行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　　※
※　在混亂母數種類中使用發散性測度藉觀局部敏感分析　※
※　　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：　☒個別型計畫　　☐整合型計畫
計畫編號：NSC 89－2118－M－032－023
執行期間：89 年 8 月 1 日至 90 年 7 月 31 日

計畫主持人：婁國仁　　副教授

執行單位：淡江大學管理科學學系

中　華　民　國　91 年　1　月　15　日

行政院國家科學委員會專題研究計畫成果報告

在混亂母數種類中使用發散性測度藉觀局部敏感分析
Local Sensitivity Analysis using Divergence Measures Under Parametric Classes of Perturbations

計畫編號：NSC 89-2118-M-032-023

執行期限：民國 89 年 8 月 1 日至 90 年 7 月 31 日

主持人：婁國仁 淡江大學管理科學學系 副教授

研究助理：張師彰、詹力權 淡江大學管理科學學系碩士班研究生

臺北縣淡水鎮英專路 151 號
Tel: (02)2621-5656 ext. 2837 E-mail: 109880@mail.tku.edu.tw

一、中文摘要

為了研究在某些混亂的集合中，察視保有固有的、健全的特性，所以我們在這些種類之中，使用在兩個事後分配之間的局部發散性測度。對於事前機率(prior)或概似函數(likelihood)而言，所獲得的局部發散性極限值來判定，其值愈小意味著愈有健全性。本報告考慮兩種的混亂(事前機率與概似函數)，分析局部性敏感度。此外，我們亦考慮概似函數來自加權分配(weighted distribution)的情況。最後，我們提出一些量化的實例，並加以說明。

關鍵詞：局部敏感度、貝氏健全性、混亂、發散、t 分配、Gamma 分配、韋伯分配、加權分配。

Abstract

We consider 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. In this paper, two kinds of the perturbations (prior and likelihood) are considered for the local sensitivity analysis. In addition, we also consider the cases when the likelihood comes from the class of weighted distributions. Finally some numerical examples are considered which provides measures of robustness.

Keywords: Local Sensitivity, Bayesian Robustness, Perturbation, φ-Divergence, t-Distribution, Gamma Distribution, Weibull Distribution, Weighted Distribution.

二、計畫緣由與目的

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 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 to potentially conceivable directions 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, 1986, 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. Alternatively, in local analysis the effects of minor perturbations around some elicited priors are studied.

Recent papers involving global sensitivity analysis are due to Berger(1990), Srinivasan and Truszczynska (1990), Basu and Dasgupta (1992), Sivaganesan (1993) and the references therein. 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 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 (usually known as f-divergence) as defined by Csiszar (1967) to measure the variation between two posterior distributions. In Dey and Birmiwal (1994), the posterior robustness was measured using φ-divergence where the variation of posterior distribution was studied for fixed likelihood when prior distribution varies within certain arithmatic and geometric contamination classes. Delampady and Dey (1994), considered the variation of the local curvatures of the φ-divergence between posteriors when the prior varies within mixtures of symetric and unimodal classes. Currently, there is an another direction of Bayesian robustness studies are being pursued, where the contamination of priors are considered within the class of scale mixtures of normal distributions. However, in these studies, robustness studies are considered when posterior moments exist. Notable results in this direction are due to Pericchi and Smith (1992) and Choy (1996) and the references contained therein.

There are many situations where the usual random sample from a population of interest is not available, due to 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 choosen bias sample may turn out to be more informative. The method 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.

Rao (1965) first unified the concept of weighted distribution. Patil and Rao(1977) discuss how truncation models and damaged observations can give rise to weighted distributions. Patil (1981) and Bayarri and DeGroot (1992) suggest several applications of weighted distribution. Recently, Larose and Dey (1996) studied weighted distributions in the context of model selection in a Bayesian framework whereas Bayarri and Berger (1993) consider robustness issues for the weight function. Here we will study the local sensitivity of weighted distribution.

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 perturbations. The class of t-distributions with varying degrees of freedom which is useful for the robustness study of the location parameter problem whereas the class of Gamma distribution or Weibull distribution 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 perturbation under classes of perturbations of distribution families to study the local sensitivity of the posterior distribution.

