Bayesian Robustness
IMS Lecture Notes - Monograph Series (1996) Volume 29

ON LOCAL SENSITIVITY MEASURES IN BAYESIAN ANALYSIS

BY DIPAK K. DEY, SUJIT K. GHOSH AND KOU-REN LOU
University of Connecticut

The sensitivity of Bayes procedures to the choice of a prior distribution is a major concern for many Bayesians. A Bayesian analysis strongly depends on modeling assumptions which make use of both prior and likelihood and to study the impact of it on the utility function. In this paper we investigate the effects of dual perturbations (prior and/or likelihood) on the posterior inference. In particular, we develop local sensitivity measures to detect how sensitive the posterior is with respect to simultaneous perturbations in both prior and likelihood. We then apply our methodology in (generalized) linear models to study the effects on posterior distributions (or measures) using some notion of distance between probability measures. Finally, discussion and an example using real data are provided.

1. Introduction. A Bayesian analysis depends strongly on the modeling assumptions, which make use of both prior and likelihood to study the impact on the utility function. Even after fitting a standard statistical model to a given set of data, one does not feel comfortable unless some sensitivity checks are made for model adequacy. One way to measure the sensitivity of the present model is to perturb the base model in potentially conceivable directions to determine the effect of such alterations on the analysis. In many situations, it is often difficult to specify or elicit a method that would yield a convincing prior. The problem becomes more difficult for high dimensional parameters. Thus, to perform a complete Bayesian analysis, one must use some sensitivity measures to check model adequacy. Notable references are due to Berger (1984,1990,1994) and those contained therein. The sensitivity analysis or the robustness issues in Bayesian inference can be classified into two broad categories, global and local sensitivity. In global analysis one considers a class of reasonable priors and calculates the range of quantities of interest.

AMS 1991 subject classifications. primary 62A15, secondary 62F15.
Key words and phrases. Bayes Factor, Bayesian robustness, Classes of probabilities, Generalized linear models, Infinitesimal sensitivity, Perturbation, Total variation.

Alternatively, in local analysis the effects of minor perturbations around some elicited priors are studied along several conceivable directions. Recent results involving global sensitivity analysis are contained in Berger (1990). In contrast, a small but quickly growing literature on Bayesian local sensitivity has developed lately; see Basu, Jamalamadaka and Liu (1993), Gustafson and Wasserman (1993), Gustafson (1994), Ruggeri and Wasserman (1993) and Ghosh and Dey (1994).

The major advantage of local sensitivity analysis is realized particularly in multivariate problems, where the global analysis is too time consuming and often analytically intractable. In the multivariate scenario, several questions arise. For example, how sensitive is the posterior marginal density for one parameter when the prior of another parameter or the likelihood corresponding to another parameter changes? These problems can be handled in a reasonable manner through local sensitivity analysis.

To develop any reasonable sensitivity measures one needs to interpose certain basic concepts. For example, 'What classes of perturbations are to be considered?,' and 'How do we assess the discrepancies between the models generated through these perturbations?'. Although there are several ways to perturb a model, we will confine ourselves to ϵ -contamination classes. Recently there have appeared a few papers concerning the effects of dual perturbation (prior and/or likelihood) on posterior inference. Such effects were considered by Lavine (1991), Sivaganesan (1993) and later studied by Basu (1994). Sivaganesan (1993) considers ϵ-contamination classes for prior and likelihood perturbation and develops measure based on a derivative of posterior quantity of interest w.r.t. ϵ to identify those aspects which influence the robustness performance most. However, we develop local sensitivity measures based on derivatives of pseudodistance and divergence between posteriors to detect how sensitive the posterior is with respect to the simultaneous perturbations in both prior and likelihood. We observe that the local sensitivity measures are expressible in terms of certain Bayes factors.

The format of the paper is thus as follows. In Section 2 we develop definitions and notations which will be used throughout the paper. Section 3 is devoted to the development of sensitivity measures under a pseudodistance function with applications to (generalized) linear models. In Section 4, we develop local sensitivity measures using φ-divergence with applications to linear models. Finally in Section 5, we provide a real data example to demonstrate the sensitivity measures in connection to a logistic regression problem, where the detection of influential observations can be linked with the local sensitivity measures.

2. Definitions and Notations. Suppose X is a random variable with density $f(x \mid \theta)$ where θ is the parameter of interest. We allow both X and

θ to be vector valued. Let Θ be the parameter space with $\mathcal{A}$, the associated σ-field over Θ. Suppose further $\mathcal{P}$ denotes the class of probability measures on $(\Theta, \mathcal{A})$, where $P \in \mathcal{P}$ is a particular member. The marginal density of X is then defined by $m_P(x, f) = \int_\Theta f(x \mid \theta) P(d\theta)$ and the posterior probability measure corresponding to P and f is defined as $P^x(f) = f(x \mid \bullet) P / m_P(x, f)$.

Let $C(\theta)$, $\theta \in \Theta$ denote the class of conditional densities on $\mathcal{X}$ given θ, where $\mathcal{X}$ denotes the sample space. Then, under linear perturbation classes the perturbed prior and the likelihood are respectively given as

$$Q_{\epsilon_1} = (1 - \epsilon_1) P + \epsilon_1 Q, \; 0 \leq \epsilon_1 \leq 1$$

and

$$g_{\epsilon_2} = (1 - \epsilon_2) f + \epsilon_2 g, \; 0 \leq \epsilon_2 \leq 1$$

P is the elicited prior, Q is the contamination which belongs to a certain class, say Γ. Similarly, f is the elicited likelihood and g is the contamination which belongs to a certain class $\mathcal{G}$. The class of ϵ-contamination was studied in the context of Bayesian robustness by Berger and Berliner (1986) and Sivaganesan and Berger (1989), among others.

3. Local Sensitivity Measures Under a Pseudo Distance Function.

3.1. Basic Theory. In this section, to study the change in the prior and posterior measures, we introduce a pseudo-distance function, between two probability measures $P, Q \in \mathcal{P}$. Define a function $d : \mathcal{P} \times \mathcal{P} \longrightarrow [0, \infty)$ such that $d(P, Q) = \rho(P - Q) = \rho(\delta)$ (say) where δ is a signed measure. We assume that the ρ-function satisfies the conditions: (i) $\rho(c\delta) = \mid c \mid \rho(\delta)$ and (ii) $\rho(\delta_1 + \delta_2) \leq \rho(\delta_1) + \rho(\delta_2)$.

