

K. Passarin

Local robustness measures for posterior summaries

2004/10



UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

<http://eco.uninsubria.it>

In questi quaderni vengono pubblicati i lavori dei docenti della Facoltà di Economia dell'Università dell'Insubria. La pubblicazione di contributi di altri studiosi, che abbiano un rapporto didattico o scientifico stabile con la Facoltà, può essere proposta da un professore della Facoltà, dopo che il contributo sia stato discusso pubblicamente. Il nome del proponente è riportato in nota all'articolo. I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

These Working papers collect the work of the Faculty of Economics of the University of Insubria. The publication of work by other Authors can be proposed by a member of the Faculty, provided that the paper has been presented in public. The name of the proposer is reported in a footnote. The views expressed in the Working papers reflect the opinions of the Authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

© Copyright K. Passarin
Printed in Italy in May 2004
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

All rights reserved. No part of this paper may be reproduced in any form without permission of the Author.

Local robustness measures for posterior summaries*

Katia Passarin †

Abstract

This paper deals with measures of local robustness for particular Bayesian quantities, i.e. posterior summaries. We build a framework where any Bayesian quantity can be seen as a posterior functional and its sensitivity to all inputs is checked. First, we use the Gateaux derivatives to measure the impact on posterior summaries of perturbations of prior or sampling models, giving some general expressions. Such quantities capture both a 'data effect' and a 'model effect' on the functional. Secondly, we check the sensitivity to one observation in the sample, once a particular combination of prior/sampling models has been chosen. Moreover, we propose a new estimator of the Bayes factor for practical implementation. Finally, illustrative examples on sensitivity analysis are provided and discussed.

1 Introduction

Any Bayesian quantity depends strongly on the modeling assumptions and on the sample of observed data. Bayesian Robust Statistics evaluates the sensitivity of this quantity to their inputs and in recent years it has met a great development (D. Ríos Insua and F. Ruggeri, 2000). Most efforts concentrate on global robustness, in particular with respect to prior specification. Such approach consists in calculating the range of the quantity of interest as the model varies within a class of distributions. If this range is small enough for the conclusions to be clear, the quantity is declared to be robust. If not, further analysis is needed. For more details on this see Lavine (1991), Berger (1994), Basu (1999), Sivaganesan (1999, 2000), Berger *et al.* (2000), Moreno (2000) and Shyamalkumar (2000).

A second approach - named local - assesses the sensitivity to deviations only in a neighborhood of the base model. Measures of local robustness are obtained by suitable derivatives of the functional (Ruggeri and Wasserman, 1993; Sivaganesan, 1993; Dey *et al.*, 1996; Gustafson *et al.*, 1996; Moreno *et al.*, 1996; Peña and Zamar, 1997). The functional is said to be robust if the computed measure is small.

*This paper has been presented at the University of Insubria (Como) on the 22th of April 2004 and its publication in the working paper series of the Dep. of Economics of the U. of Insubria has been proposed by Prof. Mira, University of Insubria.

†Università della Svizzera Italiana, Istituto di Finanza, Facoltà di Economia, Via Buffi 13, 6900, Lugano, Switzerland. Email: katia.passarin@lu.unisi.ch

Also in this case, most contributions are only concerned with local prior influence (Gustafson, 2000).

In this paper we deal with local robustness. It is interesting to note that the same approach is used in classic robust statistics (Hampel *et al.*, 1986). However the robustness perspective slightly differ in a frequentist and in a Bayesian context. We discuss this point in Section 2, introducing the concept of functional and looking at any Bayesian quantity as a function of three distinct elements (the prior, the sampling model and the data). Such point of view constitute a simple and unified framework for robustness evaluation in Bayesian statistics. In particular we consider the posterior expectation of a function $\rho(\theta)$, named posterior summary. The first goal of this paper is to check the sensitivity of posterior summaries to one input a time, all the rest remaining stable. Different diagnostic tools for distributional assumptions -named local influence measures- are derived in Section 3. Such measures capture the impact on the functional of contaminations of the base model in different directions. The sensitivity of a Bayesian functional to observations is addressed in Section 4. Section 5 deals with the matter of implementation of local influence measures when analytical calculations are not feasible. Starting from the work of Chen and Shao (1997), we propose a new estimator for the Bayes Factor which is more efficient in terms of computational time. Illustrative examples are given in Section 6 and Section 7 concludes.

2 Frequentist and Bayesian robustness

In this section we underline some common and different features of the robustness concept in a Bayesian and in a frequentist framework.

First it is worth introducing some notation. We will use capital letters for both a probability distribution and its corresponding cumulative distribution function. Moreover, we denote with small letters the corresponding density, when it exists. We consider i.i.d. one-dimensional random variables $X = (X_1, \dots, X_n)$ generated by a reference distribution F_{θ_0} , which belongs to the set $\tilde{F} = \{F_\theta : \theta \in \Theta\}$. Each observation in sample $x = (x_1, \dots, x_n)$ takes value in a sample space $\Xi \subseteq \mathbb{R}$.

We denote by $F_n(y) = \frac{1}{n} \sum_{i=1}^n \Delta_{x_i}(y)$ the empirical distribution where $\Delta_x(y)$ is the Dirach distribution which puts mass 1 at x . In a Bayesian setting we also define $\Pi(\theta)$ and $P(\theta|x)$ to be an element respectively of the set $\tilde{\Pi}$ of all possible priors and of the set \tilde{P} of all possible posteriors on the parameter space Θ .

In frequentist statistics observed data are used to make inference on the true parameter value θ_0 , which is assumed to be a fixed constant (Cox and Hinkley, 1974; Ellison, 1996). The approach of classical robust theory based on influence functions (Hampel *et al.*, 1986) deals with estimators that can be expressed as functionals i.e.

$$T : \tilde{F} \rightarrow \mathbb{R}^k.$$

It is required that the functional does not depend on the number of sample observations ($T_n(F_n) = T(F_n)$), it converges to the asymptotic value of the estimator ($T(F_n) \xrightarrow[n \rightarrow \infty]{} T(F_{\theta_0})$) and that Fisher consistency holds ($T(F_{\theta_0}) = \theta_0$).

Measures of robustness to small deviations from the reference model are obtained by computing the influence function (IF), which is the Gateaux derivative of the functional under a locally perturbed distribution in direction of a point mass. Therefore the evaluation of robustness properties of the estimator occurs at an asymptotic level. In the sample one can calculate some empirical version of the IF such as the Empirical Influence Function and the Sensitivity Curve.

In Bayesian statistics the parameter θ is not a fixed quantity, but a random variable, whose entire probability distribution have to be computed (Ellison, 1996). Two distributions are matched with the observed data: Π that represents our knowledge *a priori* on θ and F_θ that expresses the parametric model we believe generated observations x . Using the Bayes theorem, the posterior distribution for parameter θ is obtained:

$$\begin{aligned} P(\theta|x) &= \frac{\Pi(\theta)L_F(x|\theta)}{m(x; \Pi, F_\theta)} \\ &= \frac{\tilde{P}(\theta|x)}{m(x; \Pi, F_\theta)}, \end{aligned} \quad (1)$$

where $L_F(x|\theta) = \prod_i f_\theta(x_i)$ is the likelihood and $m(x; \Pi, F_\theta) = \int \tilde{p}(\theta|x) d\theta$ is the marginal likelihood. Inferential conclusions on the value of θ are based on (1).

Any Bayesian quantity can be expressed as a functional of type

$$T_B : \tilde{F}_n \times \tilde{\Pi} \times \tilde{F} \rightarrow \Upsilon,$$

where $\tilde{F}_n = \{\text{all discrete distributions with probability } p_1, \dots, p_n \text{ at the points } x_1, \dots, x_n, p_i \geq 0, \sum_i p_i = 1\}$ and Υ is a suitable space. For example, one can be interested in the entire posterior distribution ($\Upsilon = \tilde{P}$) or in some posterior summaries ($\Upsilon = \mathbb{R}^k, k \geq 1$).

When the number of observations increases, the impact of Π on (1) disappears since the likelihood dominates the prior distribution and the posterior collapses to a point mass on the true parameter value θ_0 . Therefore, Bayesian functionals satisfy $T_B(F_n, \Pi, F_\theta) \xrightarrow[n \rightarrow \infty]{} T(F_{\theta_0})$. Asymptotic functionals do not allow to capture the sensitivity of posterior quantities to perturbations in the prior. Hence, we will work with sample-based functionals. In particular we will focus on robustness evaluation for posterior summaries of type

$$T_B(F_n, \Pi, F_\theta) = \int \rho(\theta) p(\theta|x) d\theta. \quad (2)$$

In the sequel we will in short denote T_B and $m(x)$ respectively the posterior summary and the marginal likelihood under base models Π and F_θ .

3 Sensitivity to distributional assumptions

In this section we deal with the sensitivity of a Bayesian estimator to small departures from the assumed model, either the prior or the sampling distribution. In order

to simplify the notation we will denote the posterior functional only as a function of the distribution under study, say generically distribution H , keeping the remainder unchanged. We represent these deviations through ε -contamination classes of type:

$$I_\varepsilon(H) = \left\{ H_\varepsilon = (1 - \varepsilon)H + \varepsilon C \mid 0 \leq \varepsilon \leq 1, C \in \tilde{C} \right\}. \quad (3)$$

Set (3) represents the perturbation of reference distribution H in the direction of C and ε is the contamination amount (assumed to be small in local analysis). Clearly, the wider the set of contaminating distribution \tilde{C} is, the richer the neighborhood we are considering. As in Sivaganesan (1993) and Peña and Zamar (1997), we measure the impact of such contaminations on functional (2) by the Gateaux derivative:

$$\begin{aligned} LI(C; T_B, H) &= \left[\frac{\partial T_B(H_\varepsilon)}{\partial \varepsilon} \right]_{\varepsilon=0} \\ &= \int \rho(\theta) \left[\frac{\partial p_\varepsilon(\theta|x)}{\partial \varepsilon} \right]_{\varepsilon=0} d\theta. \end{aligned} \quad (4)$$

We refer to this quantity as *local influence* (LI) of T_B when H is perturbed in the direction of C . Note that measure (4) is a sample-based quantity. We will see in a while that it captures both a 'data effect', i.e. the effect on the functional of choosing a contaminating model which is more adequate than the base one with respect to observed data, and a 'model effect', i.e. the effect on the functional value of perturbing the base model in some directions. The strong dependence of measure (4) on the sample is the reason why Sivaganesan (1993) looks at it only to compare whether a functional is more sensible to prior or sampling model specifications and does not judge about its magnitude. For this purpose we define

$$LI^*(\tilde{C}; T_B, H) = \sup_{C \in \tilde{C}} \left| \frac{LI(C; T_B, H)}{T_B(H)} \right|, \quad (5)$$

which gives the maximum relative effect on the functional as the distribution moves locally around H in different directions. Measure (5) evaluates the magnitude of the sensitivity of the functional and can be used to compare robustness properties among different functionals. In the following sections we derive local influence measures for both the prior and the sampling model.

3.1 Prior distribution

Many papers in Bayesian robustness are concerned with the assessment of the sensitivity with respect to the prior (Ruggeri and Wasserman, 1993; Gustafson *et al.*, 1996; Moreno *et al.*, 1996; Peña and Zamar, 1997). The main reason for this widespread interest is probably due to the feeling that prior knowledge formalized by the researcher is the most subjective source of the analysis. Much work has been done in the direction of global robustness. A good review on the topic is provided by Berger (1994).