三、研究模型的摘要

Suppose X denotes the observable random variable (real valued) with density $f(x|\theta)$ where θ is 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 the posterior distribution 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 investigators record with probability proportional to a weight function $w(x)$. Clearly, the recorded x is not an observation on X, but on the r.v. X^w, say, having pdf

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

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

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 the posterior distribution of θ given x corresponding to f^w and π is defined as

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

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

$$D_\varphi = D_\varphi\{\pi_\delta(\theta|x), \pi(\theta|x)\} =$$

$$\int \pi(\theta|x)\varphi\left(\frac{\pi_\delta(\theta|x)}{\pi(\theta|x)}\right) d\theta \qquad (1)$$

where we assume φ is a convex function with a bounded third derivative. In our situation, $\pi_\delta(\theta|x)$ will be the general notation for a posterior distribution.

From (1) by Taylor expansion on φ-function, the general φ-divergences between two posterior distributions $\pi_{\delta_0}(\theta|x)$ and $\pi_\delta(\theta|x)$ becomes

$$D_\varphi = \int \pi_{\delta_0}(\theta x)\varphi\left(\frac{\pi_\delta(\theta|x)}{\pi_{\delta_0}(\theta|x)}\right) d\theta$$

$$= \varphi(1) +$$

$$\frac{\varphi''(1)}{2}(\delta-\delta_0)^2 E^{\pi_{\delta_0}(\theta|x)}\left\{\left[\frac{\partial}{\partial\delta}\left(\frac{\pi_\delta(\theta|x)}{\pi_{\delta_0}(\theta|x)}\right)\right]^2\right\}$$

$$+ O((\delta-\delta_0)^3) \qquad (2)$$

where δ is a perturbation of the likelihood or prior which results to the posterior distribution $\pi_\delta(\theta|x)$. 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.

From (2), let us now define the limiting local φ-divergence which will play an important role in our studies.

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi - \varphi(1)\right] =$$

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[\int \pi_{\delta_0}(\theta|x)\cdot\varphi\left(\frac{\pi_\delta(\theta|x)}{\pi_{\delta_0}(\theta|x)}\right)d\theta - \varphi(1)\right]$$

$$= \frac{\varphi''(1)}{2}\int \pi_{\delta_0}(\theta|x)\left[\frac{\partial}{\partial\delta}\left(\frac{\pi_\delta(\theta|x)}{\pi_{\delta_0}(\theta|x)}\right)\right]^2 d\theta$$

$$= \frac{\varphi''(1)}{2} E^{\pi_{\delta_0}(\theta|x)}\left\{\left[\frac{\partial}{\partial\delta}\left(\frac{\pi_\delta(\theta|x)}{\pi_{\delta_0}(\theta|x)}\right)\right]^2\right\},$$

where δ and δ_0 are positive and less than or equal 1. Note that $\varphi''(1)$ is always positive.

Similarly, the limiting local φ-divergence for weighted distribution problem between two posterior distributions $\pi^w_{\delta_0}(\theta|x)$ and $\pi^w_\delta(\theta|x)$ is the following :

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D^w_\varphi - \varphi(1)\right] = \frac{\varphi''(1)}{2}$$

$$E^{\pi^w_{\delta_0}(\theta,x)}\left\{\left[\frac{\partial}{\partial\delta}\left(\frac{\pi^w_\delta(\theta|x)}{\pi^w_{\delta_0}(\theta|x)}\right)\right]^2\right\}$$

四、結果與討論

1. Local sensitivity measures under a class of t-distributions：

(1) Perturbation of the prior for fixed likelihood

Here, we fix the likelihood function $f(x|\theta)$ from any location family density functions 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_π^2 are known location and scale parameters respectively.

In this case, it is well known that when the degree of freedom r goes to infinity, the prior goes to the Normal distribution; and when the degree of freedom r goes to 1, the prior goes to the Cauchy distribution. Thus varying r from 1 to ∞, we can generate a class of priors within the t-family.

