Rendiconti
Accademia Nazionale delle Scienze detta dei XL
Memorie di Matematica e Applicazioni
118° (2000), Vol. XXIV, fasc. 1, pagg. 27-41

MONICA MUSIO (*)

Measuring Local Sensitivity in Bayesian Analysis: A General Setting (**)

ABSTRACT. — In a Bayesian framework, the current indexes used to study the local sensitivity of a posterior expectation, with the respect to the choice of the prior in a given class, are either restricted to locally-convex classes or in the parametric case, depend on the parametrization. We will define a general sensitivity index for posterior expectations which avoids these restrictions. In the case of locally-convex classes, we show that this index coincides with the indexex currently used in literature such as the norm of the Fréchet derivative, or the norm of the Gateaux derivative. We will develope the parametric case in detail.

Misurazione della sensitività locale nell'inferenza bayesiana: un'impostazione generale

SUNTO. — Nel contesto dell'inferenza statistica bayesiana, gli indici usualmente impiegati per misurare la cosiddetta «sensitività locale» al variare della scelta di una legge *a priori* in un'assegnata classe, sono definiti solo per classi localmente convesse e, nel caso parametrico, dipendono dalla particolare parametrizzazione. Nella presente Nota si definisce un indice generale, esente da questi difetti. Dopo aver mostrato che, nel caso di classi localmente convesse, esso coincide con gli indici abituali (quali la norma della derivata di Fréchet o quella della derivata di Gateaux), si esamina più da vicino il caso parametrico.

In a Bayesian analysis, inference depends on the prior distribution over the parameter space. Then the problem of quantifying the degree whose posterior quantities are influenced by changes in the prior, is of much concern. In global robustness the goal is to find the range of posterior quantities as the prior varies in a «reasonable» class Γ. If the range is small, the analysis is robust to misspecification of the prior. For a review and extended references on the subject see Berger (1994) [2]. A limitation of

(*) Indirizzo dell'Autrice: Université de Haute Alsace, Laboratoire de Mathématiques 4, rue des Frères Lumière - F-68093 Mulhouse Cédex (E-mail: M.Musio@univ-Mulhouse.fr)
(**) Memoria presentata il 30 novembre 1999 da Giorgio Letta, uno dei XL.
1991 *Mathematics Subject Classification*: 62 F 15, 62 F 35

the global approach is that computation is often non trivial. For this reason, in the last few years, there has been a growing interest in the local approach. The local sensitivity explores the effects on the posterior of infinitesimal perturbation around some elicited prior (see Basu, Jammalamadaka and Liu (1996) [1], Cuevas and Sanz (1988) [4], Diaconis and Freedman (1986) [6], Gustafson and Wasserman (1995) [9], Gustafson, Srinivasan and Wasserman (1995) [10], Ruggeri and Wasserman (1993) [13], Gustafson (1996) [7], Gustafson (1996) [8]). Recent papers have also dealt with sensitivity to simultaneus perturbations of priors and likelihood (see Clarke and Gustafson (1998) [3], Dey, Ghosh and Lou (1996) [5]). The idea of local sensitivity is to study the rate at which the posterior changes relative to the prior. In this sense, finding the direction in which the sensitivity is the largest, might give some important information about the elicitation process. The current indexes used to study local sensitivity in a convex class of priors are based on the notions of Frechét derivative and Gateaux derivative. In the literature, only a few cases concerning local sensitivity measures for a non-locally convex class of priors are presented. One possible approach is attributed to Basu, Jammalamadaka and Liu (1996) [1], but it is restricted to the case of parametric classes. However, as we shall show later, the index they have defined depends on the chosen parametrisation for defining the class.

The aim of this paper is to propose a general definition of local sensitivity for posterior expectations with the following properties. Firstly, for a convex class, it should coincide with the indexes currently used. Secondly it has to be an *intrinsic* characteristic of the class, i.e. the definition depends only on Γ and not on the particular way the class has been defined.

The structure of the paper is as follows: in the first section we shall focus on general definitions and notations which will be used through out the present work; in section two we shall propose a definition of local sensitivity and we shall prove that it coincides with the definitions currently used in the case of a locally convex class of priors; in section three, we shall analyse the case of a parametric class in detail.

1. - Notations and definitions

Let $X = (X_1, \ldots, X_n)$ be i.i.d. random variables obtained from a distribution with density $p(x|\theta)$ where $\theta \in \Theta$ is an unknown parameter. Let $\sigma(\Theta)$ be a σ-algebra on the parameter space Θ, $\mathcal{P}$ the set of all probability measures Π on $\sigma(\Theta)$ and Γ a class of probability measures on $(\Theta, \sigma(\Theta))$. For a prior distribution $\Pi \in \Gamma$, let Ψ be the posterior expectation of a given non-constant measurable function $f: \Theta \to \boldsymbol{R}$, indeed

$$\Psi(\Pi) = \frac{N(\Pi)}{D(\Pi)}$$

where $N(\Pi) = \int_\Theta f(\theta)\, l(\theta)\, \Pi(d\theta)$ and $D(\Pi) = \int_\Theta l(\theta)\, \Pi(d\theta)$ are supposed to be well defined. The likelihood function $l(\theta) = \prod_{i=1}^{n} p(x_i \,|\, \theta)$ is supposed to be positive for each $\theta \in \Theta$.

