LOCAL SENSITIVITY ANALYSIS IN BAYESIAN DECISION THEORY [1]

JACINTO MARTIN AND DAVID RIOS INSUA

Madrid Technical University and CNR-IAMI

We study local sensitivity in Bayesian Decision Theory, allowing for imprecision in the Decision Maker's preferences and beliefs. Our tools are based on Fréchet derivatives of operators and their norms. They allow us to detect cases in which robustness is lacking and, eventually, the most critical judgments determining choice.

1. Introduction. Our initial framework is that of Bayesian Decision Theory and Inference, see Savage (1972). Several authors suggest that Bayesian foundations place excessive demands on the Decision Maker's (DM) judgments. This motivates the development of tools to check the sensitivity of conclusions of a Bayesian analysis with respect to changes in the inputs. Berger (1994) provides an excellent review. However, most work has concentrated on sensitivity to the prior. As we suggest in our discussion to that paper, fundamental (Ríos Insua and Martín, 1995) and practical (Ríos Insua and Martín, 1994) issues suggest developing a general framework for sensitivity analysis allowing for perturbations both in preferences and beliefs.

This paper studies this question from a local perspective, that is, we study whether small perturbations in the inputs to the analysis lead to important changes in the conclusions. For this, we use Fréchet derivatives. We generalize previous results by, among others, Diaconis and Freedman (1986), Srinivasan and Trusczynska (1995), Basu et al. (1993), Sivaganesan (1993) and, specially, Ruggeri and Wasserman (1993), to the more difficult case, see Berger (1994), of imprecision in both the utility and the prior.

We think of our results in an iterative fashion. Our analyses allow us to detect cases in which robustness is lacking. Moreover, we provide procedures suggesting what additional information we must elicit from the DM to increase robustness. It is specially relevant that this information is meaningful to the DM, against somewhat skeptical opinions (Berger, 1994; Das Gupta, 1995). Hence, we could think of incorporating our results into a more general framework for sensitivity analysis in Decision Theory, see Ríos Insua (1990).

We shall use the following notation: S will designate the set of states s, endowed with a σ-field $\mathcal{B}$. P will be the prior distribution modeling the DM's beliefs, updated to the posterior $P(\cdot|x)$, when x is the result of an experiment with likelihood $l(x|s)$ over a sample space X. The DM makes decisions $a \in \mathcal{A}$, the space of alternatives. We associate a consequence $c \in \mathcal{C}$,

[1]Supported by grants from CICYT TIC95-0028 and NATO CRG-950214
AMS 1991 subject classifications. Primary 62C10; secondary 49A15.
Key words and phrases. Bayesian Decision Theory, Local Sensitivity Analysis, Fréchet Derivative.

to each pair (a, s). Preferences over consequences are modeled with a utility function u, and we associate to each alternative a its posterior expected utility

$$T(u,P,a) = \frac{\int u(a,s)l(x|s)dP(s)}{\int l(x|s)dP(s)} = \frac{N(u,P,a)}{D(P)}.$$

We maximize $T(u, P, a)$ in a, as a way of obtaining the optimal alternative.

The assessment of P and u is far from simple, and the DM may demand ways of checking the impact of u and P on the conclusions. We do this, in Section 2, with the aid of some tools from functional analysis. The first one is the Fréchet derivative of $T(u, P, a)$. We use it to approximate changes in expected utility when u and P change. We also use the Fréchet derivative norm, to bound changes in expected utility when u and/or P change.

However, we run into some problems, which we deal with in Section 3. We introduce then another sensitivity measure, in presence of classes of utility functions and prior distributions, and illustrate potential uses.

We assume that $\sup_{s\in S} l(x|s) < \infty$, and $\sup_{s\in S} |u(a,s)|l(x|s) < \infty$. All integrals are over S, which will be a Polish space.

2. Derivative of Bayes operator. Let us study the local behavior of $T(u, P, a)$, the operator associating to each pair (u, P) the posterior expected utility of a. Typically, a will be an alternative suggested as optimal. Since it remains fixed in our study, we shall designate the operator by $T(u, P)$, and call it *Bayes operator*. Similarly, $N(u, P, a)$ will be called $N(u, P)$. We compute the Fréchet derivative of $T(u, P)$:

DEFINITION 1 *The derivative of $T(u,P)$, with respect to u, is the continuous linear operator $\dot{T}_u$, over the set $\mathcal{F}$ of bounded functions on C, verifying*

$$\forall m \in \mathcal{F},\ T(u+m,P) = T(u,P) + \dot{T}_u(m) + o(\|m\|).$$

DEFINITION 2 *The derivative of $T(u,P)$, with respect to P, is the continuous linear operator $\dot{T}_P$, over the set $\mathcal{M}$ of signed measures on $(S,\mathcal{B})$, verifying*

$$\forall \delta \in \mathcal{M},\ T(u,P+\delta) = T(u,P) + \dot{T}_P(\delta) + o(\|\delta\|).$$

Previous work on local Bayesian robustness has focused mainly on Definition 2, but see Gustafson et al (1995).

DEFINITION 3 *The derivative of $T(u,P)$, with respect to u and P, is the continuous linear operator $\dot{T}_{uP}$, over $\mathcal{F}\times\mathcal{M}$, verifying*

$$\forall (m,\delta) \in \mathcal{F}\times\mathcal{M},\ T((u,P)+(m,\delta)) = T(u,P) + \dot{T}_{uP}(m,\delta) + o(\|(m,\delta)\|).$$

We shall use the following norms over the incumbent vector spaces:

- In $\mathcal{F}$, the supremum norm, i.e. $\|u\| = \sup_{c\in\mathcal{C}}|u(c)|$,
- In $\mathcal{M}$, the total variation norm, i.e. $\|P\| = \sup_{A\in\mathcal{B}}|P(A)|$.
- In $\mathcal{F}\times\mathcal{M}$, $\|(m,\delta)\|_\infty = \max\{\|m\|, \|\delta\|\}$.