Result 1. *Under a class of t-priors as above* the limiting local φ − divergence *is given as*

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi-\varphi(1)\right]=$$

$$=\frac{\varphi''(1)}{2}Var_{\pi 1\ \delta_0(\theta\ \chi)}\left\{\frac{1}{2\delta_0^2}\ln\left[1+\frac{(\theta-\mu)^2}{\sigma_\pi^2}\delta_0\right]\right.$$

$$\left.-\frac{1}{2}(1+1/\delta_0)\frac{\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}}{1+\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}\delta_0}\right\}$$

Special Case 1.1. When $\delta_0=0$, that is, the prior is from normal distribution. Then the limiting local φ -divergence reduces to

$$\lim_{\delta\to 0}\frac{1}{\delta^2}\left[D_\varphi-\varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2}Var_{\pi_N(\theta|x)}\left[\frac{1}{4}\left(\frac{\theta-\mu}{\sigma_\pi}\right)^4-\frac{1}{2}\left(\frac{\theta-\mu}{\sigma_\pi}\right)^2\right]$$

where $\pi_N(\theta|x)$ is the posterior distribution under the normal prior $N(\mu,\sigma_\pi^2)$.

Special Case 1.2. When $\delta_0=1$, that is, the prior is a Cauchy distribution. Then the limiting local φ -divergence becomes

$$\lim_{\delta\to 1}\frac{1}{(\delta-1)^2}\left[D_\varphi-\varphi(1)\right]=\frac{\varphi''(1)}{2}Var_{\pi_c(\theta|x)}$$

$$\left\{\frac{1}{2}\ln\left[1+\frac{(\theta-\mu)^2}{\sigma_\pi^2}\right]-\frac{\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}}{1+\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}}\right\}$$

,where $\pi_c(\theta|x)$ is the posterior distribution under the Cauchy prior $C(\mu,\sigma_\pi^2)$.

(2) Perturbation of the likelihood for fixed prior

Next, we fix the prior distribution $\pi(\theta)$, and perturb the likelihood function within the class of t-distributions of the form

$$f_r(x|\theta)=\frac{\Gamma(\frac{r+1}{2})}{(r\pi)^{\frac{1}{2}}\Gamma(\frac{r}{2})}\left[1+\frac{(\theta-\mu)^2}{r\sigma^2}\right]^{-\frac{r+1}{2}}.$$

$$r \geq 1$$

where σ^2 is known.

In this case, it is well known that when the degree of freedom r goes to infinity, the likelihood goes to the Normal distribution; and when the degree of freedom r goes to 1, the likelihood goes to the Cauchy distribution. Thus varying r from 1 to ∞, we can generate a class of likelihood functions.

Result 2. *Under a class of t-distributions as above* the limiting local φ -divergence *is given as*

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi-\varphi(1)\right]=$$

$$\frac{\varphi''(1)}{2}Var_{\pi 1\ \delta_0(\theta\ \chi)}\left\{\frac{1}{2\delta_0^2}\ln\left[1+\frac{(\theta-x)^2}{\sigma_\pi^2}\delta_0\right]\right.$$

$$\left.-\frac{1}{2}(1+1/\delta_0)\frac{\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}}{1+\dfrac{(\theta-\mu)^2}{\sigma_\pi^2}\delta_0}\right\}$$

Special Case 1.3. When $\delta_0=0$, that is, for the normal likelihood, the limiting local φ -divergence reduces to

$$\lim_{\delta\to 0}\frac{1}{\delta^2}\left[D_\varphi-\varphi(1)\right]=$$

$$\frac{\varphi''(1)}{2}Var_{\pi_N(\theta|x)}\left[\frac{1}{4}\left(\frac{\theta-x}{\sigma}\right)^4-\frac{1}{2}\left(\frac{\theta-x}{\sigma}\right)\right].$$

where $\pi_N(\theta|x)$ is the posterior distribution under the normal likelihood $N(\theta,\sigma^2)$.

Special Case1.4. When $\delta_0 = 1$, that is, the likelihood is from a Cauchy distribution. Then the limiting local φ-divergence becomes

$$\lim_{\delta \to 1} \frac{1}{(\delta-1)^2}\left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2} Var_{\pi_C(\theta|x)} \left\{ \frac{1}{2} \ln\left[1 + \frac{(\theta-x)^2}{\sigma^2}\right] - \frac{\frac{(\theta-x)^2}{\sigma^2}}{1 + \frac{(\theta-x)^2}{\sigma^2}} \right\}.$$

where $\pi_C(\theta|x)$ is the posterior distribution under the Cauchy likelihood $C(\theta, \sigma^2)$.

