1

Introduction to Sensitivity Analysis

1.1 MODELS AND SENSITIVITY ANALYSIS

What is a model? What model input is considered in a sensitivity analysis? What is the role of uncertainty and sensitivity analyses in model building? Main approaches to the propagation of uncertainty within and across models. Implications for model quality.

1.1.1 Definition

A possible definition of sensitivity analysis is the following: *The study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input* (Saltelli *et al.*, 2004). A related practice is 'uncertainty analysis', which focuses rather on *quantifying* uncertainty in model output. Ideally, uncertainty and sensitivity analyses should be run in tandem, with uncertainty analysis preceding in current practice.

For this definition of sensitivity analysis to be of use, it must first be made clear what is meant here by 'model', numerical or otherwise, as well as by the terms 'input' and 'output' which will be used throughout this book.

Global Sensitivity Analysis. The Primer A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola

1.1.2 Models

A view of modelling that may help to illustrate the role of sensitivity analysis in the scientific process is offered in Figure 1.1, taken from the work of biologist Robert Rosen (1991) (see also Saltelli *et al.*, 2000, pp. 3–4). On the left in Rosen's diagram we have the 'world', that is the system which forms the subject of our investigation. We have reason to believe that the system, whether natural or artificial, is governed by rules which we have the ambition to uncover, or to use to our advantage. To this end we craft or hypothesize a set of structures in a model (depicted on the right-hand side of the figure). For example, a hypothesized growth mechanism for a species contained in the world can be translated into a differential equation in a model. While our species continues growing and dying quietly in the world, following the forces of its own systemic causality (which we aim to understand), our differential equation can be solved using the rules of mathematical calculus. The intuition of Rosen is that while the species in the world obeys rules, and the differential equation in the model has 'rules' as well, whether formal or mathematical, no 'rule' whatsoever can dictate how one should map the hypothesized rules in the world onto the rules in the model. In the words of Rosen, while the world and the model are each internally 'entailed', nothing entails the world with the model. Among the reasons for this paradox is the fact that the portion of the world captured by the model is an arbitrary 'enclosure' of an otherwise open, interconnected system.[1] This is the case when the world is part of a natural system, the main concern of Rosen's inquiry. Yet experience has shown that even when the world is indeed a well-defined and closed system, for instance an artefact, an artificial device or a piece of machinery, different

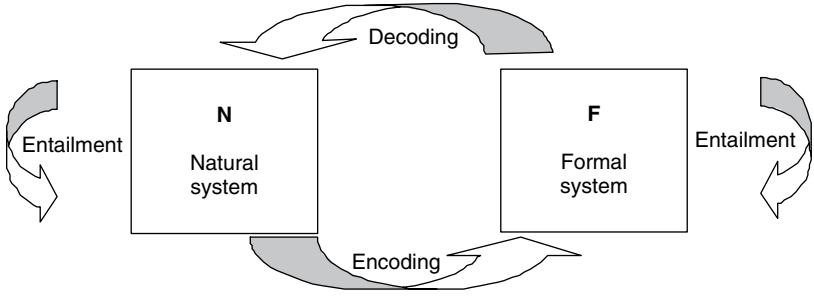

Figure 1.1 Modelling after Rosen (1991)

[1] Even more so when the purpose of a model is to learn about the nonobservable parts of a system.

modellers can generate different nonequivalent descriptions of it, that is, models whose outputs are compatible with the same set of observations but whose structures are not reconcilable with one another.

While this may be disturbing to a student accustomed to the beauty and apparent self-evidence of physical laws, practitioners of modelling have come to live with the rather unpleasant reality that more than one model may be compatible with the same set of data or evidence. Some have gone so far as to coin a word for this paradox: equifinality – Beven (1993, 2001), see also Saltelli *et al.* (2004, pp. 173–178) – meaning that different models can lead to the same end. Others refer to the phenomenon as model indeterminacy.

Since Galileo's time scientists have had to deal with the limited capacity of the human mind to create useful maps of 'world' into 'model'. The emergence of 'laws' can be seen in this context as the painful process of simplification, separation and identification which leads to a model of uncharacteristic simplicity and beauty.

1.1.3 Models and Uncertainty

What makes modelling and scientific inquiry in general so painful is uncertainty. Uncertainty is not an accident of the scientific method, but its substance.[2]

Modellers and philosophers of science have debated the issue of model indeterminacy at length (Oreskes *et al.*, 1994). Most modellers today would probably agree that a model cannot be validated, in the sense of 'be proven true'. Rather, it is more defensible and correct to say that a model has been extensively corroborated, meaning by this that the model has survived a series of tests – be they formal, of internal consistency, or relative to the model's capacity to explain or predict the 'world' in a convincing and parsimonious way.

When models fail publicly, the ensuing controversy can be devastating for the scientific parties involved.[3] Models are often used in highly polarized contexts and uncertainty may be used instrumentally. 'All parties deal with environmental information in a selective way, or even manipulate it', observed a Dutch environmental scientist (In 't Veld, 2000). Fabricated

[2] 'That is what we meant by science. That both question and answer are tied up with uncertainty, and that they are painful. But that there is no way around them. And that you hide nothing; instead, everything is brought out into the open' (Høeg, 1995).

[3] For the modelling credibility crisis in the Netherlands' RIVM Laboratories see Van der Sluijs (2002). See also Mac Lane (1988) for another example.

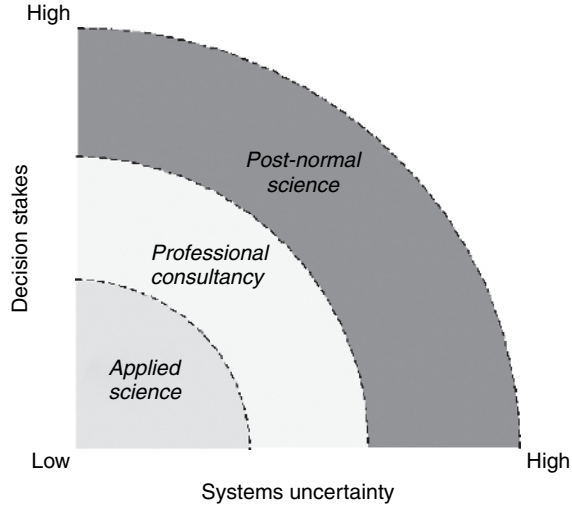

Figure 1.2 Uncertainty/stakes diagram after Funtowicz and Ravetz (1990)

uncertainty is a common concern in relation to important disputes over health or the environment (Michaels, 2005).

In short, models are part of the scientific method and hence subject to epistemological debate. A way of framing present-day debate on the scientific method is offered by Post-Normal Science (PNS, see Figure 1.2 and Funtowicz and Ravetz, 1990, 1993; Funtowicz *et al.*, 1996).

In PNS one distinguishes between three types of scientific production modes, depending on the system's uncertainties and the stakes involved. Applying this to modelling, different requirements and practices pertain:

- In applied science, when a model is written and employed within a closed consortium of experts who are the sole users of the model, e.g. when this is used to solve a circumscribed chemical kinetics problem;
- In 'consultancy' when the model is more likely to be scrutinized, e.g. as part of a cost–benefit analysis for the construction of a new road or bridge that will affect a community;
- When computing climate sensitivity in the context of global change. In this latter case we are in the domain of PNS, where science (and its models) is called on to provide evidence under circumstances of conflicting stakes and beliefs.

Like scientific theories, models may be given pedigrees which help us to judge their quality. Pedigrees take account of past usage of the model, status of its proponents, degree of acceptance by peers and so on (Van der Sluijs, 2002; Craye *et al.*, 2005). In pedigrees, model quality is more closely

associated with 'fitness for purpose' – that is, with a specific purpose – than with the model's intrinsic fabric.

A post-normal view of the modes of scientific production in relation to policy is given in Funtowicz (2004). Models as metaphors are discussed in Ravetz (2006).

1.1.4 How to Set Up Uncertainty and Sensitivity Analyses

As mentioned at the beginning of the chapter, our definition of sensitivity analysis involves models, model input and model output. We now try to define model input in relation to the nature and purpose of the model, as well as to the set-up of the uncertainty and sensitivity analyses. A model can be:

- *Diagnostic or prognostic*. In other words, we try to distinguish between models used to understand a law and models used to predict the behaviour of a system given a supposedly understood law. Models can thus range from wild speculations used to play what-if games (e.g. models for the existence of extraterrestrial intelligence) to models which can be considered accurate and trusted predictors of a system (e.g. a control system for a chemical plant).
- *Data-driven or law-driven*. A law-driven model tries to put together accepted laws which have been attributed to the system, in order to predict its behaviour. For example, we use Darcy's and Ficks' laws to understand the motion of a solute in water flowing through a porous medium. A data-driven model tries to treat the solute as a signal and to derive its properties statistically. Advocates of data-driven models like to point out that these can be built so as to be parsimonious, i.e. to describe reality with a minimum of adjustable parameters (Young *et al.*, 1996). Law-driven models, by contrast, are customarily overparametrized, as they may include more relevant laws than the amount of available data would support. For the same reason, law-driven models may have a greater capacity to describe the system under unobserved circumstances, while data-driven models tend to adhere to the behaviour associated with the data used in their estimation. Statistical models (such as hierarchical or multilevel models) are another example of data-driven models.

Many other categorizations of models are possible,[4] and the definition of model input depends on the particular model under study. For the purpose

[4] Bell *et al.* (1988) distinguish between formal (axiomatic), descriptive and normative models (rules an agent should follow to reach a target). The examples in this book are descriptive models.

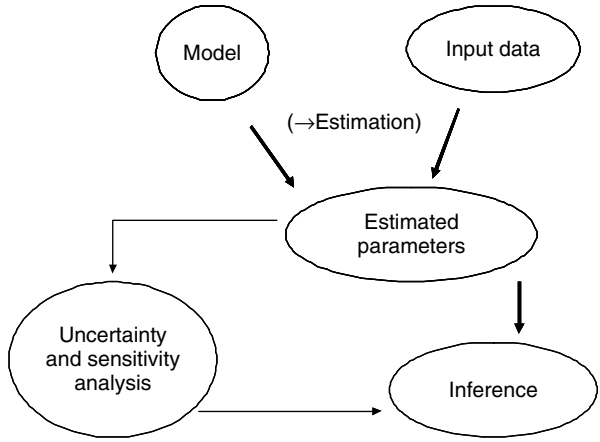

Figure 1.3 Parametric bootstrap version of uncertainty and sensitivity analyses

of uncertainty and sensitivity analyses we could liberally classify as input everything that can drive a variation in the output of the model.

Consider the scheme in Figure 1.3. Here we have observations (assumed error-free for simplicity's sake) and a model whose parameters are estimated from the data. Estimation can take different courses. Usually it is achieved by minimizing, e.g. by least squares, some measure of distance between the model's prediction and the data. At the end of the estimation step, 'best' parameter values as well as their errors are known. At this point we might consider the model 'true' and run an uncertainty analysis by propagating the uncertainty in the parameters through the model, all the way to the model output. In this case the estimated parameters become our factors.

One way of doing this is through Monte Carlo analysis, in which we look at the distribution functions of the input parameters, as derived from the estimation. For example, we may have the following scheme:

- We start from a factor $\alpha \sim N(\bar{\alpha}, \sigma_\alpha)$, which reads: after estimation α is known to be normally distributed with mean $\bar{\alpha}$ and standard deviation σ_α.
- Likewise for factors β, γ and so on. Contrary to what logic would suggest, and for the sake of simplicity, we assume that the factors are independent of each other. This issue is discussed later in the chapter.
- For each of these factors, we draw a sample from the respective distributions, i.e. we produce a set of row vectors $(\alpha^{(j)}, \beta^{(j)}, \dots)$ with $j = 1, 2, \dots, N$ in such a way that $(\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(N)})$ is a sample

from $N(\bar{\alpha}, \sigma_\alpha)$ and likewise for the distribution function of the other factors.

$$\begin{bmatrix} \alpha^{(1)} & \beta^{(1)} & \gamma^{(1)} & \dots \\ \alpha^{(2)} & \beta^{(2)} & \gamma^{(2)} & \dots \\ \dots & \dots & \dots & \dots \\ \alpha^{(N-1)} & \beta^{(N-1)} & \gamma^{(N-1)} & \dots \\ \alpha^{(N)} & \beta^{(N)} & \gamma^{(N)} & \dots \end{bmatrix} \tag{1.1}$$

- We can then compute ('run' is the conventional term) the model for all vectors $(\alpha^{(j)}, \beta^{(j)}, \dots)$ thereby producing a set of N values of a model output Y_j.[5]

$$\begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(N-1)} \\ y^{(N)} \end{bmatrix} \tag{1.2}$$

These steps constitute our uncertainty analysis. From these we can compute the average output, its standard deviation, the quantiles of its distribution, confidence bounds, plot the distribution itself and so on. It is clear that in this analysis, sometimes called a 'parametric bootstrap',[6] our inputs are the model's parameters. Having performed this uncertainty analysis we can then move on to a sensitivity analysis, in order to determine which of the input parameters are more important in influencing the uncertainty in the model output. However, we defer this step in order to continue our discussion of model input.

Note that for the purpose of the uncertainty analysis just described we consider as relevant inputs only our estimated parameters. All other types of information fed into the model, e.g. the observations, physical or mathematical constants, internal model variables (e.g. number of grid points if the model needs a mesh), are disregarded – that is, we do not allow them to vary and hence they cannot cause variation in the output.

In Figure 1.4 we have played the uncertainty analysis game differently by sampling the observations rather than the parameters. We have a limited set of observations, and we are aware that different subsets of these could

[5] Note that this model output Y_j may be different from the model output used in the estimation step.

[6] Bootstrapping is the process of repeatedly sampling 'with replacements'. For example, if we want to estimate the average sum of three Bingo chips, we could do this by extracting three random chips from the Bingo bag, computing their average, putting the chips back into the bag and extracting again. With a sufficiently large number of extractions we could determine the average sum being sought, and this strategy would be called a bootstrap of the Bingo chips.

