1 A WORKED EXAMPLE

This chapter presents an exhaustive analysis of a simple example, in order to give the reader a first overall view of the problems met in quantitative sensitivity analysis and the methods used to solve them. In the following chapters the same problems, questions, and techniques will be presented in full detail.

We start with a sensitivity analysis for a mathematical model in its simplest form, and work it out adding complications to it one at a time. By this process the reader will meet sensitivity analysis methods of increasing complexity, starting from the elementary approaches to the more quantitative ones.

1.1 A simple model

A simple portfolio model is:

$$Y = C_s P_s + C_t P_t + C_j P_j \tag{1.1}$$

where Y is the estimated risk[1] in €, C_s, C_t, C_j are the quantities per item, and P_s, P_t, P_j are *hedged* portfolios in €.[2] This means that each P_x, $x = \{s, t, j\}$ is composed of more than one item – so that the average return P_x is zero €. For instance, each hedged portfolio could be composed of an option plus a certain amount of underlying stock offsetting the option risk exposure due to

[1] This is the common use of the term. Y is in fact a return. A negative uncertain value of Y is what constitutes the risk.

[2] This simple model could well be seen as a composite (or synthetic) indicator camp by aggregating a set of standardised base indicators P_i with weights C_i (Tarantola *et al.*, 2002; Saisana and Tarantola, 2002).

Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models A. Saltelli, S. Tarantola, F. Campolongo and M. Ratto ISBN 0-470-87093-1

movements in the market stock price. Initially we assume $C_s,\ C_t,\ C_j$ = constants. We also assume that an estimation procedure has generated the following distributions for P_s, P_t, P_j:

$$\begin{aligned} P_s &\sim N(\bar{p}_s, \sigma_s), \qquad \bar{p}_s = 0, \qquad \sigma_s = 4 \\ P_t &\sim N(\bar{p}_t, \sigma_t), \qquad \bar{p}_t = 0, \qquad \sigma_t = 2 \\ P_j &\sim N(\bar{p}_j, \sigma_j), \qquad \bar{p}_j = 0, \qquad \sigma_j = 1. \end{aligned} \tag{1.2}$$

The P_xs are assumed independent for the moment. As a result of these assumptions, Y will also be normally distributed with parameters

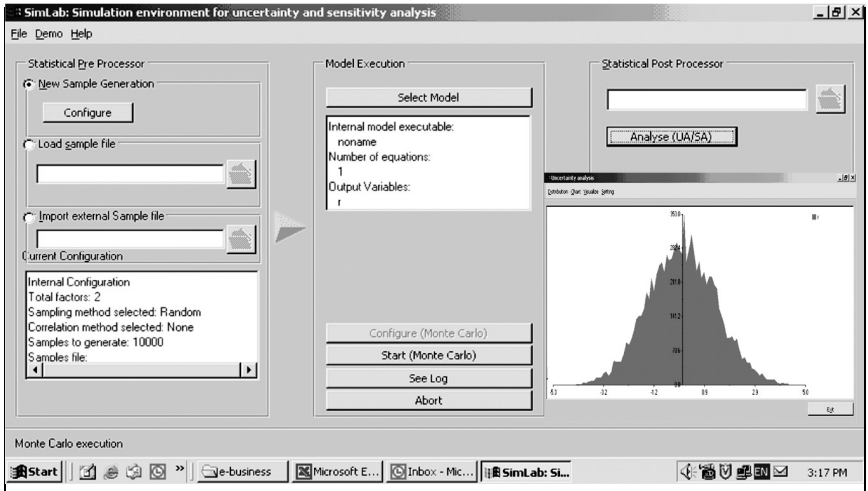
$$\bar{y} = C_s \bar{p}_s + C_t \bar{p}_t + C_j \bar{p}_j \tag{1.3}$$

$$\sigma_y = \sqrt{C_s^2 \sigma_s^2 + C_t^2 \sigma_t^2 + C_j^2 \sigma_j^2}. \tag{1.4}$$

Box 1.1 SIMLAB

The reader may want at this stage, or later in the study, to get started with SIMLAB by reproducing the results (1.3)–(1.4). This is in fact an uncertainty analysis, e.g. a characterisation of the output distribution of Y given the uncertainties in its input. The first thing to do is to input the factors P_s, P_t, P_j with the distributions given in (1.2). This is done using the left-most panel of SIMLAB (Figure 7.1), as follows:

1. Select 'New Sample Generation', then 'Configure', then 'Create New' when the new window 'STATISTICAL PRE PROCESSOR' is displayed.
2. Select 'Add' from the input factor selection panel and add factors one at a time as instructed by SIMLAB. Select 'Accept factors' when finished. This takes the reader back to the 'STATISTICAL PRE PROCESSOR' window.
3. Select a sampling method. Enter 'Random' to start with, and 'Specify switches' in the right. Enter something as a seed for random number generation and the number of executions (e.g. 1000). Create an output file by giving it a name and selecting a directory.



4. Go back to the left-most part of the SIMLAB main menu and click on 'Generate'. A sample is now available for the simulation.

5. We now move to the middle of the panel (Model execution) and select 'Configure (Monte Carlo)' and 'Select Model'. A new panel appears.

6. Select 'Internal Model' and 'Create new'. A formula parser appears. Enter the name of the output variable, e.g. 'Y' and follow the SIMLAB formula editor to enter Equation (1.1) with values of C_s, C_t, C_j of choice.

7. Select 'Start Monte Carlo' from the main model panel. The model is now executed the required number of times.

8. Move to the right-most panel of SIMLAB. Select 'Analyse UA/SA', select 'Y' as the output variable as prompted; choose the single time point option. This is to tell SIMLAB that in this case the output is not a time series.

9. Click on UA. The figure on this page is produced. Click on the square dot labelled 'Y' on the right of the figure and read the mean and standard deviation of Y. You can now compare these sample estimates with Equations (1.3–1.4).