(3) An Illustrative Example

In Berger(1985), he considered $X \sim N(0,1)$, and subjectively specified two reasonable priors $C(0,1)$ and N (0,2.19). In order to measure the inherent robustness of the Cauchy prior, he used posterior means, posterior probabilities of certain sets, Bayes risks criteria, and marginal densities. Here we provide our alternative approaches based on limiting local divergence.

In our example, we divide the problem into two parts. First, we assume that the likelihood is fixed and the prior varies within a class of t-priors. Next, we assume that the prior is fixed and the likelihood varies within a class of t-distributions.

In this part, we observe the likelihood $X \sim N(\theta, 1)$, and subjectively specify a prior median of 0 and quartiles of ± 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 N (0,2.19804). Table 1 presents values of the limiting local φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different priors using the method of numerical integration.

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., $x \le 2$), 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., $x \ge 4.5$), 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 prior. Note the dependence of robustness on the actual value of x.

For the next part, we suppose that there are four plausible candidates of likelihood from the class of t-distributions which are $N(\theta,1)$, $T_{10}(\theta,1)$, $T_5(\theta,1)$ and $C(\theta,1)$. It is assumed for simplicity that the prior is Normal $N(0,1)$. Table 2 presents 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.

The interpretation of the above table is same as in the previous case.

2. Local sensitivity measures under a class of Gamma distributions :

(1) Perturbation of the Prior for fixed likelihood

Here, we fix the likelihood function $f(x|\theta)$ 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

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

where the scale parameter β 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 β. Thus varying r from 1 to ∞, we can generate a class of priors.

Result 3. *Under a class of Gamma priors as above,* the limiting local φ-divergence *is given as*

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi-\varphi(1)\right]=$$

$$\frac{\varphi''(1)}{2}\frac{1}{\delta_o^4}Var_{\pi_1\ \delta_0(\theta\cdot x)}\left[ln\,\theta\right].$$

Special Case 2.1. When δ_0=1, that is, the prior has Exponential distribution, the limiting local φ-divergence becomes

$$\lim_{\delta\to 1}\frac{1}{(\delta-1)^2}\left[D_\varphi-\varphi(1)\right]=\frac{\varphi''(1)}{2}Var_{\pi E(\theta\ x)}\left[\ln\theta\right]$$

where $\pi_E(\theta|x)$ is the posterior distribution under

the Exponential prior $Exp(\beta)$.

Remark: Similar results are obtained, if we fix the likelihood function $f(x|\theta)$, and perturb the prior within the class of Gamma distributions with varying scale parameter r which has the form

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

where the shape parameter α is known.

(2) Perturbation of the likelihood for fixed prior

Next, we fix the prior distribution $\pi(\theta)$, and perturb the likelihood function within the class of Gamma distributions of the form

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

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

(3) An Illustrative Example

In this example, we again divide the problem into two parts in order to see the effect of the data on the scale parameter. First, we assume that the likelihood is fixed and the prior varies within a class of Gamma-priors. Next, we assume that the prior is fixed and the likelihood varies within a class of Gamma-distributions.

In this part, we observe the likelihood $X\sim Exp(\theta)$, and subjectively specify a prior 95 percentile of 2.9957. Now, we have four reasonable priors to be considered within the class of Gamma-priors with matching 95th percentile, which are Exponential prior $Exp(1)$, three Gamma-priors $Gamma$ (5,3.05555), $Gamma$ (10,5.24258), and $Gamma$ (20, 9.30624). Table 3 presents values of the limiting local φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different priors using the method of numerical integration.

For each x, the values of the limiting local divergence measures are increasing rapidly, when the shape parameters go up. For all x, it appears that these values do not change too much for each choice of prior. It seems that the actual x for each prior does not effect to much to the value of the limiting local divergence. In other word, the scale parameter does not effect too much for the observed value x. However for each x, there can be a substantial difference among those values, indicating that the answer is then not robust to reasonable variation in the prior. Note again the dependence of robustness on the actual value of x.