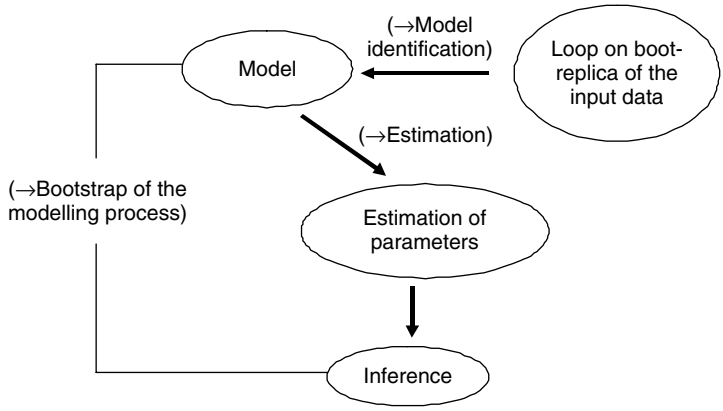

Figure 1.4 Bootstrapping of the modelling process (Chatfield, 1993)

potentially lead us to try one model rather than another to fit the data. What we can do in order to be fair to the data is to select a subset of the observations, choose a model based on these data using a pre-established model selection rule, estimate the corresponding parameters using the same sampled data, and run the model to compute Y_j. We have drawn the sample with replacement, and we can now repeat the process, identifying a potentially different (or indeed the same) model, estimating the parameters (which may differ in number from those of the previous run if the model is different), and so on for a total of N times, until we yield our desired sample for the uncertainty analysis. This approach can be called 'bootstrapping of the modelling process' (Chatfield, 1993).

The input for this uncertainty analysis is the data which have been bootstrapped, since we have assumed that all the rest (from model selection to parametric estimation) is done automatically given the data and hence adds no variation to model output.

Finally in Figure 1.5 we compare a set of plausible models with the data. Using Bayesian analysis it is possible to derive posterior probabilities for the models as well as distributions of the related parameters (Saltelli *et al.*, 2004). Once this model update and parameter estimation step is complete, a model averaging can be used in uncertainty analysis. This is done by propagating the uncertainty through the system by sampling both the model and the parameters according to their distributions, to produce a sample of model outcome Y_j. This procedure is known as Bayesian model averaging,[7] and the inputs in this case are both models and parameters, or

[7] For a thorough account of this approach see Kass and Raftery (1995) and Hoeting *et al.* (1999). See Saltelli *et al.* (2004, pp. 151–192) for related sensitivity issues.

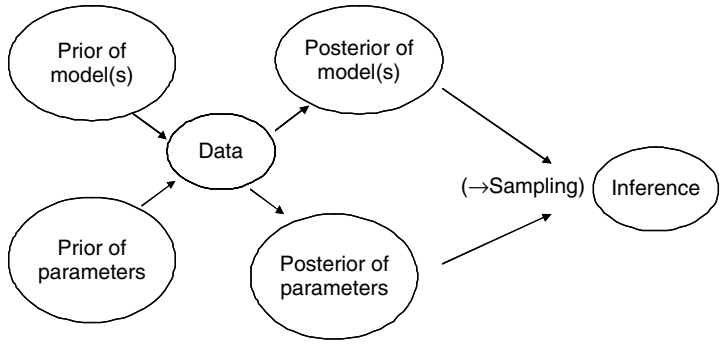

Figure 1.5 Bayesian model averaging

more precisely the probabilities of the different model representations and the distributions of the parameters. In a Monte Carlo framework, a trigger variable would be sampled to select a model according to its posterior probability, while the parameters would also be sampled and the model outcome determined. A sensitivity analysis could be executed at this point, and a question that it might address is the following: how much of the uncertainty is due to the model selection and how much to the estimation of the parameters?

1.1.5 Implications for Model Quality

The superficial illustration given above of approaches to uncertainty and sensitivity analyses has shown that what constitutes an input for the analysis depends upon how the analysis is set up. The input is that which is allowed to vary in order to study its effect on the output. A sensitivity analysis will in turn instruct the modellers as to the relative importance of the inputs in determining the output. An obvious consequence of this is that the modeller will remain ignorant of the importance of those variables which have been kept fixed. This is of course a hazard for the modeller, as a variable deemed noninfluential and kept fixed could haunt the results of the analysis at a later stage. For example, it would be unfortunate for the modeller to discover a posteriori that the mesh size had been too large, and that the number of grid points had had a dramatic effect on the model output.

It seems, therefore, that one should be as careful and objective as possible in deciding on the input for uncertainty and sensitivity analyses. Clearly, the more variables we promote to the rank of input, and allow to vary, the greater the variance to be expected in the model prediction. This could lead to a situation in which we discover that, having incorporated all uncertainties, the model prediction varies so wildly as to be of no practical use. This

trade-off has been brilliantly summarized by the econometrician Edward E. Leamer (1990):

> I have proposed a form of organized sensitivity analysis that I call 'global sensitivity analysis' in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified. Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful.

Note Leamer's emphasis on the need for 'credibility' in the selection of assumptions. The easiest way to invalidate a model is to demonstrate it fragile with respect to shaky assumptions. Note, however, that the trade-off may not be as dramatic as one might expect, and that increasing the number of input factors does not necessarily lead to an increased variance in model output. Practitioners have recorded that in most uncertainty and sensitivity analyses the input factors' importance is distributed similarly to wealth in nations, with a few factors creating almost all the uncertainty and the majority making only a negligible contribution. Hence, if the 'key' factors have been judiciously chosen, adding further variables to the analysis may add to its completeness and defensibility without adversely increasing the variance in the output.

As mentioned, the quality of a model is largely a function of its fitness for purpose. If modelling is a craft and models cannot be proven true (because of the pervasive nature of uncertainty and the difficulty of separating observation from observer and facts from values),[8] then the modeller has a moral obligation, and indeed it is in the modeller's own practical interest, to be as rigorous as possible when assessing the robustness of model inference. Doing so should produce better and more parsimonious models, and will strengthen the analyst's defence of the results in the case of scientific controversy or public policy debate.

1.2 METHODS AND SETTINGS FOR SENSITIVITY ANALYSIS – AN INTRODUCTION

What methods are available? How can a particular method be related to a problem-specific question? How can we define a factor's importance unambiguously? Suggested practices.

[8] 'Values' here mean ethical judgements. Cases in which the separation of facts and values becomes arduous are many, e.g. when models try to assess the impact of the adoption of new technologies, the relevance of environmental threats, distributional issues in economics and so on.

1.2.1 Local versus Global

As we shall learn in the following chapters, sensitivity analysis can serve a number of useful purposes in the economy of modelling. It can surprise the analyst, uncover technical errors in the model, identify critical regions in the space of the inputs, establish priorities for research, simplify models and defend against falsifications of the analysis. In the context of models used for policy assessment, sensitivity analysis can verify whether policy options can be distinguished from one another given the uncertainties in the system, and so on. What methods would one choose to perform sensitivity analysis for any or all of the above?

It is not by chance that most of the sensitivity analyses met in the literature are based on derivatives. Indeed the derivative $\partial Y_j/\partial X_i$ of an output Y_j versus an input X_i can be thought of as a mathematical definition of the sensitivity of Y_j versus X_i.

Sometimes computer programs that implement complex physical, chemical or genetic models are augmented by special routines that allow the efficient computation of large arrays of system derivatives, which are subsequently used for model calibration, model reduction or verification and model inversion (Rabitz, 1989; Turanyi, 1990; Varma *et al.*, 1999; Cacuci, 2003; Saltelli *et al.*, 2000, pp. 81–101).

The derivative-based approach has the attraction of being very efficient in computer time. The model needs to be executed few times compared to the dimension of the array of derivatives to be computed. However, it is inefficient in terms of the analyst's time. One has to intervene in the computer program, inserting ad hoc coding, to perform this operation efficiently. Yet the fatal limitation of a derivative-based approach is that it is unwarranted when the model input is uncertain and when the model is of unknown linearity. In other words, derivatives are only informative at the base point where they are computed and do not provide for an exploration of the rest of the space of the input factors. This would matter relatively little for linear systems, in which the property at a point away from the baseline can be computed quickly by linear extrapolation using first-order point derivatives, but it would matter greatly for nonlinear ones. The focus of this book is on quantitative uncertainty and sensitivity analysis in the presence of uncertain inputs. We shall therefore make use of methods based on exploring the space of the input factors, based on the consideration that a handful of data points judiciously thrown into that space is far more effective, in the sense of being informative and robust, than estimating derivatives at a single data point in the centre of the space. In this book, when we use derivatives, or rather incremental ratios such as $(Y_j(X_i + \Delta X_i) - Y_j(X_i))/\Delta X_i$, we will normally compute them at a set of different points in the space of the input factors, in order to obtain an average response of Y_j when moving

a factor X_i of a step ΔX_i at different points in the input space, i.e. for different values of $\mathbf{X}_{\sim i}$.[9]

However, in order to introduce the methods of sensitivity analysis, we shall start from derivatives, taking a very simple test case.

1.2.2 A Test Model

Imagine the model has a linear error-free form

$$Y = \sum_{i=1}^{r} \Omega_i Z_i \tag{1.3}$$

where the input factors are $\mathbf{X} = (\Omega_1, \Omega_2, \ldots . \Omega_r, Z_1, Z_2, \ldots Z_r)$.

We have dropped the subscript j of the model output Y for simplicity. Model equation (1.3) has just a single output variable. Let us assume first that the Ω's are fixed coefficients, so that the true (active) factors for model (1.3) are just the $Z_1, Z_2, \ldots Z_r$. Y could be a composite indicator, for example a sustainability index or a greenhouse gas emission index, in which the Ω's are the weights attached by experts to the individual Z-variables. For the sake of the example we consider the weights fixed, while the individual variables have been characterized as independent and distributed normally with mean zero, i.e.

$$Z_i \sim N(0, \sigma_{Z_i}) \quad i = 1, 2, \ldots, r. \tag{1.4}$$

If the model were indeed a composite indicator with 'standardized' variables[10] all σ_{Z_i}'s would be equally one.

It is easy to verify (see the Exercises) that, given the Equations (1.3, and 1.4), Y will also be normally distributed with parameters

$$\bar{y} = \sum_{i=1}^{r} \Omega_i \bar{z}_i \tag{1.5}$$

$$\sigma_Y = \sqrt{\sum_{i=1}^{r} \Omega_i^2 \sigma_{Z_i}^2}. \tag{1.6}$$

[9] Here, and in the following, $\mathbf{X}_{\sim i}$ denotes the vector of all factors but X_i.

[10] Standardization of a variable is achieved by subtracting from the variable its sample mean and dividing the result by its standard deviation.

For the sake of the example we would also like to assume that for this particular index the variables have been ordered from the less uncertain to the most uncertain, i.e.

$$\sigma_{Z_1} < \sigma_{Z_2} < \ldots < \sigma_{Z_r},$$

and that the weights Ω's are all equal and constant:

$$\Omega_1 = \Omega_2 = \ldots = \Omega_r = \text{constant}. \tag{1.7}$$

1.2.3 Scatterplots versus Derivatives

Figure 1.6 shows the scatterplots Y, Z_i that we obtain by performing a Monte Carlo experiment with our model. As already mentioned (and described in more detail in Chapter 2), Monte Carlo methods are based on sampling factors' values from their distribution. In most cases factors are assumed independent so that the samples are taken from the marginal distribution of each factor. An input sample is thus produced:

$$\mathbf{M} = \begin{bmatrix} z_1^{(1)} & z_2^{(1)} & \ldots & z_r^{(1)} \\ z_1^{(2)} & z_2^{(2)} & \ldots & z_r^{(2)} \\ \ldots & \ldots & \ldots & \ldots \\ z_1^{(N-1)} & z_2^{(N-1)} & \ldots & z_r^{(N-1)} \\ z_1^{(N)} & z_2^{(N)} & \ldots & z_r^{(N)} \end{bmatrix} \tag{1.8}$$

Computing Y for each row of matrix (1.8) using model (1.3) produces the desired output vector

$$\mathbf{Y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \ldots \\ y^{(N-1)} \\ y^{(N)} \end{bmatrix} \tag{1.9}$$

where $y^{(1)}$ is the value obtained by running Equation (1.3) with the input given by the row vector $z_1^{(1)}, z_2^{(1)}, \ldots, z_r^{(1)}$, and so on for the other rows of matrix (1.8).

With this sample of model input and output one can produce r scatterplots by projecting in turn the N values of the selected output Y (assumed here to be a scalar) against the N values of each of the r input factors. These scatterplots can be used to investigate the behaviour of models.

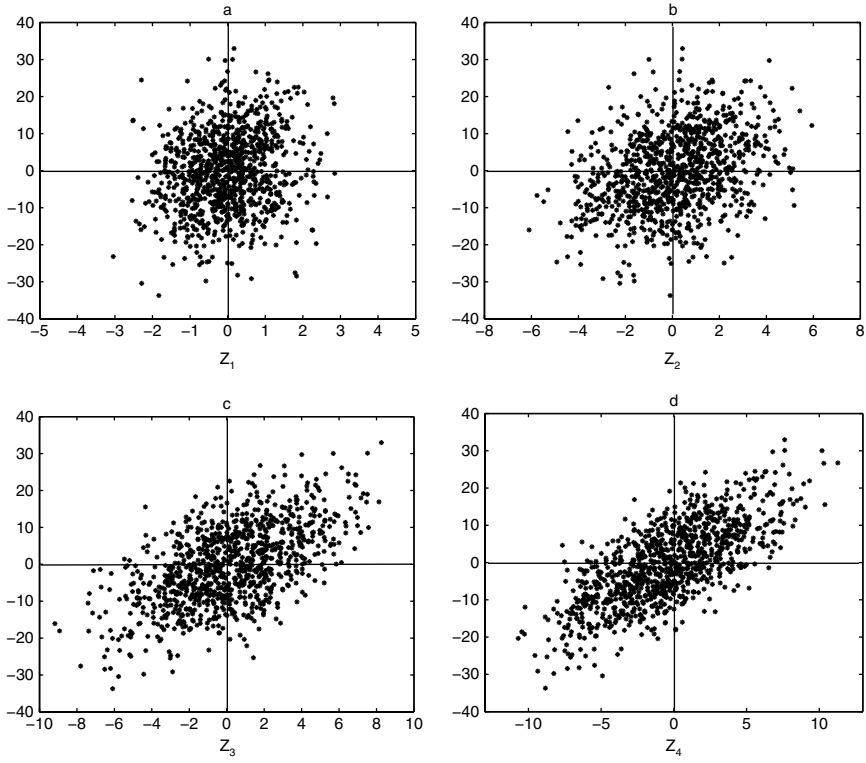

Figure 1.6 Scatterplots of Y versus $Z_1, \ldots, Z_4$. Which is the most influential factor? One can compare occupancy of quadrants I and III versus that of II and IV to decide where the positive linear relationship is stronger

The scatterplots show that Y is more sensitive to Z_4 than it is to Z_3, and that the ordering of the input factors by their influence on Y is

$$Z_4 > Z_3 > Z_2 > Z_1. \tag{1.10}$$

Such a conclusion can be drawn from Figure 1.6, as there is more shape (or a better pattern) in the plot for Z_4 than for Z_3, and so on.

