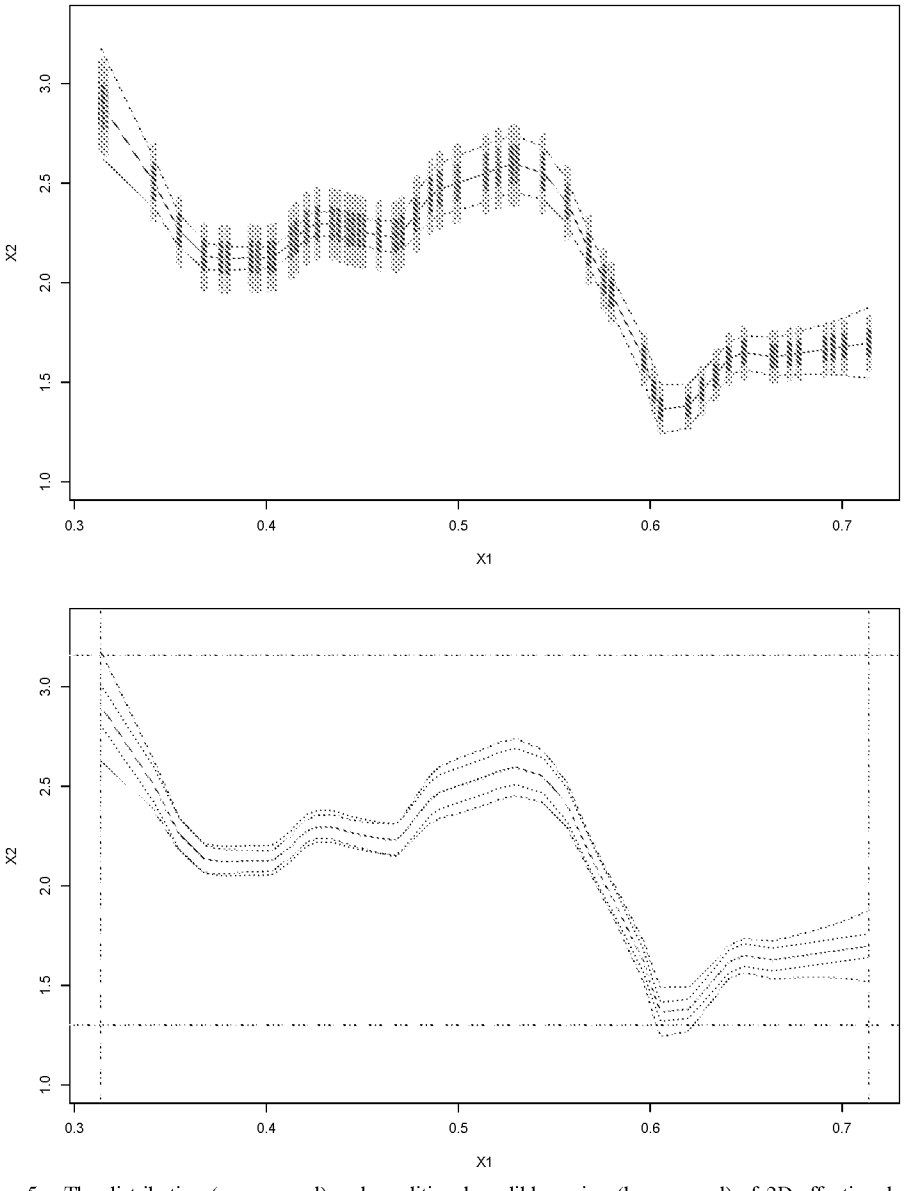

Figure 5. The distribution (upper panel) and conditional credible region (lower panel) of 2D effective dose at $p = 0.3$ in a semiparametric model.

In the upper panel of figure 5, the points with symbol '+' are those that lie between upper and lower 0.025 quantiles. In the lower panel of figure 5, we show the posterior mode of two-dimensional effective dose by a dashed line and the 95% credible region by dotted lines. The broken lines forming a frame in the lower panel indicate the margins of the experimental range of X_1 and X_2.

We notice that the posterior mode for two-dimensional effective dose is the same as the frequentist maximum likelihood estimation since the dashed line in the lower panel is almost indistinguishable from the solid line. The Bayesian credible region is again very close to the frequentist confidence region. There are some noticeable discrepancies at the margins of the

