Journal of the Korean Statistical Society (2001), **30**: *3, pp 467–480*

Local Sensitivity Analysis using Divergence Measures under Weighted Distribution

Younshik Chung[1] **and Dipak K.Dey**[2]

ABSTRACT

This paper considers the use of local φ-divergence measures between posterior distributions under classes of perturbations in order to investigate the inherent robustness of certain classes. The smaller value of the limiting local φ-divergence implies more robustness for the prior or the likelihood. We consider the cases when the likilihood comes from the class of weighted distribution. Two kinds of perturbations are considered for the local sensitivity analysis. In addition, some numerical examples are considered which provide measures of robustness.

Keywords: Local sensitivity, Bayesian robustness, Perturbation, φ-divergence, t-distribution, Gamma distribution, Weibull distribution, Weighted distribution.

1. Introduction

There are many situations where the usual random sample from a population of interest is not available, due to the data having unequal probabilities of entering the sample. Even if a random sample can be obtained, the experimenter may choose not to use it, since a carefully chosen bias sample may turn out to be more informative (e.g. Bayarri and DeGroot, 1992). The class of weighted distributions models this ascertainment bias by adjusting the probabilities of actual occurrence of events to arrive at a specification of the probabilities of the events as observed and recorded.

Suppose the random variable (or random vector) X is distributed as $f(x|\theta)$, and we are interested in performing inferences on θ. Suppose, further, that the probability that an observation x enters the sample gets multiplied by some nonnegative weight function $w(x)$. Then the observed sample is a random sample from the weighted distribution

$$f^w(x \mid \theta) = \frac{w(x)f(x \mid \theta)}{E_\theta\{w(X)\}} \tag{1.1}$$

[1]Department of Statistics, Pusan National University, Pusan, 609-735 KOREA
[2]Department of Statistics, University of Connecticut, Storrs, 06269 CT, U.S.A.

where $E_\theta\{w(X)\} = \int w(x)f(x \mid \theta)dx$. Such a random sample is termed a weighted sample. Rao(1965) first unifies the concept of weighted distribution. Since model specification is crucial in data analysis, it is important to account for possible bias in the sample by assigning a weight function $w(x)$ which specifies in a natural way the probability of recording $X = x$. Hedges and Olkin (1985) discuss how to incorporate a known weight function into a fixed-effects model for a meta analysis. In the same context, Iyengar and Greenhouse(1988) extend this approach by considering for maximum likelihood estimation of weight function. Bayarri and DeGroot (1992) suggest several applications of weighted distributions. Larose and Dey (1996) study weighted distributions in the context of model selection in a Bayesian frame work whereas Bayarri and Berger(1998) consider robustness issues for the weight function. Also, Silliman(1997) performs Bayesian estimation of weight function in the context of a hierarchical random-effects model. Recently, Chung, Dey and Jang(2000) consider the semiparametric hierarchical selection model using the weight function.

A Bayesian analysis depends strongly on the modeling assumptions which make use of both prior and likelihood. Even after fitting a standard statistical model to a given set of data, one does not feel comportable 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 to potentially conceivable direction to determine the effect of such alterations on the analysis. Often it is difficult to specify or elicit a method that would yield a convincing prior. The situation 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, 1985, 1990) and the references contained therein. Thus, 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 studies the variations of several posterior features. See Berger (1990), Basu and DasGupta (1992) and Sivaganesan (1993). Alternatively, in local analysis the effects of minor perturbations around some elicited priors are studied: see Ruggeri and Wasserman(1993), Gustafson (1994) and Dey, Ghosh and Lou(1996).

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 Bayesian robustness analysis, some researchers have used a general φ-divergence measure as defined by Csiszar (1967) to measure the variation between two posterior distributions. In Dey and Birmiwal(1994), the posterior robustness is measured using φ- divergence where the variation of posterior distribution is studied for fixed likelihood. Delampady and Dey (1994) consider the variation of the local curvatures of the φ- divergence between posteriors when the prior varies within mixtures of symetric and unimodal classes.

In this paper we consider the effect of perturbation of the standard model within a parametric family. This type of perturbations are natural when graphical or other statistical procedures indicate the possibility that the standard model may only be marginally adequate. Following Geisser(1993), we consider three different classes of pertubations. The class of t-distributions with varying degrees of freedom is useful for the robustness study of the location parameter problem, whereas the class of gamma distributions or Weibull distributions with varying the shape parameter or varying the scale parameter is useful for the robustness study of the scale family or of the shape family. We develop results using the limiting local divergence between the posterior distributions under an elicited prior and its perturbations under classes of perturbations of distributions families to study the local sensitivity of the posterior distributions.