However, if we used the straightforward derivative of Y versus Z_i for the sensitivity analysis, i.e. if we decided upon the relative importance of the Z_i's using the measure

$$S^p_{Z_i} = \frac{\partial Y}{\partial Z_i}, \tag{1.11}$$

which gives $S^p_{Z_i} = \Omega_i$ for Equation (1.3), we would have to conclude that all factors are equally important, based on Equation (1.7), irrespective of

the values in σ. This is clearly not reasonable. Note that we have used the superscript p for 'partial derivative' in Equation (1.11), and that the derivative is nonnormalized, i.e. it is based on the raw values of both input and output. Note also that the scatterplots in Figure 1.6 are more convincing than formula (1.11) as a sensitivity analysis tool. This is a rather general conclusion. Input/output scatterplots are in general a very simple and informative way of running a sensitivity analysis – we will use them often in this book, since they can provide an immediate visual depiction of the relative importance of the factors. For example, a scatterplot with little 'shape', e.g. plot Z_1 in Figure 1.6, which presents a rather uniform cloud of points over the range of the input factor on the abscissa, is an almost sure sign that the parameter is less influential than factor Z_4. We say 'almost' because there are instances in which a bidimensional scatterplot can be deceiving, leading to type II errors (nonidentification of an influential factor).[11] These are, however, very special cases, see Saltelli *et al.* (2004, pp. 160–161).

Most sensitivity analysis measures developed by practitioners aim to preserve the rich information provided by scatterplots in condensed format. The challenge for sensitivity analysis, in situations with many input factors, is how to rank the factors rapidly and automatically without having to look at many separate scatterplots. Another problem with scatterplots is that some uncertain factors might be sets, that is, groups of factors, and while compact sensitivity measures can be defined for sets, the sensitivities of sets cannot be visualized via simple two-dimensional scatterplots.[12]

1.2.4 Sigma-normalized Derivatives

Can we improve Equation (1.11) in such a way as to obtain a sensitivity measure that would rank the input factors consistently with Figure 1.6? A good possibility is

$$S_{Z_i}^{\sigma} = \frac{\sigma_{Z_i} \partial Y}{\sigma_Y \partial Z_i}. \tag{1.12}$$

[11] In sensitivity analysis, we refer to type I error when erroneously defining as important a noninfluential factor. Type II error occurs when we classify an important factor as noninfluential. It is nowadays common practice in modelling to include a third type of error: type III. This is typically a framing error, where right answers are sought for the wrong question. Sensitivity analysis is unable to help against type III errors. To make an example, if the range of plausible values for a factor taken as input for a sensitivity analysis is totally off the mark, the result of the sensitivity analysis will be of little help.

[12] In fact, one can force multidimensional scatterplots into a bidimensional plane by scanning the space of the input factors with a search curve. See Chapter 5.

This derivative is normalized by the input–output standard deviations (hence the σ in the superscript). Applied to model (1.3) this would give $S^{\sigma}_{Z_i} = (\sigma_{Z_i}/\sigma_Y)\Omega_i$. Squaring this and comparing with the square of $\sigma_Y = \sqrt{\sum_{i=1}^{r} \Omega_i^2 \sigma_{Z_i}^2}$ (Equation (1.6) above) we obtain

$$\sigma_Y^2 = \sum_{i=1}^{r} \Omega_i^2 \sigma_{Z_i}^2 \quad \text{and} \quad \left(S^{\sigma}_{Z_i}\right)^2 = \left(\frac{\sigma_{Z_i}}{\sigma_Y}\Omega_i\right)^2 \tag{1.13}$$

which gives $\sigma_Y^2 = \sigma_Y^2 \sum_{i=1}^{r} \left(S^{\sigma}_{Z_i}\right)^2$, and finally

$$\sum_{i=1}^{r} \left(S^{\sigma}_{Z_i}\right)^2 = 1. \tag{1.14}$$

Measure (1.12) is more convincing than measure (1.11), see Table 1.1: first, because the relative ordering of the Z_i's now depends on both vectors, σ and $\mathbf{\Omega}$, just as it should; and second, because the sensitivity measures are neatly normalized to one.

Note that Equation (1.12) is a measure recommended for sensitivity analysis by a guideline of the Intergovernmental Panel for Climate Change (IPCC) (1999, 2000).

1.2.5 Monte Carlo and Linear Regression

Let us return briefly to the scatterplots of Figure 1.6. As mentioned, these are the result of a Monte Carlo simulation in which a matrix such as

$$\mathbf{M} = \begin{bmatrix} z_1^{(1)} & z_2^{(1)} & \dots & z_r^{(1)} \\ z_1^{(2)} & z_2^{(2)} & \dots & z_r^{(2)} \\ \dots & \dots & \dots & \dots \\ z_1^{(N-1)} & z_2^{(N-1)} & \dots & z_r^{(N-1)} \\ z_1^{(N)} & z_2^{(N)} & \dots & z_r^{(N)} \end{bmatrix} \tag{1.15}$$

Table 1.1 Derivatives and normalized derivatives for the model (1.3, 1.4), where $r = 4$, $\Omega = (2, 2, 2, 2)$ and $\sigma = (1, 2, 3, 4)$

	$S^{a}_{Z_i}$	$\left(S^{\sigma}_{Z_i}\right)^2$
Z_1	2	0.036
Z_2	2	0.14
Z_3	2	0.31
Z_4	2	0.56

has been fed into model (1.3) to produce the desired output vector

$$\mathbf{Y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(N-1)} \\ y^{(N)} \end{bmatrix} \tag{1.16}$$

where $y^{(1)}$ is the value obtained running Equation (1.3) with the input given by the row vector $z_1^{(1)}, z_2^{(1)}, \dots, z_r^{(1)}$, and so on for the other rows of the matrix.

Note that N is the size of the Monte Carlo experiment ($N = 1000$ in Figure 1.6). N corresponds to the number of times we have computed Equation (1.3). In a sensitivity analysis experiment we shall have in general, instead of Equation (1.3), a computer program that calculates Y. Running the program to obtain a vector as Equation (1.9) is customarily the most expensive part of the analysis in terms of computer time, as the model may be complicated, while the sensitivity analysis measures are easy to compute. Thus N is referred to as the cost of the analysis. Note that computer time is not to be confused with analysis time. A derivation of the factors' uncertainty distribution such as Equation (1.4) is in practice the most time-consuming and financially expensive part of an analysis, especially when this is based on formal elicitation of expert opinion (Helton *et al.*, 2006; see also Saltelli *et al.*, 2000, pp. 101–152).

Note also that care has to be taken so that each column

$$\begin{matrix} z_1^{(1)} \\ z_1^{(2)} \\ \dots \\ z_1^{(N)} \end{matrix}$$

in matrix (1.8) is a sample from the respective distribution $Z_i \sim N(\bar{z}_i, \sigma_{Z_i})$. In general, and unless otherwise specified, we assume that the input factors are independent of each other, so that each one can be independently sampled from its marginal distributions (Equation (1.4) in the present examples).

As mentioned above, analysts would like to summarize the results in plots such as Figure 1.6 with a single number per scatterplot. This is, after all, what a sensitivity measure is intended to do. The most popular method for this is to try a simple linear regression on the data of matrix (1.8) and vector (1.9), of the form

$$y(i) = b_0 + \sum_{j=1}^{r} b_{Z_j} z_j^{(i)}, \tag{1.17}$$

where the coefficients b_0, b_{Z_j} are determined by least-square computation, based on the squared differences between the y-values produced by the regression (meta)model[13] and the actual model output produced by the Monte Carlo simulation. Because the points have been generated using a linear model, we expect that the linear regression will re-discover it, i.e. we would expect that $\hat{b}_0 \cong 0, \hat{b}_{Z_i} \cong \Omega_i, i = 1, 2, \ldots, r$, where the symbol $\cong$ means that this is what we would obtain if N were large enough and the hat denotes estimates as in standard usage.

Results for the points in Figure 1.6 ($N = 1000$) are given in Table 1.2.[14]

All available software for regression analysis will compute not only $\hat{b}_0, \hat{b}_{Z_i}$, but also their standardized equivalents $\hat{\beta}_{Z_i} = \hat{b}_{Z_i}\sigma_{Z_i}/\sigma_Y$. The β's are known as standardized regression coefficients (sometime indicated with their initial as SRC), and are in general more widely used than the raw regression coefficients b's. For our model (1.3), the regression coefficients, again for N tending to infinity, will tend to

$$\hat{\beta}_{Z_i} = \hat{b}_{Z_i}\sigma_{Z_i}/\sigma_Y \cong \Omega_i\sigma_{z_i}/\sigma_Y. \tag{1.18}$$

Comparing this formula with that previously obtained for the σ-normalized derivatives, i.e. $S^{\sigma}_{Z_i} = (\sigma_{Z_i}/\sigma_Y)\Omega_i$, we can conclude that in the special case of our model (1.3) the two measures of sensitivity coincide:

$$\hat{\beta}_{Z_i} = S^{\sigma}_{Z_i} \quad \text{for model (1.3).} \tag{1.19}$$

Table 1.2 Linear regression coefficients and standardized coefficients for the model of Equations (1.3), (1.4), where $r = 4$, $\Omega = (4, 3, 2, 1)$ and $\sigma = (2, 2, 2, 2)$, $N = 1000$

	b	$\beta^2_{Z_i}$	S_{Z_i} *Analytic (see Exercises)*
Intercept	0		
Z_1	2	0.034	0.03
Z_2	2	0.14	0.13
Z_3	2	0.31	0.3
Z_4	2	0.53	0.53

[13] Metamodels are surrogate models which are built to substitute for computationally intensive simulation models. Metamodels can be built with a variety of strategies (e.g. simple linear regression as discussed above) and purposes (e.g. to perform a sensitivity analysis). See Chapter 5.

[14] The results in Table 1.2 have been obtained with a simple piece of software for regression analysis. Yet, as explained in the next chapter, given that our model (1.3, 1.4) is linear, and the model does not contain any error term, we could have computed exact (analytic) values of the regression coefficients using only five runs and then applying the Kramer formula for a system of five equations (runs) in the five unknowns $b_0, \ldots b_4$.

As a result it will also be true for the β's that

$$\sum_{i=1}^{r} \left(\hat{\beta}_{Z_i}\right)^2 = 1 \tag{1.20}$$

when the model is linear.

The fact that the two measures coincide for our model can be generalized only to linear models and no further. If the model is nonlinear, the two measures will be different. Yet the β's will be a more robust and reliable measure of sensitivity even for nonlinear models. First of all, the β's are multidimensionally averaged measures. Unlike $S^{\sigma}_{Z_i}$, which is computed at the midpoint of the distribution of Z_i while keeping all other factors fixed at their midpoint, $\hat{\beta}_{Z_i}$ is the result of an exploration of the entire space of the input factors – the limit being in the dimension N of the sample. For small N and large r, however, the β's will be rather imprecise. Even in sensitivity analysis there is no such thing as a free meal, and one cannot expect to have explored a high-dimensionality space with a handful of points. Nevertheless a handful is better than just one. Statistical significance tests are available for the β's, so that the analysts can at least know the extent of the problem. Finally, by computing $\sum_{i=1}^{r} \left(\hat{\beta}_{Z_i}\right)^2$ or a related statistic, one will obtain a number, in general less than one, equal to the fraction of linearity of the model. More precisely, this number – known as the model coefficient of determination, and written as R^2_{Y} – is equal to the fraction of the variance of the original data which come from our model (Equations 1.3, 1.4, in this case), which is explained by the regression model of Equation (1.17). Again, this fraction should be equal to one for our model; however, to give a different example, if R^2_{Y} were instead to be of the order of 0.9, then the model would be 90% linear and one could use the β's for sensitivity analysis, at the risk of remaining ignorant of some 10% of the variance of the problem.[15]

Note that

$$\sum_{i=1}^{r} \left(\hat{\beta}_{Z_i}\right)^2 = 1 = \sum_{i=1}^{r} \left(\hat{b}_{Z_i}\sigma_{Z_i}/\sigma_{\mathrm{Y}}\right)^2, \tag{1.21}$$

so that

$$\sum_{i=1}^{r} \left(\hat{b}_{Z_i}\sigma_{Z_i}\right)^2 = \sigma^2_{\mathrm{Y}} = V(Y), \tag{1.22}$$

where $V(Y)$ indicates the variance of Y. Equation (1.22) is to highlight that both Equations (1.12) and (1.20) are variance decomposition formulas. As a

[15] This discussion holds for linear regression. More sophisticated metamodelling techniques which can overcome these shortcomings are described in Chapter 5.

sensitivity analysis tool, these formulas allow us to decompose the variance of the model output, taken as a descriptor of output uncertainty. Although most practitioners tend to agree on this usage of variance as a proxy for uncertainty, one should remember that the two things are not identical. For example, a measure of uncertainty could be defined on the basis of entropy of model output (see Saltelli *et al.*, 2000, pp. 56–57). In this book we shall use variance decomposition schemes for sensitivity analysis whenever the setting of the analysis allows it.

Wrapping up the results so far, we have seen formulas for decomposing the variance of the model output of interest according to the input factors. Yet we would like to do this for all models, independently of their degree of linearity; that is, we would like to be able to decompose the variance of Y even for models with a low R^2_Y. We want to find what is referred to in the literature as a 'model-free' approach. One such 'model-free' sensitivity measure is based on averaged partial variances, which we now move on to describe along two separate lines.

1.2.6 Conditional Variances – First Path

We have a generic model

$$Y = f(X_1, X_2, \ldots, X_k) \tag{1.23}$$

like model (1.3) above. Each X has a nonnull range of variation or uncertainty and we wish to determine what would happen to the uncertainty of Y if we could fix a factor. Imagine that we fix factor X_i at a particular value x_i^*. Let $V_{\mathbf{X}_{\sim i}}(Y \mid X_i = x_i^*)$ be the resulting variance of Y, taken over $\mathbf{X}_{\sim \mathbf{i}}$ (all factors but X_i). We call this a conditional variance, as it is conditional on X_i being fixed to x_i^*. We would imagine that, having frozen one potential source of variation (X_i), the resulting variance $V_{\mathbf{X}_{\sim i}}(Y \mid X_i = x_i^*)$ will be less than the corresponding total or unconditional variance $V(Y)$. One could therefore conceive of using $V_{\mathbf{X}_{\sim i}}(Y \mid X_i = x_i^*)$ as a measure of the relative importance of X_i, reasoning that the smaller $V_{\mathbf{X}_{\sim i}}(Y \mid X_i = x_i^*)$, the greater the influence of X_i. There are two problems with this approach. First, it makes the sensitivity measure dependent on the position of point x_i^* for each input factor, which is impractical. Second, one can design a model that for particular factors X_i and fixed point x_i^* yields $V_{\mathbf{X}_{\sim i}}(Y \mid X_i = x_i^*) > V(Y)$, i.e. the conditional variance is in fact greater than the unconditional (see the Exercises at the end of this chapter). If we take instead the average of this measure over all possible points x_i^*, the dependence on x_i^* will disappear.