Local robustness assesses effects of small prior perturbations on the functional. We consider a neighborhood of the base prior Π of type (3), with Q the contaminating distribution. The local influence of T_B when Π is perturbed in the direction of Q is given by:

$$\begin{aligned}
LI(Q; T_B, \Pi) &= \left[\frac{\partial T_B(\Pi_\varepsilon)}{\partial \varepsilon} \right]_{\varepsilon=0} \\
&= \left[\int \rho(\theta) \frac{\partial}{\partial \varepsilon} \left(\frac{L_F(x|\theta) \pi_\varepsilon(\theta)}{m(x; \Pi_\varepsilon, F_\theta)} \right) d\theta \right]_{\varepsilon=0} \\
&= \int \rho(\theta) \frac{L_F(x|\theta) [q(\theta) - \pi(\theta)]}{m(x)} d\theta + \\
&\quad + \int \rho(\theta) \frac{[m(x; Q, F_\theta) - m(x)] L_F(x|\theta) \pi(\theta)}{m(x)} d\theta \\
&= \frac{m(x; Q, F_\theta)}{m(x)} [T_B(Q) - T_B], \tag{6}
\end{aligned}$$

where $m(x; Q, F_\theta)$ and $T_B(Q)$ are respectively the marginal likelihood and the posterior summary obtained when the prior is Q . Measure (6) depends on two factors. The first is the ratio of marginal likelihoods under contaminating and base distribution respectively (Bayes factor). This can be regarded as a measure of data supporting degree for different contaminating priors that compares the researcher's subjectivity and the objectiveness of the data. If this amount is greater (smaller) than one, data may be said to support more (less) the contaminating prior than the base one. For this reason the Bayes factor can be said to capture a 'data effect' on the functional. The second factor is the difference between the functional value computed under the contaminating and the base prior respectively. It captures the effect on the functional of choosing a different model for the prior and we refer to this as 'model effect'. If the value of $T_B(Q)$ is much different from the value of T_B , the model effect turns out to be big. However, note that such effect in measure (6) is weighted by the corresponding Bayes factor. Therefore the total effect on the functional of contaminations in the direction of Q will be big itself only if Q will be supported by data more than Π . In the next section we consider the sampling model.

3.2 Sampling distribution

Another source of possible misspecification is the data-generating model. Robustness with respect to sampling model specification is referred in the literature as *model or likelihood robustness*. In most scenarios inference will depend much more heavily on the model than on the prior (see Section 2). However, few contributions in assessing likelihood robustness can be found in the literature (see Sivaganesan, 1993; Dey et al., 1996; Gustafson, 1996; Shyamalkumar, 2000).

This fact can be explained by considering the non linearity of the posterior with respect to the sampling distribution. Indeed when regarded as a function of the prior, (1) is a ratio of two linear functionals, or briefly is said to be *ratio-linear*. This is not true when considered as a function of the sampling model, as the sampling density

enters through the likelihood function. This often leads to intractable global analysis from an analytical point of view. However, in local analysis this problem can be tackled by taking the derivative with respect to the quantity of contamination ε when ε is small.

Assume we represent uncertainty about the base sampling model F_θ by (3) with G the contaminating distribution. The obtained perturbed likelihood will be differently combined with the prior according to the information G brings on θ .

If G is a distribution still governed by parameter θ , we denote the contaminating distribution by G_θ . For example G_θ can be an unimodal distribution around θ . In this case the local influence of T_B when F_θ is perturbed in the direction of G_θ is given by

$$\begin{aligned} LI(G_\theta; T_B, F_\theta) &= \left[\frac{\partial T_B(F_{\theta, \varepsilon})}{\partial \varepsilon} \right]_{\varepsilon=0} \\ &= \left[\int \rho(\theta) \frac{\partial}{\partial \varepsilon} \left(\frac{L_{F_\varepsilon}(x|\theta) \cdot \pi(\theta)}{m(x; \Pi, F_{\theta, \varepsilon})} \right) d\theta \right]_{\varepsilon=0} \\ &= \sum_j \frac{m_j(x; \Pi, F_\theta, G_\theta)}{m(x)} [T_{B,j}(F_\theta, G_\theta) - T_B], \end{aligned} \quad (7)$$

where

$$m_j(x; \Pi, F_\theta, G_\theta) = \int \tilde{p}_j(\theta|x) d\theta$$

and

$$T_{B,j}(F_\theta; G_\theta) = \frac{\int \rho(\theta) \tilde{p}_j(\theta|x) d\theta}{m_j(x; \Pi, F_\theta, G_\theta)}$$

are respectively the marginal likelihood and the posterior functional obtained when the sampling distribution is G_θ only for observation x_j and F_θ for the others, the quantity \tilde{p}_j is defined as

$$\tilde{p}_j(\theta|x) = g_\theta(x_j) L_F(x_{(-j)}|\theta) \pi(\theta),$$

and $x_{(-j)}$ is the sample x without observation x_j .

If G does not depend on θ we denote the contaminating distribution by G_η . The local influence of T_B when F_θ is perturbed in the direction of G_η is then given by:

$$\begin{aligned} LI(G_\eta; T_B, F_\theta) &= \left[\frac{\partial T_B(F_{\theta, \varepsilon})}{\partial \varepsilon} \right]_{\varepsilon=0} \\ &= \left[\int \rho(\theta) \frac{\partial}{\partial \varepsilon} \left(\frac{L_{F_\varepsilon}(x|\theta) \cdot \pi(\theta)}{m(x; \Pi, F_{\theta, \varepsilon})} \right) d\theta \right]_{\varepsilon=0} \\ &= \sum_j \frac{m_j(x; \Pi, F_\theta, G_\eta)}{m(x)} [T_B^{(-j)} - T_B], \end{aligned} \quad (8)$$

where $m_j(x; \Pi, F_\theta, G_\eta) = g_\eta(x_j) \cdot m(x_{(-j)}; \Pi, F_\theta)$ and $m(x_{(-j)}; \Pi, F_\theta)$ and $T_B^{(-j)}$ are respectively the marginal likelihood and the posterior functional under base models using sample $x_{(-j)}$. For detailed calculations see Appendix 1.

For any observation x_j the local influence measure for the sampling distribution is still a function of two factors and it captures both a 'data effect' and a 'model effect'. The Bayes factor plays the important role of increasing (decreasing) the difference when data support (do not support) the contaminating distribution more than the base distribution for observation j ('data effect'). The second factor is the difference between the value of the functional computed when model G is assumed only for observation x_j and the base functional T_B . Note that observation x_j enters in the calculation of the former value only if G depends on θ , i.e. if x_j has something to say on the parameter of interest. Otherwise, x_j cannot give any information for updating our prior knowledge and the resulting functional has the form of the base one where one observation has been dropped out. The total effect on the functional of perturbations of the sampling model turns out to be the sum of the effect for each observation.

4 Sensitivity to observations

In the previous section we assess the influence on posterior summaries of a perturbation of the assumed model in some direction. In this section we measure the influence of a given observation in the sample (*outlier robustness*). It is worth stressing the difference between model robustness and outlier robustness. Model robustness evaluates the impact on the functional of a small contamination of the base sampling model (see section 3.2). Outlier robustness evaluates the effect of moving one observation in the sample once prior and sampling distributions are fixed. In this section we still denote the Bayesian functional as a function of the distribution under study, i.e. the empirical distribution.

Little attention has been paid in Bayesian literature to the impact of outliers and mainly focused on the posterior distribution. Ramsay and Novick (1980), for example, propose to look at the rate of change of the sampling model density with respect to an observation value. A similar idea is used by West (1984) on Bayesian regression. However such approach is hardly applicable because involves derivatives which are difficult to compute apart from particular family of distributions. The same problem is addressed by Chen and Fournier (1999). Their influence measure summarizes the difference between posterior distributions computed with original data and with an additional observation. Such posterior distributions are obtained through the use of numerical techniques and therefore always applicable.

In this paper, however, we do not deal with posterior distributions directly, but with posterior summaries. Studying the sensitivity of such a quantity to observations is a well known matter in frequentist robust statistics. The right tool therefore is the Sensitivity Curve (see Hampel *et al.*, 1986), defined as

$$SC(z) = \frac{[T_B(F_n^z) - T_B(F_{n-1})]}{\frac{1}{n}}, \quad (9)$$

where $F_{n-1} = (x_1, \dots, x_{n-1})$ is the empirical distribution of the sample of $(n - 1)$ observations and $F_n^z = (x_1, \dots, x_{n-1}, z)$ is the sample in which observation z has been added. In a Bayesian context this measure captures the influence of moving

just one observation under a certain prior/sampling model combination. If this measure diverges as z becomes bigger, the functional is said to be non robust with respect to observations. Typically this curve is useful to identify observations with a large influence, such as outliers and loosely speaking an outlier is defined to be an observation that is unlikely to have been generated by the assumed sampling model. For its simple definition (9) can be implemented even when analytical calculations are not feasible by means of numerical algorithms.

In the next section we will discuss practical implementation of local sensitivity measures derived in the previous sections when analytical results are not feasible.

5 Implementation of local sensitivity measures

Posterior distribution and local influence measures are analytically tractable when conjugate prior and sampling models are assumed. However, often this is not the case and we need to use numerical procedures to compute them. Typically MCMC algorithms are used to generate a sample from complicated distributions. Local influence measures can be then easily obtained by estimating the Bayes factor and the functionals under base and contaminating distributions. In this section we concentrate on implementation of (7) by means of Metropolis-Hastings algorithm and we propose a way to speed up its computation.

Local influence measures for the sampling distribution involve the computation of Bayes factors and of posterior summaries (see Section 3.2). We first deal with the estimation of the former quantity (shortly denoted by r_j), which is given by

$$\begin{aligned} r_j &= \frac{m_j(x; \Pi, F_\theta, G)}{m(x)} \\ &= \frac{\int \tilde{p}_j(\theta|x) d\theta}{\int \tilde{p}(\theta|x) d\theta}. \end{aligned} \tag{10}$$

Different bridge estimators (Meng and Wong, 1996; Chib and Jeliazkov, 2001; Mira and Nicholls, 2001) are possible solutions. However, to compute such local influence measures we would be expected to run $n + 1$ simulations, where n is the number of observations. Clearly, the estimation procedure will take a long time when n is big.

We need a way to be more efficient in terms of computational time. A good starting point is the two-stage estimator proposed by Chen and Shao (1997). Ratio (10) can be written as

$$r_j = \frac{\int \frac{\tilde{p}_j(\theta|x)}{\xi(\theta)} \xi(\theta) d\theta}{\int \frac{\tilde{p}(\theta|x)}{\xi(\theta)} \xi(\theta) d\theta}, \tag{11}$$

where $\xi(\theta)$ is an arbitrary importance sampling density. When observations are i.i.d. from ξ , the importance density which minimizes the relative mean square error of

the estimator is given by

$$\begin{aligned}\xi_j^{opt}(\theta) &= \frac{|p_j(\theta|x) - p(\theta|x)|}{\int |p_j(\theta|x) - p(\theta|x)| d\theta} \\ &= \frac{|\tilde{p}_j(\theta|x) - r_j \cdot \tilde{p}(\theta|x)|}{\int |\tilde{p}_j(\theta|x) - r_j \cdot \tilde{p}(\theta|x)| d\theta},\end{aligned}\quad (12)$$

where $p_j = \tilde{p}_j/m_j$ and $p = \tilde{p}/m$.

The corresponding estimator \hat{r}_j^{opt} is implemented in two stages. First, a Monte Carlo estimate of (11) is computed with a random sample from an arbitrary distribution. Then a random draw from (12) can be obtained by means of a MCMC simulation. One advantage of \hat{r}_j^{opt} is that its estimate is available with a single random sample from ξ_j^{opt} rather than two samples respectively from p_j and p . However, we are still expected to generate n samples to compute (7).