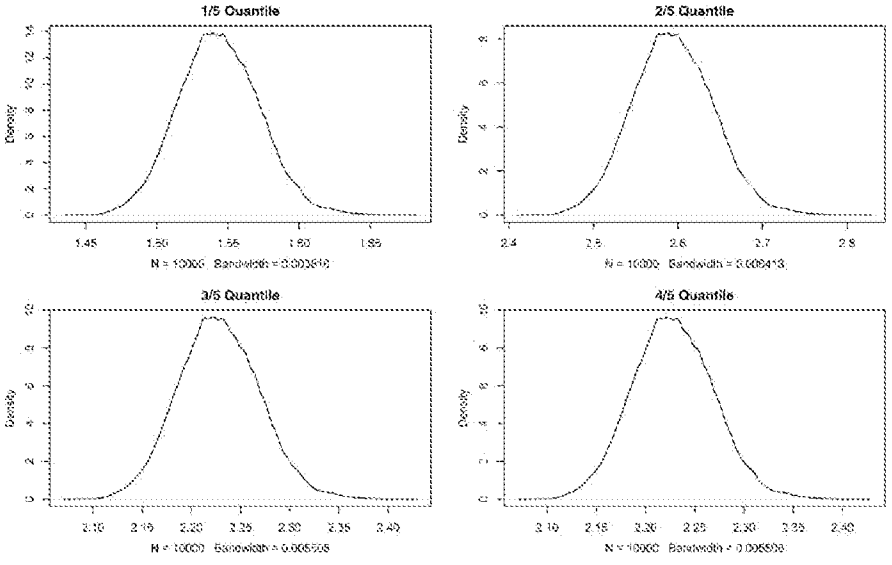

Figure 6. The conditional distribution of X_2 given X_1 for causing limb bends with $p = 0.3$.

experimental range where the non-parametric fits might suffer from the well-known boundary effects.

4. Discussion

We have proposed methods and algorithms to visualize the distribution of multi-dimensional effective dose under parametric and semiparametric logistic models. The Bayesian approach yields comparable results to the frequentist approach. For example, the Bayesian credible interval for the conditional distribution can match the frequentist pointwise confidence interval under the non-informative prior assumption. It is relatively appealing to use a Bayesian approach since no analytic derivation is needed. Current computational power facilitates the fast evaluation of complicated posterior distributions which have been cumbersome in the past.

For the semiparametric model, we study the pseudo posterior distribution after fixing the non-parametric components in the likelihood. The subsequent MCMC approach does not take into account the variability in $\hat{\gamma}(\cdot)$. Although asymptotically $\hat{\gamma}(\cdot)$ has little difference to real $\gamma(\cdot)$ (with a bias of order $O(h^2)$ for a bandwidth h), in a finite sample this approach limits the inference. We hope our tentative solution could attract further research attention on this topic.

Acknowledgements

This work was partly funded by the University of Wisconsin Sea Grant Institute under grants from the National Sea Grant College Program, National Oceanic and Atmospheric Administration, US Department of Commerce, and from the State of Wisconsin. Federal grant number NA16RG2257, project number R/NI-31. We thank the US Navy's Deep Submergence Biomedical Development program (NAVSEA) for supporting this research with funding through the OPNAV N773B Submarine Rescue and Diving Programs.

References

[1] Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B., 1995, *Bayesian Data Analysis* (London; New York: Chapman & Hall).

[2] Finney, D.J., 1978, *Statistical Methods in Biological Assays* (3rd edn) (London: Charles Griffin).

[3] Li, J., Nordheim, E.V., Zhang, C.M. and Lehner, C.E., 2006, Estimation and confidence region for multi-dimensional effective dose. Manuscript.

[4] Freeman, P.R., 1970, Optimal bayesian sequential estimation of the median effective dose. *Biometrika*, **57**, 79–89.

[5] Racine, A., Grieve, A.P., Fluhler, H. and Smith, A.F.M., 1986, Bayesian method in practice: experiences in the pharmaceutical industry. *Applied Statistics*, **35**, 93–150.

[6] Fan, J. and Zhang, C.M., 2005, Assessing Prediction Error of Nonparametric Regression and Classification under Bregman Divergence. Manuscript.

[7] Li, J., 2006, Estimation techniques for multi-dimensional effective dose under parametric and semiparametric models. Ph.D. thesis, University of Wisconsin, Madison, Department of Statistics.

[8] Zeger, S.L., Karim, M.R., 1991, Generalized linear models with random effects: a Gibbs sampling approach. *Journal of the American Statistical Association*, **86**, 79–86.

[9] Le Cam, L. and Yang, G.L., 1990, *Asymptotics in Statistics: Some Basic Concepts* (New York: Springer-Verlag).

[10] Robert, C.P. and Casella, G., 2004, *Monte Carlo Statistical Methods* (Springer).

[11] Besag, J., 1974, Spatial interaction and the statistical analysis of lattice system. *Journal of the Royal Statistical Society. Series B*, **36** (2), 192–236.

[12] Besag, J., 1977, Efficiency of pseudo-likelihood estimation for simple Gaussian fields. *Biometrika*, **64**, 616–618.

[13] Bedrick, E.J., Christensen, R. and Johnson, W., 1996, A new perspective on priors for generalized linear models. *Journal of the American Statistical Association*, **91**, 1450–1460.

[14] Silverman, B.W., 1986, *Density Estimation* (London: Chapman and Hall).