We write this as $E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$. This is always lower or equal to $V(Y)$, and in fact:

$$E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) + V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) = V(Y). \tag{1.24}$$

Hence a small $E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$, or a large $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$, will imply that X_i is an important factor. Note that, by Equation (1.24), $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) \leq V(Y)$. The conditional variance $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$ is called the first-order effect of X_i on Y and the sensitivity measure:

$$S_i = \frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)}{V(Y)} \tag{1.25}$$

is known as the first-order sensitivity index of X_i on Y. S_i is a number always between 0 and 1.[16] A high value signals an important variable. And vice versa? Does a small value of S_i flag a nonimportant variable? We leave this question for later and move directly on to the second path for S_i.

1.2.7 Conditional Variances – Second Path

Let us go back to the scatterplots of Figure 1.6. We have said before that what identifies an important factor is the existence of 'shape' or 'pattern' in the points, while a uniform cloud of points is a symptom (though not a proof) of a noninfluential factor. What, then, constitutes shape? We could say that we have a pattern when the distribution of Y-points over the abscissa, i.e. over the factor X_i, is nonuniform. In other words, if the X_i axis is cut into slices, does one see differences in the distribution of Y-points over the slices (Figure 1.7)? Does the mean value of Y in each slice vary across the slices (Figure 1.8)? From Figure 1.7 (which is the same as Figure 1.6, with the addition of 'slices') and Figure 1.8 we can see that factor Z_4 is more influential than Z_1, and that the ordering of factors by importance is $Z_4 > Z_3 > Z_2 > Z_1$, according to how much the mean value of Y varies from one slice to another.

We thus suggest as a sensitivity measure the quantity:

Variation over the slices of the expected value of Y within each slice.

---

[16] Here and in the following we shall tend to use the synthetic notation S_i when the factors considered are labelled X, while we use the lengthier notation, e.g S_{Z_i} or S_{Ω_i}, when the factor has a symbol different from X.

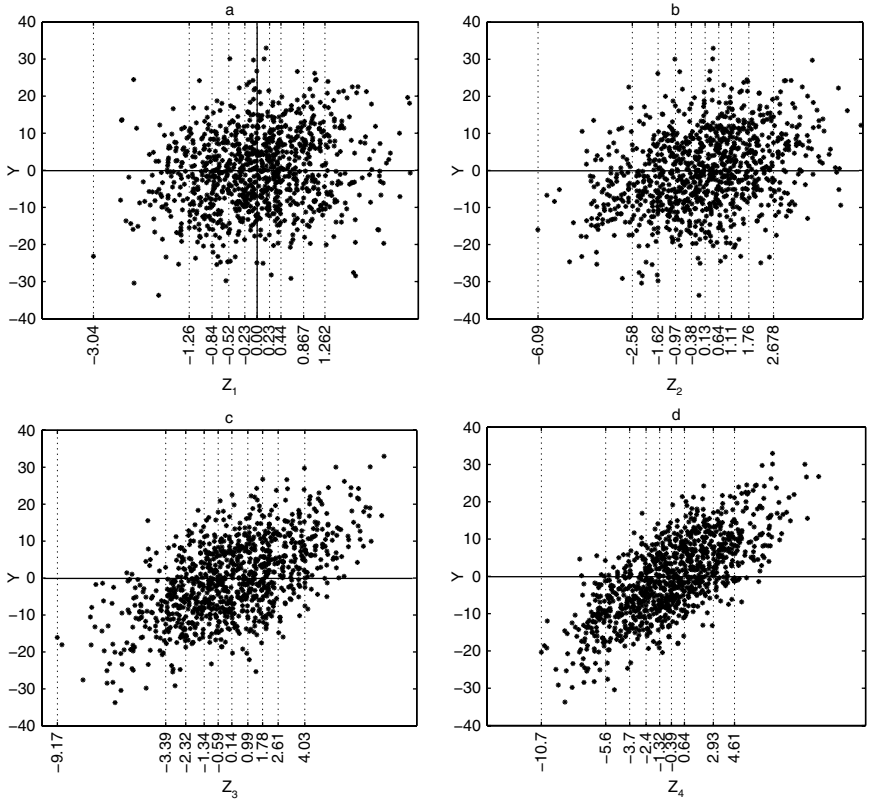

Figure 1.7 Cutting the scatterplots into slices . . .

Taking the limit of this for very thin slices we rediscover $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$. Note indeed that the expected value of Y over a very thin slice corresponds to keeping X_i fixed while averaging over all-but-X_i, which is exactly $E_{\mathbf{X}_{\sim i}}(Y \mid X_i)$. The variance operator is also easily understood.

The issue of cutting the scatterplot into slices will be taken up again in Chapter 5 in the context of metamodelling, at which point a useful approximation of the function expectation value in the slices will be introduced. We anticipate here that $E_{\mathbf{X}_{\sim i}}(Y \mid X_i)$ will be the best predictor of Y based on X_i.

1.2.8 Application to Model (1.3)

Having defined the new sensitivity measure S_i we are eager to apply it to our model of Equation (1.3). It will come as no surprise that for our well-behaved, linear model we obtain

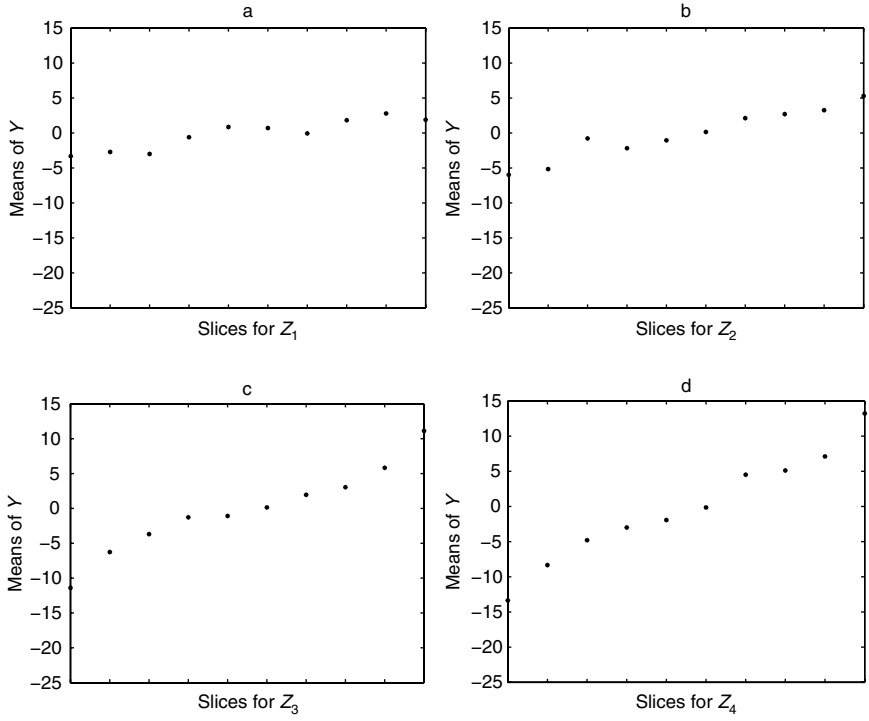

Figure 1.8 ... and taking the average within each slice. Looking at the ordinate, it is clear that Z_3 and Z_4 control more variation than Z_1 and Z_2

$$S_{Z_i} = \frac{V_{Z_i}\left(E_{\mathbf{Z}_{\sim i}}(Y \mid Z_i)\right)}{V(Y)} = \beta^2_{Z_i} \tag{1.26}$$

(See Table 1.2 for a comparison between $\beta^2_{Z_i}$ and the analytic value of S_{Z_i}.) The identity of Equation (1.26) holds for linear models, as we would expect given that S_{Z_i} is a model-free generalization of $\beta^2_{Z_i}$. For nonlinear models the two measures will differ, as we shall see in a moment. Another important difference between S_{Z_i} and $\beta^2_{Z_i}$ is that while $\sum_{i=1}^{r} \beta^2_{Z_i} = 1$ only for linear models, the relationship $\sum_{i=1}^{r} S_{Z_i} = 1$ holds for a larger class: that of additive models. By definition, a model is additive when it is possible to separate the effects of its input variables in a variance decomposition framework. For example, $Y = \sum_i Z_i^2$ is a nonlinear, additive model in the Z's; $Y = \prod_i Z_i$ is nonlinear and nonadditive.

For nonadditive models the first-order terms do not add up to one, i.e. $\sum_{i=1}^{r} S_{Z_i} \leq 1$. This is also how nonadditive models are defined. We shall turn to this presently, after first discussing the need for 'settings' in sensitivity analysis.

1.2.9 A First Setting: 'Factor Prioritization'

Experience shows that a poor definition of the objective of the sensitivity analysis (i.e. the reason we are investigating the importance of the factors, and indeed what we mean by 'importance') can lead to confused or inconclusive results. There are many statistics which can be used for sensitivity analysis, and one can easily imagine a table showing for each uncertain factor a battery of statistical measures defining the factor's importance. In general, each measure of sensitivity will produce its own ranking of factors by importance. Since this is the case, how can we tell which factor is important? To avoid this kind of confused result, it is in the analyst's best interests to define beforehand what definition of a factor's importance is relevant for the exercise in question. We call this a 'setting' (Saltelli *et al.*, 2004, pp. 49–56). A setting is a way of framing the sensitivity quest in such a way that the answer can be confidently entrusted to a well-identified measure. By way of example, we describe here the Factor Prioritization (FP) setting.

In this setting one assumes that all factors, e.g. all Z's in Equation (1.3), have a true, albeit unknown value.[17] Ideally all factors could be 'discovered' by the appropriate experiments. If all experiments have the same cost, our quest or venture is to identify which factor, once 'discovered' and fixed at its true value, would reduce $V(Y)$ the most.

One way to answer this question would be to determine or discover the value of the factors, for example, through more measurements. Yet if we were to do this we would have gone beyond uncertainty and sensitivity analyses. The challenge, therefore, is to identify the appropriate factors before any of them are measured or discovered, i.e. when the value to which each factor should be fixed is unknown. This suggests that a good contender for the title of 'most influential factor' would be that factor which, on average, once fixed, would cause the greatest reduction in variance. 'On average', in this case, means that we must average the fixing of the factor over the distribution of the factor itself. It is straightforward to see that in this setting $E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$ is the measure to use. The lower $E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$, and hence the higher $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)$, the more probable it is that factor X_i is the factor that one should measure first in order to reduce the variance most. We have thus linked the FP setting to a measure, the first-order sensitivity index S_i. This is a gamble, as we do not know

[17] In most circumstances one will have factors susceptible of determination, for which a true unknown value can be hypothesized (e.g. the failure rate of a component type, the value of an activation energy for a chemical reaction), as well as factors intrinsically uncertain (the time of failure of a specific component, the wind direction at a given time and location). These are termed epistemic and stochastic uncertainties respectively. For the purpose of illustrating the setting it is convenient to imagine all factors epistemically uncertain.

the position of the true value of a factor over its distribution. Someone actually measuring a given factor could still beat our sensitivity analysis-based guess and reduce the variance by more than we have guessed, or reduce the variance using a factor other than the one we identified via sensitivity analysis.

1.2.10 Nonadditive Models

In order to gain confidence with nonadditivity in models, we return to the input for our elementary model (1.3), $Y = \sum_{i=1}^{r} \Omega_i Z_i$, and complicate it by allowing both the Z's and the Ω's to become factors – the Ω's are no longer constants. We do this to generate a nonadditive model, as we shall see presently. The additivity of a model depends upon the characteristics of its input factors, so that it is sufficient to change a constant of the model into a factor in order to change the model from additive to nonadditive, although the model is left unchanged in the form (1.3). Our input description becomes

$$\begin{array}{ll} Z_i \sim N(\bar{z}_i, \sigma_{Z_i}) & \bar{z}_i = 0 \\ \Omega_i \sim N(\bar{\Omega}_i, \sigma_{\Omega_i}) & \bar{\Omega}_i = ic \end{array} \qquad i = 1, 2, \ldots, r. \tag{1.27}$$

The distribution of the Z's remains unchanged, while the Ω's, so far constant, become input factors with normal distribution. Their mean is not zero as it is for the Z's, but rather some number other than zero – we shall explain why in the Exercises at the end of this chapter. For the sake of the example we have made the means of the Ω's nonequivalent and equal in value to the product of the integer i (used as counter) and a positive constant c. This is simply a way to make the means of the Ω's increase, so that Equation (1.7) is no longer true. Instead

$$\bar{\Omega}_1 < \bar{\Omega}_2 < \ldots \bar{\Omega}_r. \tag{1.28}$$

The input factors for the analysis are

$$\mathbf{X} = (Z_1, Z_2, \ldots, Z_r, \Omega_1, \Omega_2, \ldots, \Omega_r) \tag{1.29}$$

and the total number of factors is $k = 2r$. We now perform another Monte Carlo experiment, sampling both the Z's and the Ω's from their respective distributions in Equation (1.27). Remember that we assume all factors independent, so each factor is sampled from its marginal distribution with no consideration of where the other factors are sampled. How the Monte Carlo sample is used to produce estimates $\hat{S}_i$ of the first-order sensitivity measures S_i is explained later in this book (see Chapter 4). We anticipate the results

Table 1.3 First-order indices S_i (analytic) and squared standardized regression coefficient β_i^2 for model (1.3, 1.27), where $r = 4$, $c = 0.5$, $\sigma = (1, 2, 3, 4)$ for both Ω_i and Z_i, and $N = 40.000$ for the regression analysis. Such a large sample was used to show the convergence between S_i and β_i^2

	S_i	$\boldsymbol{\beta}_i^2$
Z_1	0.0006	0
Z_2	0.009	0.01
Z_3	0.046	0.05
Z_4	0.145	0.14
Ω_1	0	0
Ω_2	0	0
Ω_3	0	0
Ω_4	0	0

in Table 1.3, where the squared standardized regression estimates $\hat{\beta}^2$ are also reported for comparison.