In the other part, we consider four plausible candidates of likelihood from the class of Gamma-distributions which are $Exp(\theta)$, $Gamma$ (1.5,θ), $Gamma$ (2,θ) and $Gamma$ (2.5,θ). Let us assume that the prior for θ is Exponential $Exp(1)$.Table 5 presents values of the limiting local

φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different models using the method of numerical integration.

The interpretation of the above table is the same as in the previous case.

In addition, the values of the limiting local φ-divergence for each likelihood are much smaller, when the values of the shape parameters go down.

3. Local sensitivity measures under a class of Weibull distributions :

(1) Perturbation of the Prior for fixed likelihood

Here, 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.

Result 4. *Under a class of* Weibull priors as above, the limiting local φ-divergence *is given as*

$$\lim_{\delta \to \delta_0} \frac{1}{(\delta - \delta_0)^2}\left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2}\frac{1}{\delta_o^4} Var_{\pi_{1/\delta_0}(\theta|x)}\left[(1 - b\theta^{1/\delta_0})ln\theta\right].$$

Special Case 3.1 When $\delta_o = 1$, that is, the prior is an Exponential distribution, the limiting local φ-divergence becomes

$$\lim_{\delta \to 1} \frac{1}{(\delta - 1)^2}\left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2} Var_{\pi_E(\theta|x)}\left[(1 - b\theta)ln\theta\right]$$

where $\pi_E(\theta|y)$ is the posterior distribution under the Exponential prior $Exp(b)$.

Remark: Similar results are obtained, if we fix the likelihood function $f(x|\theta)$, and perturb the prior distribution within the class of Weibull distributions with varying scale parameter r which has the form

$$\pi_r(\theta) = ar\theta^{a-1}e^{-r\theta^a}, r \geq 1, a > 0.$$

where the shape parameter a is known.

In this situation, the values of limiting local divergence at an observed data x do not always increase , when the scale parameter r increases. This is due to the fact that the shape parameter a goes down, when the scale parameter r increases, while preserving the fixed percentile.

(2) Perturbation of the likelihood for fixed prior

Next, we fix the prior distribution $\pi(\theta)$, and perturb the likelihood function within the class of Weibull distributions of the form

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

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

Result 5. *Under a class of* Weibull distributions as above, the limiting local φ-divergence *is given as*

$$\lim_{\delta \to \delta_0} \frac{1}{(\delta - \delta_0)^2}\left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2}\frac{1}{\delta_o^4} Var_{\pi 1/\delta_0(\theta|x)}\left[x^\theta\right]$$

(3) An Illustrative Example

In this example, we again consider two parts. First, we assume that the likelihood is fixed and the prior varies within a class of Weibull distributions. Next, we assume that the prior is fixed and the likelihood varies within a class of Weibull-distributions.

In this part, suppose we consider the likelihood $X \sim Weibull(\theta,1)$, and subjectively specify a prior with 95th percentile as 2.9957. Now, suppose we consider four reasonable priors with matching 95th percentile within the class of Weibull distributions, which are Exponential prior $Exp(1)$, three Weibull-priors $Weibull$ (5,0.012418), $Weibull$ (10,0.000051482) and $Weibull$ $(20,10^{-9})$. Table 5 presents values of the limiting local φ-divergence measures without the constant $\varphi''(1)/2$ for various x under different priors using the method of numerical integration.

For each x, the values of the limiting local divergence measures are increasing rapidly, when the shape parameters go up. For small x (i.e., $x \leq 2$), there can be a substantial difference among those values, indicating that the answer is then not robust to reasonable variation in the prior. For large x (i.e., $x \geq 3$), it appears that those values are small , indicating some degree of robustness with respect to the choice of the prior. For all x except 1, it appears that the values of the limiting local divergence are not varied too much, when the posterior distribution comes from Exponential prior. Note again the dependence of robustness on the actual value of x.