An example of ρ is the total variation norm, defined by $\rho(\delta) = \sup_{A \in \mathcal{A}} \mid \delta(A) \mid$.

Define

$$\lambda_1(\epsilon_1, \epsilon_2; x) = d(P^x(g_{\epsilon_2}), Q^x_{\epsilon_1}(g_{\epsilon_2})) / d(Q^x(g_{\epsilon_2}), P^x(g_{\epsilon_2})) \tag{3.1}$$

and

$$\lambda_2(\epsilon_1, \epsilon_2; x) = d(Q^x_{\epsilon_1}(f), Q^x_{\epsilon_1}(g_{\epsilon_2})) / d(Q^x_{\epsilon_1}(g), Q^x_{\epsilon_1}(f)). \tag{3.2}$$

Note that λ_1 measures the relative change in the posteriors when prior is contaminated under a contaminated likelihood. On the other hand λ_2 measures the relative change in the posteriors when likelihood is contaminated along with a contaminated prior. Observe that

$$\begin{aligned} d(P^x(g_{\epsilon_2}), Q^x_{\epsilon_1}(g_{\epsilon_2})) &= \rho(Q^x_{\epsilon_1}(g_{\epsilon_2}) - P^x(g_{\epsilon_2})) \\ &= \lambda_1 \rho(Q^x(g_{\epsilon_2}) - P^x(g_{\epsilon_2})) \\ &= \lambda_1 d(Q^x(g_{\epsilon_2}), P^x(g_{\epsilon_2})), \end{aligned}$$

where $\lambda_1 = \lambda_1(\epsilon_1, \epsilon_2; x) = \epsilon_1 m_Q(x, g_{\epsilon_2})/m_{Q_{\epsilon_1}}(x, g_{\epsilon_2})$. Similarly, it can be checked that $\lambda_2 = \lambda_2(\epsilon_1, \epsilon_2; x) = \epsilon_2 m_{Q_{\epsilon_1}}(x, g)/m_{Q_{\epsilon_1}}(x, g_{\epsilon_2})$. It follows from the definitions that $\lambda_1(\epsilon_1, \epsilon_2; x)$ and $\lambda_2(\epsilon_1, \epsilon_2; x)$ are invariant under the choice of ρ-function. It also follows that for each $Q \in \mathcal{P}$ and $g \in \mathcal{C}(\theta)$, the general form of the λ-functions, is

$$\lambda_i(\epsilon_1, \epsilon_2; x) = \frac{c_{1i}\epsilon_i + c_{2i}\epsilon_1\epsilon_2}{c_3 + c_4\epsilon_1 + c_5\epsilon_2 + c_6\epsilon_1\epsilon_2}, \ i = 1, 2;$$

where the constants c_i 's depend on the marginals $m_P(x, f)$, $m_Q(x, f)$, $m_P(x, g)$ and $\mathrm{m}_Q(x, g)$. In order to study the general sensitivity measures, we finally consider

$$\lambda(\epsilon_1, \epsilon_2; x) = \lambda_1(\epsilon_1, \epsilon_2; x) + \lambda_2(\epsilon_1, \epsilon_2; x). \tag{3.3}$$

The quantity $\lambda(\epsilon_1, \epsilon_2; x)$ in (3.3) will be called the general sensitivity measure, since several local and global diagnostics can be obtained form it. These sensitivity measures are described below.

The local sensitivity of P in the direction of Q is given by

$$s_1(P, Q; x) = \lim_{\epsilon_1 \to 0+} \frac{\lambda(\epsilon_1, 0; x)}{\epsilon_1}. \tag{3.4}$$

We may note that $s_1(P, Q; x)$ is proportional to the sensitivity measure as defined in Gustafson and Wasserman (1993), the proportionality constant being the ratio of distances between corresponding posteriors to priors. In this way our definition does not depend on the kind of 'distance' we may choose to use on the parameter space (except for the structure assumed through ρ). Also from (3.1), it follows that $s_1(P, Q; x)$ is indeed the Frechet derivative in the direction Q. It is to be noted that these directional derivatives can be turned into an elicitation tool, i.e., finding the 'direction' in which the derivative is largest (in absolute value) might indicate a particularly important direction in which to concentrate elicitation efforts. (See Berger (1994)).

In the same spirit, we can propose the local sensitivity of f in the direction of g as,

$$s_2(f, g; x) = \lim_{\epsilon_2 \to 0+} \frac{\lambda(0, \epsilon_2; x)}{\epsilon_2}. \tag{3.5}$$

The joint local sensitivity is defined as the pair $(s_1(P,Q;x), s_2(f,g;x))$. Finally, the overall local sensitivity is a function of the pair $(s_1(P,Q;x), s_2(f,g;x))$. Again, following Gustafson and Wasserman (1993), we can define local sensitivity over different classes as

$$s_1(P,\Gamma;x) = \sup_{Q\in\Gamma} s_1(P,Q;x) \tag{3.6}$$

$$s_2(f,\mathcal{G};x) = \sup_{g\in\mathcal{G}} s_2(f,g;x) \tag{3.7}$$

where Γ and $\mathcal{G}$ can be appropriately chosen.

Now, it is interesting to note that

$$s_1(P,Q;x) = \frac{m_Q(x,f)}{m_P(x,f)} \tag{3.8}$$

and similarly,

$$s_2(f,g;x) = \frac{m_P(x,g)}{m_P(x,f)}. \tag{3.9}$$

The expressions (3.8) and (3.9) show that the local sensitivity measures are expressed as ratios of marginals and in particular, (3.9) is the Bayes factor of one model with respect to the other. It also follows that $s_1(P,Q;x)/s_2(f,g;x) = m_Q(x,f)/m_P(x,g)$, the magnitude of which dictates the shape of the overall sensitivity plot, which will be evidenced through the example in Section 5.