An outline of this paper is as follows: In section 2, we define the φ-divergence and develop related notations. In section 3, we obtain the results of the limiting local φ-divergences between two posterior weighted distributions under a class of t-distribution or of gamma distributions or of Weibull distributions, which can be used as a measure of the local robustness. Finally, section 4 contains four examples to demonstrate the results obtained in section 3.

2. Definitions and Notations

Suppose x denotes the observable random variable with density $f(x|\theta)$ where θ an unknown parameter. Once a proper prior $\pi(\theta)$ is specified, then the marginal density of X corresponding to f and π is defined as

$$m(x) = \int f(x|\theta)\pi(\theta)d\theta$$

and its posterior density of θ given x corresponding to f and π is defined as

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)}.$$

In weighted distribution problem, a realization x of X under $f(x|\theta)$ enters the investigator's record with probability proportional to a weight function $w(x)$.

Clearly, the recorded x is not an observation on X, but on the random variable X^w, say, having the probability density function(pdf)

$$f^w(x|\theta) = \frac{w(x)f(x|\theta)}{E_f[w(x)]},$$

where $E_f[w(x)] = \int w(x)f(x|\theta)dx$ is the normalizing constant. The random variable X^w is called the weighted version of X and its distributions in relation to that of X is called the weighted distribution with weight function $w(x)$.

In addition, the marginal density of X^w corresponding to f^w and π is defined as

$$m^w(x) = \int f^w(x|\theta)\pi(\theta)d\theta$$

and its posterior density of θ given x corresponding to f^w and π is defined as

$$\pi^w(\theta|x) = \frac{f^w(x \mid \theta)\pi(\theta)}{m^w(x)}.$$

Following Csiszar(1967), we define the general φ-divergence between any two posterior distrubutions $\pi(\theta|x)$ and $\pi_\delta(\theta|x)$ as

$$D_\varphi = D_\varphi\{\pi_\delta(\theta|x), \pi(\theta|x)\} = \int \pi(\theta|x)\varphi(\frac{\pi_\delta(\theta|x)}{\pi(\theta|x)})d\theta, \tag{2.1}$$

where φ is assumed to be a convex function with a bounded third derivative.

There are several well-known φ-divergence measures. For example, $\varphi(x) = ln(x)$ defines the Kullback-Leibler divergences, $\varphi(x) = (\sqrt{x}-1)^2$ gives Hellinger distance, $\varphi(x) = (x-1)^2$ defines the chi-square divergence, $\varphi(x) = \frac{1}{2}|x-1|$ defines the variational distance of L_1 norm and $\varphi(x) = (x^3-1)/\lambda(\lambda+1), \lambda \neq 0, -1$ defines the power weighted divergence as studied extensively in the context of goodness of fit test in Reid and Cressie (1988).

From (2.1), by Taylor expansion on φ-function, the general φ-divergences between two posterior weighted distributions $\pi^w_{\delta_0}(\theta|x)$ and $\pi^w_\delta(\theta|x)$ becomes

$$\begin{aligned} D^w_\varphi &= \int \pi^w_{\delta_0}(\theta|x)\varphi(\frac{\pi^w_\delta(\theta|x)}{\pi^w_{\delta_0}(\theta|x)})d\theta \qquad (2.2) \\ &= \varphi(1) + \frac{\varphi''(1)}{2}(\delta-\delta_0)^2 \int \pi^w_{\delta_0}(\theta|x)\left[\frac{\partial}{\partial\delta}(\frac{\pi^w_\delta(\theta|x)}{\pi^w_{\delta_0}(\theta|x)})\right]^2 d\theta + O((\delta-\delta_0)^3) \\ &= \varphi(1) + \frac{\varphi''(1)}{2}(\delta-\delta_0)^2 E_{\pi^w_{\delta_0}(\theta|x)}\{\left[\frac{\partial}{\partial\delta}(\frac{\pi^w_\delta(\theta|x)}{\pi^w_{\delta_0}(\theta|x)})\right]^2\} + O((\delta-\delta_0)^3), \end{aligned}$$

where δ is a perturbation of the likelihood or prior which results to the posterior distribution $\pi_\delta^w(\theta|x)$, δ_0 is known, $O((\delta-\delta_0)^3)$ is the remainder term with order 3 or higher. We further assume that the differentiation with respect to δ and integration with respect to θ of the posterior p.d.f. and its derivatives are interchangeable. Note that since for all n,