We provide now the derivatives:

THEOREM 1

$$\dot{T}_u(m) = T(m,P) \tag{1}$$

$$\dot{T}_P(\delta) = \frac{1}{D(P)}[N(u,\delta) - T(u,P)D(\delta)] \tag{2}$$

$$\dot{T}_{uP}(m,\delta) = \dot{T}_u(m) + \dot{T}_P(\delta) \tag{3}$$

PROOF. (1) is immediate since $T(u,P)$ is linear in u. (2) is in Ruggeri and Wasserman (1993). For (3), consider the directional derivative:

$$\lim_{\varepsilon\to 0}\frac{T((u,P)+\varepsilon(m,\delta)) - T(u,P)}{\varepsilon} =$$

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\frac{N(u,P)+\varepsilon N(m,P)+\varepsilon N(u,\delta)+\varepsilon^2 N(m,\delta)}{D(P)+\varepsilon D(\delta)} - \frac{N(u,P)}{D(P)}\right] =$$

$$\frac{1}{D(P)}[N(m,P) + N(u,\delta) - T(u,P)D(\delta)].$$

Let us call it $\dot{T}_{uP}(m,\delta)$. It satisfies Fréchet derivative conditions:

$$\left|T((u,P)+(m,\delta)) - T(u,P) - \dot{T}_{uP}(m,\delta)\right| =$$

$$\left|\frac{N(m,\delta)}{D(P)+D(\delta)} - \frac{D(\delta)N(m,P)}{D(P)(D(P)+D(\delta))} - \right.$$

$$\left.\frac{D(\delta)N(u,\delta)}{D(P)(D(P)+D(\delta))} + \frac{D^2(\delta)N(u,P)}{D(P)^2(D(P)+D(\delta))}\right| \leq$$

$$\frac{1}{|D(P)+D(\delta)|}\left[|N(m,\delta)| + \frac{|D(\delta)|}{D(P)}|N(m,P)| + \right.$$

$$\left.\frac{|D(\delta)|}{D(P)}|N(u,\delta)| + \frac{D(\delta)^2}{D(P)^2}|N(u,P)|\right].$$

The four terms are $o(\|(m,\delta)\|_\infty)$.

1. $|N(m,\delta)| = |\int m(a,s)l(x|s)d\delta(s)| \leq \sup|m(a,s)| \sup l(x|s) \int |d\delta(s)|$ $\leq 2 \sup l(x|s) ||m|| ||\delta|| \leq K_1 ||(m,\delta)||^2_\infty$, with $K_1 = 2\sup l(x|s)$ and observing that $||m||$ and $||\delta||$ are smaller than $||(m,\delta)||_\infty$, and, in general, $\int |d\delta(s)| \leq 2||\delta||$.

2. $|D(\delta)| \leq 2\sup l(x|s)||\delta|| \leq K_1||(m,\delta)||_\infty$.
$|N(m,P)| \leq \int |m(a,s)| l(x|s) dP(s) \leq D(P)\, ||m|| \leq D(P)\, ||(m,\delta)||_\infty$.
Hence, $\dfrac{|D(\delta)|}{D(P)} |N(m,P)| \leq K_1 ||(m,\delta)||^2_\infty$.

3. $|N(u,\delta)| \leq \int |u(a,s)l(x|s)d\delta(s)| \leq 2\sup\ |u(a,s)l(x|s)| ||\delta||$ $\leq K_2||(m,\delta)||_\infty$, with $K_2 = 2\sup\ |u(a,s)l(x|s)|$. Therefore, $\dfrac{|D(\delta)|}{D(P)}|N(u,\delta)| \leq K_3||(m,\delta)||^2_\infty$, with $K_3 = K_1 \dfrac{K_2}{D(P)}$.

4. $\dfrac{D(\delta)^2}{D(P)^2}|N(u,P)| \leq K_4||(m,\delta)||^2_\infty$, with $K_4 = K_1^2 \dfrac{T(u,P)}{D(P)}$.

Then,

$$\left|T((u,P)+(m,\delta)) - T(u,P) - \dot{T}_{uP}(m,\delta)\right| \leq$$

$$\frac{1}{|D(P)+D(\delta)|}(2K_1 + K_3 + K_4)||(m,\delta)||^2_\infty = K_P||(m,\delta)||^2_\infty. \tag{4}$$

Since $|D(P)+D(\delta)| \to D(P)$ when $||\delta|| \to 0$, then $|T((u,P)+(m,\delta)) - T(u,P)$ $-\dot{T}_{uP}(m,\delta)|$ is $o(||(m,\delta)||_\infty)$.