Again, if p and q are respectively, the densities of P and Q, then it follows that

$$s_1(P,Q;x) = \int_\Theta f(x \mid \theta)\frac{q(\theta)}{p(\theta)}p(\theta)d\theta/m_P(x,f) = E^{P^x}[q(\theta)/p(\theta)] \tag{3.10}$$

where E^{P^x} denotes expectation under elicited posterior. The above equation is a very useful formula for the computation of the local sensitivity measure, since often we have samples from the elicited posterior using Markov Chain Monte Carlo methods. Similarly, we can show that

$$s_2(f,g;x) = E^{P^x}[g(x \mid \theta)/f(x \mid \theta)] \tag{3.11}$$

which again can be computed easily using a sampling based approach.

3.2. Local Sensitivity Measures in Linear Models. In this section, we obtain analytical expressions for the local sensitivity measures. Consider a standard linear model where

$$Y \sim N_n(X\beta, \sigma^2 I).$$

Since we are interested in studying perturbation of the prior on β, we assume that σ^2 is known. Suppose $P \sim N_p(\beta_1, V_1)$ and $Q \sim N_p(\beta_2, V_2)$ where β_1, β_2, V_1 and V_2 are known. This produces from (3.14) (with identity link) the corresponding priors for θ's as

$$\theta_i \sim N_n(X\beta_0, XV_iX^t),\ i = 1, 2. \tag{3.12}$$

The following theorem gives a closed form expression for the local sensitivity measure.

THEOREM 3.1. *For the normal linear model, i.e., under* $y \mid \beta \sim N_n(X\beta, I)$ *and under linear perturbation of the prior, the local sensitivity measure is given as*

$$\begin{aligned} s_1(P,Q;y) &= \left\{\frac{\mid I + XV_1X^t \mid}{\mid I + XV_2X^t \mid}\right\}^{1/2} \\ &\quad \exp\left\{\frac{1}{2}[\| y - X\beta_1 \|^2_{V_1} - \| y - X\beta_2 \|^2_{V_2}]\right\} \end{aligned} \tag{3.13}$$

where $\| y - X\beta_i \|^2_{V_i} = (y - X\beta_i)^t(I + XV_iX^t)^{-1}(y - X\beta_i),\ i = 1, 2.$

PROOF. The proof follows from the fact that $m_P(y, f)$ is $N_n(X\beta_1, I + XV_1X^t)$ and $m_Q(y, f)$ is $N_n(X\beta_2, I + XV_2X^t)$.□

Now, let us consider a perturbation of the likelihood. In our notation, f stands for $N_n(X\beta, I)$. A natural choice for a contaminated model is to posit a correlation structure in the error distribution and in this regard we take g to be $N_n(X\beta, c(I + J))$, where c is a positive constant and J is the matrix of 1's. For prior specification we consider P to be $N_p(\mu, A)$ where μ and A are known. The following theorem gives the local sensitivity measure for the likelihood perturbation.

THEOREM 3.2. *For the normal linear model, under linear perturbation of the likelihood function, the local sensitivity measure is given as*

$$\begin{aligned} s_2(f,g;y) &= \left\{\frac{\mid I + XAX^t \mid}{\mid c(I + J) + XAX^t \mid}\right\}^{1/2} \\ &\quad \times \exp\left\{\frac{1}{2}(y - X\mu)^t[(I + XAX^t)^{-1} \right. \\ &\quad \left. \times\ (c(I + J) + XAX^t)^{-1}](y - X\mu)\right\}. \end{aligned} \tag{3.14}$$

PROOF. The proof follows from the fact that the marginal distributions $m_p(y, f)$ is $N_n(X\mu, I + XAX^t)$ and $m_p(y, g)$ is $N_n(X\mu, c(I + J) + XAX^t)$.□

3.3. Local Sensitivity Measures and Leverages. Consider the local sensitivity measure for the regression model with linear perturbation of the prior as defined in (3.13). In this section, we show that the classical deletion approach to performing 'influential case analysis' (Cook and Weisberg (1982)) can be obtained as a special case of our local sensitivity measures. Again, suppose $P \sim N_p(\beta_1, V_1)$ and $Q \sim N_p(\beta_2, V_2)$, where we further assume that $\beta_1 = (X^tX)^{-1}X^ty$ and $V_1 = (X^tX)^{-1}$, i.e., the elicited prior is normal with mean equal to the least square estimate and with variance equal to the variance of the least square estimate. Such a choice of prior is somewhat artificial but it is natural in practical problems when we do not have any precise information. Next, we consider $\beta_2 = (X_{(j)}^tX_{(j)})^{-1}X_{(j)}^ty_{(j)}$ and $V_2 = (X_{(j)}^tX_{(j)})^{-1}$ where $y_{(j)} = (y_1, ..., y_{j-1}, y_{j+1}, ..., y_n)^t$ and $X_{(j)}$ is the matrix X with j^{th} row deleted. It follows immediately that

$$\beta_2 = \beta_1 - (1-h_{jj})^{-1}(y_j - x_j^t\beta_1)(X^tX)^{-1}x_j \tag{3.15}$$

and

$$\begin{aligned} V_2 &= (X^tX - x_jx_j^t)^{-1} \\ &= (X^tX)^{-1} + (1-h_{jj})^{-1}(X^tX)^{-1}x_jx_j^t(X^tX)^{-1}, \end{aligned} \tag{3.16}$$

where h_{jj} is the j^{th} diagonal element of the projection matrix $H = X(X^tX)^{-1}X^t$. The following theorem shows the relationship between the local sensitivity measure and the leverages which assumes that $s_1(P, Q; y)$ is a sensible measure.

THEOREM 3.3. *For the linear regression model, under linear perturbation of the prior, the local sensitivity measure under the above set up is given as*

$$s_1(P, Q; y) = \left\{\frac{2(1-h_{jj})}{2-h_{jj}}\right\}^{1/2} \exp\left\{-\frac{h_{jj}(y_j - x_j^t\beta_1)^2}{2(1-h_{jj})(2-h_{jj})}\right\} \tag{3.17}$$

where $\beta_1 = (X^tX)^{-1}X^ty$.