Local sensitivity measures which are commonly used in literature are based on the notion of functional derivatives. One criterion for measuring the local sensitivity around a fixed Π_0 uses the norm of the Fréchet derivative of Ψ with respect to Π evaluated at Π_0. Recall that $\Psi : \Gamma \to \boldsymbol{R}$ is *Fréchet differentiable* at Π_0, if there exists a linear bounded map $\dot{T}_{\Pi_0}$, defined on the vector space $\mathcal{M}$ of all signed measures, such that

$$\Psi(\Pi_0 + h) - \Psi(\Pi_0) = \dot{T}_{\Pi_0}(h) + o(\|h\|).$$

The index $\dot{T}_{\Pi_0}$ is called the *Fréchet derivative* of Ψ at Π_0. Define the local sensitivity of Ψ in $\Pi_0 \in \Gamma$ as

$$\|\dot{T}_{\Pi_0}\| = \sup_{h = \Pi - \Pi_0,\, \Pi \in \Gamma} \frac{|\dot{T}_{\Pi_0}(h)|}{\|h\|}, \tag{1}$$

where $\|\cdot\|$ is the total variation norm given by

$$\|h\| = \sup_{A \in \sigma(\Theta)} |h(A)| .$$

Another approach is to use the *norm* of the Gateaux derivative $D\Psi(\Pi_0, \Pi - \Pi_0)$ of Ψ at Π_0, defined by

$$G(\Pi_0) = \sup_{\Pi \in \Gamma} |D\Psi(\Pi_0, \Pi - \Pi_0)| .$$

Recall that the function Ψ is *Gateaux differentiable* at Π_0, in the direction $\Pi - \Pi_0$, if there exists a linear functional $D\Psi(\Pi_0, \Pi - \Pi_0)$, such that

$$D\Psi(\Pi_0, \Pi - \Pi_0) = \lim_{t \to 0} \frac{\Psi(\Pi_0 + t(\Pi - \Pi_0)) - \Psi(\Pi_0)}{t} .$$

The Gateaux derivative exists quite generally and the two notions coincide for bounded likelihood functions.

It is important to remark that the two notions above are well defined only in the case of the class of prior distributions being locally convex. The non-convex case has been explored only in the context of parametric classes. Indicate a parametric class of priors by $\Gamma_\Lambda = \{\Pi_\lambda : \lambda \in \Lambda\}$ and assume that $\Lambda \subset \boldsymbol{R}^m$. Basu, Jammalamadaka and Liu (1996) [1] have defined the local sensitivity of the function

$$\Psi(\lambda) = \frac{N(\Pi_\lambda)}{D(\Pi_\lambda)}$$

at 0 in terms of the norm of the total derivative of this function at 0, i.e.,

$$\|T_{\Psi_0}\| = \sup_{h \in \boldsymbol{R}^m} \frac{|T_{\Psi_0}(h)|}{\|h\|} .$$

It should be recalled that the function

$$\Psi : \Lambda \subset \boldsymbol{R}^m \to \boldsymbol{R}$$

is called total differentiable at $0 \in \Lambda$, if there exists a linear function T_{Ψ_0}, such that

$$\frac{\|\Psi(h) - \Psi(0) - T_{\Psi_0}(h)\|}{\|h\|} \to 0$$

as $\|h\| \to 0$ (where $\|h\|$ is the standard euclidian norm on $\boldsymbol{R}^m$). Basu, Jammalamadaka and Liu (1996) [1] have shown that

$$\|T_{\Psi_0}\| = \sqrt{\sum_{i=1}^{m} \left[\frac{\partial \Psi(0)}{\quad} \partial \lambda_i \right]^2} . \tag{2}$$

Remark 1.1: *Note that the index $\|T_{\Psi_0}\|$ not only depends on the class of priors, but also on the parametrisation defining the class.*

From these considerations it emerges that there may exist a way to define the local sensitivity for a general class of priors intrinsically, extending the notions commonly used in the context of locally convex classes.

2. - A general definition of local sensitivity

In order to define the local sensitivity of the function Ψ at Π_0 one is led to examine the behaviour of the quantity

$$\Psi(\Pi) - \Psi(\Pi_0)$$

in terms of certain measures of the 'difference' between Π and Π_0 as $\Pi \to \Pi_0$. In general there is not a natural way to evaluate this «difference», nor to take a limit. For these reasons we choose a particular way to let Π tend to Π_0 and to express the difference between Π and Π_0.

We will see later that in certain special cases these choices are natural and have a suitable statistical meaning.