$$\int \pi_{\delta_0}^w(\theta|X)\frac{\partial^n}{\partial\delta^n}\left(\frac{\pi_\delta^w(\theta|X)}{\pi_{\delta_0}^w(\theta|X)}\right)d\theta = \frac{\partial^n}{\partial\delta^n}\int \pi_{\delta_0}^w(\theta|x)\frac{\pi_\delta^w(\theta|x)}{\pi_{\delta_0}^w(\theta|x)}d\theta = 0,$$

therefore the term $(\delta-\delta_0)\varphi'(1)\int \pi_{\delta_0}^w(\theta|x)\frac{\partial}{\partial\delta}\left(\frac{\pi_\delta^w(\theta|x)}{\pi_{\delta_0}^w(\theta|x)}\right)d\theta$ is vanished.

From (2.2), let us now define the limiting local φ-divergence between two posterior weighted distributions $\pi_{\delta_0}^w(\theta|x)$ and $\pi_\delta^w(\theta|x)$ as

$$\begin{aligned}\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi^w-\varphi(1)\right] &= \lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[\int \pi_{\delta_0}^w(\theta|x)\varphi\left(\frac{\pi_\delta^w(\theta|x)}{\pi_{\delta_0}^w(\theta|x)}\right)d\theta-\varphi(1)\right]\\ &= \frac{\varphi''(1)}{2}\lim_{\delta\to\delta_0}E_{\pi_{\delta_0}^w(\theta|x)}\left[\frac{\partial}{\partial\delta}\left(\frac{\pi_\delta^w(\theta|x)}{\pi_{\delta_0}^w(\theta|x)}\right)\right]^2. \qquad (2.3)\end{aligned}$$

3. Weighted Distribution Families

3.1. Perturbation of Likelihood

In this section, we will only invesigate the local sensitivity measures for the perturbation of likelihood within distribution families in weighted distributions. We consider a fixed prior $\pi(\theta)$ and perturb the weighted likelihood within a distribution family of the form

$$f_r^w(x|\theta)=\frac{w(x)}{E_{f_r}[w(x)]}f_r(x|\theta)$$

where $w(x)$ is the fixed weight function and $f_r(x|\theta)$ is the unweighted distributions which could be t-distribution, gamma distribution or Weibull distribution as follows:

First, we fix the prior distribution $\pi(\theta)$, and perturb the likelihood function $f_r^w(x|\theta)$ from the unweighted distribution $f_r(x|\theta)$ which is a member of the class of t-distributions of the form

$$f_r(x|\theta)=\frac{\Gamma(\frac{r+1}{2})}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\left[1+\frac{(x-\theta)^2}{r\sigma^2}\right]^{-\frac{r+1}{2}}, \quad r\geq 1 \qquad (3.1)$$

where σ^2 is known. In this case, it is well known that when the degree of freedom r goes to infinity, the unweighted likelihood goes to the normal distribution. Thus varying r from 1 to ∞, we can generate a class of likelihood functions.

Second, we fix the prior distribution $\pi(\theta)$ and perturb the likelihood function $f_r^w(x|\theta)$ from the unweighted distribution $f_r(x|\theta)$ which is a member of the class of gamma distributions of the form

$$f_r(x|\theta) = \frac{\theta^r}{\Gamma(r)} x^{r-1} e^{-\theta x}, \quad r \geq 1, \quad \theta > 0. \tag{3.2}$$

In this case, it is well known that when the shape parameter r goes to 1, the unweighted likelihood goes to the exponential distribution with unknown scale paramter θ. Thus varying r from 1 to ∞, we can generate a class of likelihood functions.

Finally, we fix the prior distribution $\pi(\theta)$, and perturb the likelihood function $f_r^w(x|\theta)$ from the unweighted distribution $f_r(x|\theta)$ which is a member of the class of Weibull distribution of the form

$$f_r(x|\theta) = r\theta x^{\theta-1} e^{-rx^\theta}, \quad r \geq 1. \tag{3.3}$$

In this case, when the scale parameter r varies from 1 to ∞, we can generate a class of likelihood functions.