PROOF. First it follows that

$$\begin{aligned} \frac{|I + XV_1X^t|}{|I + XV_2X^t|} &= \frac{|X^tX - x_jx_j^t|}{|X^tX|}\frac{|2X^tX|}{|2X^tX - x_jx_j^t|} \\ &= \frac{2(1-h_{jj})}{2-h_{jj}}. \end{aligned} \tag{3.18}$$

Now, it can be shown that

$$\| y - X\beta_1 \|_{V_1}^2 = y^t(I-H)(I+H)^{-1}(I-H)y = y^t(I-H)y.$$

To calculate $\| y - X\beta_2 \|^2_{V_2}$, first we observe that

$$I + XV_2X^t = I + H + u_ju_j^t.$$

where

$$u_j = (1 - h_{jj})^{-1/2} X(X^tX)^{-1}x_j.$$

Thus,

$$(I + XV_2X^t)^{-1} = (I + H)^{-1} - \frac{(I + H)^{-1}u_ju_j^t(I + H)^{-1}}{1 + u_j^t(I + H)^{-1}u_j}$$

Observing that $(I + H)^{-1} = I - \frac{1}{2}H$, which implies

$$(I + XV_2X^t)^{-1} = (I - \frac{1}{2}H) - \frac{(1 - h_{jj})u_ju_j^t}{2(2 - h_{jj})}.$$

Finally, after some algebra, it follows that

$$\| y - X\beta_2 \|^2_{V_2} = y^t(I - H)y + \frac{h_{jj}(y_j - x_j^t\beta_1)^2}{(1 - h_{jj})(2 - h_{jj})}. \tag{3.19}$$

Combining (3.18) and (3.19), the proof follows from (3.13). □

It is interesting to observe that $s_1(P, Q; y)$ is a decreasing function of h_{jj}. Since $0 \leq h_{jj} \leq 1$, $s_1(P, Q; y)$ is also between 0 and 1.

3.4. Applications in Generalized Linear Models. Suppose $y_1, ..., y_n$ are independent observations, where y_i has the exponential density

$$f(y_i \mid \theta_i) = \exp\{y_i\theta_i + b(\theta_i) + c(y_i)\}, \ i = 1, ..., n. \tag{3.20}$$

Density in (3.20) is parameterized by the canonical parameter θ_i. The $b(\cdot)$ and $c(\cdot)$ are known functions. The θ_i's are related to the regression coefficients by the equation

$$\theta_i = \theta(\eta_i), \ i = 1, ..., n, \tag{3.21}$$

where $\eta_i = x_i^t\beta$, $x_i^t = (x_{i1}, ..., x_{ip})$ is a $1 \times p$ vector denoting the i^{th} row of the $n \times p$ matrix of covariates X, $\beta = (\beta_1, ..., \beta_p)^t$ is a $p \times 1$ vector of regression coefficients, and θ is a monotonic differentiable function. Any model given by (3.20) and (3.21) is called a generalized linear model (GLM). The density in (3.20) includes a large class of regression models such as logistic and probit regression, Poisson regression, etc. (see McCullagh and Nelder (1989)). Suppose our elicited prior on β is $N_p(\beta_0, V)$ where β_0 and V are the known mean vector and covariance matrix respectively, so that we can consider a linear perturbation of the prior $N_p(\beta_0, V)$. It follows immediately that

$$\eta \sim N_n(X\beta_0, XVX^t)$$

and thus the prior for θ can be obtained as

$$\begin{aligned}\pi(\theta) &= (2\pi)^{-n/2} \mid XVX^t \mid^{-1/2} \\ \text{(3.22)} \qquad & \exp\{-\frac{1}{2}(\theta^{-1}(\cdot) - X\beta_0)^t(XVX^t)^{-1}(\theta^{-1}(\cdot) - X\beta_0)\} \prod_{j=1}^{n}(\theta_j')^{-1}(\cdot).\end{aligned}$$

Thus, for given β_0, V and the link function $\theta(\cdot)$, we can explicitly evaluate p. We can now obtain the perturbation of $\pi(\theta)$ from (3.14) knowing the linear perturbation of $N_p(\beta_0, V)$ and hence $s_1(P, Q; x)$ can be obtained from (3.8). A closed form expression for $s_1(P, Q; x)$ is not possible in this case. However, a Monte Carlo estimate of $s_1(P, Q; x)$ can be obtained using sampling based approach which is performed in Section 5.

4. Local Sensitivity Measures under φ-Divergence. In this section, instead of a pseudo-distance d, we consider φ -divergence between two probability measures . Formally the φ-divergence between two probability measures P and Q is defined as

$$D_\varphi(P, Q) = \int \varphi\left(\frac{dP}{dQ}\right) dQ \tag{4.1}$$

where φ is a smooth convex function such that $\varphi(1) = 0$. Several well-known divergence measures, e.g., Kullback-Liebler, Hellinger distance, Chi-squared distance, etc. can be obtained by the appropriate choice of the φ-functions. Sensitivity diagnostics based on φ-divergence were studied by Gustafson and Wasserman (1993) and Ghosh and Dey (1994).

Now, to define several sensitivity measures based on the φ-divergence, we assume interchange of limit and integral. We define the local sensitivity measure for the prior perturbations as

$$s_{11}(x) = \lim_{\epsilon_1 \to 0+} \frac{D_\varphi(Q_{\epsilon_1}^x(f), P^x(f))}{D_\varphi(Q_{\epsilon_1}, P)} \tag{4.2}$$

$$= \frac{\frac{\partial^2}{\partial \epsilon_1^2} D_\varphi(Q_{\epsilon_1}^x(f), P^x(f)) \mid_{\epsilon_1=0}}{\frac{\partial^2}{\partial \epsilon_1^2} D_\varphi(Q_{\epsilon_1}, P) \mid_{\epsilon_1=0}}. \tag{4.3}$$

Equation (4.3) follows from (4.1) by using Taylor's expansion. The quantity $s_{11}(x)$ gives the ratio of the local curvature of φ-divergences between posterior and prior.

Similarly, to define the sensitivity measure for the likelihood perturbation, we consider

$$s_{22}(x) = \lim_{\epsilon_2 \to 0+} \frac{D_\varphi(P^x(g_{\epsilon_2}), P^x(f))}{D_\varphi(g_{\epsilon_2}, f)} \tag{4.4}$$

$$(4.5) \qquad = \frac{\frac{\partial^2}{\partial\epsilon_2^2} D_\varphi(P^x(g_{\epsilon_2}), P^x(f)) \mid_{\epsilon_2=0}}{\frac{\partial^2}{\partial\epsilon_2^2} D_\varphi(g_{\epsilon_2}, f) \mid_{\epsilon_2=0}}.$$

Finally, a local sensitivity measure which captures both perturbations can be defined as

$$(4.6) \qquad s_{12}(x) = \lim_{\substack{\epsilon_1\to 0+\\ \epsilon_2\to 0+}} \frac{D_\varphi(P^x(g_{\epsilon_2}), P^x(f))}{D_\varphi(Q_{\epsilon_1}, P)}$$

$$(4.7) \qquad = \frac{\frac{\partial^2}{\partial\epsilon_2^2} D_\varphi(P^x(g_{\epsilon_2}), P^x(f)) \mid_{\epsilon_2=0}}{\frac{\partial^2}{\partial\epsilon_1^2} D_\varphi(Q_{\epsilon_1}, P) \mid_{\epsilon_1=0}}.$$

Now, in the following theorem, we give the formula for the local sensitivity measures.