We introduce a set of pairs

$$C_{\Pi_0} = ((\Pi(t), \delta(t)))$$

defined in the interval $I = [0, \alpha]$ $\alpha > 0$, where:

1) $\Pi(\cdot)$ is a curve in Γ such that $\Pi(0) = \Pi_0$ and $\Pi(t) \neq \Pi_0$ for $t \neq 0$;

2) $\delta : I \to R$ is a continuous function tending to 0 as $t \to 0$.

The latter function is a manner to quantify the «difference» between $\Pi(t)$ and Π_0. We suppose that $\lim_{t\to 0} (\delta(t)/t) = l$ where $l \neq 0$, so that, in an intuitive way, t and $\delta(t)$ have the same order of «magnitude».

Whenever the corresponding limits exist, one defines *the local sensitivity of Ψ at Π_0 along the curve $\Pi(\cdot)$ relative to* $\delta(\cdot)$ by

$$L(\Pi(t), \delta(t)) = \lim_{t \to 0} \frac{\Psi(\Pi(t)) - \Psi(\Pi_0)}{\delta(t)} . \tag{3}$$

and *the local sensitivity of Ψ at Π_0 relative to C_{Π_0}* by

$$S(C_{\Pi_0}, \Gamma) = \sup_{(\Pi(t),\, \delta(t)) \in C_{\Pi_0}} |L(\Pi(t), \delta(t))| . \tag{4}$$

The index $S(C_{\Pi_0}, \Gamma)$ reflects the maximum possible change at Π_0 of the posterior expectation Ψ for a given «difference measure« $\delta(t)$.

We shall examine now different examples depending on the nature of the class Γ. We will use subscripts (G, GN, ect.) to identify $L(\Pi(\cdot), \delta(\cdot))$ and $S(C_{\Pi_0}, \Gamma)$ in each class.

2.1. *Γ convex class.*

Given a probability measure $\Pi_0 \in \Gamma$, for each $\Pi \in \Gamma$, the segment between Π and Π_0 is completely contained in the class Γ. A natural path to go from Π_0 to Π is then given by the line with parametrization

$$\Pi(t) = \Pi_0 + t(\Pi - \Pi_0) = (1-t)\, \Pi_0 + t\Pi$$

for $t \in [0, 1]$ and an appropriate difference measure is $\delta(t) = t$. Thus

$$C_{\Pi_0} = ((\Pi(t), \delta(t)))$$

where

$$\Pi(t) = (1-t)\, \Pi_0 + t\Pi$$

for $t \in [0, 1]$ and $\delta(t) = t$. One may now relate $L_G(\Pi(t), t)$ and $S_G(C_{\Pi_0}, \Gamma)$ to the Gateaux derivative. One has

$$L_G(\Pi(t), t) = \lim_{t \to 0} \frac{\Psi(\Pi(t)) - \Psi(\Pi_0)}{t} = \lim_{t \to 0} \frac{1}{t} \left[\frac{N(\Pi_0 + t(\Pi - \Pi_0))}{D(\Pi_0 + t(\Pi - \Pi_0))} - \frac{N(\Pi_0)}{D(\Pi_0)} \right]$$

which is a linear function of Π, in fact

$$L_G(\Pi(t), t) = \frac{N(\Pi)\, D(\Pi_0) - D(\Pi)\, N(\Pi_0)}{D(\Pi_0)^2}\,. \tag{5}$$

One can easily recognize that $L_G(\Pi(t), t)$ is the Gateaux derivative of the function Ψ at Π_0 (see Diaconis and Freedman [6]). Concerning $S_G(C_{\Pi_0}, \Gamma)$, we have

$$S_G(C_{\Pi_0}, \Gamma) = S_G(\Pi_0, \Gamma) = \sup_{\Pi \in \Gamma} \left| \frac{N(\Pi)\, D(\Pi_0) - D(\Pi)\, N(\Pi_0)}{D(\Pi_0)^2} \right| \tag{6}$$

which is the norm of the Gateaux derivative.

2.2. *Γ convex class with a given norm*

Let Γ be a set of probability measures contained in a normed vector space $\boldsymbol{V}$, where $\boldsymbol{V}$ is a subspace of the vector space $\mathcal{M}$ of all signed measures over Θ. In this case we may choose $C_{\Pi_0} = (\Pi(t), \delta(t))$, where

$$\Pi(t) = \Pi_0 + t(\Pi - \Pi_0) \qquad \text{for } t \in [0, 1]$$

and

$$\delta(t) = \|\Pi(t) - \Pi_0\| = \|t(\Pi - \Pi_0)\|\,.$$

We have

$$L_{GN}(\Pi(t), \delta(t)) = \frac{N(\Pi)\, D(\Pi_0) - D(\Pi)\, N(\Pi_0)}{\|\Pi - \Pi_0\| D(\Pi_0)^2} \tag{7}$$

and

$$S_{GN}(C_{\Pi_0}, \Gamma) = S_{GN}(\Pi_0, \Gamma) = \sup_{\Pi \in \Gamma} \left| \frac{N(\Pi)\, D(\Pi_0) - D(\Pi)\, N(\Pi_0)}{\|\Pi - \Pi_0\|\, D(\Pi_0)^2} \right|\,. \tag{8}$$