Moreover, $\dot{T}_{uP}$ is bounded and linear, so we have the result. □

In principle, we could think of using the Fréchet derivative norm as a local sensitivity measure of $T(u,P)$ with respect to changes in u, P or (u,P), since the change in expected utility due to using (u,P) instead of $(u+m, P+\delta)$ may be approximated from the definitions of derivative and norm of an operator. For example, if we consider perturbations in both the utility and the prior, since $||\dot{T}_{uP}|| = \sup_{(m,\delta)\in\mathcal{F}\times\mathcal{M}}(|\dot{T}_{uP}(m,\delta)|/||(m,\delta)||_\infty)$, we have

$$|T(u+m,P+\delta) - T(u,P)| \leq ||\dot{T}_{uP}||\ ||(m,\delta)||_\infty + o(||(m,\delta)||_\infty). \tag{5}$$

The computation of norms is relatively straightforward when no constraints are placed on the perturbations on P and u. First, note that $||\dot{T}_u|| = 1$. Hence, in this case the norm is independent of u, suggesting the same local robustness for all utility functions. This would suggest adding constraints to $\mathcal{F}$, as we do in Section 3. To perturb probabilities, we consider the set $\Delta \subset \mathcal{M}$ of signed measures δ with 0 mass ($\delta(S) = 0$). Therefore, $(P + \delta)(S) = 1$. However, we cannot guarantee that $P + \delta$ is a probability distribution. We deal with this issue in Section 3. Then, introducing $h(s) = l(x|s)(u(a,s) - T(u,P))$, $\bar{h} = \sup_{s \in S} h(s)$, $\underline{h} = \inf_{s \in S} h(s)$, Ruggeri and Wasserman (1993) prove that when P is a non atomic distribution, $\|\dot{T}_P\| = (\bar{h} - \underline{h})/D(P)$. This result is actually valid for a general P, when l and u are continuous, and S is convex. Finally, we easily have $\|\dot{T}_{uP}\|_\infty = 1 + (\bar{h} - \underline{h})/D(P)$.

Note that, in this case, the norm allows us to study local sensitivity with respect to (u, P), decomposing it in sensitivity with respect to u and with respect to P: the norm depends both on u and P, so both effects add up. Typically, *we shall observe more sensitivity in general studies than in analyses with respect to the prior only or the utility only.*

3. Local sensitivity of Bayes operator. The above results are interesting from a mathematical point of view and may be used in infinitesimal sensitivity studies. However, their application in Decision Theory demands some care for, at least, two reasons:

- $P + \delta$ is not necessarily a probability distribution, since δ is a signed measure.
- The sensitivity study is meaningful only if we normalise utilities: it is not the same a change of 1, when u varies between 0 and 1, than when it varies between 0 and 1000. Then, we have the same problem as above, since $u + m$ might not be normalized.

Besides, in most applications we shall have some information about u and P, which allow us to constrain them to certain classes $\mathcal{U}$ and Γ. Hence, we should entertain sensitivity studies over classes of utilities and priors. However, the general problem of computing Fréchet derivative norms with respect to classes is difficult, so we shall use another local sensitivity measure.

First, rather than working with neighborhoods of (u, P), we shall use neighborhoods of the origin as follows: for each neighborhood E of (u, P) within the class, consider E' such that $(m, \delta) \in E' \iff (u, P) + (m, \delta) \in E$. The sensitivity measure we use is the supremum of the absolute value of the derivative in the neighborhoods of interest. Note that: $\forall (m, \delta) \in E'$

$$(6)\quad |T(u+m, P+\delta) - T(u,P)| \leq \sup_{(m',\delta') \in E'} |\dot{T}_{uP}(m', \delta')| + o(\|(m,\delta)\|_\infty).$$

Compare this inequality with (5).

In fact, we shall limit our attention to the following classes:

$\Gamma_Q = \{r : \underline{p}_j \leq r(A_j) \leq \bar{p}_j,\ j = 1,\ldots,n\}$, where $A_1,\ldots,A_n$ is a measurable partition of S and $\sum_{j=1}^n \underline{p}_j \leq 1 \leq \sum_{j=1}^n \bar{p}_j$.

$\mathcal{U}_K = \{v : \mathcal{C} \longrightarrow \mathbb{R} :\ v_{i-1} \leq v(c) \leq v_i,\ \forall c \in C_i,\ i = 1,\ldots,k\}$, where $C_1,\ldots,C_k$ is a measurable partition of $\mathcal{C}$, for an appropriate σ-field in $\mathcal{C}$, $v_0 = 0$, and $v_k = 1$.

They are associated with the most popular methods in prior (quantile method) and utility (probability equivalent method) elicitation. Moreover, they lead to relatively simple computations. Note though that the study would be similar for other classes.

For computational reasons, we must constrain the neighborhoods of interest as follows. For $u \in \mathcal{U}_K$, let $m = \varepsilon(v - u)$ with $v \in \mathcal{U}_K$ and $\varepsilon \in [0,1]$. In such a way:

- $u + m = (1-\varepsilon)u + \varepsilon v \in \mathcal{U}_K$.
- We bound the norm of elements in the neighborhood, since $\|m\| \leq \varepsilon \max_{i=1,\ldots,k}\{\max\{\sup_{c\in C_i}[v_i - v(c)],\ -\inf_{c\in C_i}[v_{i-1} - v(c)]\}\}$.

For similar reasons, for $P \in \Gamma_Q$, we take $\delta = \varepsilon(Q - P)$ with $Q \in \Gamma_Q$ and $\varepsilon \in [0,1]$. This type of neighborhoods is very popular in sensitivity to the prior studies, and leads to ε-contaminated classes, see Berger (1994).

