Chapter 2
Setup of Sensitivity Analysis

This section is devoted to the most important step in sensitivity analysis, the formulation of the sensitivity question. Scientists have developed myriads of models in different disciplines, and there are myriads of sensitivity analysis methods waiting to be used to explore the content of those models.

In operations research, models include linear programming models, influence diagrams, fault-tree-event-tree models, optimization models, and simulation models, to name only a few.

In general, we abstract from the particular form of the mathematical model. The model itself is a black box that processes a set of numbers and maps them onto another set of numbers (Fig. 2.1). The next equation (called Eq. 2.1) offers a visualization of the equations of a quantitative model used in financial mathematics to

$$
\begin{aligned}
E(t) &= A(t)[N(x^+) - [\frac{H(t)}{A(t)}]^{\frac{2r}{\sigma^2}+1} N(y^+) - e^{-r(T-t)} K \left[N(x^-) - [\frac{H}{A(t)}]^{\frac{2r}{\sigma^2}-1} N(y^-) \right] \\
D(t) &= A(t)[N(-x^+) - [\frac{H(t)}{A(t)}]^{\frac{2r}{\sigma^2}+1} N(y^+) + e^{-r(T-t)} K \left[N(x^-) - [\frac{H}{A(t)}]^{\frac{2r}{\sigma^2}-1} N(y^-) \right] \\
x^{\pm} &= \frac{\ln A - \ln K + (r \pm \frac{e^2}{2})\Delta T}{\sigma\sqrt{\Delta T}}; x^{\pm} = \frac{2 \ln H - \ln A - \ln K + (r \pm \frac{e^2}{2})\Delta T}{\sigma\sqrt{\Delta T}}
\end{aligned}
\tag{2.1}
$$

Even if we know the analytical expression of the equations in this case, we are unable to infer the behavior of the quantities of interest (i.e., model output; the default probability) as a function of the model inputs (i.e., the variables named r or σ) based on the sole intuition.

Return to the model in Eq. 2.1. The numbers that feed into the computer code have been given different names, which often depend on the field of application. For instance, in statistics-related literature, they are called factors or covariates, while in

E. Borgonovo, *Sensitivity Analysis*, International Series in Operations Research & Management Science 251, DOI 10.1007/978-3-319-52259-3_2

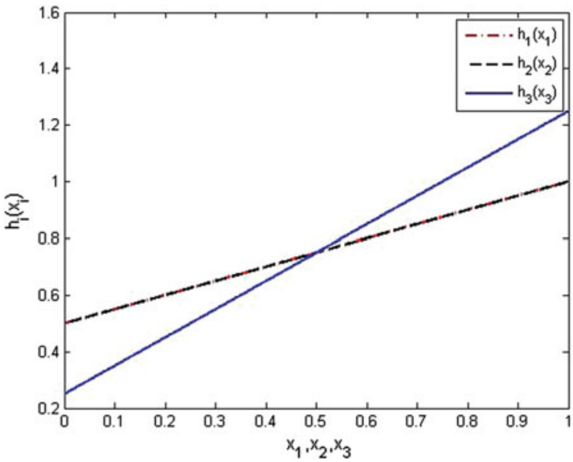

Fig. 2.1 Scientific model as a black box processing some inputs and transforming them into corresponding outputs of interest

engineering papers they are often called parameters. In economics, they are referred to as exogenous variables. Similarly, calculations of the code can be referred to as model outputs, dependent variables, or endogenous variables.

In this monograph, we use the term *model inputs*, and we denote them the symbol x. We denote model outputs with the letter y, and we denote the input–output mapping (i.e., the model that maps model inputs onto model outputs) with $g(x)$. We adopt a more formal approach to the necessary functional spaces in the following chapters, but at this point we say something more about some of the input–output mappings features. First, we need to distinguish between the models with deterministic output and the models with stochastic output. We say that a model produces a deterministic output if it is such that any time the model inputs are fixed at x^0, the corresponding value of the output is $y^0 = g(x^0)$. In other words, by fixing x at x^0, we always obtain the same value for the output. We say that a model produces a stochastic output if it generates a random value of y every time x is fixed at x^0. Thus, given x^0, the output of a stochastic model is the conditional distribution of Y given that X is fixed at x^0. Of course, a deterministic model can be seen as a special case of a stochastic model, where the output is a Dirac-δ function centered at x^0. Examples of models that can be treated as deterministic models include linear programming models, event trees, and decision trees with known probabilities. Queuing models are examples of models with a stochastic output. We also have models that produce time-dependent or time-independent outputs. Similarly, the inputs can be time-dependent (time series). In this monograph, we focus on deterministic models with time-independent inputs and outputs.

Moreover, we focus on methods that fall into the category of *independent and model-free* (Saltelli 2002b). Consider, for example, the family of linear programming models. The seminal works of Dantzig, Koopmans, and others have made linear programming a central decision-support modeling technique in applications ranging from economics to industrial engineering. A relevant question is the stability of the solution given uncertainty in some of the model inputs. Wendell's tolerance sensitivity analysis responds to this question by determining the region in which variations in model inputs do not cause the optimal solution to change [see Wendell (1984, 1985)]. As such, Wendell's approach answers this highly relevant question in linear programming and obtains analytical results for a highly complicated problem. The analytical results are model-specific. Conversely, other sensitivity methods, such as partial derivatives (which may have been the first sensitivity analysis method ever developed), apply (at least conceptually) to any differentiable input–output mapping.