REMARK 2.1: *$L_{GN}(\Pi(t), \delta(t))$ depends only on the direction of Π. In fact, each point $\tilde{\Pi} = \Pi_0 + \mu(\Pi - \Pi_0)$ of the line through Π_0 and Π gives the same value for $L_{GN}(\Pi(t), \delta(t))$. For this reason, it can be considered as a more coherent index than $L_G(\Pi(t), t)$.*

It can be highlighted that, if the Fréchet derivative of the function Ψ exists at Π_0, we have $S_{GN}(\Pi_0, \Gamma) = \|\dot{T}_{\Pi_0}\|$ (see 1).

2.3. *Γ class with a given norm.*

Let Γ be a set of probability measures contained in a normed vector space $\boldsymbol{V}$ where $\boldsymbol{V}$ is a subspace of the vector space $\mathcal{M}$ of all signed measures over Θ. In this context, the notion of derivative $\dot{\Pi}(0)$ of the curve $\Pi(\cdot)$ for $t=0$ is well defined, in fact

$$\dot{\Pi}(0) = \lim_{t\to 0} \frac{\Pi(t) - \Pi_0}{t}$$

where $\Pi(0) = \Pi_0$, or equivalently

$$\Pi(t) - \Pi_0 = t\dot{\Pi}(0) + t\varepsilon(t) \qquad \text{where } \|\varepsilon(t)\| \to 0 \text{ as } t \to 0 .$$

We choose, as C_{Π_0}, the set of pairs

$$((\Pi(t), \delta(t)))$$

defined on the interval $[0, 1]$, where $\Pi(t)$ is differentiable at $t=0$ and

$$\delta(t) = \|\Pi(t) - \Pi_0\| .$$

The expression of $L_N(\Pi(t), \delta(t))$ is given by the following theorem:

THEOREM 2.1: *Let $\Pi(t)$ be a curve in Γ differentiable at 0 such that $\dot{\Pi}(0) \neq 0$. If $N(\cdot)$ and $D(\cdot)$ are continuous functionals of Π, then*

$$L_N(\Pi(t), \delta(t)) = \lim_{t\to 0} \frac{\Psi(\Pi(t)) - \Psi(\Pi_0)}{\|\Pi(t) - \Pi_0\|} = \tag{9}$$

$$= \frac{N(\dot{\Pi}(0))\, D(\Pi_0) - D(\dot{\Pi}(0))\, N(\Pi_0)}{\|\dot{\Pi}(0)\|\, D(\Pi_0)^2} .$$

PROOF: Write

$$\Psi(\Pi(t)) - \Psi(\Pi_0) = \frac{N(\Pi_0 + t\dot{\Pi}(0) + t\varepsilon(t))}{D(\Pi_0 + t\dot{\Pi}(0) + t\varepsilon(t))} - \frac{N(\Pi_0)}{D(\Pi_0)} .$$

The result follows from the continuity and linearity of $N(\cdot)$ and $D(\cdot)$. ■

REMARK 2.2: *$L_N(\Pi(t), \delta(t))$ depends only on $\dot{\Pi}(0)$, that is, two different differentiable curves of Γ starting at Π_0 with the same derivative $k = \dot{\Pi}(0) \neq 0$ give the same value for $L_N(\Pi(t), \delta(t))$. Then, in a given normed space, the value of $L_N(\Pi(t), \delta(t))$ does only depend on $\Pi(t)$ through $\dot{\Pi}(0)$.*

On behalf of this dependence, we may introduce the *set of tangent vectors at Π_0 in Γ*, defined in the following way

$$V(\Gamma)_{\Pi_0} = \{k \in V - \{0\} : \exists \Pi(\cdot) \in C_{\Pi_0}, \dot{\Pi}(0) = k\} . \tag{10}$$

For each $k \in V(\Gamma)_{\Pi_0}$, the expression

$$L_N(\Pi(t), \delta(t)) = \frac{N(k)\, D(\Pi_0) - D(k)\, N(\Pi_0)}{\|k\|\, D(\Pi_0)^2} = L_N(k)$$

gives the sensitivity on Π_0 for all the curves $\Pi(\cdot)$ such that $\dot{\Pi}(0) = k$.