Let us initially assume that $C_s < C_t < C_j$, i.e. we hold more of the less volatile items (but we shall change this in the following). A sensitivity analysis of this model should tell us something about the relative importance of the uncertain factors in Equation (1.1) in determining the output of interest Y, the risk from the portfolio.

According to first intuition, as well as to most of the existing literature on SA, the way to do this is by computing derivatives, i.e.

$$S_x^d = \frac{\partial Y}{\partial P_x}, \text{ with } x = s, t, j \tag{1.5}$$

where the superscript 'd' has been added to remind us that this measure is in principle dimensioned ($\partial Y/\partial P_x$ is in fact dimensionless, but $\partial Y/\partial C_x$ would be in €). Computing S_x^d for our model we obtain

$$S_x^d = C_x, \text{ with } x = s, t, j. \tag{1.6}$$

If we use the S_x^ds as our sensitivity measure, then the order of importance of our factors is $P_j > P_t > P_s$, based on the assumption $C_s < C_t < C_j$. S_x^d gives us the increase in the output of interest Y per unit increase in the factor P_x. There seems to be something wrong with this result: we have more items of portfolio j but this is the one with the least volatility (it has the smallest standard deviation, see Equation (1.2)). Even if $\sigma_s \gg \sigma_t, \sigma_j$, Equation (1.6) would still indicate P_j to be the most important factor, as Y would be locally more sensitive to it than to either P_t or P_s.

Sometime local sensitivity measures are normalised by some reference or central value. If

$$y^0 = C_s p_s^0 + C_t p_t^0 + C_j p_j^0. \tag{1.7}$$

then one can compute

$$S_x^l = \frac{p_x^0}{y^0} \frac{\partial Y}{\partial P_x}, \text{ with } x = s, t, j. \tag{1.8}$$

Applying this to our model, Equation (1.1), one obtains:

$$S_x^l = C_x \frac{p_x^0}{y^0}, \text{ with } x = s, t, j. \tag{1.9}$$

In this case the order of importance of the factors depends on the relative value of the C_xs weighted by the reference values p_x^0s. The superscript 'l' indicates that this index can be written as a logarithmic ratio if the derivative is computed in p_x^0.

$$S_x^l = \frac{p_x^0}{y^0} \frac{\partial Y}{\partial P_x}\bigg|_{y^0, p_x^0} = \frac{\partial \ln(Y)}{\partial \ln(P_x)}\bigg|_{y^0, p_x^0}. \tag{1.10}$$

S_x^l gives the fractional increase in Y corresponding to a unit fractional increase in P_x. Note that the reference point p_s^0, p_t^0, p_j^0 might be made to coincide with the vector of the mean values $\bar{p}_s$, $\bar{p}_t$, $\bar{p}_j$, though this would not in general guarantee that $\bar{y} = Y(\bar{p}_s, \bar{p}_t, \bar{p}_j)$, even though this is now the case (Equation (1.3)). Since $\bar{p}_s, \bar{p}_t, \bar{p}_j = 0$ and $\bar{y} = 0$, S_x^l collapses to be identical to S_x^d.

Also S_x^l is insensitive to the factors' standard deviations. It seems a better measure of importance than S_x^d, as it takes away the dimensions and is normalised, but it still offers little guidance as to how the uncertainty in Y depends upon the uncertainty in the P_xs.

A first step in the direction of characterising uncertainty is a normalisation of the derivatives by the factors' standard deviations:

$$\begin{aligned} S_s^\sigma &= \frac{\sigma_s}{\sigma_y} \frac{\partial Y}{\partial P_s} = C_s \frac{\sigma_s}{\sigma_y} \\ S_t^\sigma &= \frac{\sigma_t}{\sigma_y} \frac{\partial Y}{\partial P_t} = C_t \frac{\sigma_t}{\sigma_y} \\ S_j^\sigma &= \frac{\sigma_j}{\sigma_y} \frac{\partial Y}{\partial P_j} = C_j \frac{\sigma_j}{\sigma_y} \end{aligned} \tag{1.11}$$

where again the right-hand sides in (1.11) are obtained by applying Equation (1.1). Note that S_x^d and S_x^l are truly local in nature, as they

Table 1.1 S_x^σ measures for model (1.1) and different values of C_s, C_t, C_j (analytical values).

Factor	$C_s, C_t, C_j =$ 100, 500, 1000	$C_s, C_t, C_j =$ 300, 300, 300	$C_s, C_t, C_j =$ 500, 400, 100
P_s	0.272	0.873	0.928
P_t	0.680	0.436	0.371
P_j	0.680	0.218	0.046

need no assumption on the range of variation of a factor. They can be computed numerically by perturbing the factor around the base value. Sometimes they are computed directly from the solution of a differential equation, or by embedding sets of instructions into an existing computer program that computes Y. Conversely, S_x^σ needs assumptions to be made about the range of variation of the factor, so that although the derivative remains local in nature, S_x^σ is a hybrid local–global measure.

Also when using S_x^σ, the relative importance of $P_s,\ P_t,\ P_j$ depends on the weights $C_s,\ C_t,\ C_j$ (Table 1.1). An interesting result concerning the S_x^σs when applied to our portfolio model comes from the property of the model that $\sigma_y = \sqrt{C_s^2\sigma_s^2 + C_t^2\sigma_t^2 + C_j^2\sigma_j^2}$; squaring both sides and dividing by σ_y^2 we obtain

$$1 = \frac{C_s^2\sigma_s^2}{\sigma_y^2} + \frac{C_t^2\sigma_t^2}{\sigma_y^2} + \frac{C_j^2\sigma_j^2}{\sigma_y^2}. \tag{1.12}$$

Comparing (1.12) with (1.11) we see that for model (1.1) the squared S_x^σ give how much each individual factor contributes to the variance of the output of interest. If one is trying to assess how much the uncertainty in each of the input factors will affect the uncertainty in the model output Y, and if one accepts the variance of Y to be a good measure of this uncertainty, then the squared S_x^σ seem to be a good measure. However beware: the relation $1 = \sum_{x=s,t,j} (S_x^\sigma)^2$ is not general; it only holds for our nice, well hedged financial portfolio model. This means that you can still use S_x^σ if the input have a dependency structure (e.g. they are correlated) or the model is non-linear, but it is no longer true that the

squared S_x^σ gives the exact fraction of variance attributable to each factor.