THEOREM 4.1. *Under linear perturbation the local sensitivity measures based on the* φ divergence *are given as*

$$(4.8) \qquad s_{11}(x) = \frac{Var_{P^x(f)}(\frac{dQ}{dP})}{Var_P(\frac{dQ}{dP})} = \left\{\frac{m_Q(x,f)}{m_P(x,f)}\right\}^2 \frac{Var_{P^x(f)}(\frac{dQ^x(f)}{dP^x(f)})}{Var_P(\frac{dQ}{dP})},$$

$$(4.9) \qquad s_{22}(x) = \frac{Var_{P^x(f)}(\frac{g}{f})}{Var_f(\frac{g}{f})} = \left\{\frac{m_P(x,g)}{m_P(x,f)}\right\}^2 \frac{Var_{P^x(f)}(\frac{dP^x(g)}{dP^x(f)})}{Var_f(\frac{g}{f})},$$

and

$$(4.10) \qquad s_{12}(x) = \frac{Var_{P^x(f)}(\frac{g}{f})}{Var_P(\frac{dP}{dQ})} = \left\{\frac{m_P(x,g)}{m_P(x,f)}\right\}^2 \frac{Var_{P^x(f)}(\frac{dP^x(g)}{dP^x(f)})}{Var_P(\frac{dQ}{dP})}.$$

PROOF. The proof of (4.8) follows from Dey and Birmiwal (1994). Proofs for (4.9) and (4.10) are similar which is given in Ghosh and Dey (1994).

It can be noticed that the above measures are free from the choice of the φ-functions.

It is interesting to note that all the local sensitivity measures as defined above depend on certain variance ratios and the ratios of marginals.

4.1. Results for Linear Models. In this section we revisit the linear models problem as described in Section 3.3. In order to calculate $s_{11}(x)$ under linear perturbation, we use equation (4.3). We note that for any two probability measures P_1 and P_2, $Var_{P_1}(\frac{dP_2}{dP_1})$ can be calculated by using a very simple formula (involving only first order moments of density ratio)

$$(4.11) \qquad Var_{P_1}(\frac{dP_2}{dP_1}) = E_{P_2}(\frac{dP_2}{dP_1}) - 1.$$

Thus, from (4.8) $s_{11}(x)$ reduces to

$$s_{11}(x) = s_1^2(P, Q; x)\frac{E_{Q^x}(\frac{dQ^x}{dP^x}) - 1}{E_Q(\frac{dQ}{dP}) - 1}. \tag{4.12}$$

It follows from Berger (1985) that, if we assume $Y \sim N_n(X\beta, I)$, $P \sim N_P(\beta_1, V_1)$ and $Q \sim N_P(\beta_2, V_2)$, then $P^x \sim N_P(\mu_x^P, V_x^P)$ where

$$\mu_x^P = (X^tX + V_1^{-1})^{-1}(V_1^{-1}\beta_1 + X^ty), V_x^P = (X^tX + V_1^{-1})^{-1}.$$

Similarly $Q^x \sim N_P(\mu_x^Q, V_x^Q)$ where

$$\mu_x^Q = (X^tX + V_2^{-1})^{-1}(V_2^{-1}\beta_2 + X^ty), V_x^Q = (X^tX + V_2^{-1})^{-1}.$$

Then we apply the following lemma twice.

LEMMA 4.1. *For $P \sim N_P(\mu_1, A_1)$, $Q \sim N_P(\mu_2, A_2)$ it follows that*

$$E_Q(\frac{dQ}{dP}) = \frac{| A_1 |}{| A_2 |^{1/2} | 2A_1 - A_2 |^{1/2}} \exp\{(\mu_1 - \mu_2)^t(2A_1 - A_2)^{-1}(\mu_1 - \mu_2)\}$$

provided $2A_1 - A_2$ is positive definite.

It can be obtained that

$$s_{11}(x) = s_1^2(P, Q; x)\frac{\frac{|V_1|}{|V_2|^{1/2}|2V_1 - V_2|^{1/2}} \exp\{\| \beta_1 - \beta_2 \|_*^2\} - 1}{\frac{|V_x^P|}{|V_x^Q|^{1/2}|2V_x^P - V_x^Q|^{1/2}} \exp\{\| \mu_x^P - \mu_x^Q \|_*^2\} - 1} \tag{4.13}$$

where $\| \beta_1 - \beta_2 \|_*^2 = (\beta_1 - \beta_2)^t(2V_1 - V_2)^{-1}(\beta_1 - \beta_2)$,

$$\| \mu_x^P - \mu_x^Q \|_*^2 = (\mu_x^P - \mu_x^Q)^t(2V_x^P - V_x^Q)^{-1}(\mu_x^P - \mu_x^Q)$$

and $s_1(P, Q; x)$ is obtained from (3.13).

5. An Illustrative Example. In this section we consider some binomial count data and compute the local sensitivity measures as described in previous sections. The data (from Lindsey (1993)) describes the effects of salinity and temperature on the proportion of eggs on English sole hatching. The covariates are salinity (x_1) and temperature (x_2). Here y_i refers to the number of eggs hatched, and follows a $Bin(n_i, p_i)$, $i = 1, ..., 72$. We identify this model as f according to the above notation. Using the canonical link, we consider $\theta_i = \log(p_i/(1 - p_i)) = x_i^t\beta$ where $x_i^t = (1, x_{i1}, x_{i2})$ and $\beta = (b_0, b_1, b_2)^t$. To choose prior measures P and Q, we consider $P \sim N(\beta_1, V_1)$ and $Q \sim N(\beta_2, V_2)$ where β_i and V_i are respectively the mean vector and covariance matrix obtained by using logistic and probit regression models (using GLIM). An alternative way to implement binomial regression is to use a transformation based on normal approximation. It is well known that for a smooth function h, as n_i becomes large

$$h(\widehat{p}_i) \mid \beta \sim N(x_i^t\beta,\ (h'(\widehat{p}_i))^2\widehat{p}_i(1-\widehat{p}_i)/n_i)$$

where $\widehat{p}_i = y_i/n_i$, $i = 1,\dots,n(=72)$. By choosing $h(p) = \log(p/(1-p))$ and $h(p) = \Phi^{-1}(p)$, where Φ is the c.d.f. of the standard normal random variable, we can have two plausible choices of contaminated model g.