Then one has

$$S_N(C_{\Pi_0}, \Gamma) = S_N(\Pi_0, \Gamma) = \sup_{k \in V(\Gamma)_{\Pi_0}} |L_N(k)| \; . \tag{11}$$

In the case where Γ is a convex class one introduces the line parametrized by

$$\Pi(t) = \Pi_0 + t(\Pi - \Pi_0)$$

where $\Pi \in \Gamma$, that verifies the condition $\dot{\Pi}(0) = \Pi - \Pi_0 \neq 0$. Thus one has

$$L_{GN}(\Pi(t), \delta(t)) = L_N(\Pi(t), \delta(t)) \; .$$

In the particular case of the curves

$$\Pi(t) = \Pi_0 + t(\Pi - \Pi_0)$$

we characterise the sensitivity at Π_0 using the set of *restricted tangent vectors at Π_0 in Γ* defined by

$$VR(\Gamma)_{\Pi_0} = (h \in V - \{0\} : \exists \tau > 0,\ \forall t \in [0, \tau],\ \Pi_0 + th \in \Gamma) \; . \tag{12}$$

Then

$$S_{GN}(\Pi_0, \Gamma) = \sup_{k \in VR(\Gamma)_{\Pi_0}} |L_{GN}(k)| \; . \tag{13}$$

Remark 2.3: *Let Π_0 be specified. Then $V(\Gamma)_{\Pi_0} \cup \{0\}$ and $VR(\Gamma)_{\Pi_0} \cup \{0\}$ are not, in general, vector spaces, for instance if Π_0 in in the boundary of Γ.*

We have

$$S_{GN}(\Pi_0, \Gamma) \leqslant S_N(\Pi_0, \Gamma)$$

because $S_{GN}(\Pi_0, \Gamma)$ requires a maximisation over $VR(\Gamma)_{\Pi_0}$, whereas $S_N(\Pi_0, \Gamma)$ requires a maximisation over $V(\Gamma)_{\Pi_0}$, and clearly $VR(\Gamma)_{\Pi_0} \subset V(\Gamma)_{\Pi_0}$. Moreover, as a consequence of the following two theorems, we will see that these two indexes are equal.

THEOREM 2.2: *If* $N(\cdot)$ *and* $D(\cdot)$ *are two continuous functionals of* Π, *and if* $VR(\Gamma)_{\Pi_0}$ *is dense in* $V(\Gamma)_{\Pi_0}$, *then*

$$S_{GN}(\Pi_0, \Gamma) = S_N(\Pi_0, \Gamma).$$

PROOF: $S_N(\Pi_0, \Gamma) = \sup_{\Pi \in V(\Gamma)_{\Pi_0}} L_N(\Pi(t), \delta(t))$. Then, for each $\varepsilon > 0$ there exists $h \in V(\Gamma)_{\Pi_0}$ such that

$$L_N(h) \in \left[S_N(\Pi_0, \Gamma) - \frac{\varepsilon}{2}, S_N(\Pi_0, \Gamma) \right].$$

The continuity of $L_N(\Pi(t), \delta(t))$ is a consequence of the continuity of the functionals $N(\cdot)$ and $D(\cdot)$. Thus, there exists $\eta > 0$ such that, for $a \in V(\Gamma)_{\Pi_0}$, $\|a - h\| < \eta$, we have

$$\|L_N(a) - L_N(h)\| < \frac{\varepsilon}{2}.$$

By the density of $VR(\Gamma)_{\Pi_0}$, it contains c such that $\|c - h\| < \eta$. Hence

$$\|L_N(c) - L_N(h)\| < \frac{\varepsilon}{2}$$

and

$$L_N(c) \geqslant S_N(\Pi_0, \Gamma) - \varepsilon.$$

Since

$$L_N(c) = L_{GN}(c)$$

one has

$$S_{GN}(\Pi_0, \Gamma) \geqslant S_N(\Pi_0) - \varepsilon$$

for each $\varepsilon > 0$, then

$$S_{GN}(\Pi_0, \Gamma) \geqslant S_N(\Pi_0, \Gamma).$$

Conversely, one clearly has

$$S_{GN}(\Pi_0, \Gamma) \leqslant S_N(\Pi_0, \Gamma). \quad \blacksquare$$

THEOREM 2.3: *Let* $\mathbf{V}$ *be a normed vector space,* Γ *a subset of* $\mathbf{V}$, $VR(\Gamma)_{\Pi_0}$ *and* $V(\Gamma)_{\Pi_0}$ *defined as above. Then, if* Γ *is convex, the set* $VR(\Gamma)_{\Pi_0}$ *is dense in* $V(\Gamma)_{\Pi_0}$.

PROOF: One has

$$\Pi(t) - \Pi(0) = t\dot{\Pi}(0) + t\varepsilon(t) \qquad \text{where } \|\varepsilon(t)\| \to 0 \text{ as } t \to 0 .$$

From the convexity it follows that

$$h = \frac{\Pi(t) - \Pi(0)}{t} = \dot{\Pi}(0) + \varepsilon(t)$$

is a vector of $VR(\Gamma)_{\Pi_0}$. Then, each element $k = \dot{\Pi}(0)$ of $V(\Gamma)_{\Pi_0}$ is the limit of a sequence of elements of $VR(\Gamma)_{\Pi_0}$, and the theorem is proved. ■

We now carry out the complete computation of the sensitivity index $S(C_{\Pi_0}, \Gamma)$ in the parametric case. We will take advantage of the geometric structure induced by the parametric class Γ on $V(\Gamma)_{\Pi_0}$, to solve the problem of finding the supremum of the posterior expectation over the space $V(\Gamma)_{\Pi_0}$.