In order to run a single MCMC simulation we propose to use an importance sampling density with a form similar to the optimal one, but which does not depend on j . Such a density is given by

$$\xi^*(\theta) = \frac{|\tilde{p}^*(\theta|x) - r^* \cdot \tilde{p}(\theta|x)|}{\int |\tilde{p}^*(\theta|x) - r^* \cdot \tilde{p}(\theta|x)| d\theta},\quad (13)$$

where $\tilde{p}^*(\theta|x) = \frac{1}{n} \sum_{j=1}^n \tilde{p}_j(\theta|x)$ and $r^* = \frac{\int \tilde{p}^*(\theta|x) d\theta}{\int \tilde{p}(\theta|x) d\theta}$. Figure 1 compares density (13) with the posterior densities p and p'_j 's. The sampling density displays fatter tails which is a crucial characteristic for a good importance sampling. The corresponding modified two-stages estimator is given by

$$\hat{r}_j^* = \frac{\sum_{i=1}^{n_\xi} \frac{\tilde{p}_j(\theta_i|x)}{\xi^*(\theta_i|x)}}{\sum_{i=1}^{n_\xi} \frac{\tilde{p}(\theta_i|x)}{\xi^*(\theta_i|x)}},\quad (14)$$

where $[\theta_i]_{i=1}^{n_\xi}$ is the output of a MCMC simulation for (13). We tested the performance of the new estimator by running $K = 30$ independent simulations of length s ($s = 1000, 2000, \dots, 5000$) under the normal sampling model. For each chain we estimate (14) and we compute its mean value with the corresponding confidence interval. Figure 2 shows that estimator (14) behaves well with a mean value of \hat{r}_j^* close to the analytical value and smaller variability with increasing number of simulations.

To estimate the local influence measure for the sampling distribution, we still need to compute T_B and $T_{B,j}(F_\theta, G)$. The former quantity can be obtained by running a MCMC simulation for posterior p . The latter can be obtained using importance sampling technique with ξ^* as importance density. Finally, measure (7) can be written as

$$LI(G; T_B, F_\theta) = \sum_j \frac{m_j(x; \Pi, F_\theta, G)}{m(x)} [T_{B,j}(F_\theta, G) - T_B]$$

$$\begin{aligned}
&= \sum_j r_j \left[\int \rho(\theta) p_j(\theta|x) d\theta - \int \rho(\theta) p(\theta|x) d\theta \right] \\
&= \sum_j r_j \cdot \left[\frac{m_\xi}{m_j} \cdot \int \rho(\theta) \frac{\tilde{p}_j(\theta|x)}{\tilde{\xi}^*(\theta|x)} \xi^*(\theta|x) d\theta - \int \rho(\theta) p(\theta|x) d\theta \right] \\
&= \sum_j \left[r_\xi \cdot \int \rho(\theta) \frac{\tilde{p}_j(\theta|x)}{\tilde{\xi}^*(\theta|x)} \xi^*(\theta|x) d\theta - r_j \cdot \int \rho(\theta) p(\theta|x) d\theta \right], \quad (15)
\end{aligned}$$

where $r_\xi = \frac{m_\xi}{m}$. Denoting by $[\theta_s]_{s=1}^{n_p}$ and $[\theta_i]_{i=1}^{n_\xi}$ respectively the samples from $p(\theta|x)$ and from $\xi^*(\theta)$, the ratio r_ξ can be estimated using optimal Meng and Wong's bridge estimator given by

$$\hat{r}_\xi^{t+1} = \frac{\frac{1}{n_p} \sum_{s=1}^{n_p} \frac{\tilde{\xi}^*(\theta_s)}{n_\xi \cdot \tilde{\xi}^*(\theta_s) + n_p \cdot \hat{r}_\xi^t \cdot \tilde{p}(\theta_s)}}{\frac{1}{n_\xi} \sum_{i=1}^{n_\xi} \frac{\tilde{p}(\theta_i)}{n_\xi \cdot \tilde{\xi}^*(\theta_i) + n_p \cdot \hat{r}_\xi^t \cdot \tilde{p}(\theta_i)}}.$$

An estimate of (15) is then obtained as

$$\widehat{LI}(G; T_B, F_\theta) = \sum_{j=1}^n \left[\hat{r}_\xi \left(\frac{1}{n_\xi} \sum_{i=1}^{n_\xi} \rho(\theta_i) \frac{\tilde{p}_j(\theta_i|x)}{\tilde{\xi}^*(\theta_i|x)} \right) - \hat{r}_j^* \left(\frac{1}{n_p} \sum_{s=1}^{n_p} \rho(\theta_s) \right) \right].$$

In the next section we will provide some examples of how performing a Bayesian sensitivity analysis.

6 Examples of local sensitivity analyses

In the following simple examples we perform sensitivity analyses of the functional of interest. We keep the same notation as in previous sections. We first consider the Bayes estimator given by the mean of the posterior distribution. For this example we simulate a sample of $n = 3$ observations from a standard univariate normal given by (0.5375, 1.4221, 1.0946). Then we consider a Bayesian regression model using real data. In both case we perform conjugate analyses in order to obtain analytical results.

6.1 Posterior mean

The posterior mean is a frequently used estimator of the parameter of interest. We now illustrate how a sensitivity analysis on this functional can be carried out. We assume that prior Π is $N(\theta_0, \sigma_0^2)$ with $\theta_0 = 0.5$ and $\sigma_0^2 = 1$. Moreover sampling distribution F_θ is $N(\theta, \sigma^2)$ with $\sigma^2 = 0.2$. The posterior mean and the marginal likelihood can be computed analytically and turn out to be respectively

$$T_B = \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2} \bar{x} + \frac{\sigma^2}{n\sigma_0^2 + \sigma^2} \theta_0.$$

and

$$m(x) = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-1)}{2}} (n\sigma_0^2 + \sigma^2)^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 \right\} \exp \left\{ -\frac{n(\theta_0 - \bar{x})^2}{2(n\sigma_0^2 + \sigma^2)} \right\}.$$

First, we assume to be not very confident about the value of prior mean θ_0 . We express our uncertainty through the set of possible contaminating prior distribution $\tilde{Q} = \{N(\lambda, \sigma_0^2) : \lambda \in [-4.5, 5.5]\}$. In this case the local influence measure is given by (6) with

$$T_B(Q) = \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2} \bar{x} + \frac{\sigma^2}{n\sigma_0^2 + \sigma^2} \lambda$$

and

$$m(x; Q, F_\theta) = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-1)}{2}} (n\sigma_0^2 + \sigma^2)^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 \right\} \exp \left\{ -\frac{n(\lambda - \bar{x})^2}{2(n\sigma_0^2 + \sigma^2)} \right\}.$$

Table 1 and Figure 3 show such a measure for different values of σ_0^2 . The magnitude of LI decreases with increasing prior variances, meaning that flatter priors are less influenced by perturbations. The two factors of measure (6) are displayed in Figure 4. The effect on the functional of choosing prior Q instead of prior Π ('model effect') is linear and smaller with decreasing prior precision. Moreover, priors with α_0 around the value of the sample mean ($\bar{x} = 1.01$) appear to be more adequate than Π for small value of σ_0^2 . As long as the base prior becomes flatter, the Bayes factor approaches to 1 for all possible contaminating distributions.

We turn now to the sampling model. We account for perturbations of the base distribution in the direction of flatter ones. The chosen contaminating set is $\tilde{G}_\theta = \{N(\theta, \eta^2) : \eta^2 \in [0.2, 2]\}$. Clearly this contamination is quite restrictive, but it leads to analytical results. LI measure for the sampling model is given by (7) with

$$m_j(x; \Pi, F_\theta, G_\theta) = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-2)}{2}} (\sigma^2 \eta^2 + (n-1)\eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i \neq j} (x_i - \bar{x}_{(-j)})^2 - \frac{\sigma^2 (x_j - \theta_0)^2}{2(\sigma^2 \eta^2 + (n-1)\eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)} \right\} \cdot \exp \left\{ -\frac{(n-1)\eta^2 (\bar{x}_{(-j)} - \theta_0)^2 + (n-1)\sigma_0^2 (\bar{x} - \bar{x}_{(-j)})^2}{2(\sigma^2 \eta^2 + (n-1)\eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)} \right\},$$

where $\bar{x}_{(-j)}$ is the mean of the sample without observation x_j . Calculations can be found in Appendix 2.

Table 2 and Figure 5 show measures (7) for different values of σ^2 . LI measure is very small when $\sigma^2 = 0.2$, which corresponds to the value of the sample variance, and LI^* shows its minimum value which is around 0.009. As long as σ^2 moves away

from 0.2, LI^* increases up to around 0.065. To better understand such a result, each row of Figure 6 plots the two factors of measure (7) for observation j ($j = 1, 2, 3$). The 'model effect' on the functional is increasing with increasing variance of the contaminating model, but it is no longer linear as in the prior case. When $\sigma^2 = 0.1$ or $\sigma^2 = 0.2$, data support at least few contaminating models more than the base one. This is not true in other cases where the the Bayes factor declines rapidly. Therefore the plot of the Bayes factor helps also to check whether the assumed sampling model is reasonable with respect to the data we have in the hand.

Comparing now the two bold columns in Table 1 and Table 2, we conclude that with these data the posterior mean is more sensible to perturbations in the prior model specification ($LI^* (\tilde{Q}; T_B, \Pi) = 0.0702 > LI^* (\tilde{G}_\theta; T_B, F_\theta) = 0.0096$). However both measures are small and the estimate is judge locally robust with respect to our distributional assumptions.

Finally Figure 7 plots the $SC(z)$. We let observation z move in the range $[-5, 5]$. The effect of an extreme observation on the posterior mean with a normal prior/normal sampling model combination is linear and therefore potentially unbounded. Hence, it is crucial to assess whether some extreme observations are present in the sample. We expect that in such a case measure (7) increases since data would support sampling models with higher variance more than the base one and model effect would also display a greater value. In order to investigate this point we introduce the observation $x_4 = -5$ in the sample and we compute LI measures again. Results given in Table 3 support our hypothesis. Therefore in presence of outliers measure (7) takes into account the fact that the normal distribution becomes inadequate.

6.2 Linear Bayesian Regression

We now consider the Bayesian linear model $y = X\beta + u$. For simplicity, we assume that the error distribution F is a $N(0, \sigma^2 I)$ with known variance σ^2 . We further adopt a normal prior distribution $\Pi(\beta)$ of type $N(\beta_0, \sigma^2 \Sigma_0)$. Under the assumed models, the Bayes estimator of β is given by

$$\hat{\beta}_{Bayes} = (\Sigma_0^{-1} + X'X)^{-1} (\Sigma_0^{-1}\beta_0 + X'y).$$

If \tilde{Q} is the family $\{N(\alpha_0, \sigma^2 \Sigma_0) : \alpha_0^{\text{inf}} \leq \alpha_0 \leq \alpha_0^{\text{sup}}\}$ that accounts for uncertainty in the prior mean, measure (6) is given by

$$LI(Q; T_B, \Pi) = \exp \left\{ -\frac{(\alpha_0 - \beta_0)' [\Sigma_0^{-1} - \Sigma_0^{-1} V' \Sigma_0^{-1}] (\alpha_0 - \beta_0)}{2\sigma^2} \right\} \cdot \left[(X'X + \Sigma_0^{-1})^{-1} \Sigma_0^{-1} (\alpha_0 - \beta_0) \right]. \quad (16)$$

Furthermore, assuming a contaminating family \tilde{G} for the sampling distribution

of type $\{N(0, c^2) : c^{\text{inf}} \leq c^2 \leq c^{\text{sup}}\}$, measure (7) becomes

$$LI(G; T_B, F) = \sum_{j=1}^n \left(\frac{c^2 |V|}{\sigma^2 |V_j|} \right)^{-\frac{1}{2}} \cdot \exp \left\{ - \frac{\left(\frac{\sigma^2}{c^2} - 1 \right) y_j^2 + \widehat{\beta}'_{Bayes} V^{-1} \widehat{\beta}_{Bayes}}{2\sigma^2} \right\} \\ \cdot \exp \left\{ \frac{\widehat{\beta}_{Bayes}^{(j)'} V_j^{-1} \widehat{\beta}_{Bayes}^{(j)}}{2\sigma^2} \right\} \cdot \left(\widehat{\beta}_{Bayes}^{(j)} - \widehat{\beta}_{Bayes} \right), \quad (17)$$

where c^2 is the variance of the contaminating distribution, $V_j = \left[X'_{(-j)} X_{(-j)} + \frac{\sigma^2}{c^2} x_j x_j' + \Sigma_0^{-1} \right]^{-1}$ and $\widehat{\beta}_{Bayes}^{(j)} = V_j \cdot \left(X'_{(-j)} y_{(-j)} + \frac{\sigma^2}{c^2} x_j y_j + \Sigma_0^{-1} \beta_0 \right)$ are respectively the posterior variance and mean when distribution G is assumed only for observation j , x_j' is the row of matrix X corresponding to observation j , $X_{(-j)}$ and $y_{(-j)}$ are respectively matrix X and vector y without observation j . For detailed calculations see Appendix 3.