Note that for each case, r denotes the perturbation of the likelihood function. From now on, we use r as the perturbation instead of δ for notational simplicity. Then for each case, $f_r^w(x|\theta)$ can be written as

$$f_r^w(x|\theta) = \frac{w(x)h_x(r)}{\int w(x)h_x(r)\,dx}, \tag{3.4}$$

where $h_x(r) = \begin{cases} [1 + \frac{(x-\theta)^2}{\sigma^2 r}]^{-\frac{r+1}{2}}, & \text{for t-distibution} \\ x^{r-1}e^{-\theta x}, & \text{for gamma distribution} \\ x^{\theta-1}e^{-rx^\theta}, & \text{for Weibull distribution.} \end{cases}$

Note that $\frac{\partial h_x(r)}{\partial r}$ can be expressed as

$$\frac{\partial h_x(r)}{\partial r} = h_x(r) l_{x|\theta}(r)$$

where

$$l_{x|\theta}(r) = \begin{cases} \frac{r+1}{2}\frac{\frac{(x-\theta)^2}{\sigma^2 r^2}}{1+\frac{(x-\theta)^2}{\sigma^2 r}} - \frac{1}{2}log[1 + \frac{(x-\theta)^2}{\sigma^2 r}], & \text{for t-distribution} \\ logx, & \text{for gamma distribution} \\ -x^\theta, & \text{for Weibull distribution.} \end{cases}$$

Theorem 3.1. *Suppose that the weight function $f_r^w(x|\theta)$ satisfies the condition above. Then the limiting local φ-divergence is given as*

$$\lim_{r\to r_0}\frac{1}{(r-r_0)^2}[D_\varphi^w-\varphi(1)] = \lim_{r\to r_0}\frac{1}{(r-r_0)^2}\left[\int \pi_{r_0}^w\varphi\left(\frac{\pi_r^w(\theta|x)}{\pi_{r_0}(\theta|x)}\right)d\theta-\varphi(1)\right]$$
$$= \frac{\varphi''(1)}{2}Var_{\pi_{r_0}^w}(\theta|x)\left[l_{x|\theta}(r_0)-E_{f_{r_0}^w}[l_{x|\theta}(r_0)]\right] \quad (3.5)$$

where r_0 is known and $\pi_{r_0}^w(\theta|x)$ is the posterior weighted distribution with the likelihood function $f_{r_0}^w(x|\theta)$ and $l_{x|\theta}(r)$ is defined as before.

Proof. Recall that $f_r^w(x|\theta)=\frac{w(x)h_x(r)}{\int w(x)h_x(r)dx}$. Therefore,

$$\frac{\partial f_r^w(x|\theta)}{\partial r} = f_r^w(x|\theta)\left[l_{x|\theta}(r)-\int f_r^w(x|\theta)l_{x|\theta}(r)dx\right] \quad (3.6)$$

and then

$$\frac{\partial\pi_r^w(\theta|x)}{\partial r} = \frac{\frac{\partial}{\partial r}f_r^w(x|\theta)\pi(\theta)\int f_r^w(x|\theta)\pi(\theta)d\theta - f_r^w(x|\theta)\pi(\theta)\int\frac{\partial}{\partial r}f_r^w(x|\theta)\pi(\theta)dr}{(\int f_r^w(x|\theta)\pi(\theta)d\theta)^2}$$
$$= \pi_r^w(\theta|x)\left[l_{x|\theta}(r)-\frac{\int w(x)h_x(r)l_{x|r}(r)dx}{\int w(x)h_x(r)dx}\right]$$
$$- \pi_r^w(\theta|x)*\int\pi_r^w(\theta|x)\left[l_{x|\theta}(r)-\frac{w(x)h_x(r)l_{x|\theta}(r)dx}{\int w(x)h_x(r)dx}\right]d\theta$$
$$= \pi_r^w(\theta|x)\left[k_r(\theta,x)-E_{\pi_r^w(\theta|x)}k_r(\theta,x)\right], \quad (3.7)$$

where $k_r(\theta,x)=l_{x|\theta}(r)-\int l_{x|\theta}(r)f_r^w(x|\theta)dx$. Therefore,

$$E_{\pi_{r_0}^w(\theta|x)}\left(\frac{\partial}{\partial r}\frac{\pi_r^w(\theta|x)}{\pi_{r_0}^w(\theta|x)}\right)^2 = E_{\pi_{r_0}^w(\theta|x)}\left[k_{r_0}(\theta,x)-E_{\pi_{r_0}^w(\theta|x)}k_{r_0}(\theta,x)\right]^2$$
$$= Var_{\pi_{r_0}^w(\theta|x)}\left[k_{r_0}(\theta,x)\right]. \quad (3.8)$$

This completes the proof.

The following three corollaries are easily obtained from Theorem 3.1. Therefore, their proofs are omitted here.