It is evident from Table 1.3 that while $\hat{S}_{Z_i}$ are still greater than zero, the $\hat{S}_{\Omega_i}$ are practically zero. Furthermore

$$\sum_{i=1}^{k} \hat{S}_{X_i} = \sum_{i=1}^{r} \hat{S}_{Z_i} + \sum_{i=1}^{r} \hat{S}_{\Omega_i} < 1. \tag{1.30}$$

We had already anticipated that for a nonadditive model the sum of the first-order indices would be less than one.

However, it might seem puzzling that the Ω input factors seem to have no influence. In fact, it is not difficult to understand why S_{Ω_i} must be zero (Figure 1.9).

Let us go back to our definition of S_i, Equation (1.25):

$$S_i = \frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right)}{V(Y)}, \tag{1.31}$$

and let us compute it for Ω_i,

$$S_{\Omega_i} = \frac{V_{\Omega_i}\left(E_{\mathbf{X}_{\sim \Omega_i}}(Y \mid \Omega_i)\right)}{V(Y)}. \tag{1.32}$$

We focus on the inner expectation $E_{\mathbf{X}_{\sim \boldsymbol{\Omega}_i}}(Y|\Omega_i)$ which we now have to write explicitly as $E_{\mathbf{X}_{\sim \Omega_i}}(Y \mid \Omega_i = \omega_i^*)$ in order to remind ourselves that we have fixed Ω_i.

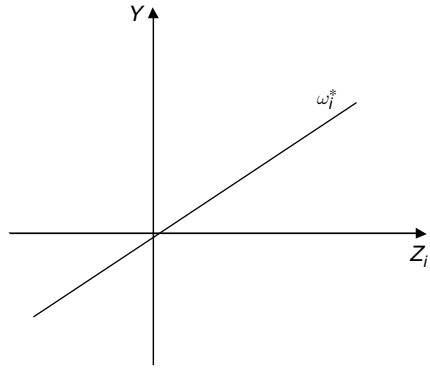

Figure 1.9 Y versus Z_i for fixed values of $\mathbf{X}_{\sim Z_i}$

Note that $E_{\mathbf{X}_{\sim\Omega_i}}$ now means that the mean is taken over all Z_j's, including Z_i, and over all Ω_j's but Ω_i.

Figure 1.9 shows the plot of Y versus Z_i for a fixed nonzero value of ω_i^* of input factor Ω_i in the case that all the remaining Ω_j's, with $j \neq i$, are fixed to zero. This straight line will be shifted up or down vertically when the Ω_j's, with $j \neq i$, are fixed to values other than zero.

Positive and negative values of Y will hence be equally probable and equally distributed, so that $E_{\mathbf{X}_{\sim\Omega_i}}(Y \mid \Omega_i = \omega_i^*)$ will be zero. Figure 1.10 shows how this emerges from Monte Carlo generated scatterplots of Y versus Z_i and Y versus Ω_i. It is clear that if $E_{\mathbf{X}_{\sim\Omega_i}}(Y \mid \Omega_i = \omega_i^*)$ is zero for any value ω_i^*, its variance over all possible values of ω_i^* will also be zero, so that both $V_{\Omega_i}(E_{\mathbf{X}_{\sim\Omega_i}}(Y \mid \Omega_i))$ and S_{Ω_i} will be zero for all factors Ω_i.

We now understand that the measure S_{Ω_i} is zero, but we retain the belief that factors Ω_i should have some influence, especially since this is suggested by the conical pattern evident in Figure 1.10. It seems therefore that there may be a problem with our sensitivity measure. A regression coefficient $\hat{\beta}_{\Omega_i}$ would produce a straight horizontal line through the horizontal conical plot of Y versus Ω_i in Figure 1.10. However, it is clear from the shape of this plot that variable Ω_j is influential. A possible interpretation is that $\hat{\beta}_{\Omega_i}$ fails as a sensitivity measure in this case. Does the fact that S_{Ω_i} is zero imply that also S_{Ω_i} fails?

Indeed it is unfair to say that β_{Ω_i} fails in Figure 1.10. β_{Ω_i} is a linear measure, so clearly it should not be used on a nonlinear model. S_{Ω_i}, however, is a model-free measure, and must be applicable to nonlinear models. Indeed this is the case, and we can say that if S_{Ω_i} is zero, this means that Ω_i has no effect on Y 'at the first order' (recall that we have thus far discussed first-order sensitivity indices). The reader familiar with experimental design

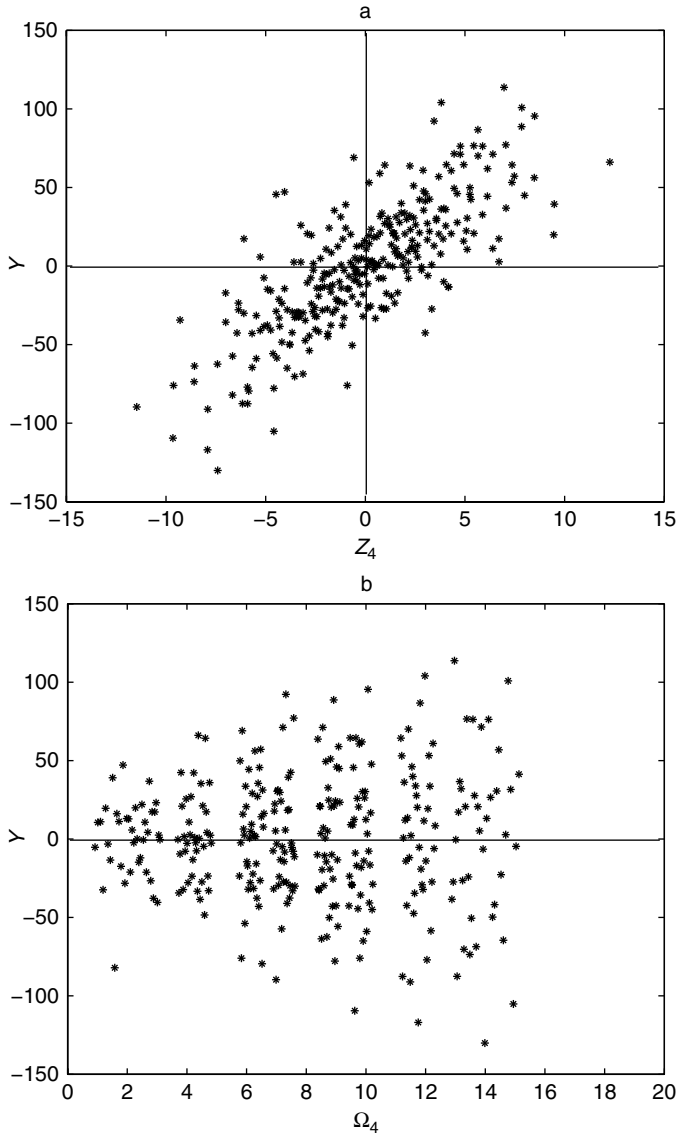

Figure 1.10 Scatterplots of Y versus Z_4 and versus Ω_4 for model (1.3, 1.27), at sample size $N = 1000$. The first-order sensitivity index for Z_4 is greater than zero while that for Ω_4 is zero

will already have guessed that the effect of Ω_i must be captured by some higher-order effect, as we now proceed to discuss.

1.2.11 Higher-order Sensitivity Indices

We continue our game with conditioned variances by playing with two factors instead of one. Take for instance

$$\frac{V(E(Y \mid Z_i, Z_j))}{V(Y)}, \tag{1.33}$$

with $i \neq j$. We have dropped the indices of both the E and V operators. Indeed we do not need them if we accept the convention that the argument conditioning the inner operator, Z_i, Z_j in this case, is also the set over which we apply the outer operator, i.e. the variance is taken over Z_i, Z_j (we should have written V_{Z_i,Z_j}). By default, the inner operator, the average E, must be taken over all-but-(Z_i, Z_j). What would happen if we could compute (1.33), with $i \neq j$, and compare it with the corresponding measure for the individual factors Z_i, Z_j? We would observe that

$$\frac{V(E(Y \mid Z_i, Z_j))}{V(Y)} = S_{Z_i} + S_{Z_j} \quad \text{for} \quad i \neq j, \tag{1.34}$$

while

$$\frac{V(E(Y \mid \Omega_i, \Omega_j))}{V(Y)} = 0, \tag{1.35}$$

and

$$\frac{V(E(Y \mid Z_i, \Omega_i))}{V(Y)} > S_{Z_i} + S_{\Omega_i}. \tag{1.36}$$

We anticipate from Chapter 4 that, given two generic factors X_i, X_j, the following result holds:

$$V(E(Y \mid X_i, X_j)) = V_i + V_j + V_{ij}, \tag{1.37}$$

where

$$\begin{aligned} V_i &= V(E(Y \mid X_i)) \\ V_j &= V(E(Y \mid X_j)) \\ V_{ij} &= V(E(Y \mid X_i, X_j)) - V_i - V_j. \end{aligned} \tag{1.38}$$

The term V_{ij} is the interaction term between factors X_i, X_j. It captures that part of the response of Y to X_i, X_j that cannot be written as a superposition of effects separately due to X_i, and X_j. Recalling our previous examples of $Y = \sum_i Z_i^2$ (a nonlinear, additive model) and $Y = \prod_i Z_i$ (nonlinear, nonadditive), the latter model will have nonzero second-order terms such as V_{ij}, while the former model will not.

Looking at Equations (1.37, 1.39) and remembering that for our model all S_{Ω_i} are zero, we are now ready to grasp the results of Equations (1.34–1.36) (see also Table 1.4).

- Equation (1.34) holds because the interaction term between Z_i and Z_j is zero, which is evident from the form of Equation (1.3).
- Equation (1.35) holds because the S_{Ω_i} and S_{Ω_j} as well as their interaction term are zero.
- Equation (1.36) can be rewritten as

$$\frac{V(E(Y \mid Z_i, \Omega_i))}{V(Y)} = S_{Z_i} + S_{\Omega_i} + S_{Z_i,\Omega_i},$$

where $S_{\Omega_i} = 0$, $S_{Z_i,\Omega_i} = V_{Z_i,\Omega_i}/V(Y)$ and the term V_{Z_i,Ω_i} is the only type of nonzero second-order term in model (1.3).

If we now sum all the nonzero first-order and and second-order terms we get

Table 1.4 First- and second-order indices for model (1.3, 1.27, analytic), where $r = 4$, $c = 0.5$, $\sigma = (1, 2, 3, 4)$ for both Ω_i and Z_i

Factor	S_i, S_{ij}	*Factor*	S_{ij}	*Factor*	S_{ij}
Z_1	0.0006	Z_1, Ω_2	0	Z_3, Ω_3	0.183
Z_2	0.009	Z_1, Ω_3	0	Z_3, Ω_4	0
Z_3	0.046	Z_1, Ω_4	0	Z_4, Ω_1	0
Z_4	0.145	Z_2, Z_3	0	Z_4, Ω_2	0
Ω_1	0	Z_2, Z_4	0	Z_4, Ω_3	0
Ω_2	0	Z_2, Ω_1	0	Z_4, Ω_4	0.578
Ω_3	0	Z_2, Ω_2	0.036	Ω_1, Ω_2	0
Ω_4	0	Z_2, Ω_3	0	Ω_1, Ω_3	0
Z_1, Z_2	0	Z_2, Ω_4	0	Ω_1, Ω_4	0
Z_1, Z_3	0	Z_3, Z_4	0	Ω_2, Ω_3	0
Z_1, Z_4	0	Z_3, Ω_1	0	Ω_2, Ω_4	0
Z_1, Ω_1	0.002	Z_3, Ω_2	0	Ω_3, Ω_4	0

$$\sum_{i=1}^{r} \left(S_{Z_i} + S_{Z_i \Omega_i} \right) = 1. \tag{1.39}$$

This means that even for a nonadditive model we have found a way to recover (that is, to understand) 100% of the variance of Y. Thus variance-based sensitivity measures provide a theoretical framework whereby – provided one has the patience to compute all interaction terms – one can achieve a full understanding of the model's sensitivity pattern. Patience is indeed required, as in principle a model can have interactions of even higher order. Again anticipating one result from Chapter 4, a full analysis of a model with k factors is composed of

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>j} S_{ijl} + \ldots + S_{123\ldots k} = 1. \tag{1.40}$$

Model (1.3) can only have nonzero terms up to the second order, and this can be seen 'by inspection', as the structure of the model is very simple. In practical applications the subject model of our analysis will be a computer program, and the only way to ascertain whether an interaction exists or not will be to estimate it numerically. The problem is that the series development of Equation (1.40) has as many as $2^k - 1$ terms. For $k = 3$ this gives just 7 terms, i.e. $S_1, S_2, S_3 . S_{12}, S_{23}, S_{13}, S_{123}$; for $k = 10$ it gives 1023, too many to look at in practice.

In fact, the variance-based analysis can help us in these circumstances, by computing for each factor a 'total effect' term, which we describe next.

1.2.12 Total Effects

What is a total effect term? Let us again use our extended model (1.3, 1.27), and ask what we would obtain if we were to compute $V\left(E\left(Y \mid \mathbf{X}_{\sim \Omega_i}\right)\right)/V(Y)$. We are conditioning now on all factors but Ω_i. In other words

$$\frac{V(E(Y|\mathbf{X}_{\sim \Omega_i}))}{V(Y)} = \frac{V\left(E\left(Y \mid \Omega_1, \Omega_2, \ldots, \Omega_{i-1}, \Omega_{i+1}, \ldots \Omega_r, Z_1, Z_2, \ldots, Z_r\right)\right)}{V(Y)}. \tag{1.41}$$

By analogy with our discussion of second-order terms, Equation (1.41) should include all terms of any order that do not include factor Ω_i. As the sum of all possible sensitivity terms must be 1, the difference

$$\left(1-\frac{V(E(Y|\mathbf{X}_{\sim\Omega_i}))}{V(Y)}\right)$$

must be made up of all terms of any order that include ω_i. For our model, which has only first- and second-order terms, this gives

$$\left(1-\frac{V(E(Y|\mathbf{X}_{\sim\Omega_i}))}{V(Y)}\right)=S_{\Omega_i}+S_{Z_i\Omega_i}\quad\text{(See Table 1.5)}\tag{1.42}$$

To consider a different example, for a generic three-factor model, one would have

$$S_{T1}=\left(1-\frac{V\left(E\left(Y\mid\mathbf{X}_{-1}\right)\right)}{V(Y)}\right)=S_1+S_{12}+S_{13}+S_{123}\tag{1.43}$$

and

$$S_{T2}=S_2+S_{12}+S_{23}+S_{123}$$
$$S_{T3}=S_3+S_{13}+S_{23}+S_{123},$$

where S_{Ti} denotes the total effect of factor X_i. We recall that we tend to use the synthetic notation $(S_i, S_{Ti}, V_i, S_{ij})$ when the factors considered are labelled X, while we use the lengthier notation $(S_{Z_i}, S_{TZ_i}, V_{\Omega_i}, S_{Z_i\Omega_i})$ when the factor has a symbol other than X.