Relative measures of local influence are given respectively by

$$LI^* \left(\widetilde{Q}; T_B; \Pi \right) = \sup_{Q \in \widetilde{Q}} \left| \text{diag}^{-1} \left(\widehat{\beta}_{Bayes} \right) \cdot LI(Q; T_B, \Pi) \right|$$

and

$$\sqrt{LI^* \left(\widetilde{G}; T_B; F \right)} = \sup_{G \in \widetilde{G}} \left| \text{diag}^{-1} \left(\widehat{\beta}_{Bayes} \right) \cdot LI(G; T_B, F) \right|.$$

Bayesian estimation and local influence measures in the normal linear model are now illustrated. We use the same data set employed by Ramsay and Novick (1980). These are observations on 29 children on 3 psychological variables: a test of verbal intelligence (VI), a test of performance intelligence (PI) and $\sin^{-1}(\sqrt{p_i})$, where p_i is the proportion correct on a dichotic listening task (DL). We regress DL on remaining variables including a constant term. β_1 and β_2 are the coefficient corresponding to VI and PI respectively, whereas β_3 is the intercept. We also adopt the same values for both prior parameters and sampling variance which have been discussed at length by the authors. Analytical Bayes estimate of regression coefficients $\widehat{\beta}_{Bayes}$ equals $(0.7458, -0.0734, 38.3505)'$.

Plots of measure (16) and (17) are shown in Figure 8 and 9. Each component of contaminating prior mean α_0 varies in the range $(-2, 2)$ with respect to the corresponding component of β_0 . The impact on the Bayes estimate of contaminations in the prior is negligible. However, this is probably more a proof of the disappearing impact of the prior as the number of observations increases than a sign of robustness itself. Contaminating variance c^2 moves in the range $(\sigma^2, 10 \cdot \sigma^2)$. Perturbations of the sampling distribution play an important role on the estimates. The effect seems more pronounced for intercept β_3 , but relative measures of Table 4 reveal a stronger impact for β_2 . The size of LI^* measure for the sampling model is not negligible at all. In this case a small contamination with a flatter normal distribution leads to quite a big effect on coefficient estimates. The big size of LI measure for the sampling distribution reveals that the normal model does not fit data very well.

We now concentrate on the sensitivity to observations. We move the value of the first two regressors in the range¹ (65, 135) as represented by asterisks in Figure 10

¹This interval represents the theoretical values of the regressors.

and we look at the effect on the estimates. Figure 11 measures whether the added observation is an influential point through the Cook's distance. As the value moves away from the mean value of the regressors ($\overline{VI} = 99.75$ and $\overline{PI} = 104.89$), the added point becomes more and more influential. The same pattern is found in Figure 12 where the SC of β is displayed. Coefficient estimates are strongly dependent on the value of just one observation. In normal regression, hence, coefficients turn out to be so sensible that we do not necessary have to observe "extreme" value before estimates are influenced.

7 Conclusive remarks

In this paper we construct a framework to perform the sensitivity analysis of any Bayesian quantity to all inputs. Past literature on the field checked the sensitivity mainly to the prior distribution only. In our framework the sensitivity to all inputs is considered, giving the whole picture of robustness properties of the functional itself. We concentrate on posterior summaries and we measure the impact of perturbations of prior or sampling models in different direction by local influence measures. Such impact is the product of two effects: a 'data effect', i.e. the effect on the functional of choosing a contaminating model which is more adequate than the base one with respect to observed data, and a 'model effect', i.e. the effect on the functional value of perturbing the base model in some directions. In some special cases we also derive analytical formulations for these quantities. Local influence measure for the prior model decreases with flatter (less informative) prior and with increasing number of observations. However, the latter is probably simply an effect of the disappearing impact of the prior as the number of observations increases.

Then we check the sensitivity of a Bayesian functional to observations by means of the Sensitivity Curve. Typically this curve is useful to identify observations with a large influence, such as outliers and loosely speaking an outlier is defined to be an observation that is unlikely to have been generated by the assumed sampling model. Therefore when the influence on the functional of a single observation is potentially unbounded, it is crucial to determine whether some outliers are present in the sample. We show that the local influence measure for the sampling model can be used for this purpose. In this case, indeed, it assumes huge values revealing that base sampling model is very sensible to perturbations and hence probably inadequate for the presence of some outlying observations. Further research in this direction includes the computation of measure of sensitivity to more than one input a time.

Finally we deal with the issue of practical implementation. We concentrate on the local influence measure for the sampling model and we propose a new estimator for the Bayes factor which speeds up computations. Such estimator performs well, giving precise estimates with small confidence intervals. Further developments could be also in more specific and efficient estimators for the quantities involved in local influence measures.

References

- [1] Basu, S. (1999), Posterior sensitivity to the sampling distribution and the prior: more than one observation, *Ann. Inst. Statist. Math.*, **51**, 3, 499-513.
- [2] Berger, J. O. (1994), An Overview of Robust Bayesian Analysis (with discussion), *Test*, **3**, 1, 5-124.
- [3] Berger, J. O., Rios Insua, J., Ruggeri, F. (2000), Bayesian robustness, in *Robust Bayesian Analysis* (D. Rios Insua and F. Ruggeri, eds), Springer-Verlag, New York.
- [4] Chen, M. H., Shao, Q. M. (1997), On Monte Carlo methods for estimating ratios of normalizing constants, *The Annals of Statistics*, **25**, 1563-1594.
- [5] Chen, Y., Fournier, D. (1999), Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries, *Canadian Journal of Fisheries and Aquatic Sciences*, **56**, 1525-1533.
- [6] Chib, S., Jeliazkov, I. (2001), Marginal likelihood from the Metropolis-Hastings output, *Journal of the American Statistical Association*, **96**, 270-281.
- [7] Cox, D. R., Hinkley, D. V. (1974), *Theoretical Statistics*, Chapman & Hall, London.
- [8] Dey, D. K., Ghosh, S. K., Lou, K. (1996), On local sensitivity measures in Bayesian analysis (with discussion), in *Bayesian Robustness* (J. O. Berger et al., eds), IMS Lecture Notes - Monograph Series, **29**, Hayward: IMS.
- [9] Ellison, A. M. (1996), An introduction to Bayesian inference for ecological research and environmental decision-making , *Ecological Applications*, **6**, 1036-1046.
- [10] Gustafson, P. (1996), Model influence functions based on mixtures, *The Canadian Journal of Statistics*, **24**, 4, 535-548.
- [11] Gustafson, P., Srinivasan, C., Wasserman, L. (1996), Local Sensitivity Analysis, *Bayesian Statistics 5*, Oxford University Press, Oxford.
- [12] Gustafson, P. (2000), Local robustness in Bayesian analysis, in *Robust Bayesian Analysis* (D. Rios Insua and F. Ruggeri, eds), Springer-Verlag, New York.
- [13] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel, W. A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.
- [14] Lavine, M. (1991), Sensitivity in Bayesian statistics: the prior and the likelihood, *Journal of the American Statistical Association*, **86**, 414, 396-399.
- [15] Meng, X., Wong, W. (1996), Simulating ratios of normalizing constants via simple identity: a theoretical exploration, *Statistica Sinica*, **6**, 831-860.

- [16] Mira, A., Nicholls, G. (2001), Bridge Estimation of the probability density at a point, *Technical report* n. 7, University of Insubria, Varese, Italy.
- [17] Moreno, E., Martínez, C., Cano, J. A. (1996), Local robustness and influence of contamination classes of prior distributions (with discussion), in *Bayesian Robustness* (J. O. Berger et al., eds), IMS Lecture Notes - Monograph Series, **29**, 137-154. Hayward: IMS.
- [18] Peña, D., Zamar, R. (1997), A simple diagnostic tool for local prior sensitivity, *Statistics and Probability Letters*, **36**, 205-212.01-907.
- [19] Ríos Insua, D., Ruggeri, F. (2000), *Robust Bayesian Analysis*, Springer-Verlag, New York.
- [20] Ruggeri, F., Wasserman, L. (1993), Infinitesimal sensitivity of posterior distributions, *The Canadian Journal of Statistics*, **21**, 2, 195-203.
- [21] Shyamalkumar, N., D. (1996), Bayesian robustness with asymptotically nice classes of priors, *Technical Report*, **96-22**, Dept. Stat., Purdue University.
- [22] Shyamalkumar, N., D. (2000), Likelihood Robustness, in *Robust Bayesian Analysis* (D. Ríos Insua and F. Ruggeri, eds), Springer-Verlag, New York.
- [23] Sivaganesan, S. (1993), Robust Bayesian diagnostics, *Journal of Statistical and Planning Inference*, **35**, 2, 171-188.
- [24] Sivaganesan, S. (1999), A likelihood based robust Bayesian summary, *Statistics and Probability Letters*, **43**, 5-12.
- [25] Sivaganesan, S. (2000), Global and local robustness approaches: uses and limitation, in *Robust Bayesian Analysis* (D. Ríos Insua and F. Ruggeri, eds), Springer-Verlag, New York.

8 Appendix 1

Consider a linear perturbation of the sampling distribution of type (3) with G the contaminating distribution. The perturbed posterior density is given by

$$p_\varepsilon(\theta|x) = \frac{\pi(\theta) \cdot L_{F_\varepsilon}(x|\theta)}{m(x; \Pi, F_{\theta, \varepsilon})},$$

and its derivative

$$\begin{aligned} \left[\frac{\partial p_\varepsilon(\theta|x)}{\partial \varepsilon} \right]_{\varepsilon=0} &= \left[\frac{\left(\pi(\theta) \frac{\partial L_{F_\varepsilon}(x|\theta)}{\partial \varepsilon} \right) m(x; \Pi, F_{\theta, \varepsilon}) - (\pi(\theta) \cdot L_{F_\varepsilon}(x|\theta)) \left(\frac{\partial m(x; \Pi, F_{\theta, \varepsilon})}{\partial \varepsilon} \right)}{m(x; \Pi, F_{\theta, \varepsilon})^2} \right]_{\varepsilon=0} \\ &= \frac{\left(\pi(\theta) \cdot \sum_{j=1}^n (g(x_j) - f_\theta(x_j)) \prod_{i \neq j} f_\theta(x_i) \right)}{m(x; \Pi, F_\theta)} - \\ &\quad \frac{(\pi(\theta) \cdot L_F(x|\theta)) \left(\sum_{j=1}^n (m_j(x; \Pi, F_\theta, G) - m(x; \Pi, F_\theta)) \right)}{m(x; \Pi, F_\theta)^2} \\ &= \sum_{j=1}^n p_j(\theta|x) \frac{m_j(x; \Pi, F_\theta, G)}{m(x; \Pi, F_\theta)} - n \cdot p(\theta|x) \\ &\quad - \pi(\theta|x) \sum_{j=1}^n \frac{m_j(x; \Pi, F_\theta, G)}{m(x; \Pi, F_\theta)} + n \cdot p(\theta|x) \\ &= \sum_{j=1}^n \frac{m_j(x; \Pi, F_\theta, G)}{m(x; \Pi, F_\theta)} [p_j(\theta|x) - p(\theta|x)], \end{aligned}$$

where

$$p_j(\theta|x) = \frac{\pi(\theta) \cdot g(x_j) \cdot \prod_{i \neq j} f_\theta(x_i)}{m_j(x; \Pi, F_\theta, G)}$$

is the posterior obtained when a sampling distribution G is adopted only for observation j and

$$m_j(x; \Pi, F_\theta, G) = \int g(x_j) \cdot \prod_{i \neq j} f_\theta(x_i) \pi(\theta) d\theta$$

is the corresponding marginal likelihood.