We then have:

THEOREM 2 *Let* $\rho = T(u,P)$, $\varepsilon_u > 0$, $\mathcal{F}_{\varepsilon_u} = \{m : m = \varepsilon(v-u),\ v \in \mathcal{U}_K,\ \varepsilon \leq \varepsilon_u\}$. *Then:*

$$\sup_{\mathcal{F}_{\varepsilon_u}} |\dot{T}_u(m)| = \varepsilon_u \max\left\{\sum_{i=1}^k (v_i - \rho)P(B_i|x), -\sum_{i=1}^k (v_{i-1} - \rho)P(B_i|x)\right\},$$

where $B_i = \{s \in S : (a,s) = c \in C_i\}$.

PROOF. For $\varepsilon \leq \varepsilon_u$, we have:

$$\sup_{v\in\mathcal{U}_K} |\dot{T}_u(\varepsilon(v-u))| = \sup_{v\in\mathcal{U}_K} |\int \varepsilon(v-u)(a,s)dP(s|x)| =$$

$$\varepsilon_u \sup_{v\in\mathcal{U}_K} |\int (v(a,s) - \rho)dP(s|x)| = \varepsilon_u \sup_{v\in\mathcal{U}_K} \left|\sum_{i=1}^k \int_{B_i} (v(a,s) - \rho)dP(s|x)\right| =$$

$$\varepsilon_u \max\left\{\begin{array}{l} \sup\limits_{v\in\mathcal{U}_K} \sum\limits_{i=1}^k \int_{B_i} (v(a,s) - \rho)dP(s|x), \\ \\ -\inf\limits_{v\in\mathcal{U}_K} \sum\limits_{i=1}^k \int_{B_i} (v(a,s) - \rho)dP(s|x) \end{array}\right\}.$$

Since

$$(v_{i-1} - \rho)P(B_i|x) \leq \int_{B_i} (v(a,s) - \rho)dP(s|x) \leq (v_i - \rho)P(B_i|x),$$

we get the result, noting that there are functions $v \in \mathcal{U}_K$ for which the infimum and the supremum are attained. □

In the case of imprecision only in the utility function, $\dot{T}_u$ measures the change in expected utility exactly. Hence, it is not only a local sensitivity measure. In particular, when $\varepsilon_u = 1$ we get a global sensitivity measure over $\mathcal{U}_K$.

An important consequence of Theorem 4 is that for those cases deemed too sensitive, i.e. when we believe that we must reduce the supremum of the derivative, the result suggests which subsets of $\mathcal{C}$ demand more elicitation efforts: the supremum is decomposed in terms relative to the sets C_i. We can even suggest a direction in which information should be elicited. We illustrate this fundamental idea in Examples 1 and 2 below.

Theorem 3 deals with the case of imprecision in the prior distribution only. Ruggeri and Wasserman (1993, Thm. 6) provide a related result. Their prior P needs to be non atomic, and their class requires quantiles assessed precisely. Also, we do not use the norm, but the supremum of the derivative.

THEOREM 3 *Given $P \in \Gamma_Q$, let $\Delta_{\varepsilon_P} = \{\delta : \delta$ signed measure such that $\delta = \varepsilon(Q - P), Q \in \Gamma_Q, \varepsilon \leq \varepsilon_P\}$. Then,*

$$\sup_{\Delta_{\varepsilon_P}} |\dot{T}_P(\delta)| = \frac{\varepsilon_P}{D(P)} \max\{H_1, -H_2\},$$

where H_1, H_2 are, respectively, the optimal values of linear programming problems

$$\begin{array}{cc} \max \sum_{j=1}^{n} p_j \bar{h}_j & \min \sum_{j=1}^{n} p_j \underline{h}_j \\ s.t. & s.t. \\ \sum_{j=1}^{n} p_j = 1 & \sum_{j=1}^{n} p_j = 1 \\ \underline{p}_j \leq p_j \leq \bar{p}_j & \underline{p}_j \leq p_j \leq \bar{p}_j \end{array}$$

with $\bar{h}_j = \sup_{s \in A_j} h(s)$ and $\underline{h}_j = \inf_{s \in A_j} h(s)$.

PROOF: For $\delta = \varepsilon(Q - P)$, with $Q \in \Gamma_Q$, $\varepsilon \leq \varepsilon_P$, and $\rho = T(u, P)$, we have

$$|\dot{T}_P(\delta)| = \frac{1}{D(P)} |N(u, \varepsilon(Q - P)) - \rho D(\varepsilon(Q - P))| =$$

$$= \frac{\varepsilon}{D(P)}|N(u,Q) - \rho D(Q)| = \frac{\varepsilon}{D(P)}|\int (u(a,s) - \rho) l(x|s) dQ(s)| =$$

$$= \frac{\varepsilon}{D(P)}|\int h(s) dQ(s)| = \frac{\varepsilon}{D(P)}\left|\sum_{j=1}^{n} \int_{A_j} h(s) dQ(s)\right|.$$

Then,
$\sup_{\Delta_{\varepsilon_P}} |\dot{T}_P(\delta)| =$

$$\frac{\varepsilon_P}{D(P)} \max\left\{\sup_{Q \in \Gamma_Q} \sum_{j=1}^{n} \int_{A_j} h(s) dQ(s), - \inf_{Q \in \Gamma_C} \sum_{j=1}^{n} \int_{A_j} h(s) dQ(s)\right\}.$$

The result follows now by the well-known fact that the extreme Q's are discrete. The search of the sup (inf) leads to the problem giving H_1 (H_2). □

Again, the result suggests for which A_j we should assess additional information so as to reduce sensitivity. Let us see a forecasting example, taken from Ríos Insua et al. (1995).