Table 1.5 First-order and total effects for model (1.3, 1.27, analytic), where $r=4$, $c=0.5$, $\sigma=(1,2,3,4)$ for both Ω_i and Z_i

	S_i		S_{Ti}
Z_1	0.0006	Z_1	0.003
Z_2	0.009	Z_2	0.045
Z_3	0.046	Z_3	0.229
Z_4	0.145	Z_4	0.723
Ω_1	0	Ω_1	0.002
Ω_2	0	Ω_2	0.036
Ω_3	0	Ω_3	0.183
Ω_4	0	Ω_4	0.578

We have argued in a series of works (Saltelli *et al.*, 2004, and references therein) that a good, synthetic, though nonexhaustive characterization of the sensitivity pattern for a model with k factors is given by the total set of first-order terms plus the total effects. For a system with 10 factors this makes 20 terms rather than 1023.

One last observation about the total effect terms is the following. For the algebraic rule already mentioned in Equation (1.24) we have

$$E_{X_i}\left(V_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) + V_{X_i}\left(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)\right) = V(Y),$$

and hence

$$S_{Ti} = 1 - \frac{V\left(E(Y \mid \mathbf{X}_{\sim i})\right)}{V(Y)} = \frac{E\left(V(Y \mid \mathbf{X}_{\sim i})\right)}{V(Y)}. \tag{1.44}$$

Equipped with this new sensitivity measure, the total effect, we are now ready to introduce another useful 'setting' for sensitivity analysis.

1.2.13 A Second Setting: 'Factor Fixing'

One use of sensitivity analysis is to simplify models. If a model is used systematically in a Monte Carlo framework, so that input uncertainties are always propagated through the output, it might be useful to ascertain which of the input factors can be fixed anywhere in their range of variation without appreciably affecting a specific output of interest Y. This could help to simplify a model in a greater sense, since we might be able to condense (lump) an entire section of our model if all factors entering that section are noninfluential. From the preceding discussion it will be clear that $S_i = 0$ is a necessary but insufficient condition for fixing factor X_i. This factor might be involved in interactions with other factors such that, although its first-order term is zero, there might be nonzero higher-order terms. This is exactly what happened with our factors Ω_i in the model (1.3, 1.27).

Imagine now that a factor X_i is truly noninfluential. Let us compute $V_{X_i}(Y \mid \mathbf{X}_{\sim i} = \mathbf{x}^*_{\sim i})$, where we have fixed a point $\mathbf{x}^*_{\sim i}$ in the multidimensional space $\mathbf{X}_{\sim i}$. If factor X_i is noninfluential, then $V_{X_i}(Y \mid \mathbf{X}_{\sim i} = \mathbf{x}^*_{\sim i})$ must be zero, as the value of Y is totally determined by $\mathbf{X}_{\sim i}$ and there will be no variance over X_i. Averaging over non-X_i will not change the result, so that $E_{\mathbf{X}_{\sim i}}\left(V_{X_i}(Y \mid \mathbf{X}_{\sim i})\right)$ must be zero as well. Based on our convention of not indicating the conditioning argument, we can also write this as $E(V(Y \mid \mathbf{X}_{\sim i})) = 0$. These considerations prove that if X_i is noninfluential, then $S_{Ti} = 0$ by Equation (1.44) above.

Conversely, if $S_{Ti} = 0$ for factor X_i, then $E(V(Y \mid \mathbf{X}_{\sim i})) = 0$. As the variance can only be a positive number, the fact that the mean $E(V(Y \mid \mathbf{X}_{\sim i}))$ is zero implies that $V(Y \mid \mathbf{X}_{\sim i} = \mathbf{x}^*_{\sim i})$ is identically zero for any value of $\mathbf{x}^*_{\sim i}$, which proves that X_i is noninfluential – there is no point in the hyperspace of $\mathbf{X}$ where X_i has an effect. This demonstrates that $S_{Ti} = 0$ is a necessary and sufficient condition for X_i being noninfluential.

Note that the model simplification underpinned by the 'factor fixing' setting can become very important when models need to be audited, for example in the face of a scientific controversy or for use in policy assessment. In these situations one might wish to optimize the 'relevance' R of a model, defined as the ratio (Beck *et al.*, 1997):

$$R = \frac{\text{number of factors that truly induce variations in the output of interest}}{\text{total number of factors in the model}}.$$

This approach would guard against the criticism that an overly complex model was being used by one party to obfuscate or discourage investigation.

The concepts of parsimony or simplicity in the context of modelling are illustrated by the works of Peter C. Young (Young *et al.*, 1996; Young, 1999a), who recommends the use of data-driven models, in which a minimum of parameters are inferred directly from the data, as an alternative to law-driven, usually overparametrized models. To give an example, the hydrogeology of a catchment area can be modelled with a complex model based on Darcy's laws or with a low-order model based on direct interpretation of precipitation and runoff time series. Such a parsimonious description of the system can also be thought of as a complement to a law-driven model. More generally, for models to be used in impact assessment or other regulatory settings, it might be advisable to have a back-of-the-envelope version of the general model for the purpose of negotiating assumptions and inferences with stakeholders. Sensitivity analysis may be instrumental in deriving such a simplified model.

The foregoing discussion of possible settings for sensitivity analysis allows us to make a few more observations on the rationale for sensitivity analysis.

1.2.14 Rationale for Sensitivity Analysis

Possible motivations for sensitivity analysis are:

- Model corroboration. Is the inference robust? Is the model overly dependent on fragile assumptions?

- Research prioritization. Which factor is most deserving of further analysis or measurement? ⇒ factor prioritization setting.
- Model simplification. Can some factors or compartments of the model be fixed or simplified? ⇒ factor fixing setting.
- Identifying critical or otherwise interesting regions in the space of the input factors. Identifying factors which interact and which may thus generate extreme values. This is important, for example in system reliability.
- Prior to parameter estimation, to help set up the (actual or numerical) experiment in those conditions in which the sensitivity of the output to the factor to be estimated is the greatest.

To illustrate the last point, imagine that one has actual measurements against which to compare model predictions. Ideally, predictions and measurements can feed into an estimation step. Yet before this is done, it is worth investigating what drives, for instance, the sum of the squared differences between model prediction and actual measurements. Only factors with this type of influence are good candidates for the estimation step. In this way the analyst can decide which experimental conditions are more interesting for the subsequent estimation (Saltelli *et al.*, 2004, pp.151–191).

We have already mentioned that uncertainty and sensitivity analyses can be run in tandem to ascertain whether different policies (e.g. strategies to alleviate an environmental problem) are indeed different from one another when compared in the overall space of the uncertainties. An example of such an analysis is found in Saltelli *et al.* (2000 pp. 385–397).

It is worth noting in this case that high uncertainty in the inference is not synonymous with low quality in the resulting assessment. Though uncertain, the assessment might still allow policy A to be distinguished from policy B (implying high quality) while the opposite is also possible, i.e. that the model might not allow these options to be distinguished even with only moderate uncertainties in the inference (implying a low-quality assessment). On a similar ground, when confronted with a plurality of stakeholders' views and beliefs as to how an issue should be tackled or framed, we may use sensitivity analysis to ascertain whether – within the latitude of the different framings and assumptions – we still can reach some robust inference, i.e. a high-quality assessment. We would call such an inference – or the resulting preferred policy – 'socially' robust, as it is compatible with such a plurality of viewpoints. On the contrary, we might find that the different framings give rise to such great latitude in the resulting inference that no robust policy can be identified.

Another general consideration with respect to the global, explorative nonparametric methods for the sensitivity analysis just described is that these have a better chance of being resilient towards type II errors than local (derivative-based) methods. The possibility of important factors

being overlooked or dangerous or critical combinations of input factors neglected decreases with the level of exploration of the space of the input factors (Farrell, 2007). The attention paid in global methods to interaction effects is also a protection against type II errors. In Saltelli *et al.* (2005) we show that, for even a relatively simple and well-studied chemical reactor system, global sensitivity analysis and attention to the interactions can lead to the identification of a larger 'runaway' portion in the space of the input factors than could previously be identified.

Some of the motivations just described would demand being able to apportion uncertainty not only among factors, but also among sets of factors, for example to distinguish data uncertainty from experts' uncertainty, system uncertainty from policy option uncertainty and so on. We offer a few tools for this in the following.

1.2.15 Treating Sets

An additional interesting feature of variance-based methods is that they allow for a concise treatment of the sensitivity of sets of factors. Referring again to model (1.3, 1.27), we can imagine computing a variance measure conditioned on a subset of the input factors, e.g. on the set $\mathbf{\Omega}$, $S_{\mathbf{\Omega}} = V(E(Y \mid \mathbf{\Omega}))/V(Y)$. From the description in the previous sections it is easy to understand that $S_{\mathbf{\Omega}}$ will include all first-order terms related to $\mathbf{\Omega}$ plus second- and higher-order product terms including only members of $\mathbf{\Omega}$. We already know that these are all zero. We can likewise compute $S_{\mathbf{Z}} = V(E(Y \mid \mathbf{Z}))/V(Y)$ for the set $\mathbf{Z}$. This similarly contains all nonzero first-order terms plus the null second- and higher-order terms internal to $\mathbf{Z}$. Finally we can compute

$$S_{\mathbf{\Omega},\mathbf{Z}} = V(Y) - S_{\mathbf{\Omega}} - S_{\mathbf{Z}}, \tag{1.45}$$

which will contain all cross-product terms not involved in $S_{\mathbf{\Omega}}, S_{\mathbf{Z}}$. Going back to our example of Equation (1.3) as a composite indicator with weights $\mathbf{\Omega}$ given by experts and variables $\mathbf{Z}$ coming from statistical offices, with Equation (1.45) we have apportioned variance between data and experts and an interaction between the two.

Similarly, we could share the uncertainty in Y among the couples $\mathbf{A}_i = (\Omega_i, Z_i)$ and apply

$$\sum_i S_{\mathbf{A}_i} + \sum_i \sum_{j>i} S_{\mathbf{A}_i\mathbf{A}_j} + \sum_i \sum_{j>i} \sum_{l>j} S_{\mathbf{A}_i\mathbf{A}_j\mathbf{A}_l} + \ldots = 1. \tag{1.46}$$

As we already know that for our model all cross-product terms with $i \neq j$ are zero, this can be reduced to the convenient

$$\sum_{i}^{r} S_{\mathbf{A}_i} = 1, \tag{1.47}$$

in which uncertainty is divided among 'themes', each theme comprising an indicator and its weight. It is easy to imagine similar applications. For example, one could divide uncertainty among observational data, estimation, model assumptions, model resolution and so on.

1.2.16 Further Methods

So far we have discussed the following tools for sensitivity analysis:

- derivatives and sigma-normalized derivatives;
- regression coefficients (standardized);
- variance-based measures;
- scatterplots.

We have shown the equivalence of sigma-normalized coefficients $S_i^{\sigma} = \sigma_{Z_i} \partial Y / \sigma_Y \partial X_i$, regression coefficients β_i and variance-based first-order sensitivity indices S_i for linear models, as well as how S_i is a model-free extension of the variance decomposition scheme to models of unknown linearity. We have discussed how nonadditive models can be treated in the variance-based sensitivity framework. We have also indicated that scatterplots are a powerful tool for sensitivity analysis and shown how S_i can be interpreted in relation to the existence of 'shape' in an X_i versus Y scatterplot.

At a greater level of detail (Ratto *et al.*, 2007) one can use modern regression tools (such as state-space filtering methods) to interpolate points in the scatterplots, producing very reliable $E(Y \mid X_i = x_i^*)$ curves. The curves can then be used for sensitivity analysis. Their shape is more evident than that of dense scatterplots (compare Figure 1.7 with Figure 1.8). Furthermore, one can derive the first-order sensitivity indices directly from those curves, so that an efficient way to estimate S_i is to use state-space regression on the scatterplots and then take the variances of these.

In general, for a model of unknown linearity, monotonicity and additivity, variance-based measures constitute a good means of tackling settings such as factor fixing and factor prioritization. We shall discuss one further setting before the end of this chapter, but let us first consider whether there are alternatives to the use of variance-based methods for the settings so far described.

Why might we need an alternative? The main problem with variance-based measures is computational cost. Estimating the sensitivity coefficients takes many model runs (see Chapter 4). Accelerating the computation of sensitivity indices of all orders – or even simply of the S_i, S_{Ti} couple – is the most intensely researched topic in sensitivity analysis (see the filtering approach just mentioned). It can reasonably be expected that the estimation of these measures will become more efficient over time.

At the same time, and if only for screening purposes, it would be useful to have methods to find approximate sensitivity information at lower sample sizes. One such method is the Elementary Effect Test.

1.2.17 Elementary Effect Test

The Elementary Effect Test is simply an average of derivatives over the space of factors. The method is very simple. Consider a model with k independent input factors $X_i, i = 1, 2, \ldots, k$, which varies across p levels. The input space is the discretized p-level grid Ω. For a given value of $\mathbf{X}$, the elementary effect of the ith input factor is defined as

$$EE_i = \frac{[Y(X_1, X_2, \ldots, X_{i-1}, X_i + \Delta, \ldots X_k) - Y(X_1, X_2, \ldots, X_k)]}{\Delta}, \quad (1.48)$$

where p is the number of levels, Δ is a value in $\{1/(p-1), \ldots, 1-1/(p-1)\}$, $X = (X_1, X_2, \ldots X_k)$ is any selected value in Ω such that the transformed point $(\mathbf{X} + \mathbf{e}_i\Delta)$ is still in Ω for each index $i = 1, \ldots, k$, and $\mathbf{e}_i$ is a vector of zeros but with a unit as its ith component. Then the absolute values of the EE_i, computed at r different grid points for each factor, are averaged

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r} |EE_i^j| \quad (1.49)$$

and the factors ranked according to the obtained mean μ_i^*.