Corollary 3.1. Under the class of t-distributions, the limiting local φ-divergence between two posterior weighted distributions is given as

$$\lim_{r\to r_0}\frac{1}{(r-r_0)^2}\left[D_\varphi^w-\varphi(1)\right]=\frac{\varphi''(1)}{2}Var_{\pi_{r_0}^w(\theta|x)}\left[l_{x|\theta}(r_0)-E_{f_{r_0}^w(x|\theta)}[l_{x|\theta}(r_0)]\right]$$

$$= \frac{\varphi''(1)}{2}\{Var_{\pi_{r_0}^w(\theta|x)}\left[l_{x|\theta}(r_0)\right]+Var_{\pi_{r_0}^w(\theta|x)}\left[E_{f_{r_0}^w(x|\theta)}[l_{x|\theta}(r_0)]\right]\}$$
$$- \varphi''(1)Cov_{\pi_{r_0}^w(\theta|x)}\left(l_{x|\theta}(r_0),E_{f_{r_0}^w(x|\theta)}[l_{x|\theta}(r_0)]\right),$$

where r_0 is known and $\pi^w_{r_0}(\theta|x)$ is defined as before and

$$l_{x|\theta}(r_0) = \frac{r_0+1}{2}\frac{\frac{(x-\theta)^2}{\sigma^2 r_0^2}}{1+\frac{(x-\theta)^2}{\sigma^2 r_0}} - \frac{1}{2}log\left[1+\frac{(x-\theta)^2}{\sigma^2 r_0}\right].$$

Corollary 3.2. Under the class of gamma distributions, the limiting local $\varphi-$ divergence between two posterior weighted distributions is given as

$$\lim_{r\to r_0}\frac{1}{(r-r_0)^2}\left[D^w_\varphi - \varphi(1)\right] = \frac{\varphi''(1)}{2}Var_{\pi^w_{r_0}(\theta|x)}\left[lnx - E_{f^w_{r_0}(x|\theta)}[lnx]\right].$$

Corollary 3.3. Under the class of Weibull distributions, the limiting local φ-divergence between two posterior weighted distributions is given as

$$\lim_{r\to r_0}\frac{1}{(r-r_0)^2}\left[D^w_\varphi - \varphi(1)\right] = \frac{\varphi''(1)}{2}Var^w_{\pi_{r_0}}(\theta|x)\left[x^\theta - E_{f^w_{r_0}(x|\theta)}[x^\theta]\right].$$

3.2. Perturbation of Prior

We consider a likelihood $f^w(x|\theta)$ of the form

$$f^w(x|\theta) = \frac{w(x)}{E_f[w(x)]}f(x|\theta)$$

where $w(x)$ is the fixed weight function and $f(x|\theta)$ is the unweighted distributions and perturb the prior which could be again a t-distribution, gamma distribution, or Weibull distribution.

First, we fix the likelihood function $f^w(x|\theta)$ from any location family density function with location parameter θ and perturb the prior distribution within the class of t-distributions with varying degree of freedom r which has the form

$$\pi_r(\theta) = \frac{\Gamma(\frac{r+1}{2})}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\left[1+\frac{(\theta-\mu)^2}{r\sigma^2_\pi}\right]^{-\frac{r+1}{2}}, r\geq 1,$$

where μ and σ^2_π are known location and scale parameters respectively. In this case, it has the similar interpretation like the perturbation of likelihood in section 3.1.

Next, we fix the likelihood function $f^w(x|\theta)$ where θ is a scale parameter from arbitrary distribution on $(0,\infty)$, and perturb the prior within the class of gamma distributions with varying shape parameter r which has the form

$$\pi_r(\theta) = \frac{\beta^r}{\Gamma(r)}\theta^{r-1}e^{-\beta\theta}, r\geq 1, \beta>0,$$

where the scale parameter β is assumed to be known. Also, it has the same interpretation like the perturbation of likelihood in section 3.1.

Finally, we fix the likelihood function $f(x|\theta)$, where θ is a shape parameter from arbitrary distribution on $(0,\infty)$and perturb the prior distribution within the class of Weibull distributions with varying shape parameter r which has the form

$$\pi_r(\theta) = br\theta^{r-1}e^{-b\theta^r}, r \geq 1, b > 0,$$

where the scale parameter b is assumed to be known. In this case it is well known that when the shape parameter r goes to 1, the prior reduces to the exponential distribution with known scale parameter b. Thus varying r from 1 to ∞, we can generate a class of priors.