In order to describe the graphical plots of λ_1/ϵ_1 and λ_2/ϵ_2, we need to compute the marginals, since λ_i/ϵ_i, $i = 1, 2$, are expressible as a function of all the marginals. Note that $m_P(y, g)$ and $m_Q(y, g)$ have closed form analytical expressions for both choices of g . On the contrary $m_P(y, f)$ does not have any closed form expression. We use a Monte Carlo estimate of $m_P(y, f)$ given as

$$\begin{aligned} m_P(y,f) &= \int f(y \mid \beta) dP(\beta) \\ &\approx \frac{1}{N}\sum_{t=1}^{N} f(y \mid \beta^{(t)}) \end{aligned}$$

where $\beta^{(t)}$ are samples from the prior P. In this example, we take $N = 2,500$ samples. In the expression for $f(y \mid \beta^{(t)})$ we also used Stirling's approximation for factorial whenever $\binom{n_i}{y_i}$ is large.

We then plot λ_1/ϵ_1 and λ_2/ϵ_2 functions for the given data against ϵ_1 and ϵ_2 in Figure 1, when g comes from logistic model. Figure 1 indicates monotonicity of all the functions in terms of ϵ_1 and ϵ_2, the amounts of contamination. For any pair of ϵ_1 and ϵ_2 values, the sensitivity measures can be obtained from the graphs. To obtain the local sensitivity measure, it follows that for the full model $s_1(P, Q; x) = 9.43$ and $s_2(f, g; x) = 16.94$, when g comes from a logistic model. From the magnitude of the local sensitivity measures (using Jeffreys scale of evidence, Jeffreys (1961), also see Kass and Raftery (1994)) it is clear that both choices of Q and g are sensitive. From the data, we suspect that the observations 25, 26 and 27 are potential outliers. We then delete these observations and plot λ_i/ϵ_i, $i = 1, 2$ in Figures 1a, 1b and 1c to note the changes after deletion. Table 1 displays the values of the local sensitivity measures for the full model and after deletion of potential outliers. It follows that the deletions have significant effects on the local sensitivity measures. Similar graphs are drawn in Figure 2 where g comes from probit model. Also, the effects of $25th$, $26th$ and $27th$ observations are studied through Figures 2a, 2b and 2c. Local sensitivity measures are displayed in Table 1.