EXAMPLE 1. Consider the following model to forecast accidents of a certain company. Let X_k, n_k, D_k be, respectively, the number of accidents, the number of workers and the accident history of the company in period k. Assume that X_k follows a Poisson distribution with parameter $n_k\lambda$, with n_k known, and λ modeling the accident proneness of a worker.

In a specific case, experts provided prior information about λ leading to the class:

$$\Gamma = \{Q : Q(0, a_1] = p_1, Q(a_1, a_2] = p_2, Q(a_2, a_3] = p_3, Q(a_3, \infty) = p_4\}$$

with $p_i = .25, \forall i$, $a_1 = .38, a_2 = .58, a_3 = .98$. For computational convenience, we associate to that information a gamma prior with parameters $p = 1.59$ and $a = 2.22$.

Table 1 provides the range of the predictive mean $E(X_k|D_k)$, when the prior ranges over Γ, computed with the aid of a result in Ruggeri (1990).

TABLE 1

Year	$\underline{E}[X_k\|D_k]$	$\bar{E}[X_k\|D_k]$
1988	17.8	206.4
1989	17.3	201.2
1990	17.1	198.9

These results suggest lack of robustness, since expected forecasts vary widely when the prior ranges in Γ.

To increase robustness, we may appeal to our local sensitivity measure, which will suggest where to center additional elicitation efforts. The results are in Table 2, which includes the decomposition of the sensitivity measure, with $\bar{g}_i = \bar{h}_i$ or $-\underline{h}_i$:

TABLE 2

Year	$\sup_{\delta \in \Delta_{\varepsilon_0}} \lvert\dot{E}^\pi(\delta)\rvert$	$\bar{g}_1, \bar{g}_2, \bar{g}_3, \bar{g}_4$
1988	$25.15\ \varepsilon_0$	$25.15, -2.18 \times 10^{-23}, -3.24 \times 10^{-69}, 0$
1989	$27.45\ \varepsilon_0$	$27.45, 8.59 \times 10^{-19}, 2.07 \times 10^{-57}, 6.24 \times 10^{-153}$
1990	$23.22\ \varepsilon_0$	$23.22, -5.18 \times 10^{-83}, -3.23 \times 10^{-223}, 0$

For example, if $\varepsilon_0 = .2$, the imprecision in the expected forecast for 1989 is 5.49 accidents, which might be considered too big. In all years, the suprema is due to the first quartile (I_1). Hence, we extract additional information about such quartile from the expert. We divide it in three subintervals, $I_{11} = [0, 0.15)$, $I_{12} = [0.15, 0.2)$ and $I_{13} = [0.2, 0.38)$ with probabilities 0.1, 0.05 and 0.1, respectively. Hence, we may keep the same gamma prior. Then, for example, for 1989, the new value of $\sup_{\delta \in \Delta_{\varepsilon_0}} |\dot{E}^\pi(\delta)|$ is 7.76 ε_0, which is a considerable reduction.

This obviously has an impact on the upper and lower bounds of the predictive expectation:

TABLE 3

Year	$\underline{E}[X_k\|D_k]$	$\bar{E}[X_k\|D_k]$
1988	53.4	85.4
1989	52.0	79.8
1990	51.4	77.9

Note the reduction in ranges from Table 1 to Table 3, suggesting much more robustness. □

For the case in which there are changes in both the utility and the prior, we have the following two results, depending on the corresponding classes of priors (Δ_{ε_P} or Δ'_{ε_P}):

THEOREM 4 *With $\rho = T(u, P)$,*

$$\sup_{\mathcal{F}_{\varepsilon_u} \times \Delta_{\varepsilon_P}} |\dot{T}_{uP}(m, \delta)| = \max \left\{ \begin{array}{c} \varepsilon_u \sum_{i=1}^{k} (v_i - \rho) P(B_i|x) + \frac{\varepsilon_P}{D(P)} H_1, \\ \\ -\varepsilon_u \sum_{i=1}^{k} (v_{i-1} - \rho) P(B_i|x) - \frac{\varepsilon_P}{D(P)} H_2 \end{array} \right\}$$

with H_1, H_2 as in Theorem 3.

PROOF. The result follows from the expression of the derivative and Theorems 2 and 3. □

THEOREM 5 *Let $\Delta'_{\varepsilon_P} = \{\delta : \delta \in \Delta, \|\delta\| \leq \varepsilon_P\}$ and $\rho = T(u, P)$. We have:*

$$\sup_{\mathcal{F}_{\varepsilon_u} \times \Delta'_{\varepsilon_P}} |\dot{T}_{uP}(m,\delta)| = \max \left\{ \begin{array}{c} \varepsilon_u \sum_{i=1}^{k} (v_i - \rho) P(B_i|x) + \varepsilon_P \dfrac{\bar{h} - \underline{h}}{D(P)}, \\ \\ -\varepsilon_u \sum_{i=1}^{k} (v_{i-1} - \rho) P(B_i|x) + \varepsilon_P \dfrac{\bar{h} - \underline{h}}{D(P)} \end{array} \right\}$$

PROOF. Immediate, using the reasoning in Theorem 4 and taking into account Theorem 2 and Ruggeri and Wasserman (1993, Thm. 3). □

Note the interest of Theorems 4 and 5: if we consider excessive the value of $\sup |\dot{T}_{uP}(m,\delta)|$, we may see where do we need to refine the information to reduce the supremum. Moreover, it separates *utility* and *prior* effects, suggesting additional elicitation efforts.