Using S_x^σ we see from Table 1.1 that for the case of equal weights (= 300), the factor that most influences the risk is the one with the highest volatility, P_s. This reconciles the sensitivity measure with our expectation.

Furthermore we can now put sensitivity analysis to use. For example, we can use the S_x^σ-based SA to build the portfolio (1.1) so that the risk Y is equally apportioned among the three items that compose it.

Let us now imagine that, in spite of the simplicity of the portfolio model, we chose to make a Monte Carlo experiment on it, generating a sample matrix

$$\mathbf{M} = \begin{matrix} p_s^{(1)} & p_t^{(1)} & p_j^{(1)} \\ p_s^{(2)} & p_t^{(2)} & p_j^{(2)} \\ \dots & \dots & \dots \\ p_s^{(N)} & p_t^{(N)} & p_j^{(N)} \end{matrix} = [\mathbf{p}_s, \mathbf{p}_t, \mathbf{p}_j]. \tag{1.13}$$

M is composed of N rows, each row being a trial set for the evaluation of Y. The factors being independent, each column can be generated independently from the marginal distributions specified in (1.2) above. Computing Y for each row in **M** results in the output vector **y**:

$$\mathbf{y} = \begin{matrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(N)} \end{matrix} \tag{1.14}$$

An example of scatter plot (Y vs P_s) obtained with a Monte Carlo experiment of 1000 points is shown in Figure 1.1. Feeding both **M** and **y** into a statistical software (SIMLAB included), the analyst might then try a regression analysis for Y. This will return a model of the form

$$y^{(i)} = b_0 + b_s p_s^{(i)} + b_t p_t^{(i)} + b_j p_j^{(i)} \tag{1.15}$$

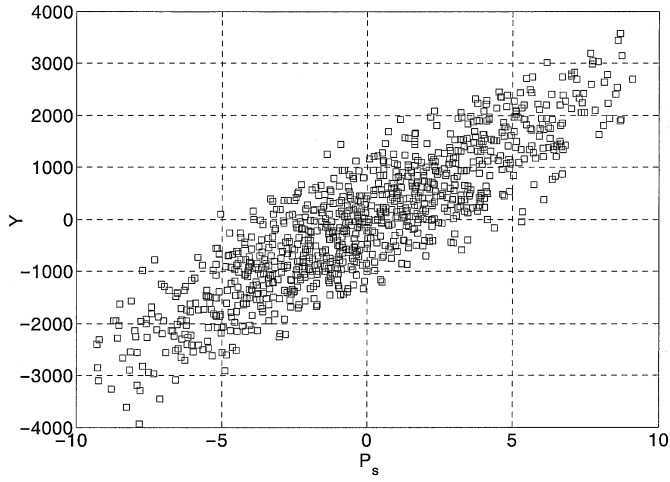


Figure 1.1 Scatter plot of Y vs. P_s for the model (1.1) $C_s = C_t = C_j = 300$. The scatter plot is made of $N = 1000$ points.

where the estimates of the b_xs are computed by the software based on ordinary least squares. Comparing (1.15) with (1.1) it is easy to see that if N is at least greater than 3, the number of factors, then $b_0 = 0$, $b_x = C_x$, $x = s, t, j$.

Normally one does not use the b_x coefficients for sensitivity analysis, as these are dimensioned. The practice is to computes the standardised regression coefficients (SRCs), defined as

$$\beta_x = b_x \sigma_x / \sigma_y. \tag{1.16}$$

These provide a regression model in terms of standardised variables

$$\tilde{\mathbf{y}} = \frac{\mathbf{y} - \bar{y}}{\sigma_y}; \quad \tilde{\mathbf{p}}_x = \frac{\mathbf{p}_x - \bar{p}_x}{\sigma_x} \tag{1.17}$$

i.e.

$$\tilde{\mathbf{y}} = \frac{\hat{\mathbf{y}} - \bar{y}}{\sigma_y} = \sum_{x=s,t,j} \beta_x \frac{\mathbf{p}_x - \bar{p}_x}{\sigma_x} = \sum_{x=s,t,j} \beta_x \tilde{\mathbf{p}}_x \tag{1.18}$$

where $\hat{\mathbf{y}}$ is the vector of regression model predictions. Equation (1.16) tells us that the β_xs (standardised regression coefficients)

for our portfolio model are equal to $C_x\sigma_x/\sigma_y$ and hence for linear models $\beta_x = S_x^\sigma$ because of (1.11). As a result, the values of the β_xs can also be read in Table 1.1.

Box 1.2 SIMLAB

You can now try out the relationship $\beta_x = S_x^\sigma$. If you have already performed all the steps in Box 1.1, you have to retrieve the saved input and output samples, so that you again reach step 9. Then:

10. On the right most part of the main SIMLAB panel, you activate the SA selection, and select SRC as the sensitivity analysis method.
11. You can now compare the SRC (i.e. the β_x) with the values in Table 1.1.

We can now try to generalise the results above as follows: for linear models composed of independent factors, the squared SRCs and S_x^σs provide the fraction of the variance of the model due to each factor.