Then for each class, the posterior weighted distribution is expressed as

$$\pi_r^w(\theta|x) = \frac{f^w(x|\theta)\pi_r(\theta)}{\int f^w(x|\theta)\pi_r(\theta)d\theta} = \frac{f^w(x|\theta)q_\theta(r)}{\int f^w(x|\theta)q_\theta(r)d\theta} \tag{3.9}$$

where $q_\theta(r) = \begin{cases} [1+\frac{(\theta-\mu)^2}{r\sigma_\pi^2}]^{-\frac{r+1}{2}}, & \text{for t-distribution} \\ \theta^{r-1}e^{-\beta\theta}, & \text{for gamma distribution} \\ \theta^{r-1}e^{-b\theta^r}, & \text{for Weibull distribution.} \end{cases}$

Then $\frac{\partial}{\partial r}q_\theta(r) = q_\theta(r)l_{\theta|x}(r)$, where

$$l_{\theta|x}(r) = \begin{cases} \frac{r+1}{2}\frac{\frac{(\theta-\mu)^2}{\sigma_\pi^2 r^2}}{1+\frac{(\theta-\mu)^2}{\sigma_\mu^2 r}} - \frac{1}{2}log[1+\frac{(\theta-\mu)^2}{\sigma_\pi^2 r}], & \text{for t-distribution} \\ log\theta, & \text{for gamma distribution} \\ [1-b\theta^r]log\theta, & \text{for Weibull distribution.} \end{cases}$$

Theorem 3.2. *Under the above prior, the limiting local φ-divergence is given as*

$$\begin{aligned} \lim_{r\to r_0}\frac{1}{(r-r_0)^2}[D_\varphi - \varphi(1)] &= \lim_{r\to r_0}\frac{1}{(r-r_0)^2}\left[\int \pi_{r_0}^w(\theta|x)\varphi\left(\frac{\pi_r^w(\theta|x)}{\pi_{r_0}(\theta|x)}\right)d\theta - \varphi(1)\right] \\ &= \frac{\varphi''(1)}{2}Var_{\pi_{r_0}^w(\theta|x)}\left[l_{\theta|x}(r_0)\right] \end{aligned} \tag{3.10}$$

where r_0 is known and $\pi_{r_0}^w(\theta|x)$ is the posterior weighted distribution with the likelihood function $f_{r_0}^w(x|\theta)$ and $l_{\theta|x}(r_0)$ is defined as before.

Proof. It follows that

$$\begin{aligned} \frac{\partial}{\partial r}\pi_r^w(\theta|x) &= l_{\theta|x}(r)\pi_r^w(\theta|x) - \pi_r^w(\theta|x)\int l_{\theta|x}(r)\pi_r^w(\theta|x)d\theta \\ &= \pi_r^w(\theta|x)\left[l_{\theta|x}(r) - E_{\pi_r^w(\theta|x)}l_{\theta|x}(r)\right]. \end{aligned} \tag{3.11}$$

Therefore the result is easily obtained.

4. An Illustrative Example

We consider four plausible candidates of likelihood from the class of t-destributions which are $N(\theta,1), T_{10}(\theta,1), T_5(\theta,1)$ and $C(\theta,1)$ under the suitable weight function $w(x)$. Silliman (1997) considered $w(x) = |x|^{\alpha}$ as weight function. For our example, we set $w(x) = |x|$ and x^2 in Table 4.1 and 4.2, respectively. Since there does not exist moments of Cauchy distribution, we can not find the weighted distribution with $w(x) = x^2$. Thus Table 4.2 has the values of $T_3(\theta,1)$ instead of Cauchy $C(\theta,1)$. It is assumed for simplicity that the prior is Normal $N(0,1)$.

Remark 4.1. Within a class of t distributions, when $r_0 = \infty$, that is, for normal unweighted likelihood, $l_{x|\theta}(r_0)$ is easily obtained from corollary 3.1 as follows;

$$l_{x|\theta}(r_0) = \left[\frac{1}{4}\left(\frac{\theta-x}{\sigma}\right)^4 - \frac{1}{2}\left(\frac{\theta-x}{\sigma}\right)^2\right] \tag{4.1}$$