EXAMPLE 2. Suppose $S = \{s_1, s_2, s_3\}$. Let a be such that $(a, s_1) = -1000$ pts., $(a, s_2) = 0$ pts. and $(a, s_3) = 3000$ pts. Suppose we have assessed:

$$0 \leq u(-1000) \leq 0.1 \leq u(0) \leq 0.9 \leq u(3000) \leq 1,$$

$$0.3 \leq p_1 \leq 0.4,\ 0.4 \leq p_2 \leq 0.6,\ 0.1 \leq p_3 \leq 0.2, \text{with } p_i = P(s_i),\ i = 1,2,3.$$

Initially, we adopt a utility function u_0 and a prior P_0 such that

$$u_0(-1000) = 0,\ u_0(0) = 0.25,\ u_0(3000) = 1,$$

$$P_0(s_1) = \frac{1}{3},\ P_0(s_2) = \frac{1}{2},\ P_0(s_3) = \frac{1}{6}.$$

Suppose also that $l(x|s_i) = 1/3$, $i = 1,2,3$, for the observed x. Then, $\rho = T(u_0, P_0) = 0.296$.

Due to Theorem 4,

$$\sup_{\mathcal{F}_{\varepsilon_u} \times \Delta_{\varepsilon_P}} |\dot{T}_{uP}(m,\delta)| = \max \left\{ \begin{array}{c} \varepsilon_u \sum_{i=1}^{3} (v_i - \rho) P_0(s_i|x) + \dfrac{\varepsilon_P}{D(P)} H_1 \\ \\ -\varepsilon_u \sum_{i=1}^{3} (v_{i-1} - \rho) P_0(s_i|x) - \dfrac{\varepsilon_P}{D(P)} H_2 \end{array} \right\}$$

with $v_0 = 0,\ v_1 = 0.1,\ v_2 = 0.9,\ v_3 = 1$ and

$$H_1 = \max \sum_{j=1}^{3} p_j \bar{h}_j \quad \text{s.t.} \quad \sum_{j=1}^{3} p_j = 1,\quad 0.4 \le p_1 \le 0.6,\quad 0.3 \le p_2 \le 0.4,\quad 0.1 \le p_3 \le 0.2$$

$$H_2 = \min \sum_{j=1}^{3} p_j \underline{h}_j \quad \text{s.t.} \quad \sum_{j=1}^{3} p_j = 1,\quad 0.4 \le p_1 \le 0.6,\quad 0.3 \le p_2 \le 0.4,\quad 0.1 \le p_1 \le 0.2$$

with $\bar{h}_j = \underline{h}_j = l(x|s_j)(u_0(a, s_j) - \rho), j = 1, 2, 3$, since the A_j's are singletones. We get

$$\bar{h}_1 = \underline{h}_1 = -0.0971,\ \bar{h}_2 = \underline{h}_2 = -0.0138,\ \bar{h}_3 = \underline{h}_3 = 0.2361.$$

Then,

$$\text{(7)} \qquad \sum_{i=1}^{3} (v_i - \rho) P_0(s_i|x) = -0.0638 + 0.3042 + 0.1180 = 0.3584,$$

$$\sum_{i=1}^{3} (v_{i-1} - \rho) P_0(s_i|x) = -0.0915,$$

and

$$H_1 = 0.0028,\quad H_2 = -0.0387.$$

Hence,

$$\text{(8)} \sup_{\mathcal{F}_{\varepsilon_u} \times \Delta_{\varepsilon_P}} |\dot{T}_{uP}(m, \delta)| = \max\{0.3584\varepsilon_u + 0.0028\varepsilon_P,\ 0.0915\varepsilon_u + 0.0387\varepsilon_P\}.$$

Note, first, the interaction between the utility and the prior as in (8). The most influential term is 0.3584 ε_u, assuming $\varepsilon_u = \varepsilon_P$. Then, we should probably concentrate further elicitation efforts in the utility function. The most influential term in (7) is 0.3042, due to $u(0)$. Hence, efforts should concentrate on the assessment of $u(0)$. Moreover, we should pay more attention to the upper bound 0.9, reducing it, if possible. □

The introduction of classes of utilities and/or priors is possible when we have partial information about u and P. If this is not the case, we may still do the study as a particular case, choosing neighborhoods of (u, P) as before, this time with Γ_C and $\mathcal{U}_K$ including all priors and all utility functions bounded between 0 and 1, respectively. Denoting those classes by Γ_0 and $\mathcal{U}_0$, this is equivalent to assuming $n = k = 1$ in the previous study, and corresponds to sensitivity for ε-contaminated classes, when the contaminating classes are the class of all priors over S and the class of all utility functions bounded between 0 and 1.

COROLLARY 1 *(of Theorem 2) Let* $\rho = T(u,P)$ *and* $\varepsilon_u > 0$. $\mathcal{F}^0_{\varepsilon_u} = \{m : m = \varepsilon(v-u),\ \varepsilon \le \varepsilon_u,\ v \in \mathcal{U}_0\}$. *We have*

$$\sup_{\mathcal{F}^0_{\varepsilon_u}} |\dot{T}_u(m)| = \max\{\varepsilon_u(1-\rho), \varepsilon_u\rho\}.$$

COROLLARY 2 *(of Theorem 3) Let* $\Delta^0_{\varepsilon_P} = \{\delta : \delta = \varepsilon(Q-P),\ \varepsilon \le \varepsilon_P,\ Q \in \Gamma_0\}$. *We have*

$$\sup_{\Delta^0_{\varepsilon_P}} |\dot{T}_P(\delta)| = \max\left\{\frac{\varepsilon_P}{D(P)}\bar{h}, -\frac{\varepsilon_P}{D(P)}\underline{h}\right\}.$$