For the standardised regression coefficients, these results can be further extended to the case of non-linear models as follows. The quality of regression can be judged by the model coefficient of determination R_y^2. This can be written as

$$R_y^2 = \frac{\sum_{i=1}^{N} (\hat{y}^{(i)} - \bar{y})^2}{\sum_{i=1}^{N} (y^{(i)} - \bar{y})^2} \tag{1.19}$$

where $\hat{y}^{(i)}$ is the regression model prediction. $R_y^2 \in [0, 1]$ represents the fraction of the model output variance accounted for by the regression model. The β_xs tell us how this fraction of the output

variance can be decomposed according to the input factors, leaving us ignorant about the rest, where this rest is related to the non-linear part of the model. In the case of the linear model (1.1) we have, obviously, $R_y^2 = 1$.

The β_xs are a progress with respect to the S_x^σ; they can always be computed, also for non-linear models, or for models with no analytic representation (e.g. a computer program that computes Y). Furthermore the β_xs, unlike the S_x^σ, offer a measure of sensitivity that is multi-dimensionally averaged. While S_x^σ corresponds to a variation of factor x, all other factors being held constant, the β_x offers a measure of the effect of factor x that is averaged over a set of possible values of the other factors, e.g. our sample matrix (1.13). This does not make any difference for a linear model, but it does make quite a difference for non-linear models.

Given that it is fairly simple to compute standardised regression coefficients, and that decomposing the variance of the output of interest seems a sensible way of doing the analysis, why don't we always use the β_xs for our assessment of importance?

The answer is that we cannot, as often R_y^2 is too small, as e.g. in the case of non-monotonic models.[3]

1.2 Modulus version of the simple model

Imagine that the output of interest is no longer Y but its absolute value. This would mean, in the context of the example, that we want to study the deviation of our portfolio from risk neutrality. This is an example of a non-monotonic model, where the functional relationship between one (or more) input factor and the output is non-monotonic. For this model the SRC-based sensitivity analysis fails (see Box 1.3).

[3] Loosely speaking, the relationship between Y and an input factor X is monotonic if the curve $Y = f(X)$ is non-decreasing or non-increasing over all the interval of definition of X. A model with k factors is monotonic if the same rule applies for all factors. This is customarily verified, for numerical models, by Monte Carlo simulation followed by scatter-plots of Y versus each factor, one at a time.

Box 1.3 SIMLAB

Let us now estimate the coefficient of determination R_y^2 for the modulus version of the model.

1. Select 'Random sampling' with 1000 executions.
2. Select 'Internal Model' and click on the button 'Open existing configuration'. Select the internal model that you have previously created and click on 'Modify'.
3. The 'Internal Model' editor will appear. Select the formula and click on 'Modify'. Include the function 'fabs()' in the Expression editor. Accept the changes and go back to the main menu.
4. Select 'Start Monte Carlo' from the main model panel to generate the sample and execute the model.
5. Repeat the steps in Box 1.2 to see the results. The estimates of SRC appear with a red background as the test of significance is rejected. This means that the estimates are not reliable. The model coefficient of determination is almost null.

Is there a way to salvage our concept of decomposing the variance of Y into bits corresponding to the input factors, even for non-monotonic models? In general one has little a priori idea of how well behaved a model is, so that it would be handy to have a more robust variance decomposition strategy that works, whatever the degree of model non-monotonicity. These strategies are sometimes referred to as 'model free'.

One such strategy is in fact available, and fairly intuitive to get at. It starts with a simple question. If we could eliminate the uncertainty in one of the P_x, making it into a constant, how much would this reduce the variance of Y? Beware, for unpleasant models fixing a factor might actually increase the variance instead of reducing it! It depends upon where P_x is fixed.

The problem could be: how does $V_y = \sigma_y^2$ change if one can fix a generic factor P_x at its mid-point? This would be measured by $V(Y|P_x = \bar{p}_x)$. Note that the variance operator means in this case that while keeping, say, P_j fixed to the value $\bar{p}_j$ we integrate over P_s, P_t.

$$V(Y|\bar{P}_j = \bar{p}_j) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} N(\bar{p}_s, \sigma_s)\, N(\bar{p}_t, \sigma_s)\, [(C_s P_s + C_t P_t + C_j \bar{p}_j) - (C_s \bar{p}_s + C_t \bar{p}_t + C_j \bar{p}_j)]^2 \mathrm{d}P_s \mathrm{d}P_t. \tag{1.20}$$

In practice, beside the problem already mentioned that $V(Y|P_x = \bar{p}_x)$ can be bigger than V_y, there is the practical problem that in most instances one does not know where a factor is best fixed. This value could be the true value, which is unknown at the simulation stage.

It sounds sensible then to average the above measure $V(Y|P_x = \bar{p}_x)$ over all possible values of P_x, obtaining $E(V(Y|P_x))$. Note that for the case, e.g. $x = j$, we could have written $E_j(V_{s,t}(Y|P_j))$ to make it clear that the average operator is over P_j and the variance operator is over P_s, P_t. Normally, for a model with k input factors, one writes $E(V(Y|X_j))$ with the understanding that V is over $\mathbf{X}_{-j}$(a $(k-1)$ dimensional vector of all factors but X_j) and E is over X_j.

$E(V(Y|P_x))$ seems a good measure to use to decide how influential P_x is. The smaller the $E(V(Y|P_x))$, the more influential the factor P_x is. Textbook algebra tells us that

$$V_y = E(V(Y|P_x)) + V(E(Y|P_x)) \tag{1.21}$$

i.e. the two operations complement the total unconditional variance. Usually $V(E(Y|P_x))$ is called the *main effect* of P_x on Y, and $E(V(Y|P_x))$ the residual. Given that $V(E(Y|P_x))$ is large if P_x is influential, its ratio to V_y is used as a measure of sensitivity, i.e.

$$S_x = \frac{V(E(Y|P_x))}{V_y} \tag{1.22}$$

S_x is nicely scaled in [0, 1] and is variously called in the literature the *importance measure*, *sensitivity index*, *correlation ratio* or *first*

Table 1.2 S_x measures for model Y and different values of C_s, C_t, C_j (analytical values).