Next, we suppose that there are four plausible candidates of likelihood from the class of Weibull-distributions which are $Weibull(\theta,1)$ $Weibull(\theta,1.5)$ and $Weibull(\theta,2.5)$. It is assumed that the prior is Exponential $Exp(1)$. Table 6 presents values of the limiting local φ-divergence measure without the constant $\varphi''(1)/2$ for various x under different models using the method of numerical integration.

The interpretation of the above table is the same as in the previous case.

In addition, the values of the limiting local divergence for each likelihood are much smaller, when the values of the scale parameters go down.

4. Local sensitivity measures under weighted distribution families:

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, $f_r(x|\theta)$ is the unweighted distributions which could be t-distribution, Gamma distribution, and Weibull distribution before respectively.

Result 6. *Under a class of* t-distributions, the limiting local φ-divergence between two posterior weighted distributions *is given as*

$$\lim_{\delta \to \delta_0} \frac{1}{(\delta - \delta_0)^2} \left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2} \left\{ Var_{\pi_{1}^{w}\,\delta_0(\theta|x)} \left[l_{x|\theta}(\delta_0)\right] + Var_{\pi_{1}^{w}\,\delta_0(\theta|x)} \left[\frac{E_{f_1\,\delta_0}\left[w(x) l_{x|\theta}(\delta_0)\right]}{E_{f_1\,\delta_0}[w(x)]}\right] \right\} -$$

$$\varphi''(1) Cov_{\pi_{1}^{w}\,\delta_0(\theta|x)} \left(l_{x|\theta}(\delta_0), \frac{E_{f_1\,\delta_0}\left[w(x) l_{x|\theta}(\delta_0)\right]}{E_{f_1\,\delta_0}[w(x)]} \right)$$

where $\delta_0 = 1/r_0$, r_0 is known. $\pi^w(\theta|x)$ is posterior weighted distribution, and

$$l_{x|\theta}(\delta_0) = \frac{1}{2\delta_0^2} ln\left[1 + \frac{(\theta - x)^2}{\sigma^2}\delta_0\right] - \frac{1}{2}\left(1 + \frac{1}{\delta_0}\right) \frac{\frac{(\theta - x)^2}{\sigma^2}}{1 + \frac{(\theta - x)^2}{\sigma^2}\delta_0}$$

Result 7. *Under a class of Gamma distributions* the limiting local φ-divergence between two posterior weighted distributions *is given as*

$$\lim_{\delta \to \delta_0} \frac{1}{(\delta - \delta_0)^2} \left[D_\varphi - \varphi(1)\right] =$$

$$\frac{\varphi''(1)}{2} \cdot \frac{1}{\delta_0^4} \cdot Var_{\pi_{1}^{w},\delta_0(\theta|x)} \left[\frac{E_{f_1|\delta_0}\left[w(x) \ln x\right]}{E_{f_1\,\delta_0}[w(x)]}\right]$$

Result 8. *Under a class of Weibull distributions* the limiting local φ-divergence between two posterior weighted distributions *is given as*

$$\lim_{\delta\to\delta_0}\frac{1}{(\delta-\delta_0)^2}\left[D_\varphi-\varphi(1)\right]=$$

$$\frac{\varphi''(1)}{2}\cdot\frac{1}{\delta_0^4}\cdot Var_{\pi^\omega_{/y_{\delta_0}}}\left(\theta|x\right)\left[x^\theta-\frac{E_{f/y_{\delta_0}}\left[x^\theta\omega(x)\right]}{E_{f/y_{\delta_0}}\left[\omega(x)\right]}\right].$$

五、計畫成果自評

From the previous section, it follows that there is a duality relationship between perturbation of the likelihood and the prior. Similar phenomenon was also observed in Pericchi and Smith (1992).

Suppose we observe the likelihood function with scale parameter and suppose there are several reasonable priors within different classes of distributions. then it is uncomparable in different classes of distributions using the limiting local divergence measures. Similar results hold for the likelihood with shape parameter.

In this paper we have considered the local robustness for the prior or the likelihood using φ-divergence measures under t-family. Gamma family and Weibull family. Similar studies can be performed to other classes of perturbations. Several other measures can be introduced to explore the sensitivity analysis.

In this research. we have reached the goals in my proposal.