3. - Γ PARAMETRIC CLASS

Let Γ be a parametric class, defined by the $\mathcal{C}^1$ map

$$\Phi : \Lambda \to \Gamma$$

such that

$$\lambda = (\lambda_1, \ldots, \lambda_m) \to \Pi = \Phi(\lambda_1, \ldots, \lambda_m) .$$

Here, Γ is a set of probability measures contained in a normed vector subspace $\boldsymbol{V}$ of the vector space $\mathcal{M}$ of all signed measures over Θ, and Λ is an open set of $\boldsymbol{R}^m$.

Suppose that $0 \in \Lambda$. Let $\Pi_0 = \Phi(0)$ and assume that the *rank* of Φ at 0 is equal to m, which means that the vectors of $\boldsymbol{V}$

$$\frac{\partial \Phi}{\partial \lambda_1}(0), \ldots, \frac{\partial \Phi}{\partial \lambda_m}(0)$$

are linear independent. It is a known result of differential geometry that the set $V(\Gamma)_{\pi_0} \cup \{0\}$ is a vector space generated by these vectors. Thus, each element $h \in V(\Gamma)_{\pi_0} \cup \{0\}$ can be written as follows

$$h = \sum_{i=1}^{m} a_i \frac{\partial \Phi}{\partial \lambda_i}(0)$$

where $(a_1, \ldots, a_m) \in \boldsymbol{R}^m$. Generally, the class Γ is non-convex. The local sensitivity at Π_0 defined in Section 2.3 is equal to

$$S_N(\Pi_0, \Gamma) = \sup_{k \in V(\Gamma)_{\Pi_0}} \frac{|D(\Pi_0)\, N(k) - N(\Pi_0)\, D(k)|}{\|k\| D(\Pi_0)^2} .$$

In order to compute explicity $S_N(\Pi_0, \Gamma)$, we have to find

$$\sup_{k \in V(\Gamma)_{\Pi_0}} \frac{|D(\Pi_0)\, N(k) - N(\Pi_0)\, D(k)|}{\|k\|\, D(\Pi_0)^2}$$

under the constraint

$$\left\| \sum_{i=1}^{m} \frac{\partial \Phi}{\partial \lambda_i}(0) \right\| = 1 .$$

Assume that the norm $\|\cdot\|$ is associated with an inner product $\langle\, , \rangle$. Let $\{I_i\}_{0 \leqslant i \leqslant m}$ be an orthonormal basis of $V(\Gamma)_{\Pi_0} \cup \{0\}$. Then, one can write the linear operator L such that

$$L(k) = D(\Pi_0)\, N(k) - N(\Pi_0)\, D(k)$$

as

$$L(k) = \langle k, B \rangle ,$$

and the coordinates of B in the basis $\{I_i\}_{0 \leqslant i \leqslant m}$ are given by

$$\langle I_i, B \rangle = L(I_i) = D(\Pi_0)\, N(I_i) - N(\Pi_0)\, D(I_i) .$$

In order to compute explicity $S_N(\Pi_0, \Gamma)$, we will use the following proposition:

PROPOSITION 3.1: *Let $\boldsymbol{A}$ be a finite dimensional Hilbert space and $B \in \boldsymbol{A}$. Then*

$$\sup_{\|k\| = 1,\, k \in \boldsymbol{A}} |\langle k, B \rangle| = \|B\| .$$

The hypotheses of the Proposition 3.1 are satisfied for $\boldsymbol{A} = V(\Gamma)_{\Pi_0} \cup \{0\}$ and we have

$$S_N(\Pi_0, \Gamma) = \frac{\|B\|}{D(\Pi_0)^2} = \frac{\sqrt{\sum_{i=1}^{m} L(I_i)^2}}{D(\Pi_0)^2}$$

or equivalently

$$S_N(\Pi_0, \Gamma) = \frac{\sqrt{\sum_{i=1}^{m} [D(\Pi_0)\, N(I_i) - N(\Pi_0)\, D(I_i)]^2}}{D(\Pi_0)^2} .$$

REMARK 3.1: *It follows directly from the intrinsic definition of $S_N(\Pi_0)$ that a change of parametrisation induces a change of the vectors I_i $i = 1, \ldots, m$ but not of the value of $S_N(\Pi_0, \Gamma)$.*