COROLLARY 3 *(of Theorem 4) Let* $\rho = T(u,P)$. *We have*

$$\sup_{\mathcal{F}^0_{\varepsilon_u}\times\Delta^0_{\varepsilon_P}} |\dot{T}_{uP}(m,\delta)| = \max\left\{\varepsilon_u(1-\rho) + \frac{\varepsilon_P}{D(P)}\bar{h}, \varepsilon_u\,\rho - \frac{\varepsilon_P}{D(P)}\underline{h}\right\}.$$

4. Conclusions. We have provided a framework for general local sensitivity analysis in Bayesian Decision Theory. The framework is general in that it allows perturbations both in the beliefs and preferences of the DM. It allows the detection of cases lacking robustness, and, eventually suggests which additional information may be assessed to increase robustness of decisions. Additional examples in the medical context may be seen in Martín and Muller (1995). Note though that we do not see these tools as stand-alone, but rather as complementary to other tools, that could be eventually incorporated into a more general, sensitivity analysis based, scheme for Decision Analysis, see Ríos Insua (1990). Our main point here is to show that the use of these measures lead to meaningful quantities in terms of differences of expected utilities, facilitating the assessment of additional information.

Pragmatically, this, in part, would mitigate the asymptotic shortcomings mentioned in Gustafson et al (1995). On a different stand, another way of alleviating asymptotic problems is by reducing the neighborhoods of interest, as an attempt to eliminate unreasonable priors, see e.g. Sivaganesan (1995). In that sense, note that the absolute value of K_P in the expression of the infinitesimum (4) increases at the rate of the derivative norm. For example, for $\dot{T}_{uP}$, K_P is a function of $\sup_{s\in S} l(x|s)$, $\sup_{s\in S} l(x|s)^2$ and

$\sup_{s \in S} |u(a,s)| l(x|s)^2$. This could help us to reduce the neighborhoods of interest, in computing the values of ε_u and ε_P in Section 3. This points out to the very important issue of estimating the infinitesima, which we have considered here dismissable, as a way to identify neighborhoods where the approximation is acceptable.

Many other issues still remain. For example, we lack formal methods to find out whether an error is big or small. The extension of computations in Section 3 to other classes should be of interest, including shape constraints, such as monotonicity of the utility functions and unimodality of the prior distributions.

It should be also important to study similar issues for the operator *difference of posterior expected utilities* of two alternatives. This operator is specially relevant in robustness studies, since it is basic when looking for nondominated actions.

Acknowledgements. This research started when the authors were visiting ISDS at Duke University. The final version was prepared while the second author was visiting CNR-IAMI. We are grateful to discussions with other members of UPM Decision Analysis Group, Katja Ickstadt, Fabrizio Ruggeri and the referees.

REFERENCES

BASU, S., JAMMALAMADAKA, S. AND LIU, W. (1993) Qualitative robustness and stability of the posterior distribution and posterior quantities, Technical Report 239, Dep. Stats., Univ. California at Sta. Barbara.

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

DAS GUPTA, A. (1995) Discussion of *Local sensitivity analysis*, in Bernardo, Berger, Dawid, Smith (eds) *Bayesian Statistics 5.*

DIACONIS, P. AND FREEDMAN, D. (1986) On the consistency of Bayes estimates, *Ann. Statist.* **14** 1-67.

GUSTAFSON, P., SRINIVASAN, C., AND WASSERMAN, L. (1995) Local sensitivity analysis, in Bernardo, Berger, Dawid, Smith (eds) *Bayesian Statistics 5.*

MARTIN, J. AND MULLER, P. (1995) Sensitivity analysis in medical decision problems, Technical Report, Dept. of AI, UPM.

RIOS INSUA, D. (1990) *Sensitivity Analysis in Multiobjective Decision Making*, Springer.

RIOS INSUA, D. AND MARTIN, J. (1994) Robustness issues under imprecise beliefs and preferences, *J. Statist. Plann. Inference* **48** 383-389.

RIOS INSUA, D. AND MARTIN, J. (1995) On the foundations of robust Bayesian Statistics, Technical Report, Dept. of Statistics, Purdue University.

RIOS INSUA, S., MARTIN, J., RIOS INSUA, D. AND RUGGERI, F. (1995) Bayesian forecasting for accident proneness evaluation, Technical Report, Dept. of AI, UPM.

RUGGERI, F. (1990) Posterior ranges of functions of parameters under priors with specified quantiles, *Comm. Statist. Theory and Methods* **19** 127-144.

RUGGERI, F. AND WASSERMAN, L. (1993) Infinitesimal Sensitivity of Posterior Distributions, *Can. Jour. Statist.* **21** 195-203.

SAVAGE, L. J. (1972) *The Foundations of Statistics*, Dover.

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

SIVAGANESAN, S. (1995) Multi-dimensional Priors: Global and Local Robustness, presented at Second International Workshop on Bayesian Robustness.