After the report is completed, the empirical results will be submitted to the related conferences and journals.

References

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

(2) Basu, S., Jammalamadaka, S.R. and Liu (1993)."Local posterior robustness with parametric priors: maximum and average sensitivity." Technical Report 239, Department of Statistics and Applied Probability, University of California, Santa Barbara.

(2) Bayarri, M.J. and Berger, J.O. (1993). "Robust Bayesian analysis of selection models." Technical Report No. 93-6, Department of Statistics, Purdue University.

(3) 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.

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

(5) Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis* (2nd Edition). Springer-Verlag, N.Y.

(6) Berger, J.O. (1986). Comment on: "On the consistency of Bayes estimates" by Diaconis, P. and Freedman, D. *Ann. Statist.*, 14, 30-37.

(7) Berger, J.O. (1990). "Robust Bayesian analysis: sensitivity to the prior." *J. Statist. Plan. Inference*, 25, 303-328.

(8) Choy, S.T.B. (1996). "Robust Bayesian analysis using scale mixture of normals distributions."Ph.D. dissertation, Department of Mathematics, Imperial College London.

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

(10) Delampady, M. and Dey, D.K. (1994). "Bayesian robustness for multiparameter problems." *J. Statist. Planning and Inf.*, 40, 375-382

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

(12) Dey, D.K., Ghosh, S.K. and Lou, K.R. (1996). "On local sensitivity measures in Bayesian analysis.", *IMS monograph series*, vol.29, 21-39.

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

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

(15) Gustafson, P. and Wasserman, L. (1993). "Local sensitivity diagnostics for Bayesian inference." Technical Report 574, Department of Statistics, Carnegie Mellon University.

(16)Larose. D. and Dey, D.K. (1996). "Weighted distributions viewed in the context of model selection: a Bayesian perspective." *Test.* To appear

(17)Patil. G.P. (1981). "Studies in statistical ecology involving weighted distributions." *Proceedings of the ISI Golden Jubilee International Conference on Statistics: Applications and New Directions*}.Calcutta. 478-503.

(18)Patil. G.P. and Rao, C.R. (1977). "Weighted distributions and a survey of their applications." In *Applications of Statistics*, P.R. Krishnaiah(ed.). Amsterdam: North Holland Publishing Co., 383-405.

(19)Pericchi. L.R. and Smith. A.F.M. (1992). "Exact and approximate posterior moments for a normal local parameter." *Joural of the Royal Staistical Society*}. Ser. B. 54. 793-804.

(20)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.

(21)Reid. T.R.C. and Cressie. N.A.C. (1988). "Goodness of fit statistics for discrete multivariate data." Springer Verlag.

(22)Ruggeri. F. and Wasserman. L. (1993). "Infinitesimal sensitivity of posterior distributions." *Canad. J. Statist.*. Vol.21. No. 2. 195-203.

(23)Sivaganesan. S. (1993). "Robust Bayesian diagnostics." *J. Statist. Plan. Inf.*. 35. 171-188.

(24)Srinivasan. C. and Truszcynska. H. (1990). "On the ranges of posterior quantities." Technial Report 294. Department of Statistics. University of Kentucky.

Table 1: Local sensitivity under different priors of t-distributions

	Normal prior	**T_{10} prior**	**T_5 prior**	**Cauchy prior**
x = 0	0.0145276	0.0100620	0.00788721	0.00320824
x = 1	0.0719450	0.0380081	0.0238755	0.00464302
x = 2	0.652840	0.302226	0.167166	0.0155271
x = 3	4.10058	1.67824	0.827474	0.0425573
x = 4	18.0456	6.01343	2.42554	0.0526621
x = 4.5	34.1138	9.86071	3.47360	0.0478026
x = 5	60.9925	14.8642	4.49378	0.0410030
x = 6	170.142	27.0991	5.91405	0.0291381
x = 10	3292.86	50.0911	4.74467	0.0102072
x = 15	36307.0	34.2500	2.48926	0.00450993
x = 20	201606	21.6289	1.48682	0.00255203