REMARK 3.2: *This expression of $S_N(\Pi_0, \Gamma)$ does not match, in general, the expression of the index $\|T_{\Psi_0}\|$ (see (1.3)). In fact, the former depends only on the class Γ and not on the parametrisation chosen to define the class. We find the same result only if the parametrisation is such that the vectors*

$$\left\{ \frac{\partial \Phi}{\partial \lambda_i}(0) \right\}_{1 \leqslant i \leqslant m}$$

are orthonormal. In fact, in this case we have

$$\|T_{\Psi_0}\| = \sqrt{\sum_{i=1}^{m} \left[\frac{\partial(\Psi \circ \Phi)(0)}{\partial \lambda_i} \right]^2}$$

and

$$\frac{\partial(\Psi \circ \Phi)(0)}{\partial \lambda_i} = \frac{\partial}{\partial \lambda_i} \left[\frac{N \circ \Phi(0)}{D \circ \Phi(0)} \right] =$$

$$= \frac{N(\partial \Phi / \partial \lambda_i)\, D(\Phi(0)) - D(\partial \Phi / \partial \lambda_i)\, N(\Phi(0))}{D(\Phi(0))^2} .$$

Thus we find

$$\|T_{\Psi_0}\| = \frac{\sqrt{\sum_{i=1}^{m} \left[N\left(\frac{\partial \Phi}{\partial \lambda_i} \right) D(\Pi_0) - D\left(\frac{\partial \Phi}{\partial \lambda_i} \right) N(\Pi_0) \right]^2}}{D(\Pi_0)^2} = S_N(\Pi_0, \Gamma) .$$

Suppose, for instance, that $\Theta \subset \boldsymbol{R}^2$ and assume the following prior density

$$\pi = \Phi(\lambda_1, \lambda_2) = \frac{1}{2\pi} \exp\left\{ -\frac{1}{2} [(\theta_1 - \lambda_1)^2 + (\theta_2 - \lambda_2)^2] \right\}$$

the two vectors

$$\frac{\partial \Phi}{\partial \lambda_1} = (\theta_1 - \lambda_1)\, \Phi(\lambda_1, \lambda_2)$$

$$\frac{\partial \Phi}{\partial \lambda_2} = (\theta_2 - \lambda_2)\, \Phi(\lambda_1, \lambda_2)$$

give a basis of $V(\Gamma)_{\pi_0}$. Let $\{I_1, I_2\}$ be the following orthonormalized basis associated to

$$\left\{ \frac{\partial \Phi}{\partial \lambda_1}, \frac{\partial \Phi}{\partial \lambda_2} \right\}:$$

$$I_1 = \sqrt{8\pi}\,\frac{\partial \Phi}{\partial \lambda_1}$$

$$I_2 = \sqrt{8\pi}\,\frac{\partial \Phi}{\partial \lambda_2}\,.$$

Then

$$S_N(\pi_0, \Gamma) = \frac{\sqrt{8\pi \sum_{i=1}^{2}\left[N\left(\frac{\partial \Phi}{\partial \lambda_i} \right) D(\Pi_0) - D\left(\frac{\partial \Phi}{\partial \lambda_i} \right) N(\Pi_0) \right]^2}}{D(\Pi_0)^2} = \sqrt{8\pi}\,\|T_{\Psi_0}\|\,.$$

4. - Discussion

The main novelty of this work is the introduction of a quite general definition of local sensitivity. This definition can be applied to any class of priors (convex or not, parametric or non parametric) and it gives an intrinsec characteristic of the class under consideration. The resulting expression has a natural statistical interpretation: we regard the sensitivity index $S(C_{\Pi_0}, \Gamma)$ as a measure on the posterior expectation of the effect of small changes in Π_0. The calculation is carried out completely in the parametric case. More work will be needed to obtain explicit expression in the other cases. Solutions of this problem are already available in the literature. For results in this area we remaind to Gustavson (1996) [7], Gustavson (1996) [8], Musio (1997) [11], Musio (1999) [12]. In Musio (1997) [11] the calculation of the sensitivity index is solved in other non-convex cases i.e. classes with specified moments or for classes defined by k independent non-linear equations. Most of the recent work on local sensitivity indices has addressed the problem of weird behavior in the case where the sample size increases to infinity. Local measures tend to diverge to infinity as the sample size grows. Gustavson (1996) shows that if we do not consider the whole posterior distribution but we restrict to a posterior expectation, we can obtain sensitivity measures with a more satisfactory asymptotic behaviour. In the context of ε-contamination classes, Sivaganesan (1996) [14] gives some sufficient conditions for the convergence to zero of local sensitivity for posterior expectation Ψ. In this case the local sensitivity of Ψ with respect to the prior

$$\pi \in \Gamma = \{\pi_\varepsilon(\theta) = (1-\varepsilon)\,\pi_0 + \varepsilon q,\ q \in \mathcal{Q}\}$$

based on a sample of size n is defined as

$$LS(n) = \sup_q \left[\frac{d}{d\varepsilon} \Psi(\pi) \right]_{\varepsilon = 0} .$$

Here π is the density of Π w.r.t. the Lebesgue measure, and θ is one dimensional. The index $LS(n)$ is a particular case of the definition 2.5 for the choices $\pi(\varepsilon) = \pi_0 + \varepsilon(q - - \pi_0)\ q \in \mathcal{Q}$, and $\delta(\varepsilon) = \varepsilon$. In our more general context we can easily extende the Sivaganesan's convergence therem. Under the same regularity assumptions ([14]), guaranteeing the asymptotic convergence of the posterior w.r.t. each $\pi \in \Gamma$, we have the following theorem:

THEOREM 4.1: *If $\dot{\pi}(0)$ is bounded then $S_N(\pi_0, \Gamma)$ is $O(n^{1/2})$.*

The proof follows easily from the convergence of $LS(n)$, if we remark that

$$S_N(\pi_0, \Gamma) = LS(n) \frac{1}{\|\dot{\pi}(0) - \pi_0\|} .$$

Than the asymptotic behavior of $S_N(\pi_0, n)$ follows from the behavior of $LS(n)$.