Factor	$C_s, C_t, C_j =$ 100, 500, 1000	$C_s, C_t, C_j =$ 300, 300, 300	$C_s, C_t, C_j =$ 500, 400, 100
P_s	0.074	0.762	0.860
P_t	0.463	0.190	0.138
P_j	0.463	0.048	0.002

order effect. It can be always computed, also for models that are not well-behaved, provided that the associate integrals exist. Indeed, if one has the patience to calculate the relative integrals in Equation (1.20) for our portfolio model, one will find that $S_x = (S_x^\sigma)^2 = \beta_x^2$, i.e. there is a one-to-one correspondence between the squared S_x^σ, the squared standardised regression coefficients and S_x for linear models with independent inputs. Hence all what we need to do to obtain the S_xs for the portfolio model (1.1) is to square the values in Table 1.1 (see Table 1.2). A nice property of the S_xs when applied to the portfolio model is that, for whatever combination of C_s, C_t, C_j, the sum of the three indices S_s, S_t, S_j is one, as one can easily verify (Table 1.2). This is not surprising, as the same was true for the β_x^2 when applied to our simple model. Yet the class of models for which this nice property of the S_xs holds is much wider (in practice that of the *additive models*[4]).

S_x is a good model-free sensitivity measure, and it always gives the expected reduction in the variance of the output that one would obtain if one could fix an individual factor.

As mentioned, for a system of k input uncertain factors, in general $\sum_{i=1}^{k} S_i \leq 1$.

Applying S_x to model $|Y|$, modulus of Y, one gets the estimations in Table 1.3. with SIMLAB.

We can see that the estimates of the expected reductions in the variance of $|Y|$ are much smaller than for Y. For example, in the case of $C_s, C_t, C_j = 300$, fixing P_s gives an expected variance reduction of 53% for $|Y|$, whilst the reduction of the variance for Y is 76%.

[4] A model $Y = f(X_1, X_2, \ldots, X_k)$ is additive if f can be decomposed as a sum of k functions f_i, each of which is a function only of the relative factor X_i.

Table 1.3 Estimation of S_xs for model $|Y|$ and different values of C_s, C_t, C_j.

Factor	$C_s, C_t, C_j =$ 100, 500, 1000	$C_s, C_t, C_j =$ 300, 300, 300	$C_s, C_t, C_j =$ 500, 400, 100
P_s	0	0.53	0.69
P_t	0.17	0.03	0.02
P_j	0.17	0	0

Given that the modulus version of the model is non-additive, the sum of the three indices S_s, S_t, S_j is less than one. For example, in the case $C_s, C_t, C_j = 300$, the sum is 0.56. What can we say about the remaining variance that is not captured by the S_xs? Let us answer this question not on the modulus version of model (1.1) but – for didactic purposes – on the slightly more complicated a six-factor version of our financial portfolio model.

Box 1.4 SIMLAB

Let us test the functioning of a variance-based technique with SIMLAB, by reproducing the results in Table 1.3

1. Select the 'FAST' sampling method and then 'Specify switches' on the right. Select 'Classic FAST' in the combo box 'Switch for FAST'. Enter something as seed and a number of executions (e.g. 1000). Create a sample file by giving it a name and selecting a directory.
2. Go back to the left-most part of the SIMLAB main menu and click on 'Generate'. A FAST-based sample is now available for the simulation.
3. Load the model with the absolute value as in Box 1.3 and click on 'Start (Monte Carlo)'.
4. Run the SA: a pie chart will appear reporting the estimated of S_x obtained with FAST. You can also see the tabulated values, which might be not as close to those reported in

Table 1.3 due to sampling error (the sample size is 1000). Try again with larger sample sizes and using the Sobol method, an alternative to the FAST method.

1.3 Six-factor version of the simple model

We now revert to model (1.1) and assume that the quantities C_xs are also uncertain. The model (1.1) now has six uncertain inputs. Let us assume

$$\begin{aligned} C_s &\sim N(250, 200) \\ C_t &\sim N(400, 300) \\ C_j &\sim N(500, 400)\,. \end{aligned} \tag{1.23}$$

The three distributions have been truncated at percentiles [11, 99.9], [10.0, 99.9] and [10.5, 99.9] respectively to ensure that $C_x > 0$.

There is no alternative now to a Monte Carlo simulation: the output distribution is in Figure 1.2, and the S_xs, as from Equation (1.22), are in Table 1.4. The S_xs have been estimated using a large

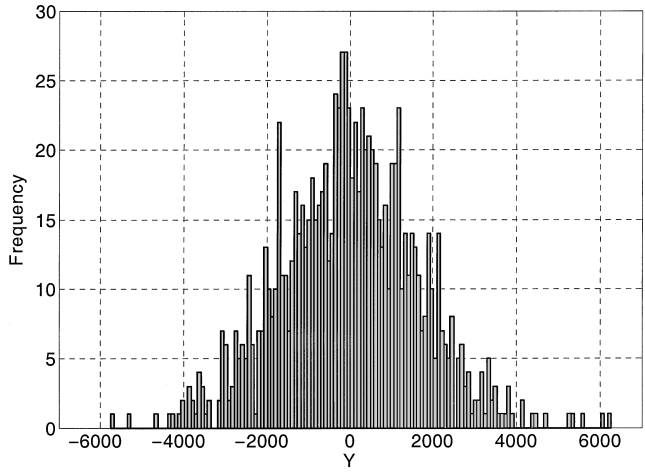


Figure 1.2 Output distribution for model (1.1) with six input factors, obtained from a Monte Carlo sample of 1000 elements.