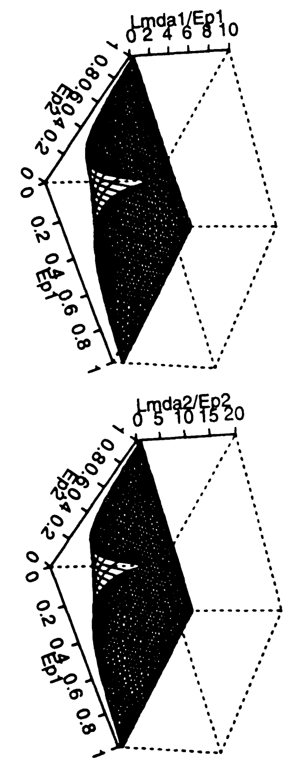

Figure 1: Full Model

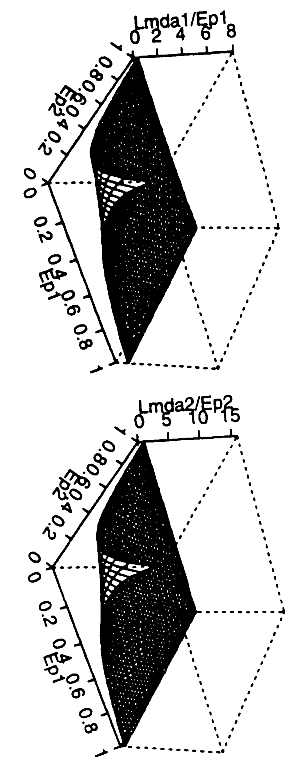

Figure 1(a): Obsn.25 Deleted

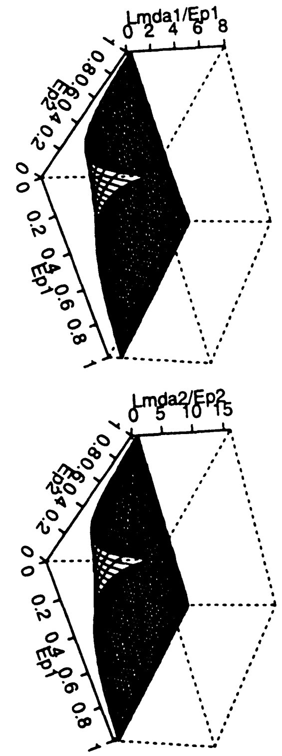

Figure 1(b): Obsn.26 Deleted

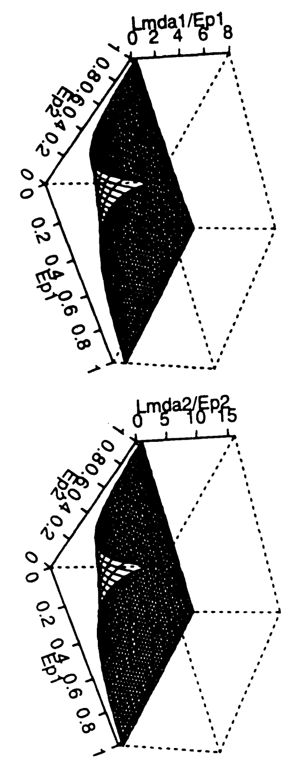

Figure 1(c): Obsn.27 Deleted

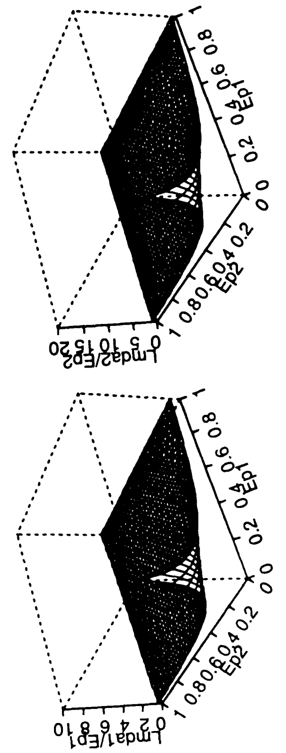

Figure 2: Full Model

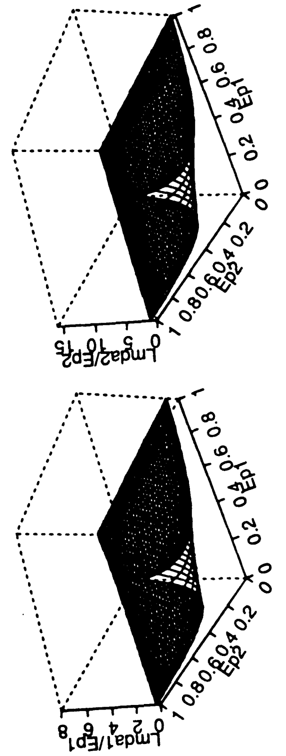

Figure 2(a): Obsn.25 Deleted

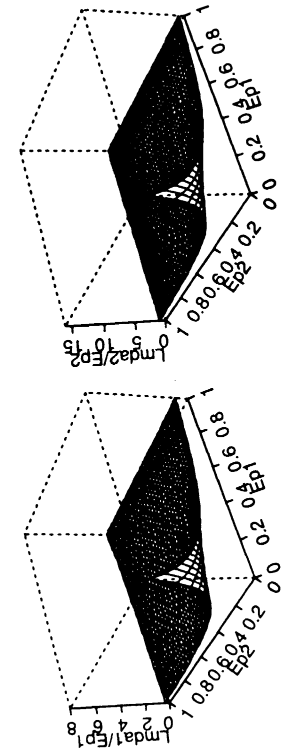

Figure 2(b): Obsn.26 Deleted

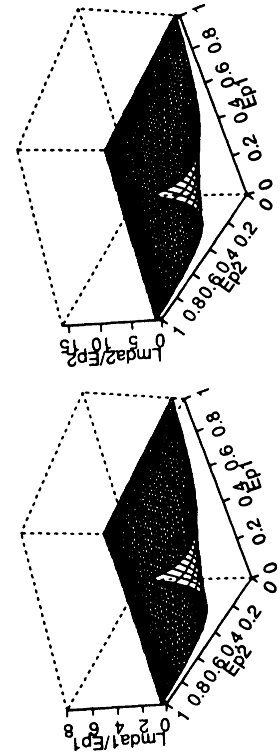

Figure 2(c): Obsn.27 Deleted

TABLE 1.
Local Sensitivity Measures for the Binomial Count Data

Models	g: logistic		g: probit	
	s_1	s_2	s_1	s_2
Full Model	9.43	16.94	9.41	16.91
Obs. 25 deleted	7.36	14.83	7.36	14.77
Obs. 26 deleted	7.58	14.75	7.58	14.73
Obs. 27 deleted	7.41	14.09	7.43	14.12

It follows from Table 1 that observation 25, 26 and 27 are all influential with respect to the change in local sensitivity measures.

A little study of Table 1 shows that the sensitivity measures s_1 and s_2 remain almost invariant when we move the transformation function h from logistic to probit. Similar feature might be observed if we vary h over a class of smooth functions giving rise to normal approximation. Let $\mathcal{G}$ be the class of densities g, obtained through the h function. Then $s_2(f, \mathcal{G}; x)$ as defined in (3.7) has value closed to as observed in Table 1. Thus, we conclude that whatever approximation (using h-function) we might use, the binomial model remains sensitive to such choice of likelihoods (obtained via normal approximation). This feature would have been overlooked if we just perturbed the prior, thus establishing the need for dual perturbation as a check for model adequacy.

Acknowledgements. The authors would like to thank Dr. Fabrizio Ruggeri and the two referees for many helpful suggestions.

REFERENCES

BASU, S. (1994). Posterior sensitivity to the sampling distribution and the prior: more than one observation. Research Report No.66, Department of Mathematical Sciences, University of Arkansas.

BASU, S., JAMMALAMADAKA, S.R. AND LIU, W. (1993). Local posterior robustness with parametric priors: maximum and average sensitivity. Technical Report No.239, Department of Statistics and Applied Probability, University of California, Santa Barbara.

BERGER, J.O. (1984). The robust Bayesian viewpoint (with discussion). *In Robustness in Bayesian Statistics* (J. Kadane ed.): North-Holland, Amsterdam.

BERGER, J.O. (1990). Robust Bayesian analysis: sensitivity to the prior. *J. Statist. Plann. Inference* **25** 303-328.

BERGER, J.O. (1994). An overview of robust Bayesian analysis (with discussion). *Test* **3** 5-124.

BERGER, J.O. AND BERLINER, M. (1986). Robust Bayes and empirical Bayes analysis with ϵ-contaminated priors. *Ann. Statist.* **14** 461-486.

COOK, R.D. AND WEISBERG, S. (1982). *Residuals and Influence in Regression.* London: Chapman and Hall.

DEY, D. K. AND BIRMIWAL, L. (1994). Robust Bayesian analysis using entropy and divergence measures. *Statist. and Prob. Letters* **20** 287-294.

GHOSH, S.K. AND DEY, D.K. (1994). Sensitivity diagnostics and robustness issues in Bayesian inference. Technical Report No. 94-30, Department of Statistics, University of Connecticut.

GUSTAFSON, P. (1994). Local sensitivity in Bayesian statistics. Ph.D. dissertation, Department of Statistics, Carnegie Mellon University.

GUSTAFSON, P. AND WASSERMAN, L. (1993). Local sensitivity diagnostics for Bayesian inference. Technical Report No.574, Department of Statistics, Carnegie Mellon University.