The measure of local influence of the functional to the sampling model is therefore

given by

$$\begin{aligned}
LI(G; T_B, F_\theta) &= \int \rho(\theta) \left[\frac{\partial p_\varepsilon(\theta|x)}{\partial \varepsilon} \right]_{\varepsilon=0} d\theta \\
&= \int \rho(\theta) \cdot \sum_{j=1}^n \left[\frac{m_j(x; \Pi, F_\theta, G)}{m(x; \Pi, F_\theta)} \cdot (p_j(\theta|x) - p(\theta|x)) \right] d\theta \\
&= \sum_{j=1}^n \frac{m_j(x; \Pi, F_\theta, G)}{m(x; \Pi, F_\theta)} \int \rho(\theta) \cdot (p_j(\theta|x) - p(\theta|x)) d\theta. \quad (18)
\end{aligned}$$

Expression (18) takes different forms according to the information G brings on the parameter of interest. If G is a distribution still governed by parameter θ , we denote the contaminating distribution by G_θ . Local influence measure of T_B is then given by:

$$LI(G_\theta; T_B, F_\theta) = \sum_{j=1}^n \frac{m_j(x; \Pi, F_\theta, G_\theta)}{m(x; \Pi, F_\theta)} \cdot \left(T_B^{(j)}(F_\theta, G_\theta) - T_B(F_\theta) \right)$$

where $m_j(x; \Pi, F_\theta, G_\eta) = \int g_\theta(x_j) \cdot \prod_{i \neq j} f_\theta(x_i) \pi(\theta) d\theta$.

If G depends on a different known parameter η ($\eta \neq \theta$), the contaminating distribution is denoted by G_η and (18) turns out to be

$$LI(G_\eta; T_B, F_\theta) = \sum_j \frac{m_j(x; \Pi, F_\theta, G_\eta)}{m(x)} \left(T_B^{(-j)} - T_B \right)$$

where $m_j(x; \Pi, F_\theta, G_\eta) = g_\eta(x_j) \cdot \int \prod_{i \neq j} f_\theta(x_i) \pi(\theta) d\theta$ and $T_B^{(-j)}$ is the posterior functional under base models using sample x without observation x_j .

9 Appendix 2

Assume a prior Π and a sampling model F_θ to be respectively $N(\theta_0, \sigma_0^2)$ and $N(\theta, \sigma^2)$. We need to compute the marginal likelihood $m(x) = \int L_F(x|\theta) \pi(\theta) d\theta$ where $L_F(x|\theta)$ is the likelihood under the reference sampling model. It is well known that in this case the posterior is $N(\theta_{post}; \sigma_{post}^2)$ with $\theta_{post} = \frac{n\sigma_0^2\bar{x} + \sigma^2\theta_0}{n\sigma_0^2 + \sigma^2}$ and $\sigma_{post}^2 = \frac{\sigma^2\sigma_0^2}{n\sigma_0^2 + \sigma^2}$. Our quantity of interest turns out to be:

$$\begin{aligned} m(x) &= \int \pi(\theta) \cdot L_F(x|\theta) d\theta \\ &= \int (2\pi\sigma_0^2)^{-\frac{1}{2}} (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma_0^2} (\theta - \theta_0)^2 - \frac{1}{2\sigma^2} \sum_i (x_i - \theta)^2 \right\} d\theta \\ &= (2\pi)^{-\frac{(n+1)}{2}} (\sigma^2)^{-\frac{n}{2}} (\sigma_0^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 \right\} \cdot \int \exp \{A(\theta)\} d\theta. \end{aligned}$$

Let's work with the exponent of the integrand term, given by $A(\theta) = -\frac{1}{2\sigma_0^2} (\theta - \theta_0)^2 - \frac{n}{2\sigma^2} (\bar{x} - \theta)^2$. We have

$$\begin{aligned} A(\theta) &= -\frac{1}{2} \left[\frac{(\theta^2 + \theta_0^2 - 2\theta\theta_0)}{\sigma_0^2} + \frac{n(\bar{x}^2 + \theta^2 - 2\theta\bar{x})}{\sigma^2} \right] \\ &= -\frac{1}{2\sigma_0^2\sigma^2} [\sigma^2\theta^2 + \sigma^2\theta_0^2 - 2\sigma^2\theta\theta_0 + n\sigma_0^2\bar{x}^2 + n\sigma_0^2\theta^2 - 2n\sigma_0^2\theta\bar{x}] \\ &= -\frac{1}{2} \underbrace{\frac{\sigma^2 + n\sigma_0^2}{\sigma_0^2\sigma^2}}_{=\sigma_{post}^{-2}} \left[\theta^2 - 2 \underbrace{\left(\frac{\sigma^2\theta_0 + n\sigma_0^2\bar{x}}{\sigma^2 + n\sigma_0^2} \right)}_{=\theta_{post}} \theta + \frac{\sigma^2\theta_0^2 + n\sigma_0^2\bar{x}^2}{\sigma^2 + n\sigma_0^2} \right]. \end{aligned}$$

Adding and subtracting θ_{post}^2 we get

$$\begin{aligned} A(\theta) &= -\frac{\sigma_{post}^{-2}}{2} \left[(\theta - \theta_{post})^2 + \frac{\sigma^2\theta_0^2 + n\sigma_0^2\bar{x}^2}{\sigma^2 + n\sigma_0^2} - \theta_{post}^2 \right] \\ &= -\frac{1}{2} \left(\frac{\theta - \theta_{post}}{\sigma_{post}} \right)^2 - \frac{1}{2} \frac{n}{\sigma^2 + n\sigma_0^2} (\theta_0 - \bar{x})^2. \end{aligned}$$

Therefore substituting into $m(x)$ we have

$$\begin{aligned} m(x) &= (2\pi)^{-\frac{(n+1)}{2}} (\sigma^2)^{-\frac{n}{2}} (\sigma_0^2)^{-\frac{1}{2}} (2\pi\sigma_{post}^2)^{\frac{1}{2}} \\ &\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 - \frac{1}{2} \frac{n}{\sigma^2 + n\sigma_0^2} (\theta_0 - \bar{x})^2 \right\} \end{aligned}$$

$$\begin{aligned}
& \underbrace{\int (2\pi\sigma_{post}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 \right] \right\} d\theta}_{=1} \\
&= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-1)}{2}} (\sigma^2 + n\sigma_0^2)^{-\frac{1}{2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - \bar{x})^2 - \frac{1}{2} \frac{n}{\sigma^2 + n\sigma_0^2} (\theta_0 - \bar{x})^2 \right\}.
\end{aligned}$$

Consider now the class of contaminating distribution

$$\tilde{G}_\theta = \{N(\theta, \eta^2) : \eta^2 \in [\sigma^2, 10 \cdot \sigma^2]\}.$$

We need to compute the marginal likelihood in the case where contaminating model G is assumed only for observation j . We denote with $L_{F,G}^{(j)}(\theta|x)$ the likelihood function in this case. The marginal likelihood is now given by

$$\begin{aligned}
m_j(x; \Pi, F_\theta, G) &= \int L_{F,G}^{(j)}(\theta|x) \pi(\theta) d\theta \\
&= (2\pi\sigma_0^2)^{-\frac{1}{2}} (2\pi\sigma^2)^{-\frac{(n-1)}{2}} (2\pi\eta^2)^{-\frac{1}{2}} \\
&\quad \cdot \int \exp \left\{ -\frac{1}{2\sigma_0^2} (\theta - \theta_0)^2 - \frac{1}{2\sigma^2} \sum_i (x_i - \theta)^2 - \frac{1}{2\eta^2} (x_j - \theta)^2 \right\} d\theta \\
&= (2\pi)^{-\frac{(n+1)}{2}} (\sigma^2)^{-\frac{(n-1)}{2}} (\sigma_0^2\eta^2)^{-\frac{1}{2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i \neq j} (x_i - \bar{x}_{(j)})^2 \right\} \cdot \int \exp \{B_j(\theta)\} d\theta.
\end{aligned}$$

Working again with the exponent of the integrand term we have:

$$\begin{aligned}
B_j(\theta) &= -\frac{1}{2\sigma_0^2} (\theta - \theta_0)^2 - \frac{(n-1)}{2\sigma^2} (\bar{x}_{(j)} - \theta)^2 - \frac{1}{2\eta^2} (x_j - \theta)^2 \\
&= -\frac{1}{2} \left[\frac{(\theta^2 + \theta_0^2 - 2\theta\theta_0)}{\sigma_0^2} + \frac{(n-1) (\bar{x}_{(j)}^2 + \theta^2 - 2\theta\bar{x}_{(j)})}{\sigma^2} + \frac{(x_j^2 + \theta^2 - 2\theta x_j)}{\eta^2} \right] \\
&= -\frac{1}{2\sigma_0^2\sigma^2\eta^2} [\sigma^2\eta^2\theta^2 + \sigma^2\eta^2\theta_0^2 - 2\sigma^2\eta^2\theta\theta_0 + (n-1)\eta^2\sigma_0^2\bar{x}_{(j)}^2 \\
&\quad + (n-1)\eta^2\sigma_0^2\theta^2 - 2(n-1)\eta^2\sigma_0^2\theta\bar{x}_{(j)} + \sigma^2\sigma_0^2x_j^2 + \sigma^2\sigma_0^2\theta^2 - 2\sigma^2\sigma_0^2\theta x_j] \\
&= -\frac{1}{2\sigma_0^2\sigma^2\eta^2} [(\sigma^2\eta^2 + (n-1)\eta^2\sigma_0^2 + \sigma^2\sigma_0^2) \cdot \theta^2 + \sigma^2\sigma_0^2x_j^2 + \sigma^2\eta^2\theta_0^2 \\
&\quad + (n-1)\eta^2\sigma_0^2\bar{x}_{(j)}^2 - 2(\sigma^2\eta^2\theta_0 + (n-1)\eta^2\sigma_0^2\bar{x}_{(j)} + \sigma^2\sigma_0^2x_j) \cdot \theta]
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2} \underbrace{\frac{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2}{\sigma_0^2 \sigma^2 \eta^2}}_{=\sigma_{post,j}^{-2}} \left[\theta^2 - 2 \underbrace{\left(\frac{\sigma^2 \eta^2 \theta_0 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)} + \sigma^2 \sigma_0^2 x_j}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2} \right)}_{=\theta_{post,j}} \cdot \theta \right. \\
&\quad \left. + \frac{\sigma^2 \sigma_0^2 x_j^2 + \sigma^2 \eta^2 \theta_0^2 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)}^2}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2} \right].
\end{aligned}$$

Adding and subtracting $\theta_{post,j}^2$ we get

$$\begin{aligned}
&= -\frac{\sigma_{post,j}^{-2}}{2} \left[(\theta - \theta_{post,j})^2 + \frac{\sigma^2 \sigma_0^2 x_j^2 + \sigma^2 \eta^2 \theta_0^2 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)}^2}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2} - \theta_{post,j}^2 \right] \\
&= -\frac{1}{2} \left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 - \frac{\sigma_{post,j}^{-2}}{2} \cdot \left(\frac{\sigma^2 \sigma_0^2 x_j^2 + \sigma^2 \eta^2 \theta_0^2 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)}^2}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2} - \theta_{post,j}^2 \right) \\
&= -\frac{1}{2} \left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 - \frac{\sigma_{post,j}^{-2}}{2} \cdot \left(\frac{(\sigma^2 \sigma_0^2 x_j^2 + \sigma^2 \eta^2 \theta_0^2 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)}^2)}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)^2} \right. \\
&\quad \left. \cdot \frac{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)^2} - \frac{(\sigma^2 \eta^2 \theta_0 + (n-1) \eta^2 \sigma_0^2 \bar{x}_{(j)} + \sigma^2 \sigma_0^2 x_j)^2}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)^2} \right) \\
&= -\frac{1}{2} \left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 - \frac{\sigma_{post,j}^{-2}}{2} \cdot \frac{\sigma_0^2 \sigma^2 \eta^2}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)^2} \\
&\quad \cdot \left(\sigma^2 (x_j - \theta_0)^2 + (n-1) \eta^2 (\bar{x}_{(j)} - \theta_0)^2 + (n-1) \sigma_0^2 (\bar{x}_{(j)} - x_j)^2 \right) \\
&= -\frac{1}{2} \left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 - \frac{1}{2} \cdot \frac{\left(\sigma^2 (x_j - \theta_0)^2 + (n-1) \eta^2 (\bar{x}_{(j)} - \theta_0)^2 \right)}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2} \\
&\quad - \frac{1}{2} \cdot \frac{(n-1) \sigma_0^2 (\bar{x}_{(j)} - x_j)^2}{\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2}.
\end{aligned}$$