Table 1.4 Estimates of first order effects S_x for model (1.1) with six input factors.

Factor	S_x
P_s	0.36
P_t	0.22
P_j	0.08
C_s	0.00
C_t	0.00
C_j	0.00
Sum	0.66

number of model evaluations (we will come back to this in future chapters; see also Box 1.5).

How is it that all effects for C_x are zero? All the P_x are centred on zero, and hence the conditional expectation value of Y is zero regardless of the value of C_x, i.e. for model (1.1) we have:

$$E(Y|C_x = c_x^*) = E(Y) = 0, \quad \text{for all } c_x^* \tag{1.24}$$

and as a result, $V(E(Y|C_x)) = 0$. This can also be visualised in Figure 1.3; inner conditional expectations of Y can be taken averaging along vertical 'slices' of the scatter plot. In the case of Y vs. C_s (lower panel) it is clear that such averages will form a perfectly horizontal line on the abscissas, implying a zero variance for the averaged Ys and a null sensitivity index. Conversely, for Y vs. P_s (upper panel) the averages along the vertical slices will form an increasing line, implying non-zero variance for the averaged Ys and a non-null sensitivity index.

As anticipated the S_xs do not add up to one. Let us now try a little experiment. Take two factors, say P_s, P_t, and estimate our sensitivity measure on the pair i.e. compute $V(E(Y|P_s, P_t))/V_y$. By definition this implies taking the average over all factors except P_s, P_t, and the variance over P_s, P_t. We do this (we will show how later) and call the results $S^c_{P_s P_t}$, where the reason for the superscript c will be clear in a moment. We see that $S^c_{P_s P_t} = 0.58$, i.e.

$$S^c_{P_s P_t} = \frac{V(E(Y|P_s, P_t))}{V_y} = S_{P_s} + S_{P_t} = 0.36 + 0.22. \tag{1.25}$$

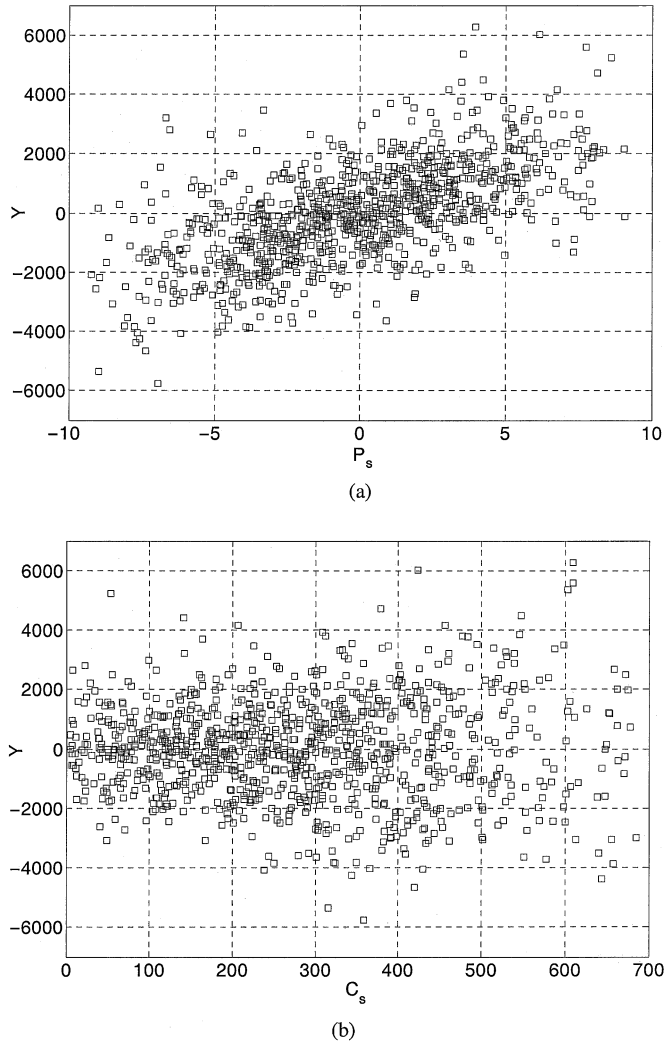


Figure 1.3 Scatter plots for model (1.1): (a) of Y vs. P_s, (b) of Y vs. C_s. The scatter plots are made up of $N = 1000$ points.

This seems a nice result. Let us try the same game with C_s, P_s. The results show that now:

$$S^c_{C_s P_s} = \frac{V(E(Y|C_s, P_s))}{V_y} = 0.54 > S_{C_s} + S_{P_s} = 0.36. \quad (1.26)$$

Table 1.5 Incomplete list of pair-wise effects S_{xz} for model (1.1) with six input factors.

Factors	S_{xz}
P_s, C_s	0.18
P_t, C_t	0.11
P_j, C_j	0.05
P_s, C_t	0.00
P_s, C_j	0.00
P_t, C_s	0.00
P_t, C_j	0.00
P_j, C_s	0.00
P_j, C_t	0.00
Sum of first order terms (Table 1.4)	0.66
Grand sum	1

Let us call $S_{C_s P_s}$ the difference

$$S_{C_s P_s} = \frac{V(E(Y|C_s, P_s))}{V_y} - \frac{V(E(Y|C_s))}{V_y} - \frac{V(E(Y|P_s))}{V_y}$$
$$= S^c_{C_s P_s} - S_{C_s} - S_{P_s}. \tag{1.27}$$

Values of this measure for pairs of factors are given in Table 1.5.

Note that we have not listed all effects of the type P_x, P_y, and C_x, C_y in this table as they are all null.