An additional relevant distinction is the difference between *sensitivity* and *sensitivity analysis*. The term sensitivity is often used as a synonym for *dependence*. Many investigators are interested in the sensitivity of a given property or quantity to changes in another property or quantity. In several works, sensitivity is determined through experiments. These exercises are characterized by high scientific validity, especially when the dependence in question refers to long-standing and unsolved problems. However, studying the sensitivity of a property or quantity to another quantity is different from performing a systematic sensitivity analysis of a computer code.

Thus, a distinction must be made about whether the model developer is interested in sensitivity or in sensitivity analysis. The US Environmental Protection Agency (EPA) defines sensitivity as *The degree to which the model outputs are affected by changes in selected input parameters* US EPA (2009, p. 46), while sensitivity analysis is defined as *The computation of the effect of changes in input values or assumptions (including boundaries and model functional form) on the outputs* US EPA (2009, p. 46).

Thus, much of what we can gather from a sensitivity analysis depends on the first conceptual step: the formulation of the sensitivity question. A poor formulation may lead an analyst to use an inappropriate method and, thereby, obtain only a partially informative (if not wrong) answer to the question at hand. If the question is clearly stated, we can readily and solidly identify the best method that can be used to answer it. Along these lines, researchers have developed the concept of sensitivity analysis settings, which originated with the works of Saltelli and Tarantola (Saltelli 2002b; Saltelli and Tarantola 2002). In their literature review, Borgonovo and Plischke (2016) identify five sensitivity analysis settings that have been uncovered in the extant research.

The first setting is *model input prioritization* [factor prioritization in Saltelli et al. (2004)]. In this setting, the goal is to identify the key drivers of model behavior. The search for a key driver of model behavior assumes different meanings depending on whether the analysis is performed locally or globally, and on whether we are in a pre-decision or post-decision setting. In a local sensitivity analysis, the modeler varies the model inputs around a predetermined value of interest. Perturbations in the model inputs can be small (infinitesimal) or finite. In a global sensitivity analysis, the modeler assigns ranges to the model inputs and may specify corresponding probability distributions. The modeler is then interested in assessing the behavior of the model, as the inputs span the entire model input space. Thus, a model input can be a key driver of model behavior for small perturbations but not a key driver when variations in the entire model input space are considered. Key global drivers are also the model inputs on which to focus resources in further data collection.

Moreover, information about key drivers has different meanings depending on whether we are in a pre-decision or post-decision phase. In a pre-decision phase, it is important to know which model input can cause the preferred strategy to change. In a post-decision phase, we are interested in the factors on which to *focus managerial attention during implementation* (Eschenbach 1992).

A second setting is *model input fixing* [factor fixing in Saltelli et al. (2004)]. In this setting, which is typical in screening exercises, we are interested in determining the model inputs that can be fixed to their nominal value. Such inputs can be safely excluded from further information collection or modeling efforts, at least in the first phase of a scientific investigation.

A third setting is *model structure* (Borgonovo 2010a). In this setting, an analyst is interested in analyzing the structure of the model and in understanding whether interactions are present among the model inputs. This question has to do with the structure of the model, and it can be appreciated both locally and globally.

A fourth setting is sign of change (Borgonovo 2010a). This setting can be linked to the seminal work of Samuelson (1941) on sensitivity analysis in economics, which introduces the well-known methodology of comparative statics. In comparative statics, the analyst is interested in the sign (or direction) of the change in the model output. As Samuelson states: *If no more than this could be said, the economist would be truly vulnerable to the gibe that he is only a parrot taught to say "supply and demand." Simply to know that there are efficacious "laws" determining equilibrium tells us nothing of the character of these laws. In order for the analysis to be useful it must provide information concerning the way in which our equilibrium quantities will change as a result of changes in the parameters taken as independent data* (Samuelson 1941, p. 97)". The first part of Samuelson's statement warns that simply developing a model and then passively accepting its results is unsatisfactory. A model needs to be shaken to let its content emerge. The second part suggests a need to understand whether an increase in model inputs gives rise to an increase in the model's output or vice versa. Samuelson obtains this information analytically through the use of partial derivatives. In this monograph, we show that the information can also be obtained globally through a different technical instrument (see Sect. 22.6).

The last setting, which applies to several operations research problems, is the *stability setting* (Borgonovo and Plischke 2016). In the stability setting, the analyst is interested in determining whether perturbations in the model inputs may cause the preferred alternative to change. This setting applies in a local sense if one is interested in small model input perturbations or in a global sense if one is interested in determining the region over which variations in the model inputs do not cause the preferred alternative to change. Stability sensitivity analysis is important in all optimization problems. As discussed above, in linear programming, Wendell's tolerance sensitivity approach is best suited for providing a consistent answer to this question given simultaneous variations in model inputs (Wendell 1985; Ravi and Wendell 1985). This question, however, also applies to decision-support models expressed in the form of influence diagrams or decision trees.

The five settings presented here accompany methods that may be transversal to several problems and models. Of course, they do not preclude the existence of other relevant sensitivity questions that have already been asked or might be asked in the future. This monograph cannot claim exhaustiveness in that respect. However, the relevant intuition should be clear. Before applying a sensitivity method, it is necessary to distinctly specify the goal of the analysis. Do we, for example, wish to understand whether model inputs are involved in interactions or whether a certain model input is a key uncertainty driver? These are two distinct questions and, depending on the problem, deriving answers may require different approaches.

In the next chapter, we start with a simple, widely used group of sensitivity methods—one-factor-at-a-time methods.

http://www.springer.com/978-3-319-52257-9

Sensitivity Analysis
An Introduction for the Management Scientist
Borgonovo, E.
2017, XIV, 294 p. 68 illus., Hardcover
ISBN: 978-3-319-52257-9