Therefore substituting in $m_j(x; \Pi, F_\theta, G)$ we get

$$\begin{aligned}
m_j(x; \Pi, F_\theta, G) &= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-1)}{2}} (\sigma_0^2 \eta^2)^{-\frac{1}{2}} (2\pi \sigma_{post,j}^2)^{\frac{1}{2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i \neq j} (x_i - \bar{x}_{(j)})^2 - \frac{1}{2} \cdot \frac{\sigma^2 (x_j - \theta_0)^2}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)} \right\} \\
&\quad \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(n-1) \eta^2 (\bar{x}_{(j)} - \theta_0)^2 + (n-1) \sigma_0^2 (\bar{x}_{(j)} - x_j)^2}{(\sigma^2 \eta^2 + (n-1) \eta^2 \sigma_0^2 + \sigma^2 \sigma_0^2)} \right\} \\
&\quad \cdot \underbrace{\int (2\pi \sigma_{post,j}^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\left(\frac{\theta - \theta_{post,j}}{\sigma_{post,j}} \right)^2 \right] \right\} d\theta}_{=1}
\end{aligned}$$

$$\begin{aligned}
&= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{(n-2)}{2}} (\sigma^2\eta^2 + (n-1)\eta^2\sigma_0^2 + \sigma^2\sigma_0^2)^{-\frac{1}{2}} \\
&\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i \neq j} (x_i - \bar{x}_{(j)})^2 - \frac{1}{2} \cdot \frac{\sigma^2 (x_j - \theta_0)^2}{(\sigma^2\eta^2 + (n-1)\eta^2\sigma_0^2 + \sigma^2\sigma_0^2)} \right\} \\
&\quad \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(n-1)\eta^2 (\bar{x}_{(j)} - \theta_0)^2 + (n-1)\sigma_0^2 (\bar{x}_{(j)} - x_j)^2}{(\sigma^2\eta^2 + (n-1)\eta^2\sigma_0^2 + \sigma^2\sigma_0^2)} \right\}.
\end{aligned}$$

10 Appendix 3

Consider the Bayesian linear regression model where a normal distribution is assumed both for the error model F and for the prior Π . The posterior distribution of regression coefficients turns out to be normal with mean

$$E(\beta|y, X) = (\Sigma_0^{-1} + X'X)^{-1} (\Sigma_0^{-1}\beta_0 + X'y)$$

and variance

$$Var(\beta|y, X) = \sigma^2 (\Sigma_0^{-1} + X'X)^{-1}.$$

The Bayes estimator β_{Bayes} for regression coefficients is given by $E(\beta|y, X)$, which is a posterior functional of type(2). We also denote by V the quantity $(\Sigma_0^{-1} + X'X)^{-1}$. Therefore measures of local influence of the functional to prior and sampling model perturbations are respectively given by

$$\begin{aligned} LI(Q; T_B, \Pi) &= \left[\frac{\partial T_B(\Pi_\varepsilon)}{\partial \varepsilon} \right]_{\varepsilon=0} \\ &= \int \beta \cdot \left[\frac{\partial}{\partial \varepsilon} p(\beta|y, X, \Pi_\varepsilon, F) \right]_{\varepsilon=0} d\beta \\ &= \frac{m(y, X; Q, F)}{m(y, X)} [T_B(Q) - T_B], \end{aligned}$$

and

$$\begin{aligned} LI(G; T_B, F) &= \left[\frac{\partial T_B(F_\varepsilon)}{\partial \varepsilon} \right]_{\varepsilon=0} \\ &= \int \beta \cdot \left[\frac{\partial}{\partial \varepsilon} p(\beta|y, X, \Pi, F_\varepsilon) \right]_{\varepsilon=0} \beta \\ &= \sum_{j=1}^n \frac{m_j(y, X; \Pi, F, G)}{m(y, X)} \int \beta \cdot [p_j(\beta|y, X) - p(\beta|y, X)] d\beta \\ &= \sum_{j=1}^n \frac{m_j(y, X; \Pi, F, G)}{m(y, X)} [T_B^{(j)}(F, G) - T_B], \end{aligned}$$

where $m_j(y, X; \Pi, F, G) = \int L_{F,G}^{(j)}(y|X, \beta) \pi(\beta) d\beta$ and $p_j(\beta|y, X) = \frac{\pi(\beta) \cdot L_{F,G}^{(j)}(y|X, \beta)}{m_j(y, X; \Pi, F, G)}$.

Both measures can be solved analytically only performing a conjugate analysis. Suppose that the uncertainty about the prior distribution on β is represented by the family $\tilde{Q} = \{N(\alpha_0, \sigma^2 \Sigma_0) : \alpha_0^{\inf} \leq \alpha_0 \leq \alpha_0^{\sup}\}$. The posterior derived with such a prior is still normal with mean $\beta_{Bayes}^* = (X'X + \Sigma_0^{-1})^{-1} (X'y + \Sigma_0^{-1}\alpha_0)$ and covariance matrix $\sigma^2 V^* = \sigma^2 (X'X + \Sigma_0^{-1})^{-1} = \sigma^2 V$. The corresponding marginal likelihood is given by

$$m(y, X; Q, F) = (2\pi\sigma^2)^{-\frac{k+n}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{A^*}{2} \right\} (2\pi\sigma^2)^{\frac{k}{2}} |V^*|^{\frac{1}{2}}$$

with $A^* = \sigma^{-2} (y'y + \alpha_0' \Sigma_0^{-1} \alpha_0 - \beta_{Bayes}^{*'} V_{23}^{-1} \beta_{Bayes}^*)$.

Under this assumption the local influence for the prior becomes

$$\begin{aligned}
LI(Q; T_B, \Pi) &= \frac{\exp\left\{-\frac{A^*}{2}\right\}}{\exp\left\{-\frac{A}{2}\right\}} (\beta_{Bayes}^* - \beta_{Bayes}) \\
&= \exp\left\{-\frac{(\alpha_0 - \beta_0)' \Sigma_0^{-1} (\alpha_0 - \beta_0)}{2\sigma^2}\right\} \\
&= \exp\left\{+\frac{(\beta_{Bayes}^* - \beta_{Bayes})' V^{-1} (\beta_{Bayes}^* - \beta_{Bayes})}{2\sigma^2}\right\} (\beta_{Bayes}^* - \beta_{Bayes}) \\
&= \exp\left\{-\frac{(\alpha_0 - \beta_0)' [\Sigma_0^{-1} - \Sigma_0^{-1} V' \Sigma_0^{-1}] (\alpha_0 - \beta_0)}{2\sigma^2}\right\} [V \Sigma_0^{-1} (\alpha_0 - \beta_0)].
\end{aligned}$$

Let's now consider the perturbation of the sampling distribution. We will denote by x'_j ($1 \times k$) the row j of matrix X corresponding to observation j and with $X_{(-j)}$ ($(n-1) \times k$) and $y_{(-j)}$ respectively the matrix X and the vector y where the observation j has been dropped out. Assuming a contaminating family of type $\tilde{G} = \{N(0, c^2) : c^{\text{inf}} \leq c^2 \leq c^{\text{sup}}\}$ the marginal likelihood $m_j(y, X; \Pi, F, G)$ is given by

$$\begin{aligned}
m_j(y, X; \Pi, F, G) &= \int L_{F,G}^{(j)}(y|X, \beta) \pi(\beta) d\beta \\
&= (2\pi\sigma^2)^{-\frac{(k+n-1)}{2}} (2\pi c^2)^{-\frac{1}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \tilde{B}\right\}.
\end{aligned}$$

The terms \tilde{B} is given by

$$\begin{aligned}
\tilde{B} &= \sigma^{-2} (\beta - \beta_0)' \Sigma_0^{-1} (\beta - \beta_0) + c^{-2} (y_j - x'_j \beta)' (y_j - x'_j \beta) \\
&\quad + \sigma^{-2} (y_{(-j)} - X_{(-j)} \beta)' (y_{(-j)} - X_{(-j)} \beta) \\
&= \sigma^{-2} \beta' \Sigma_0^{-1} \beta - 2\sigma^{-2} \beta' \Sigma_0^{-1} \beta_0 + \sigma^{-2} \beta_0' \Sigma_0^{-1} \beta_0 + c^{-2} y_j^2 - 2c^{-2} \beta' x_j y_j \\
&\quad + c^{-2} \beta' x_j x'_j \beta + \sigma^{-2} y'_{(-j)} y_{(-j)} - 2\sigma^{-2} \beta' X'_{(-j)} y_{(-j)} + \sigma^{-2} \beta' X'_{(-j)} X_{(-j)} \beta \\
&= \beta' [\sigma^{-2} (X'_{(-j)} X_{(-j)}) + c^{-2} (x_j x'_j) + \sigma^{-2} \Sigma_0^{-1}] \beta \\
&\quad - 2\beta' (\sigma^{-2} X'_{(-j)} y_{(-j)} + c^{-2} x_j y_j + \sigma^{-2} \Sigma_0^{-1} \beta_0) \\
&\quad + \sigma^{-2} \beta_0' \Sigma_0^{-1} \beta_0 + c^{-2} y_j^2 + \sigma^{-2} y'_{(-j)} y_{(-j)} \\
&= \underbrace{\sigma^{-2} \beta_0' \Sigma_0^{-1} \beta_0 + c^{-2} y_j^2 + \sigma^{-2} y'_{(-j)} y_{(-j)} - \sigma^{-2} \mathbf{m}_j V_j^{-1} \mathbf{m}_j}_{=\tilde{B}_j} \\
&\quad + \sigma^{-2} (\beta - \beta_{Bayes}^{(j)})' V_j^{-1} (\beta - \beta_{Bayes}^{(j)}) \\
&= \tilde{B}_j + \sigma^{-2} (\beta - \beta_{Bayes}^{(j)})' V_j^{-1} (\beta - \beta_{Bayes}^{(j)}),
\end{aligned}$$

where

$$\beta_{Bayes}^{(j)} = \left[X'_{(-j)} X_{(-j)} + \frac{\sigma^2}{c^2} x_j x'_j + \Sigma_0^{-1} \right]^{-1} \left(X'_{(-j)} y_{(-j)} + \frac{\sigma^2}{c^2} x_j y_j + \Sigma_0^{-1} \beta_0 \right),$$

and

$$V_j = \left[X'_{(-j)} X_{(-j)} + \frac{\sigma^2}{c^2} x_j x'_j + \Sigma_0^{-1} \right]^{-1}.$$

Marginal $m_j(y, X; \Pi, F, G)$ becomes

$$m_j(y, X; \Pi, F, G) = (2\pi\sigma^2)^{-\frac{(n-1)}{2}} (2\pi c^2)^{-\frac{1}{2}} |\Sigma_0|^{-\frac{1}{2}} |V_j|^{\frac{1}{2}} \exp \left\{ -\frac{\tilde{B}_j}{2} \right\},$$

and the corresponding posterior distribution turns out to be a $N\left(\beta_{Bayes}^{(j)}, \sigma^2 V_j\right)$.

Therefore $T_B^{(j)}(F, G_\beta) = \beta_{Bayes}^{(j)}$.