Trying to make sense of this result, one might ponder that if the combined effect of two factors, i.e. $V(E(Y|C_s, P_s))/V_y$, is greater than the sum of the individual effects $V(E(Y|C_s))/V_y$ and $V(E(Y|P_s))/V_y$, perhaps this extra variance describes a synergistic or co-operative effect between these two factors. This is in fact the case and $S_{C_s P_s}$ is called the *interaction* (or two-way) effect of C_s, P_s on Y and measures that part of the effect of C_s, P_s on Y that exceeds the sum of the first-order effects. The reason for the superscript c in $S^c_{C_s P_s}$ can now be explained: this means that the effect measured by $S^c_{C_s P_s}$ is *closed* over the factors C_s, P_s, i.e. by it we capture all the effects that include only these two factors. Clearly if there were a non-zero interaction between C_s and a third factor, say C_t, this would not be captured by $S^c_{C_s P_s}$. We see from

Tables 1.4–5 that if we sum all first-order with all second-order effects we indeed obtain 1, i.e. all the variance of Y is accounted for.

This is clearly only valid for our financial portfolio model because it only has interaction effects up to the second order; if we were to compute higher order effects, e.g.

$$S_{C_s P_s P_t} = S^c_{C_s P_s P_t} - S_{C_s P_s} - S_{P_s P_t} - S_{C_s P_t} - S_{C_s} - S_{P_s} - S_{P_t} \quad (1.28)$$

they would all be zero, as one may easily realise by inspecting Equation (1.1). $S^c_{C_s P_s P_t}$ on the other hand is non-zero, and is equal to the sum of the three second-order terms (of which only one differs from zero) plus the sum of three first-order effects. Specifically

$$S^c_{C_s P_s P_t} = 0.17 + 0.02 + 0.35 + 0.22 = 0.76.$$

The full story for these partial variances is that for a system with k factors there may be interaction terms up to the order k, i.e.

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>j} S_{ijl} + \ldots S_{12\ldots k} = 1 \quad (1.29)$$

For the portfolio model with $k = 6$ all terms above the second order are zero and only three second-order terms are nonzero.

This is lucky, one might remark, because these terms would be a bit too numerous to look at. How many would there be? Six first order, $\binom{6}{2} = 15$ second order, $\binom{6}{3} = 20$ third order, $\binom{6}{4} = 15$ fourth order, $\binom{6}{5} = 6$ fifth order, and one, the last, of order $k = 6$. This makes 63, just equal to $2^k - 1 = 2^6 - 1$, which is the formula to use. This result seems to suggest that the S_i, and their higher order relatives S_{ij}, S_{ijl} are nice, informative and model free, but they may become cumbersomely too many for practical use unless the development (1.29) quickly converges to one. Is there a recipe for treating models that do not behave so nicely?

For this we use the so-called total effect terms, whose description is given next.

Let us go back to our portfolio model and call $\mathbf{X}$ the set of all factors, i.e.

$$\mathbf{X} \equiv (C_s, C_t, C_j, P_s, P_t, P_j) \tag{1.30}$$

and imagine that we compute:

$$\frac{V(E(Y|\mathbf{X}_{-C_s}))}{V_y} = \frac{V(E(Y|C_t, C_j, P_s, P_t, P_j))}{V_y} \tag{1.31}$$

(the all-but-C_s notation has been used). It should now be apparent that Equation (1.31) includes all terms in the development (1.29), of any order, that do not contain the factor C_s. Now what happens if we take the difference

$$1 - \frac{V(E(Y|\mathbf{X}_{-C_s}))}{V_y}? \tag{1.32}$$

The result is nice; for our model, where only a few higher-order terms are non-zero, it is

$$1 - \frac{V(E(Y|\mathbf{X}_{-C_s}))}{V_y} = S_{C_s} + S_{C_s P_s} \tag{1.33}$$

i.e. the sum of all non-zero terms that include C_s. The generalisation to a system with k factors is straightforward:

$$1 - \frac{V(E(Y|\mathbf{X}_{-i}))}{V_y} = \text{sum of all terms of any order that include the factor } X_i.$$

Note that because of Equation (1.21), $1 - V(E(Y|\mathbf{X}_{-i}))/V_y = E(V(Y|\mathbf{X}_{-i}))/V_y$. We indicate this as S_{Ti} and call it the total effect term for factor X_i. If we had computed the S_{Ti} indices for a three-factor model with orthogonal inputs, e.g. our modulus model of Section 1.2, we would have obtained, for example, for factor P_s:

$$S_{TP_s} = S_{P_s} + S_{P_s P_t} + S_{P_s P_j} + S_{P_s P_t P_j} \tag{1.34}$$

and similar formulae for P_t, P_j. For the modulus model, all terms in (1.34) could be non-zero. Another way of looking at the measures $V(E(Y|X_i))$, $E(V(Y|\mathbf{X}_{-i}))$ and the corresponding indices S_i, S_{Ti} is in terms of top- and bottom-marginal variances. We have already said that $E(V(Y|X_i))$ is the average output

Table 1.6 Estimates of the main effects and total effect indices for model (1.1) with six input factors.

Factor	S_x	S_{Tx}
P_s	0.36	0.57
P_t	0.22	0.35
P_j	0.08	0.14
C_s	0.00	0.19
C_t	0.00	0.12
C_j	0.00	0.06
Sum	0.66	1.43

variance that would be left if X_i could be known or could be fixed. Consequently $V(E(Y|X_i))$ is the expected reduction in the output variance that one would get if X_i could be known or fixed. Michiel J. W. Jansen, a Dutch statistician, calls this latter a top marginal variance. By definition the total effect measure $E(V(Y|\mathbf{X}_{-i}))$ is the expected residual output variance that one would end up with if all factors but X_i could be known or fixed. Hence the term, still due to Jansen, of bottom marginal variance. For the case of independent input variables, it is always true that $S_i \leq S_{Ti}$, where the equality holds for a purely additive model.