For instance if we assume that $p(x|\theta) = N(\theta, \sigma^2)$, $\pi(\theta) = N(\mu, \tau^2)$, and $\Gamma = = \{Q_{a, \tau^2}\}$, then the hypotheses of the theorem above is satisfied and the $\lim_{n \to \infty} S_N(\pi_0, \Gamma) = 0$. From a different way, we find the same result as Gustafson, Srinivasan and Wasserman, (1995) [10].

A problem related to sensitivity is the one of calibration, that is, the interpretation of the numerical value found. The index $S_N(\Pi_0, \Gamma)$ depends on the prior Π_0, on the choice of $\delta(t)$ and on the choice of the class Γ. If $\delta(t)$ is specified, for instance if we take $\delta(t) = \|\Pi(t) - \Pi_0\|$, Γ_1 and Γ_2 are classes of priors, we call the ratio

$$\frac{S_N(\Pi_0, \Gamma_1)}{S_N(\Pi_0, \Gamma_2)} .$$

the corresponding *relative sensitivity*. From a mathematical viewpoint, one is not able to interpret the values of the sensitivity index absolutely, but only in a relative sense.

REFERENCES

[1] BASU S., JAMMALAMADAKA S. R., LIU W., *Local Posterior Robustness with Parametric Priors: Maximum and Average Sensitivity*. In *Maximum Entropy and Bayesian Methods* (G. Heidbreder, ed.) Kluwer, Dordecht (1996).

[2] BERGER J. O., *An Overview of Robust Bayesian Analysis (with discussion)*, Test, 3 (19949, 5-26.

[3] Clarke B. - Gustafson P., *On the Overall Sensitivity of the Posterior Distribution to its Inputs*, J. Statist. Plann. Inference, 71 (1998), 137-150.
[4] Cuevas A. - Sanz P., *On Differentiability Properties of Bayes Operators*, in *Bayesian Statistics 3* (editors J. M. Bernardo, M. H. DeGroot, D. V. Lindley, A. F. M. Smith), pp. 569-577. Oxford University Press (1998).
[5] Dey D. K. - Ghosh S. K. - Lou K., *On Local Sensitivity Measure in Bayesian Analysis, IMS monograph* the proceedings of the second international workshop on Bayesian Robustness, Rimini (1996).
[6] Diaconis P. - Freedman D., *On the Consistency of Bayes Estimates*, Ann. Statist., 14 (1986), 1-67.
[7] Gustafson P., *Local Sensitivity of Posterior Expectations*, Ann. Statist., 24 (1996), 174-195.
[8] Gustafson P., *Local Sensitivity of Inferences to Prior Marginals*, J. Amer. Statist. Assoc., **91** (1996) 774-781.
[9] Gustafson P. - Wasserman L., *Local Sensitivity Diagnostics for Bayesian Inference*, Ann. Statist., 23 (1995), 2153-2167.
[10] Gustafson P. - Srinivasan C. - Wasserman L., *Local Sensitivity Analysis, In Bayesian Statistics 5*, J. M. Bernardo, et. al. (Eds.), pp. 218-238. Oxford University Press (1995).
[11] Musio M., *Un Indice generale di Sensitività nell'Analisi della Robustezza Bayesiana e suo Calcolo mediante la Geometria Euclidea. Tesi di Dottorato*, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma «La Sapienza» (1997).
[12] Musio M., *Computing Bayesian Local Sensitivity Through Geometry*, préprint (1998).
[13] Ruggeri F. - Wasserman L., *Infinitesimal Sensitivity of Posterior Distributions*, Canad. J. of Statist., 21 (1993), 195-203.
[14] Sivaganesan S., *Multi-Dimensional Priors: Global and Local Robustness, IMS monograph* the proceedings of the second international workshop on Bayesian Robustness, Rimini (1996)
[15] Wasserman L., *Recent Methodological Advances in Robust Bayesian Inference* (with discussion), in *Bayesian Statistics 4*, J.M. Bernardo, et. al. (Eds.), pp. 483-502, Oxford University Press (1992).

Direttore responsabile: Prof. A. Ballio - Autorizz. Trib. di Roma n. 7269 dell'8-12-1959
«Monograf» - Via Collamarini, 5 - Bologna