Under this assumption the local influence for the sampling turns out to be

$$\begin{aligned} LI(G; T_B, F) &= \sum_{j=1}^n \left[\frac{(2\pi\sigma^2)^{-\frac{(k+n-1)}{2}} (2\pi c^2)^{-\frac{1}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{\tilde{B}_j}{2} \right\} (2\pi\sigma^2)^{\frac{k}{2}} |V_j|^{\frac{1}{2}}}{(2\pi\sigma^2)^{-\frac{k+n}{2}} |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{A^*}{2} \right\} (2\pi\sigma^2)^{\frac{k}{2}} |V|^{\frac{1}{2}}} \right. \\ &\quad \left. \cdot \left(\beta_{Bayes}^{(j)} - \beta_{Bayes} \right) \right] \\ &= \sum_{j=1}^n \left[\left(\frac{c^2 |V|}{\sigma^2 |V_j|} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{(\tilde{B}_j - A)}{2} \right\} \right] \left(\beta_{Bayes}^{(j)} - \beta_{Bayes} \right) \\ &= \sum_{j=1}^n \left[\left(\frac{c^2 |V|}{\sigma^2 |V_j|} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{\left(\frac{\sigma^2}{c^2} - 1 \right) y_j^2}{2\sigma^2} \right\} \right. \\ &\quad \left. \cdot \exp \left\{ -\frac{\beta'_{Bayes} V^{-1} \beta_{Bayes} - \beta_{Bayes}^{(j)'} V_j^{-1} \beta_{Bayes}^{(j)}}{2\sigma^2} \right\} \left(\beta_{Bayes}^{(j)} - \beta_{Bayes} \right) \right]. \end{aligned}$$

Table 1: Relative local sensitivity measures of the posterior mean with respect to the prior model with different prior precisions.

	σ_0^2			
	0.5	1	10	100
T_B	0.9571	0.9857	1.0146	1.0177
$LI^* \left(\tilde{Q}; T_B, \Pi \right)$	0.1270	0.0702	0.0148	0.0029
λ for $LI^* \left(\tilde{Q}; T_B, \Pi \right)$	1.6	1.8	3.9	5.5
λ for $\max m(x; Q, F_\theta)/m(x)$	1	1	1	1

Table 2: Relative local sensitivity measures of the posterior mean with respect to the sampling model with different sampling precisions.

	σ^2			
	0.1	0.2	1	4
T_B	1.0014	0.9857	0.8885	0.7220
$LI^* \left(\tilde{G}_\theta; T_B, F_\theta \right)$	0.0650	0.0096	0.0544	0.0651
η^2 for $LI^* \left(\tilde{G}_\theta; T_B, F_\theta \right)$	1.0	0.6	4.0	13.6

Table 3: Relative local sensitivity measures of the posterior mean with respect to the base prior and sampling models. Contaminated sample.

T_B	-0.4395
$LI^* \left(\tilde{Q}; T_B, \Pi \right)$	0.2303
λ for $LI^* \left(\tilde{Q}; T_B, \Pi \right)$	-1.1
$LI^* \left(\tilde{G}_\theta; T_B, F_\theta \right)$	$7.41 \cdot 10^{25}$
η^2 for $LI^* \left(\tilde{G}_\theta; T_B, F_\theta \right)$	2

Table 4: Relative local sensitivity measures of regression coefficient estimates with respect to the base prior and sampling models.

component	1	2	3
$LI^* \left(\tilde{Q}; T_B; \Pi \right)$	$2.2 \cdot 10^{-19}$	$2.9 \cdot 10^{-18}$	$1.0 \cdot 10^{-18}$
α_0 for $LI^* \left(\tilde{Q}; T_B; \Pi \right)$	-1.69	-1.69	41
$LI^* \left(\tilde{G}; T_B; F \right)$	42.93	458.18	13.14
c^2 for $LI^* \left(\tilde{G}; T_B; F \right)$	360	360	360

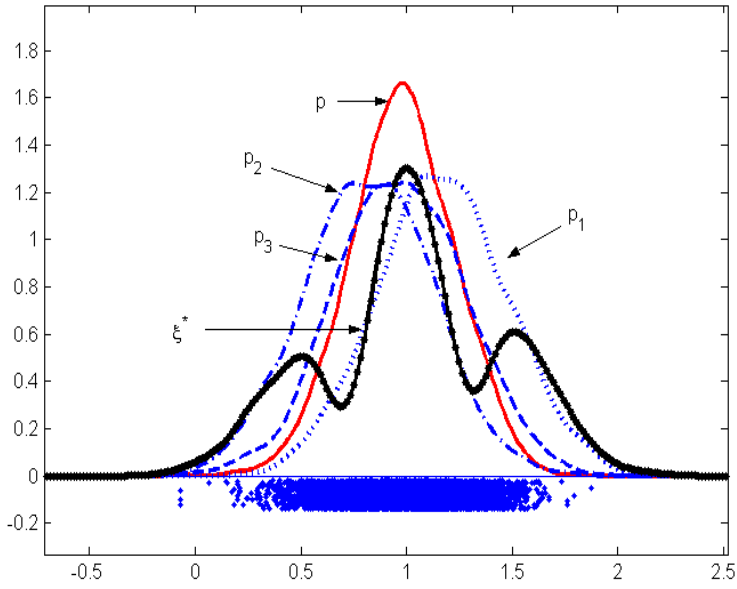


Figure 1: Importance sampling density ξ^* and posterior densities p and p_j 's.

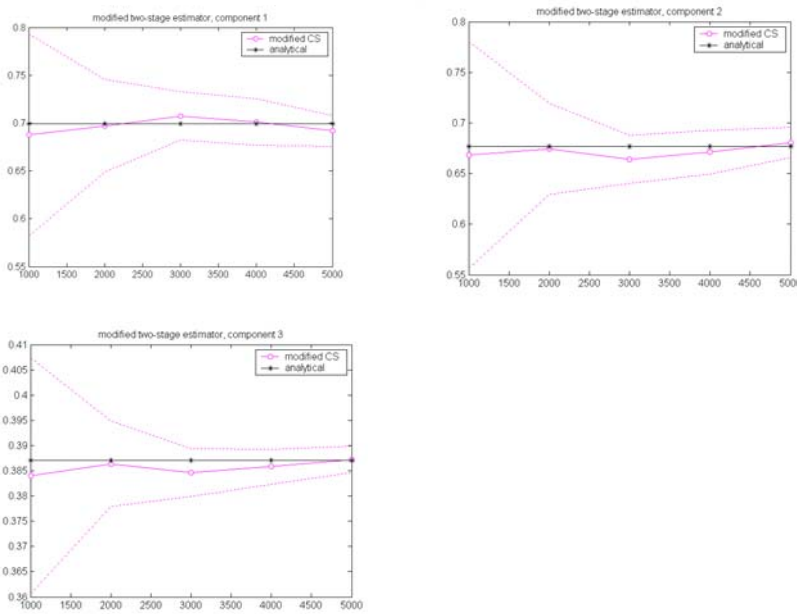


Figure 2: Analytical and estimated value of r_j ($j = 1, 2, 3$) with confidence interval.

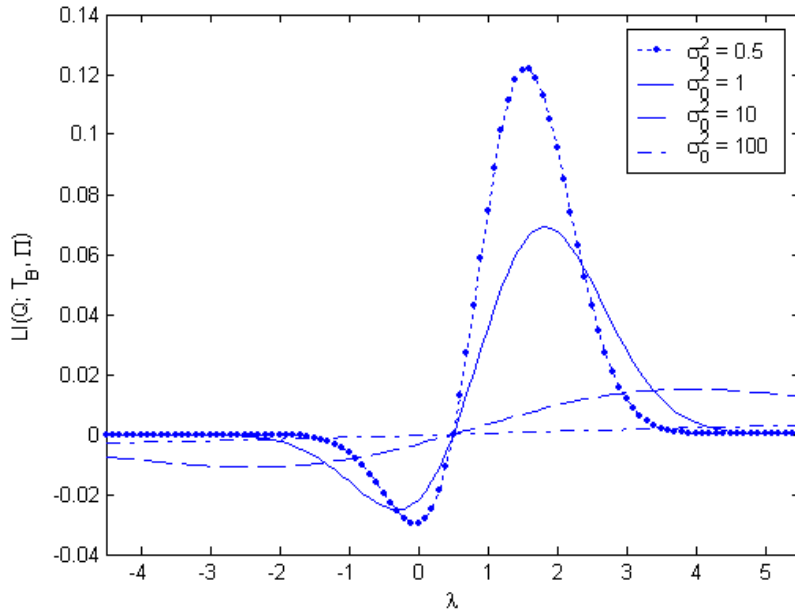


Figure 3: $LI(Q; T_B, \Pi)$ measure for the posterior mean with different values of prior variance σ_0^2 .

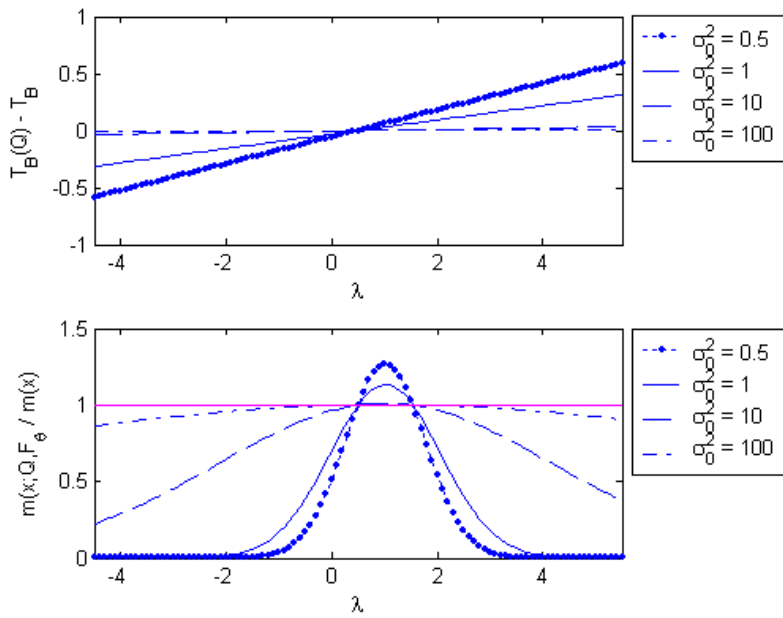


Figure 4: Difference $T_B(Q) - T_B$ and ratio $\frac{m(x; Q, F_\theta)}{m(x)}$ for different values of prior variance σ_0^2 .

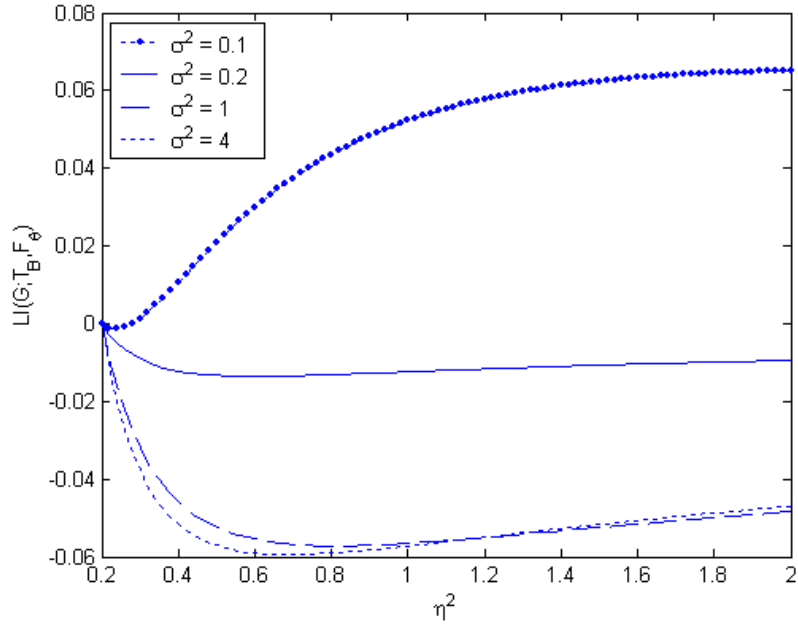


Figure 5: $LI(G; T_B, F_\theta)$ measure for the posterior mean with different values of sampling variance σ^2 .