JEFFREYS, H. (1961). *Theory of Probability* (3rd edn.). Oxford University Press, London.

KASS, R.E. AND RAFTERY, A.E. (1995). Bayes factors, *J. Amer. Statist. Assoc.* **90** 773-795.

LAVINE, M. (1991). Sensitivity in Bayesian Statistics: the prior and the likelihood. *J. Amer. Statist. Assoc.* **86** 396-399.

LINDSEY, J.K. (1993). *Models for Repeated Measurements.* Clarendon Press, Oxford.

MCCULLAGH, P. AND NELDER, J.A. (1989). *Generalized Linear Models,* Second Edition, London: Chapman and Hall.

RUGGERI, F. AND WASSERMAN, L. (1993). Infinitesimal sensitivity of posterior distributions. *Canad. J. Statist* **21** 195-203.

SIVAGANESAN, S. (1993). Robust Bayesian diagnostics. *J. Statist. Plann. Inference* **35** 171-188.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CONNECTICUT
U-120, 196 AUDITORIUM ROAD
STORRS, CT 06269.
SUJITG@STAT.UCONN.EDU

On Local Sensitivity Measures in Bayesian Analysis

discussion by
ISABELLA VERDINELLI
University of Rome, Italy
and Carnegie Mellon University, Pittsburgh, PA, U.S.A.

I enjoyed reading this interesting paper addressing the use of general measures of local sensitivity. These measures allow one to consider the effect on posterior inference of perturbations, not only in the prior distribution, but also in the likelihood. I liked, in particular, the simple representations (3.8) and (3.9), that express the local sensitivity measures as Bayes factors of one model with respect to the other. The authors also show how to consider these local sensitivity measures for Generalized Linear Models and do an application to a data set to detect model adequacy.

I believe it is a good idea to extend Bayesian robustness to allow for uncertainty in the model specification. One particular problem, arising when perturbing the likelihood, is that the model's parameters need to maintain their meaning when the likelihood is perturbed. In fact, the requirement made in section 2 of the paper is that the base likelihood, f, and the contaminated likelihood, g, are compatible in the sense that they both belong to the same location-scale family. This is quite reasonable, but it appears as if, when examining the data set in section 5 this condition was overlooked.

The base model f of section 5 is the logistic model. The two possible alternatives considered, g_1 and g_2 say, are the normal regression model, after a logit transformation, and the probit model. These three models are not in the same location-scale family and the meaning of the respective parameters does not appear to be the same. This has an effect on the diagnostic proposed to detect sensitivity to model specifications. The perturbed likelihood is a mixture either of the models f and g_1, or of the models f and g_2. The local diagnostic, defined through the Fréchet derivative of f in the direction of g, depends on all the parameters involved in the two models, and I do not quite understand whether the values of the diagnostics for the different models proposed are in fact comparable.

Consider again formula (3.8) that represents the local sensitivity of P in the direction of Q as the ratio $m_Q(x, f)/m_P(x, f)$. Note that that this particular ratio is not quite a Bayes factor. In fact since

$$s_1(P, Q, x) = \frac{m_Q(x, f)}{m_P(x, f)} = \frac{\int \prod f(x_i|\theta)\ Q(d\theta)}{\int \prod f(x_i|\theta)\ P(d\theta)},$$

it can be thought of as a Bayes factor for the prior distributions. Although it might be reasonable to extend the concept of *model* to the set of all the assumptions made, so that in the Bayesian framework a model would include both the likelihood and the prior distribution, still the interpretation of the above ratio is unclear.

Other authors (e.g. Gustafson and Wasserman, 1993) derived different expressions for (3.8) that also involved a distance. While it seems an advantage that this diagnostics does not depend on the type of distance chosen, I still wonder if the interpretation of this measure of sensitivity to variation in the prior distribution might be somewhat easier when a particular distance is considered as well.

My next comment concerns the implementation of the Monte Carlo sampling in the example. The authors chose to approximate integrals of the type $I = \int f(x|\theta)P(d\theta)$ by the Monte Carlo estimate $\hat{I} = N^{-1}\sum_{t=1}^{N} f(x|\theta^{(t)})$ where $\theta^{(t)}, t = 1, 2, \ldots N$ are sampled from the prior distribution P. This results in an inefficient estimate of I. A better way to proceed would be to compute instead:

$$\hat{I} = \left[\frac{1}{N}\sum_{t=1}^{N}\frac{h(\theta^{(t)})}{f(x|\theta^{(t)})P(\theta^{(t)})}\right]^{-1},$$

where $h(\cdot)$ is a density and $\theta^{(t)}, t = 1, 2, \ldots N$ are sampled from the posterior distribution, as in Gelfand and Dey (1994).

My last question is about the interpretation of the plots. They all look very similar to me, and I do not understand if these pictures actually help us to see the effect of deleting observations from the sample.

ADDITIONAL REFERENCES

GELFAND, A. AND DEY, D. (1994). Bayesian Model Choice; Asymptotics and Exact Calculations. *Journal of the Roy. Statist. Society*, Ser. B, **56**, 501-514.

REJOINDER

DIPAK K. DEY, SUJIT K. GHOSH AND KUO-REN LOU

We thank the discussant for her kind and generally positive remarks. We agree that the notion of compatibility of likelihood is very tricky. In this paper we mention location and scale family as an example of compatibility. However, in the illustrative example, the compatibility of f is made with g_1 and g_2 through the one-one transformation, i.e. through the function h of the binomial proportion p.

We think that a model in a Bayesian framework is defined through the likelihood and the prior. Thus $s_1(P, Q, x)$ is really a ratio of marginal likelihoods and can be thought of as Bayes factor for the prior distributions and Jeffreys' scale of evidence can be used for the model selection. Regarding the difference in definition between ours with Gustufson and Wasserman (1993), we want our sensitivity measures to be free from a particular choice of distance, so that we can calibrate our measure.

The Monte Carlo estimates of the marginal distribution mentioned in this paper are definitely less stable than the estimator proposed in Gelfand and Dey (1994), but it is done to avoid a Metropolis step to keep the computation simple. We agree that the nature of the plots are similar, e.g. monotonicity, but the scales are totally different. See Table 1 for clarification.

Finally, we think deeper exploration of compatibility of models and different types of perturbations of the models may help us to develop better understanding of robustness problem.