In order to compute efficiently, a well-chosen strategy is needed for moving from one effect to the next, so that the input space is explored with a minimum of points (see Chapter 3).

Leaving aside computational issues for the moment, μ^* is a useful measure for the following reasons:

1. It is semi-quantitative – the factors are ranked on an interval scale;
2. It is numerically efficient;
3. It is very good for factor fixing – it is indeed a good proxy for S_{Ti};
4. It can be applied to sets of factors.

Due to its semi-quantitative nature the μ^* can be considered a screening method, especially useful for investigating models with many (from a few dozen to 100) uncertain factors. It can also be used before applying a variance-based measure to prune the number of factors to be considered. As far as point (3) above is concerned, μ^* is rather resilient against type II errors, i.e. if a factor is deemed noninfluential by μ^* it is unlikely to be identified as influential by another measure.

1.2.18 Monte Carlo Filtering

While μ^* is a method of tackling factor fixing at lower sample size, the next method we present is linked to an altogether different setting for sensitivity analysis. We call this 'factor mapping' and it relates to situations in which we are especially concerned with a particular portion of the distribution of output Y. For example, we are often interested in Y being above or below a given threshold. If Y were a dose of contaminant, we might be interested in how much (how often) a threshold level for this contaminant is being exceeded. Or Y could be a loss (e.g. financial) and we might be interested in how often a maximum admissible loss is being exceeded. In these settings we tend to divide the realization of Y into 'good' and 'bad'. This leads to Monte Carlo filtering (MCF, see Saltelli *et al.*, 2004, pp. 151–191 for a review). In MCF one runs a Monte Carlo experiment producing realizations of the output of interest corresponding to different sampled points in the input factor space, as for variance-based or regression analysis. Having done this, one 'filters' the realizations, e.g. elements of the Y-vector. This may entail comparing them with some sort of evidence or for plausibility (e.g. one may have good reason to reject all negative values of Y). Or one might simply compare Y against thresholds, as just mentioned. This will divide the vector Y into two subsets: that of the well-behaved realizations and that of the 'misbehaving' ones. The same will apply to the (marginal) distributions of each of the input factors. Note that in this context one is not interested in the variance of Y as much as in that part of the distribution of Y that matters – for example, the lower-end tail of the distribution may be irrelevant compared to the upper-end tail or vice versa, depending on the problem. Thus the analysis is not concerned with which factor drives the variance of Y as much as with which factor produces realizations of Y in the forbidden zone. Clearly, if a factor has been judged noninfluential by either μ^* or S_{Ti}, it will be unlikely to show up in an MCF. Steps for MCF are as follows:

- A simulation is classified as either B, for behavioural, or $\overline{B}$, for nonbehavioural (Figure 1.11).
- Thus a set of binary elements is defined, allowing for the identification of two subsets for each X_i: one containing a number n of elements denoted

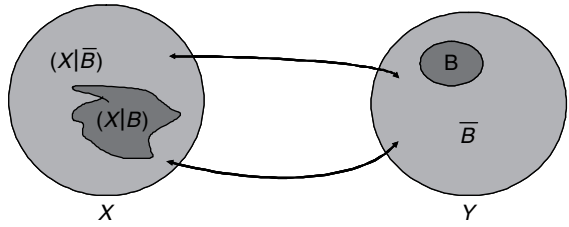

Figure 1.11 Mapping behavioural and nonbehavioural realizations with Monte Carlo filtering

$(X \mid B)$ and a complementary set $\left(X \mid \overline{B}\right)$ containing the remaining $\overline{n} = N - n$ simulations (Figure 1.11).

- A statistical test can be performed for each factor independently, analysing the maximum distance between the cumulative distributions of the $(X \mid B)$ and $\left(X \mid \overline{B}\right)$ sets (Figure 1.12).

If the two sets are visually and statistically[18] different, then X_i is an influential factor in the factor mapping setting.

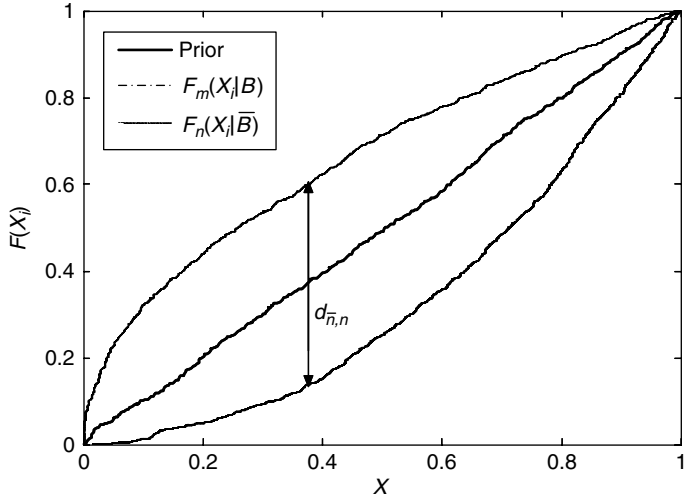

Figure 1.12 Distinguishing between the two sets using a test statistic

[18] Smirnov two-sample test (two-sided version) is used in Figure 1.12 (see Saltelli *et al.*, 2004, pp. 38–39).

1.3 NONINDEPENDENT INPUT FACTORS

Throughout this introductory chapter we have systematically assumed that input factors are independent of one another. The main reason for this assumption is of a very practical nature: dependent input samples are more laborious to generate (although methods are available for this; see Saltelli *et al.*, 2000) and, even worse, the sample size needed to compute sensitivity measures for nonindependent samples is much higher than in the case of uncorrelated samples.[19] For this reason we advise the analyst to work on uncorrelated samples as much as possible, e.g. by treating dependencies as explicit relationships with a noise term.[20] Note that when working with the MCF just described a dependency structure is generated by the filtering itself. The filtered factors will probably correlate with one another even if they were independent in the original unfiltered sample. This could be a useful strategy to circumvent the use of correlated samples in sensitivity analysis. Still there might be very particular instances where the use of correlated factors is unavoidable. A case could occur with the parametric bootstrap approach described in Figure 1.3. After the estimation step the factors will in general be correlated with one another, and if a sample is to be drawn from these, it will have to respect the correlation structure.

Another special instance when one has to take factors' dependence into consideration is when analyst A tries to demonstrate the falsity of an uncertainty analysis produced by analyst B. In such an adversarial context, A needs to show that B's analysis is wrong (e.g. nonconservative) even when taking due consideration of the covariance of the input factors as explicitly or implicitly framed by B.

1.4 POSSIBLE PITFALLS FOR A SENSITIVITY ANALYSIS

As mentioned when discussing the need for settings, a sensitivity analysis can fail if its underlying purpose is left undefined; diverse statistical tests and measures may be thrown at a problem, producing a range of different factor rankings but leaving the researcher none the wiser as to which

[19] Dependence and correlation are not synonymous. Correlation implies dependence, while the opposite is not true. Dependencies are nevertheless described via correlations for practical numerical computations.

[20] Instead of entering X_1 and X_2 as correlated factors one can enter X_1 and X_3, with X_3 being a factor describing noise, and model X_2 as a function of X_1 and X_3.

ranking to believe or privilege. Another potential danger is to present sensitivity measures for too many output variables Y. Although exploring the sensitivity of several model outputs is sound practice for testing the quality of the model, it is better, when presenting the results of the sensitivity analysis, to focus on the key inference suggested by the model, rather than to confuse the reader with arrays of sensitivity indices relating to intermediate output variables. Piecewise sensitivity analysis, such as when investigating one model compartment at a time, can lead to type II errors if interactions among factors of different compartments are neglected. It is also worth noting that, once a model-based analysis has been produced, most modellers will not willingly submit it to a revision via sensitivity analysis by a third party.

This anticipation of criticism by sensitivity analysis is also one of the 10 commandments of applied econometrics according to Peter Kennedy:

> Thou shall confess in the presence of sensitivity. Corollary: Thou shall anticipate criticism [$\cdots$] When reporting a sensitivity analysis, researchers should explain fully their specification search so that the readers can judge for themselves how the results may have been affected. This is basically an 'honesty is the best policy' approach, advocated by Leamer, (1978, p. vi) (Kennedy, 2007).

To avoid this pitfall, an analyst should implement uncertainty and sensitivity analyses routinely, both in the process of modelling and in the operational use of the model to produce useful inferences.

Finally the danger of type III error should be kept in mind. Framing error can occur commonly. If a sensitivity analysis is jointly implemented by the owner of the problem (which may coincide with the modeller) and a practitioner (who could again be a modeller or a statistician or a practitioner of sensitivity analysis), it is important to avoid the former asking for just some 'technical help' from the latter upon a predefined framing of the problem. Most often than not the practitioner will challenge the framing before anything else.

1.5 CONCLUDING REMARKS

1. We have just shown different settings for sensitivity analysis, such as:

   - factor prioritization, linked to S_i;
   - factor fixing, linked to S_{Ti} or μ^*;
   - factor mapping, linked to MCF;
   - metamodelling (hints).

The authors have found these settings useful in a number of applications. This does not mean that other settings cannot be defined and usefully applied.

2. We have discussed fitness for purpose as a key element of model quality. If the purpose is well defined, the output of interest will also be well identified. In the context of a controversy, this is where attention will be focused and where sensitivity analysis should be concentrated.
3. As discussed, a few factors often account for most of the variation. Advantage should be taken of this feature to simplify the results of a sensitivity analysis. Group sensitivities are also useful for presenting results in a concise fashion.
4. Assuming models to be true is always dangerous. An uncertainty/sensitivity analysis is always more convincing when uncertainty has been propagated through more than just one model. Using a parsimonious data-driven and a less parsimonious law-driven model for the same application can be especially effective and compelling.
5. When communicating scientific results transparency is an asset. As the assumptions of a parsimonious model are more easily assessed, sensitivity analysis should be followed by a model simplification.

The reader will find in this and the following chapters didactic examples for the purpose of familiarization with sensitivity measures. Most of the exercises will be based on models whose output (and possibly the associated sensitivity measures) can be computed analytically. In most practical instances the model under analysis or development will be a computational one, without a closed analytic formula.

Typically, models will involve differential equations or optimization algorithms involving numerical solutions. For this reason the best available practices for numerical computations will be presented in the following chapters. For the Elementary Effects Test, we shall offer numerical procedures developed by Campolongo *et al.* (1999b, 2000, 2007). For the variance-based measures we shall present the Monte Carlo based design developed by Saltelli (2002) as well as the Random Balance Designs based on Fourier Amplitude Sensitivity Test (FAST-RBD, Tarantola *et al.*, 2006, see Chapter 4). All these methods are based on true points in the space of the input factors, i.e. on actual computations of the model at these points. An important and powerful class of methods will be presented in Chapter 5; such techniques are based on metamodelling, e.g. on estimates of the model at untried points. Metamodelling allows for a great reduction in the cost of the analysis and becomes in fact the only option when the model is expensive to run, e.g. when a single simulation of the model takes tens of minutes or hours or more. The drawback is that metamodelling tools such as those developed by Ratto *et al.* (2007) are less straightforward to encode than plain Monte Carlo. Where possible, pointers will be given to available software.

1.6 EXERCISES

1. Prove that
$$V(Y) = E(Y^2) - E^2(Y).$$
2. Prove that for an additive model of two independent variables X_1 and X_2, fixing one variable can only decrease the variance of the model.
3. Why in μ^* are absolute differences used rather than simple differences?
4. If the variance of Y as results from an uncertainty analysis is too large, and the objective is to reduce it, sensitivity analysis can be used to suggest how many and which factors should be better determined. Is this a new setting? Would you be inclined to fix factors with a larger first-order term or rather those with a larger total effect term?
5. Suppose X_1 and X_2 are uniform variates on the interval [0, 1]. What is the mean? What is the variance? What is the mean of $X_1 + X_2$? What is the variance of $X_1 + X_2$?
6. Compute S_i analytically for model (1.3, 1.4) with the following values: $r = 2$, $\sigma = \{1, 2\}$ and $\mathbf{\Omega} = \{2, 1\}$.
7. Write a model (an analytic function and the factor distribution functions) in which fixing an uncertain factor increases the variance.
8. What would have been the result of using zero-centred distributions for the Ω's in Equation (1.27)?

1.7 ANSWERS

1. Given a function $Y = f(\mathbf{X})$ where $\mathbf{X} = (X_1, X_2, \cdots X_k)$ and $\mathbf{X} \sim p(\mathbf{X})$ where $p(\mathbf{X})$ is the joint distribution of $\mathbf{X}$ with $\int p(\mathbf{X})\, d\mathbf{X} = 1$, then the function mean can be defined as
$$E(Y) = \int f(\mathbf{X})\, p(\mathbf{X})\, \mathrm{d}\mathbf{X},$$
and its variance as
$$\begin{aligned}\mathrm{Var}(Y) &= \int (f(\mathbf{X}) - E(Y))^2 p(\mathbf{X}) \mathrm{d}\mathbf{X} \\ &= \int f^2(\mathbf{X}) p(\mathbf{X}) \mathrm{d}\mathbf{X} + E^2(Y) - 2 \int E(Y) f(\mathbf{X}) p(\mathbf{X}) \mathrm{d}\mathbf{X} \\ &= E(Y^2) + E^2(Y) - 2E^2(Y) \\ &= E(Y^2) - E^2(Y).\end{aligned}$$
Using this formula it can easily be proven that $\mathrm{Var}(Y) = \mathrm{Var}(Y + c)$, with c an arbitrary constant. This result is used in Monte Carlo-based

variance (and conditional variance) computation by rescaling all values of Y subtracting $E(Y)$. This is done because the numerical error in the variance estimate increases with the value of Y.

2. We can write the additive model of two variables X_1 and X_2 as $Y = f_1(X_1) + f_2(X_2)$, where f_1 is only a function of X_1 and f_2 is only a function of X_2.

   Recalling that the variance of (Y) can be written as $V(Y) = E(Y^2) - E^2(Y)$, where E stands for the expectation value, and applying it to Y we obtain

$$V(Y) = E(f_1^2 + f_2^2 + 2f_1f_2) - E^2(f_1 + f_2).$$

   Given that $E(f_1f_2) = E(f_1)E(f_2)$ for independent variables, then the above can be reduced to

$$V(Y) = E(f_1^2) + E(f_2^2) - E^2(f_1) - E^2(f_2),$$

   which can be rewritten as

$$V(Y) = V(f_1) + V(f_2),$$

   which proves that fixing either X_1 or X_2 can only reduce the variance of Y.

3. Modulus incremental ratios are used in order to avoid positive and negative values cancelling each other out when calculating the average.