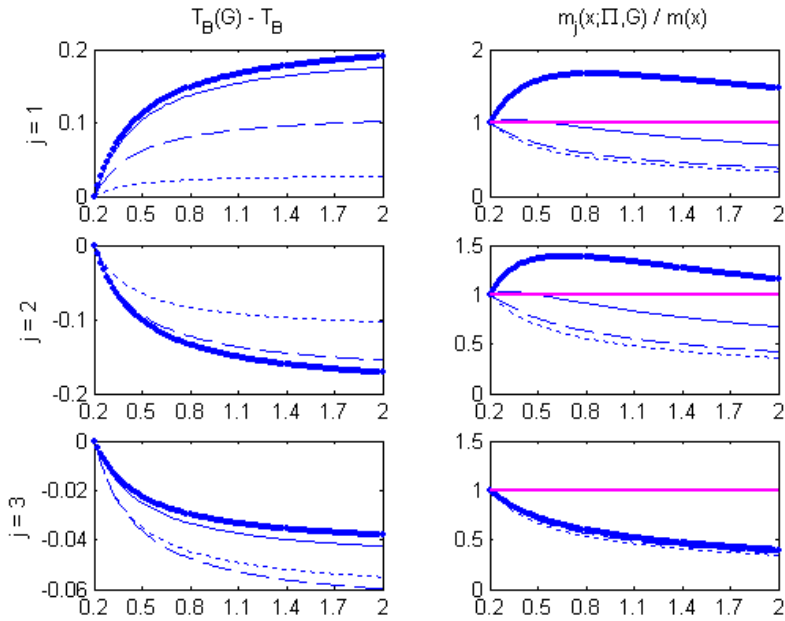


Figure 6: Difference $T_{B,j}(F_\theta, G) - T_B$ and ratio $m_j(x; \Pi, G) / m(x)$ for different values of sampling variance σ^2 .

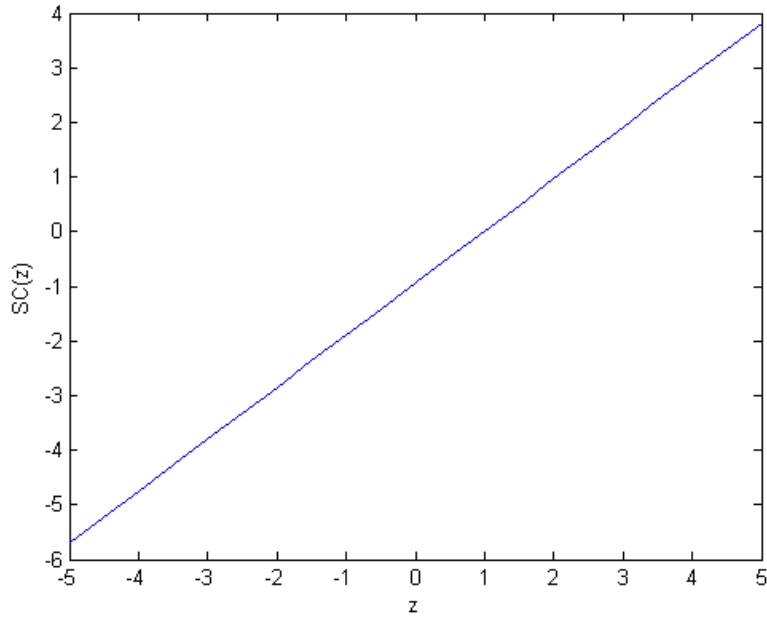


Figure 7: SC for the posterior mean under normality of both prior and sampling distributions.

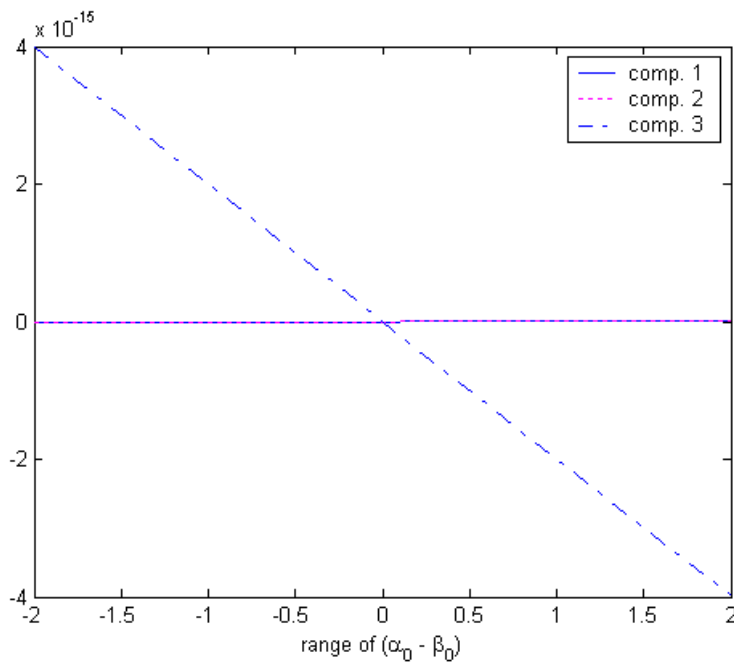


Figure 8: $LI(Q; T_B, \Pi)$ measure for regression coefficients.

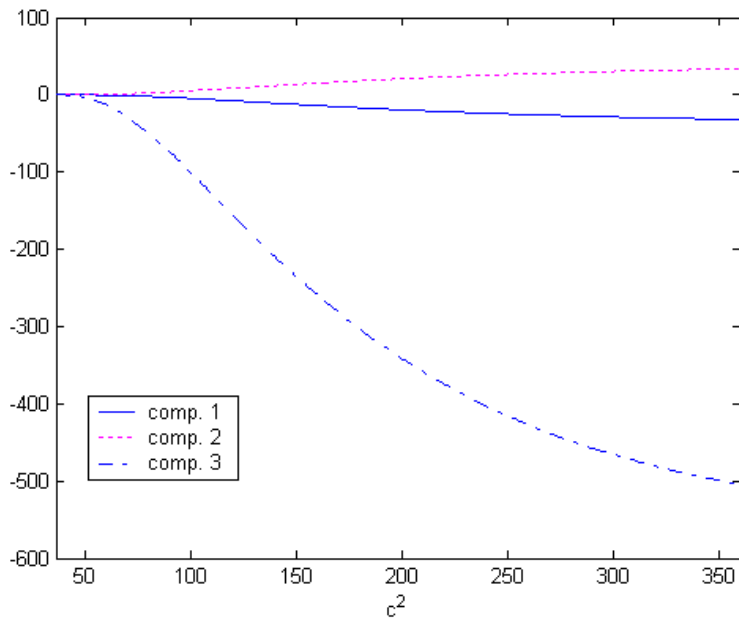


Figure 9: $LI(G; T_B, F)$ measure for regression coefficients.

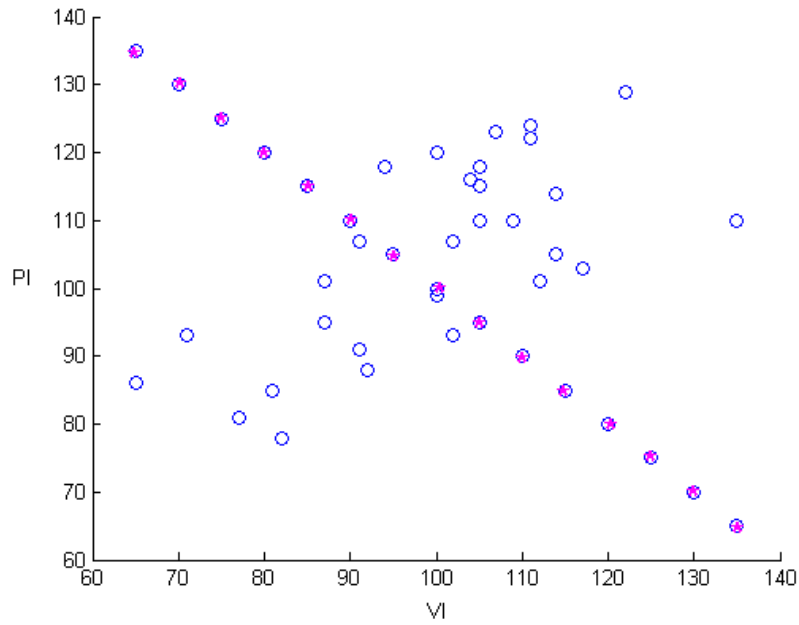


Figure 10: Scatterplot of VI towards PI . Asterisks represent the observations which have been added.

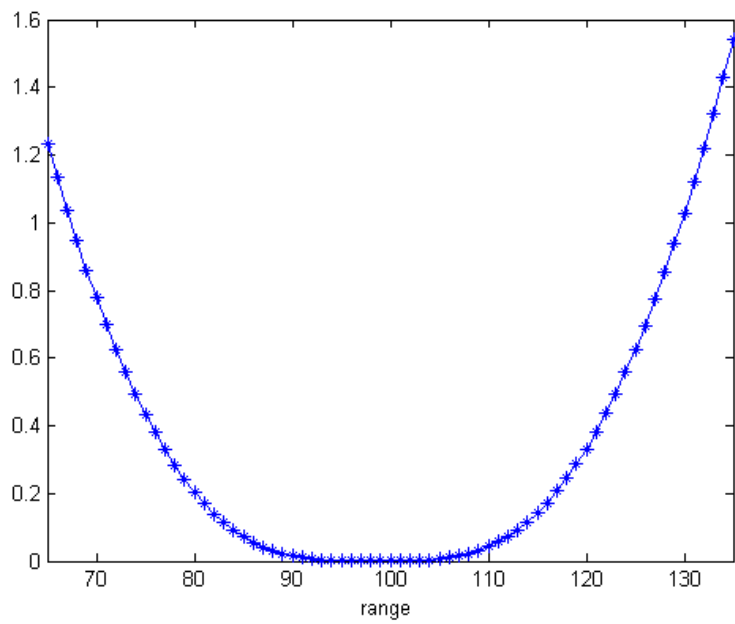


Figure 11: Cook's distance for observations which have been added.

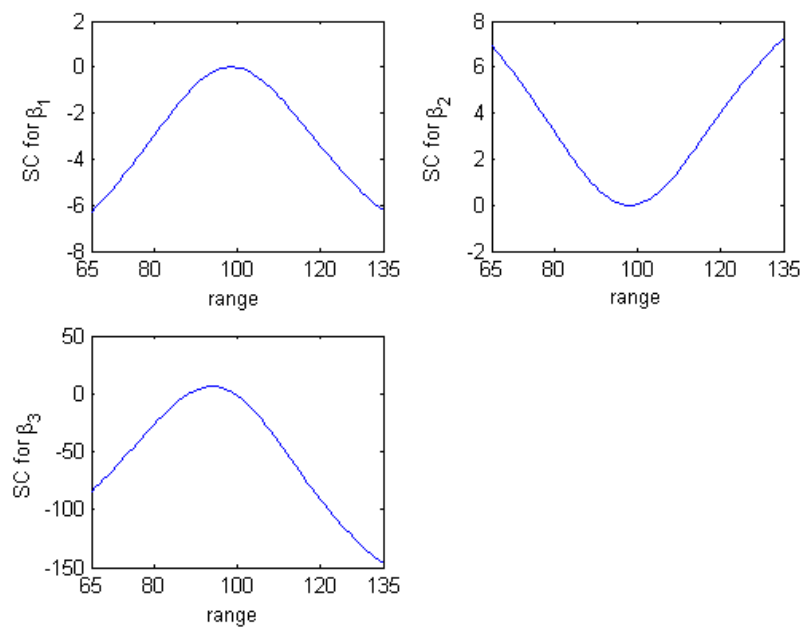


Figure 12: SC of regression coefficients moving the first two regressors in the range (65, 135).