Table 2: Local sensitivity under different likelihoods of t-distributions

	$N(\theta, 1)$	**$T_{10}(\theta, 1)$**	**$T_5(\theta,1)$**	**$C(\theta, 1)$**
X = 0	0.125000	0.0512764	0.0259029	0.00320823
X = 1	0.546869	0.216579	0.101257	0.00464300
X = 2	3.49995	1.39501	0.628775	0.0155271
X = 3	16.8595	6.58580	2.68657	0.0425572
X = 4	63.1247	22.1104	7.04881	0.0526615
X = 4.5	112.973	35.5024	9.40704	0.0478026
X = 5	193.484	52.0776	11.1553	0.0410036
X = 6	507.494	85.9700	12.0078	0.0291370
X = 10	8859.34	84.8572	5.98584	0.0102072
X = 15	94208.2	42.3906	2.75952	0.00451469
X = 20	516375	24.4609	1.57178	0.00254250

Table 3: Local sensitivity under different priors of Gamma distributions

	Exp.(1) prior	**G(5, •) prior**	**G(10, •) prior**	**G(20, •) prior**
x = 0	0.645134	113.343	952.073	7808.17
x = 1	0.644934	113.320	951.562	7811.05
x = 2	0.644830	113.333	951.848	7810.29
x = 3	0.644653	113.310	951.591	7804.62
x = 4	0.644625	113.327	951.595	7806.33
x = 5	0.644916	113.327	951.632	7803.54
x = 6	0.644868	113.326	951.612	7805.18
x = 10	0.644444	113.334	951.645	7802.54
x = 15	0.643314	113.376	951.686	7802.75
x = 20	0.640590	113.254	951.842	7805.20

Table 4: Local sensitivity under different likelihoods of Gamma distributions

	Exp.(θ)	**Gamma(1.5, θ)**	**Gamma(2, θ)**	**Gamma(2.5, θ)**
x = 0	0.644964	2.48232	6.31704	12.9041
x = 1	0.646395	2.49814	6.39199	13.0634
x = 2	0.644894	2.48210	6.32136	12.9176
x = 3	0.643987	2.48081	6.35796	13.1533
x = 4	0.644566	2.47991	6.31531	12.9180
x = 5	0.644796	2.48251	6.31661	12.8950
x = 6	0.643051	2.48404	6.32135	12.9041
x = 10	0.643355	2.48318	6.30170	12.7904
x = 15	0.645365	2.49251	6.34140	12.9506
x = 20	0.645592	2.48556	6.35551	12.9777

Table 5: Local sensitivity under different priors of Weibull distributions

	Exp.(1) prior	**W(5, •) prior**	**W(10, •) prior**	**W(20, •) prior**
x = 0.1	0.344941	121.036	4678.61	124656
x = 0.5	0.945286	543.628	9937.48	172214
x = 1	4.99187	814.287	12159.1	189754
x = 2	0.219165	69.3167	2413.04	81851.8
x = 3	0.290409	33.0278	261.355	1920.46
x = 4	0.346536	42.2777	323.555	2365.03
x = 5	0.380084	44.5102	328.808	2357.43
x = 6	0.400533	45.2019	329.684	2354.32
x = 10	0.436962	45.6615	330.131	2351.50
x = 15	0.452139	45.7056	329.846	2344.67
x = 20	0.459299	45.7114	330.338	2347.03

Table 6: Local sensitivity under different likelihoods of Weibull distributions.

	Weibull(θ,1)	**Weibull(θ,1.5)**	**Weibull(θ,2)**	**Weibull(θ,2.5)**
x = 0.1	0.0466785	0.216641	0.622116	1.37578
x = 0.5	0.0478122	0.235932	0.720015	1.68323
x = 1.5	0.479050	1.35632	2.79434	4.92591
x = 2	0.78413	2.03600	3.99965	6.70504
x = 3	1.01392	2.52011	4.80441	7.88651
x = 4	1.11075	2.71492	5.12518	8.38768
x = 5	1.16384	2.82381	5.30244	8.64919
x = 6	1.19841	2.89476	5.43117	8.81746
x = 10	1.26848	3.04228	5.66566	9.15535
x = 15	1.30651	3.12050	5.79160	9.33692
x = 20	1.32860	3.16354	5.86118	9.45449