4. It is a new setting. In Saltelli *et al.* (2004) we called it the variance cutting setting, when the objective of sensitivity analysis is the reduction of the output variance to a lower level by fixing the smallest number of input factors. This setting can be considered as relevant in, for example, risk assessment studies. Fixing the factors with the highest total effect term increases our chances of fixing, besides the first-order terms, some interaction terms possibly enclosed in the totals, thus maximizing our chances of reducing the variance (see Saltelli *et al.*, 2004).

5. Both X_1 and X_2 are uniformly distributed in $[0, 1]$, i.e.

$$p(X_1) = p(X_2) = U(0, 1).$$

   This means that $p(X_i)$ is 1 for $X_i \in [0, 1]$ and zero otherwise. Thus

$$E(X_1) = E(X_2) = \int_{x=0}^{1} p(x)x\mathrm{d}x = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}.$$

   Further:

$$Var(X_1) = Var(X_2) = \int_{x=0}^{1} p(x)\left(x - \frac{1}{2}\right)^2 \mathrm{d}x = \int_{x=0}^{1}\left(x^2 - x + \frac{1}{4}\right)\mathrm{d}x$$

$$= [\frac{x^3}{3} - \frac{x^2}{2} + \frac{1}{4}x]_0^1 = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}$$
$$E(X_1 + X_2) = E(X_1) + E(X_2) = 1,$$

as the variables are separable in the integral.

Given that $X_1 + X_2$ is an additive model (see Exercise 1) it is also true that

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) = \frac{1}{6}.$$

The same result is obtained integrating explicitly

$$Var(X_1 + X_2) = \int_{x_1=0}^{1} \int_{x_2=0}^{1} p(\mathbf{x}) (x_1 + x_2 - 1)^2 \, \mathrm{d}x_1 \mathrm{d}x_2.$$

6. Note that the model (1.3, 1.4) is linear and additive. Further, its probability density function can be written as the product of the factors' marginal distributions (independent factors). Writing the model for $r = 2$ we have

$$Y(Z_1, Z_2) = \Omega_1 Z_1 + \Omega_2 Z_2$$

with

$$Z_1 \sim N\ (0, 1) \quad \text{or equivalently } p(Z_1) = \frac{1}{\sigma_{Z_1}\sqrt{2\pi}} \mathrm{e}^{-} \frac{(Z_1)^2}{2\sigma_{Z_1}^2}$$

and a similar equation for $p(Z_2)$. Note that by definition the distributions are normalized, i.e. the integral of each $p(Z_i)$ over its own variable Z_1 is 1, so that the mean of Y can be reduced to

$$E(Y) = \Omega_1 \int_{-\infty}^{+\infty} Z_1 p(Z_1) \mathrm{d}_{Z_1} + \Omega_2 \int_{-\infty}^{+\infty} Z_2 p(Z_2) \mathrm{d}_{Z_2}.$$

These integrals are of the type $\int x \mathrm{e}^{-x^2} \mathrm{d}x$, whose primitive $-\mathrm{e}^{-x^2}/2$ vanishes at the extremes of integration, so that $E(Y) = 0$. Given that the model is additive, the variance will be

$$V(Y) = V_{Z_1} + V_{Z_2} = V\left(\Omega_1 Z_1\right) + V\left(\Omega_2 Z_2\right).$$

For either Z_1 or Z_2 it will be

$$V_{Z_i} = V\left(\Omega_i Z_i\right) = \Omega_i^2 V\ (Z_i).$$

We write

$$V(Z_i) = E(Z_i^2) - E^2(Z_i) = E\ (Z_i^2)$$

and

$$E(Z_i^2) = \frac{1}{\sqrt{2\pi}\sigma_{Z_i}} \int_{\infty}^{\infty} z_i^2 \mathrm{e}^{-z_i^2/2\sigma_{Z_i}^2} \mathrm{d}z_i.$$

The tabled form is

$$\int_{-0}^{+\infty} t^2 \mathrm{e}^{-at^2} \mathrm{d}_t = \frac{\sqrt{\pi}}{4},$$

which gives with an easy transformation

$$E(Z_i^2) = \sigma_{Z_i}^2$$

so that

$$V_{Z_i} = \Omega_i^2 \sigma_{Z_i}^2$$

and

$$V(Y) = \Omega_1^2 \sigma_{Z_1}^2 + \Omega_2^2 \sigma_{Z_2}^2$$

and

$$S_{Z_i} = \frac{\Omega_i^2 \sigma_{Z_i}^2}{V(Y)}.$$

Inserting the values $\sigma = \{1, 2\}$ and $\mathbf{\Omega} = \{2, 1\}$ we obtain $V(Y) = 8$ and $S_{Z_1} = S_{Z_2} = \frac{1}{2}$.

The result above can be obtained by explicitly applying the formula for S_i to our model:

$$S_{Z_i} = \frac{V(E(Y \mid Z_i))}{V(Y)},$$

which entails computing first $E(Y \mid Z_i = z_i^*)$. Applying this to our model $Y = \Omega_1 Z_1 + \Omega_2 Z_2$ we obtain, for example, for factor Z_1:

$$E(Y \mid Z_1 = z_1^*) = \int_{-\infty}^{+\infty} p(z_1) p(z_2) \left(\Omega_1 z_1^* + \Omega_2 z_2\right) \mathrm{d}z_1 \mathrm{d}z_2 = \Omega_1 z_1^*.$$

Hence V_{Z_1} – the variance over z_1^* of $\Omega_1 z_1^*$ – is, as before, equal to $\Omega_1^2 \sigma_{Z_1}^2$ and

$$S_{Z_i} = \frac{\Omega_i^2 \sigma_{Z_i}^2}{\Omega_1^2 \sigma_{Z_1}^2 + \Omega_2^2 \sigma_{Z_2}^2}.$$

7. We consider the model $Y = X_1 \cdot X_2$, with the factors identically distributed as

$$X_1, X_2 \sim N(0, \sigma).$$

Based on the previous exercise it is easy to see that

$$E(Y) = E(X_1 X_2) = E(X_1)E(X_2) = 0,$$

so that

$$V(Y) = E\left(X_1^2 X_2^2\right) = E\left(X_1^2\right) E\left(X_2^2\right) = \sigma^4.$$

If X_2 is fixed to a generic value x_2^*, then

$$E(X_1 x_2^*) = x_2^* E(X_1) = 0$$

as in a previous exercise, and

$$V(Y \mid X_2 = x_2^*) = V(X_1 x_2^*) = E\left(X_1^2 (x_2^*)^2\right) = (x_2^*)^2 E\left(X_1^2\right) = (x_2^*)^2 \sigma^2.$$

It is easy to see that $V(X_1 x_2^*)$ becomes bigger than $V(Y)$ whenever the modulus of x_2^* is bigger than σ.

Further, from the relation

$$V(Y \mid X_2 = x_2^*) = (x_2^*)^2 \sigma^2$$

one gets

$$E(V(Y \mid X_2)) = \sigma^4 = V(Y).$$

Given that

$$E(V(Y \mid X_2)) + V(E(Y \mid X_2)) = V(Y)$$

it must be that

$$V(E(Y \mid X_2)) = 0,$$

i.e. the first-order sensitivity index is null for both X_1 and X_2. These results are illustrated in the two figures which follow.

Figure 1.13 shows a plot of $V_{X_{\sim 2}}(Y \mid X_2 = x_2^*)$, i.e. $V_{X_1}(Y \mid X_2 = x_2^*)$ at different values of x_2^* for $\sigma = 1$. The horizontal line is the unconditional variance of Y. The ordinate is zero for $x_2^* = 0$, and becomes higher than $V(Y)$ for $x_2^* \sim 1$.

Figure 1.14 shows a scatterplot of Y versus x_1^* (the same shape would appear for x_2^*). It is clear from the plot that whatever the value of

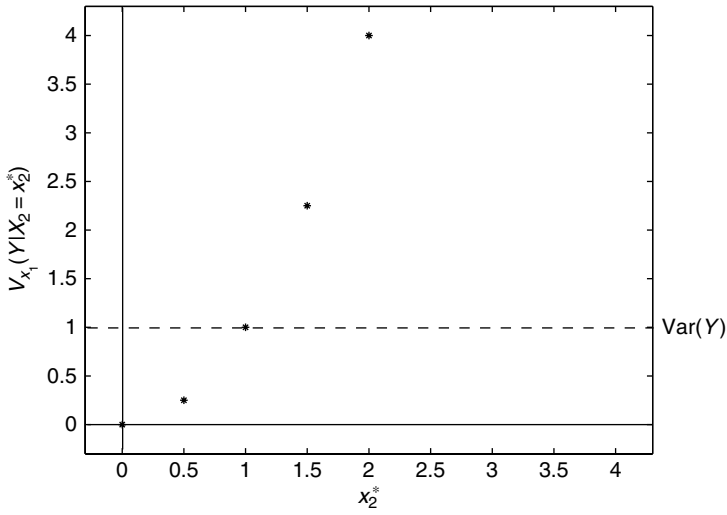

Figure 1.13 $V_{X_1}(Y \mid X_2 = x_2^*)$ at different values of x_2^*

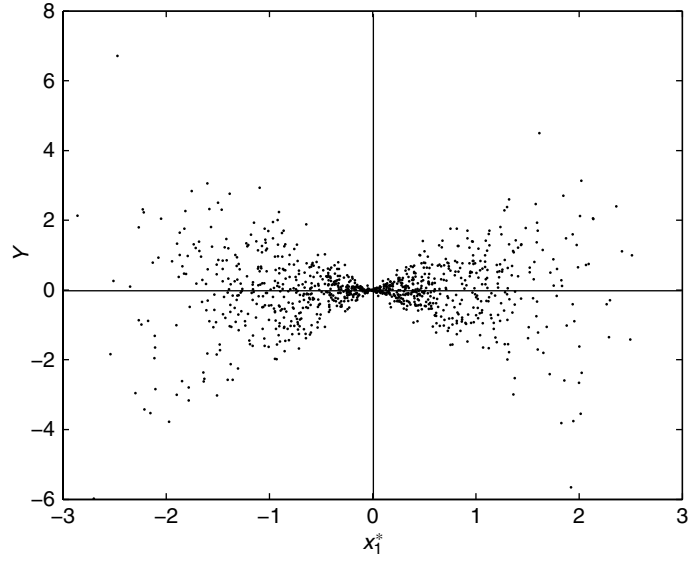

Figure 1.14 Scatterplot of Y versus x_1^*

the abscissa, the average of the points on the ordinate is zero, i.e. $E_{\mathbf{X}_{\sim 1}}(Y \mid X_1) = E_{X_2}(Y \mid X_1) = 0$. It is also clear from Figure 1.14 that even $V_{X_1}\left(E_{X_2}(Y \mid X_1)\right)$ will be zero, such that both S_1 and S_2 are zero for this model and all variance is captured by the second-order term, i.e. $S_{12} = 1$.

8. Referring to the previous exercise it is clear that if both Z_i and Ω_i are centred in zero, all first-order terms will be zero and the model will be purely interactive. In this case the only nonzero terms are the four interactions (second order) relative to the couples $(Z_1, \Omega_1), \cdots (Z_4, \Omega_4)$.

1.8 ADDITIONAL EXERCISES

1. Given the function

$$f(x) = \sin(X_1 \sin(X_2 \sin(X_3)))$$

with X_1, X_2, X_3 distributed normally with mean zero, can you guess what the first-order indices will be?
2. Consider the model $\Omega_1 Z_1 + \Omega_2 Z_2$ with Ω_1, Ω_2 as fixed constants and

$$Z_i \sim N(\mu_{Z_i}, \sigma_{Z_i}), \quad i = 1, 2$$

with

$$\mu_{Z_i} \neq 0, \quad i = 1, 2$$

and compute the variance-based sensitivity indices S_1, S_2.
3. Consider the model $Y = X_1 \cdot X_2$, where the two factors are normally distributed as

$$X_i \sim N(\mu_i, \sigma_i), \quad i = 1, 2$$

with

$$\mu_i \neq 0, \quad i = 1, 2$$

and compute the variance-based sensitivity indices S_1, S_2, and S_{12}.
4. Given a set of standardized variables $X_1, X_2, \ldots, X_k$ (all variables have thus zero mean and unit standard deviation), and a linear polynomial of the form $f(X_1, X_2, \ldots, X_k) = a_0 + \sum_{i=1}^{k} a_i X_i$, where $a_0, a_1, \cdots, a_k$ are constants, write the formula for the first-order indices S_i.
5. Repeat the previous exercise, for the case where both the $a_0, a_1, \cdots, a_k$ and the $X_1, X_2, \ldots, X_k$ are normally distributed:

$$X_i \sim N\left(\mu_{X_i}, \sigma_{X_i}\right)$$

and

$$a_i \sim N\left(\mu_{a_i}, \sigma_{a_i}\right).$$

1.9 SOLUTIONS TO ADDITIONAL EXERCISES

1. The first-order indices will be zero.
2. The solution is

$$S_{Z_i} = \frac{\Omega_i^2 \sigma_{Z_i}^2}{V(Y)}$$

as for the case with

$$\mu_{Z_i} = 0, \quad i = 1, 2$$

3. The solution is

$$S_1 = \frac{\mu_2^2 \sigma_1^2}{\left(\mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2 + \sigma_1^2 \sigma_2^2\right)}$$

and analogous formula for S_2, while

$$S_{12} = \frac{\sigma_1^2 \sigma_2^2}{\left(\mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2 + \sigma_1^2 \sigma_2^2\right)}.$$

4. It is simply

$$S_i = \frac{a_i^2}{\sum_{i=1}^{k} a_i^2},$$

i.e. each sensitivity index is proportional to the square of its coefficient.

5. The problem is additive in a_0 and in the k sets $\{a_i, X_i\}$. Using this and the results from Exercise 3 it is easy to derive the solution.

$$S_{a_i} = \frac{\mu_{X_i}^2 \sigma_{a_i}^2}{V}$$

$$S_{X_i} = \frac{\mu_{a_i}^2 \sigma_{X_i}^2}{V}$$

$$S_{a_i X_i} = \frac{\sigma_{a_i}^2 \sigma_{X_i}^2}{V}$$

$$V = \sum_{i=1}^{k} \left(\mu_{a_i}^2 \sigma_{X_i}^2 + \mu_{X_i}^2 \sigma_{a_i}^2 + \sigma_{a_i}^2 \sigma_{X_i}^2\right).$$

By putting $a_0 = 0$ the above solution can be used to compute the sensitivity indices for model (1.3, 1.27).