Table 4.1 Different likelihoods of t-distributions with weight $|x|$

	$N(\theta,1)$	$T_{10}(\theta,1)$	$T_5(\theta,1)$	$C(\theta,1)$
x=0	0.0562	0.0259	0.0149	0.0031
x=1	0.3758	0.1516	0.0734	0.0045
x=2	3.4836	1.3237	0.5889	0.0156
x=3	22.011	7.1544	2.6634	0.0418
x=4	100.41	24.092	6.4163	0.0515
x=4.5	189.42	37.304	8.0274	0.0464
x=5	328.89	51.685	8.9276	0.0396
x=6	782.84	74.054	8.8017	0.0284
x=10	3895.3	59.773	4.2992	0.0102
x=15	13802	30.101	2.0150	0.0045
x=20	46438	17.506	1.1578	0.0025

Also, when $r = 1$, that is, the likelihood is from a Cauchy distribution. Then

$$l_{x|\theta}(r_0) = \left\{\frac{\frac{(x-\theta)^2}{\sigma^2}}{1+\frac{(x-\theta)^2}{\sigma^2}} - \frac{1}{2}log\left[1+\frac{(x-\theta)^2}{\sigma^2}\right]\right\} \tag{4.2}$$

Table 4.1 and Table 4.2 present values of the limiting local φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different models. Here the calculations are performed using the method of numerical integration.

Table 4.2 Different likelihoods of t-distributions with weight $|x|^2$

	$N(\theta,1)$	$T_{10}(\theta,1)$	$T_5(\theta,1)$	$T_3(\theta,1)$
x=0	0.0329	0.0191	0.0157	0.0083
x=1	0.2759	0.1268	0.6124	0.0319
x=2	3.5197	1.3048	0.5731	0.2410
x=3	27.009	7.4419	2.6470	0.9063
x=4	133.62	24.479	6.0082	1.5766
x=4.5	256.33	36.905	7.2686	1.6735
x=5	447.81	49.824	7.8752	1.6240
x=6	1055.6	68.153	7.5439	1.3490
x=10	4625.2	51.575	3.6398	0.5662
x=15	14773	25.848	1.7147	0.2601
x=20	48197	15.045	0.9844	0.1478

For each x, the values of the limiting local divergence measures are decreasing, when the degrees of freedom go down. For small $x(i.e., x \leq 2)$, it appears that those values are small, indicating some degree of robustness with respect to the choice of the prior. For moderate to large $x(i.e., x \geq 4.5)$, however, there can be a substantial difference among those values, indicating that the answer is then not robust to reasonable variation in the prior. For all x, it appears that the values of the limiting local divergence are not varied too much, when the posterior distribution comes from Cauchy likelihood. Note the dependence of robustness on the actual value of x. The conclusion in Table 4.2 is similar to that of Table 4.1.

Finally, we observe the likelihood $X \sim N(\theta,1)$ with weight function $w(x) = |x|$ and subjectively specify a prior median of 0 and quartiles of $+1$ and -1. Now we have four reasonable priors to be considered within the class of t-priors which are Cauchy prior $C(0,1)$, two t-priors $T_5(0,1.89361)$, $T_{10}(0,2.04198)$ and Normal prior. Table 4.3 and Table 4.4 present values of limiting local φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different priors using the method of numerical integration.

Remark 4.2. Within a class of t distributions, when $r_0 = \infty$, that is, for normal prior distribution, $l_{\theta|x}(r_0)$ is easily obtained as follows;

$$l_{\theta|x}(r_0) = \left[\frac{1}{4}\left(\frac{\theta-\mu}{\sigma_\pi}\right)^4 - \frac{1}{2}\left(\frac{\theta-\mu}{\sigma_\pi}\right)^2\right]. \tag{4.3}$$

Table 4.3. Different priors of t-distributions with weight $|x|$

	$N(0,2.198)$	$T_{10}(0,2.041)$	$T_5(0,1.893)$	$C(0,1)$
x=0	0.0030	0.0030	0.0032	0.0030
x=1	0.0061	0.0059	0.0056	0.0036
x=2	0.0385	0.0255	0.0190	0.0082
x=3	0.6844	0.3104	0.1856	0.0294
x=4	4.7352	1.6529	0.8044	0.0613
x=4.5	12.061	3.1737	1.3295	0.0609
x=5	22.726	4.9982	1.7619	0.0493
x=6	93.077	9.7506	2.6315	0.0315
x=10	1984.2	22.127	2.9775	0.0078
x=15	24674	19.785	1.7902	0.0041
x=20	101029	20.188	1.3050	0.0005