SRINIVASAN, C. AND TRUSZCZYNSKA, H. (1995) Ranges of non-linear posterior quantities, to appear in *Ann. Statist.*

J. MARTIN
D. RIOS INSUA
DECISION ANALYSIS GROUP
MADRID TECHNICAL UNIVERSITY
CAMPUS DE MONTEGANCEDO
28660 MADRID, SPAIN

D. RIOS INSUA
CNR-IAMI
VIA AMPERE, 56
20131 MILANO, ITALY

Local Sensitivity Analysis In Bayesian Decision Theory

discussion by
KATJA ICKSTADT
Universität Basel

The main goal of sensitivity analysis is to discover how changes in the model inputs affect the inference. In a Bayesian setting the inputs to consider are the likelihood, the prior, the utility or loss function and the data. Most of the literature in this area deals with investigating sensitivity to the prior; only a few contributions consider sensitivity to the likelihood [e.g. Cuevas and Sanz (1988), Dey et al. (1995)]. Sensitivity to the utility or loss function has hardly been analysed at all, although those are important components in decision theory. The paper by Martín and Ríos Insua overcomes this gap. Furthermore, the authors offer an elaborate mathematical structure for studying imprecisions in both the utility function and the prior simultaneously.

Fréchet derivatives have widely been used in local sensitivity studies and Martín and Ríos Insua generalize them to their situations in a very natural way. In a Bayesian sensitivity analysis the Fréchet derivative is especially appropriate: the Bayes operator $T(u, P)$ evaluated at some initial values (u, P) is compared to all the other values $T((u, P)+(m, \delta))$ (cf. Definition 3) in a neighborhood of (u, P). Thus, in contrast to directional derivatives the Fréchet derivative approximates the differences $T((u, P)+(m, \delta)) - T(u, P)$ in all directions simultaneously.

The norms Martín and Ríos Insua have chosen are appropriate for robustness studies, since they preserve volumes. For example, the image of an element (m, δ) of the unit ball in $\mathcal{F} \times \mathcal{M}$, $||(m, \delta)||_\infty$, is an element of the unit interval $[0, 1]$. In the talk another norm in $\mathcal{F} \times \mathcal{M}$ was introduced, a convex combination of norms in $\mathcal{F}$ and in $\mathcal{M}$, i.e. of changes in u and in P. It is remarkable that the authors are able to give explicit computations (analogous to Theorems 1 and 3) for this other volume preserving norm as well.

Martín and Ríos Insua introduce two different measures for local sensitivity, the operator norm (cf. (4)) as well as the supremum of the Fréchet derivative (cf. (5)). They indicate that the second measure has the advantage of allowing sensitivity analysis with respect to certain classes of utilities and priors. However, the supremum of the Fréchet derivative is calculated for changes in the utility (cf. Theorem 2), for changes in the prior (cf. Theorem 3) and for changes in both utility and prior (cf. Theorems 4 and 5), whereas the operator norm is just determined for imprecisions in the prior and for imprecisions in both utility and prior. What can be stated about the operator norm if disturbances appear in the utility function only?

My only suggestion for improvement concerns the examples. Although the supremum of the Fréchet derivative (cf. (5)) is the more fruitful local sensitivity measure, it would be interesting to compare the results of Examples 1 and 2 to the results one would obtain applying the operator norm (cf. (4)) as a local sensitivity measure. Example 1 studies changes only in the prior. In order to emphasize the new ideas of this paper an example investigating changes only in the utility function would have been more helpful. Example 2 is a nice application in which imprecisions in both utility and prior are taken into consideration. Still, there remain many more features to explore, e.g., what difficulties arise with an uncountable set S, and how do data influence the analysis (recall, that the Bayes operator $T(u, P, a)$ depends implicitly on the data x as does the supremum of the Fréchet derivative).

I agree with the authors that future work would be welcome extending the results to other classes of priors, utilities and alternatives.

REFERENCES

CUEVAS, A. and SANZ, P. (1988). On differentiability properties of Bayes operators. In *Bayesian Statistics 3* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 569–577. Oxford University Press.

DEY, D. K., GHOSH, S. K. and LOU, K. (1995). On local sensitivity measures in Bayesian analysis (this volume).

INSTITUT FÜR STATISTIK UND ÖKONOMETRIE
UNIVERSITÄT BASEL
HOLBEINSTR. 12
4051 BASEL, SWITZERLAND

REJOINDER

J. MARTIN AND D. RIOS INSUA

We are grateful to Dr. Ickstadt for her neat discussion. We have always found fascinating how Bayesians have basically concentrated on issues regarding sensitivity to the prior. We have argued several times that robustness should be approached from a more global perspective considering joint sensitivity to various inputs to an analysis. Our initial studies with respect to the prior and the utility/loss are encouraging we believe.

Having said that we view partial sensitivity studies as first steps towards more global analysis. In that sense, and given space constraints, we had to limit the material in the paper.

As an example, Dr. Ickstadt mentions no operator norm for the case of imprecision in the utility only. With no constraints in the utility, the norm is 1, hence the result is not very useful. With constraints, we run into numerical problems as with the other cases and we have to undertake the approach in Section 3. That is why we provided no examples with sensitivity with respect to the utility only, since they are encompassed in the more interesting, and technically more difficult, example 2. Also, example 2 is relatively easy in that the set of states is finite. Numerical results are not much harder in the continuous case, since we have to appeal only to nonlinear and linear programming problems. For our settings there are many tools available for the solution of these problems. See e.g. Nemhauser et al (1989). Finally, the usual Bayesian argument support our consideration of only the observed data, and not those potentially observed. This last point might be of interest if a procedure is going to be used by several users, but leads us to the shaky ground of frequentist inference.

We view this paper as a first step for general local sensitivity analysis. Let us note that several of the discussants to Berger's (1994) review suggest that this type of joint studies deserve much more attention. We hope that more authors join us in the fun in the near future.

REFERENCES

NEMHAUSER, G., RINNOOY KAN, A. AND TOOD, J. (1989) *Optimization.* North Holland.