Also, when $r=1$, that is, the prior is a Cauchy distribution. Then

$$l_{\theta|x}(r_0)=\left\{\frac{\frac{(\mu-\theta)^2}{\sigma_\pi^2}}{1+\frac{(\mu-\theta)^2}{\sigma_\pi^2}}-\frac{1}{2}log\left[1+\frac{(\mu-\theta)^2}{\sigma_\pi^2}\right]\right\}. \tag{4.4}$$

Table 4.4 Different priors of t-distributions with weight $|x|^2$

	$N(0,2.198)$	$T_{10}(0,2.041)$	$T_5(0,1.893)$	$C(0,1)$
x=0	0.0021	0.0020	0.0021	0.0029
x=1	0.0056	0.0057	0.0049	0.0029
x=2	0.0298	0.0236	0.0158	0.0072
x=3	0.6395	0.2987	0.1725	0.0278
x=4	4.2789	1.2793	0.7392	0.0576
x=4.5	10.985	2.9782	1.2998	0.0569
x=5	20.278	4.8976	1.6739	0.0457
x=6	89.908	8.9779	1.7983	0.0311
x=10	1927.0	20.119	2.8957	0.0076
x=15	2198.6	18.256	1.6739	0.0038
x=20	9928.29	18.998	1.2096	0.0003

Again we observe that robustness is achieved for smaller values of x and under heavy tailed priors.

REFERENCES

Basu, S. and Das Gupta. A.(1992). Bayesian analysis with distribution bands: the role of the loss function. Technical Report 208, Department of Statinstics and Applied Probability, University of California, Santa Barbara.

Bayarri. M. J. and Berger. J. O.(1998). Robust Bayesian analysis of selection models. *Annals of Statistics*, **26**, 645-659.

Bayarri. M. J. and DeGroot. M.(1992). A BAD view of weighted distributions and selection models. In *Bayesian Statistics* IV, J.M. Bernardo, et. al (eds.) Oxford University Press. London.

Berger. J.O.(1984). The robust Bayesian viewpont (with discussion). *In Robustness in Baysian Statistics* (J. Kadane ed.): North-Holland. Amsterdam.

Berger. J.O.(1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer-Verlag, New York.

Berger. J.O.(1990). Robust Bayesian analysis sensitivity to the prior. *Journal of Statistical Planning and Inference*, **25**, 303-328.

Chung, Y., Dey, D.K. and Jang, J.(2000). Semiparametric hierarchical selection models for Bayesian meta analysis. *Proceedings of the Biophamaceatical section, American Statistical Association*, 1-8.

Csiszar, I.(1967). Information-type measures of difference of probability distributions and indirect observations. *Studio Scientiarum Mathematicarum Hungarica*, **2**, 299-318.

Delampady, M. and Dey, D.K.(1994). Bayesian robustness for multiparameter problems. *Journal of Statistical Planning and Inference*, **40**, 375-382.

Dey, D.K. and Birmiwal, L.(1994). Robust Bayesian analysis using entropy and divergence measures. *Statistics and Probability Letters*, **20**, 287-294.

Dey, D.K., Ghosh. S.K. and Lou, K. R.(1996). On local sensitivity measures in Bayesian analysis. In *Bayesian Robustness.* J.O Berger et. al.(eds.) IMS monograph series, **29**, 21-40.

Geisser. S(1993). *Predictive inference: an introduction*, Chapman and Hall, London.

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

Hedges, L.V. and Olkin, I.(1985). *Statistical Methods for Meta Analysis*, Academic Press, Orlando.

Iyenger, S. and Greenhouse, J.B.(1988). Selection models and the file drawer problem (with discussion). *Statistical Sciences*, **3**, 109-135.

Larose, D. and Dey, D.K.(1996). Weighted distributions viewed in the context of model selection: a Bayesian perspective. *Test*, **5**, 227-246.

Rao, C.R.(1965). On discrete distributions arising out of methods of ascertainment. In *Classical and Contagious Discrete Distributions*, G.P. Patil (ed.),Calcutta, India: Pergamon Press and Statistical Publishing Society, 320-332.

Reid, T.R.C. and Cressie, N.A.C.(1988). *Goodness of fit statistics for discrete multivariate data*, Springer Verlag, New York.

Ruggeri, F. and Wasserman, L(1993). Infinitestimal sensitivity of posterior distributions. *Canadian Journal of Statistics*, **21**, 195-203.

Silliman, N.P.(1997). Nonparametric classes of weighted functions to model publication bias. *Biometrika*, **84**, 909-918.

Sivaganesan, S(1993). Robust Bayesian diagnostics. *Journal of Statistical Planning and Inference*, **35**, 171-188.