In a series of works published since 1993, we have argued that if one can compute all the k S_i terms plus all the k S_{Ti} ones, then one can obtain a fairly complete and parsimonious description of the model in terms of its global sensitivity analysis properties. The estimates for our six-factor portfolio model are given in Table 1.6.

As one might expect, the sum of the first-order terms is less than one, the sum of the total order effects is greater than one.

Box 1.5 SIMLAB

Let us try to obtain the numbers in Table 1.6 using SIMLAB. Remember to configure the set of factors so as to include the three factors C_x with their respective distributions and truncations (see Equations (1.23)).

1. Select the 'FAST' sampling method and then 'Specify switches' on the right. Now select 'All first and total order effect calculation (Extended FAST)' in the combo box 'Switch for FAST'. Enter an integer number as seed and the cost of the analysis in terms of number of model executions (e.g. 10 000).
2. Go back to the SIMLAB main menu and click on 'Generate'. After a few moments a sample is available for the simulation.
3. Load the model as in Box 1.3 and click on 'Start (Monte Carlo)'.
4. Run the SA: two pie charts will appear reporting both the S_x and S_{Tx} estimated with the Extended FAST. You can also look at the tabulated values. Try again using the method of Sobol'.

Here we anticipate that the cost of the analysis leading to Table 1.6 is $N(k+2)$, where the cost is expressed in number of model evaluations and N is the column dimension of the Monte Carlo matrix used in the computations, say $N = 500$ to give an order of magnitude (in Box 1.5 $N = 1000/8 = 1250$). Computing all terms in the development (1.29) is more expensive, and often prohibitively so.[5] We would also anticipate, this time from Chapter 4, that a gross estimate of the S_{Tx} terms can be obtained at a lower cost using an extended version of the method of Morris. Also for this method the size is proportional to the number of factors.

1.4 The simple model 'by groups'

Is there a way to compact the results of the analysis further? One might wonder if one can get some information about the overall sensitivity pattern of our portfolio model at a lower price. In fact a nice property of the variance-based methods is that the variance

[5] The cost would be exponential in k, see Chapter 5.

Table 1.7 Estimates of main effects and total effect indices of two groups of factors of model (1.1).

Factor	S_x	S_{Tx}
$\mathbf{P} \equiv (P_s, P_t, P_j)$	0.66	1.00
$\mathbf{C} \equiv (C_s, C_t, C_j)$	0.00	0.34
Sum	0.66	1.34

decomposition (1.29) can be written for sets of factors as well. In our model, for instance, it would be fairly natural to write a variance decomposition as:

$$S_{\mathbf{C}} + S_{\mathbf{P}} + S_{\mathbf{C},\mathbf{P}} = 1 \tag{1.35}$$

where $\mathbf{C} = C_s, C_t, C_j$ and $\mathbf{P} = P_s, P_t, P_j$. The information we obtain in this way is clearly less than that provided by the table with all S_i and S_{Ti}.[6]

Looking at Table 1.7 we again see that the effect of the **C** set at the first order is zero, while the second-order term $S_{\mathbf{C},\mathbf{P}}$ is 0.34, so it is not surprising that the sum of the total effects is 1.34 (the 0.34 is counted twice):

$$\begin{aligned} S_{T\mathbf{C}} &= S_{\mathbf{C}} + S_{\mathbf{C},\mathbf{P}} \\ S_{T\mathbf{P}} &= S_{\mathbf{P}} + S_{\mathbf{C},\mathbf{P}} \end{aligned} \tag{1.36}$$

Now all that we know is the combined effect of all the amounts of hedges purchased, C_x, the combined effect of all the hedged portfolios, P_x, plus the interaction term between the two. Computing all terms in Equation 1.35 (Table 1.7) only costs $N \times 3$, one set of size N to compute the unconditional mean and variance, one for **C** and one for **P**, $S_{\mathbf{C},\mathbf{P}}$ being computed by difference using (1.35). This is less than the $N \times (6 + 2)$ that one would have needed to compute all terms in Table 1.6. So there is less information at less cost, although cost might not be the only factor leading one to decide to present the results of a sensitivity analysis by groups. For instance, we could have shown the results from the portfolio model as

$$S_{\mathbf{s}} + S_{\mathbf{t}} + S_{\mathbf{j}} + S_{\mathbf{s},\mathbf{t}} + S_{\mathbf{t},\mathbf{j}} + S_{\mathbf{s},\mathbf{j}} + S_{\mathbf{s},\mathbf{t},\mathbf{j}} = 1 \tag{1.37}$$

[6] The first-order sensitivity index of a group of factors is equivalent to the closed effect of all the factors in the group, e.g.: $S_{\mathbf{C}} = S^{c}_{C_s,C_t,C_j}$.

Table 1.8 Main effects and total effect indices of three groups of factors of model (1.1).

Factor	S_x	S_{Tx}
$\mathbf{s} \equiv (C_s, P_s)$	0.54	0.54
$\mathbf{t} \equiv (C_t, P_t)$	0.33	0.33
$\mathbf{j} \equiv (C_j, P_j)$	0.13	0.13
Sum	1	1

where $\mathbf{s} \equiv (C_s, P_s)$ and so on for each sub-portfolio item, where a sub-portfolio is represented by a certain amount of a given type of hedge. This time the problem has become additive, i.e. all terms of second and third order in (1.37) are zero. Given that the interactions are 'within' the groups of factors, the sum of the first-order effects for the groups is one, i.e. $S_{\mathbf{s}} + S_{\mathbf{t}} + S_{\mathbf{j}} = 1$, and the total indices are the same as the main effect indices (Table 1.8).

Different ways of grouping the factors might give different insights into the owner of the problem.

Box 1.6 SIMLAB

Let us estimate the indices in Table 1.7 with SIMLAB.

1. Select the 'FAST' sampling method and then 'Specify switches' on the right. Now select 'All first and total order effect calculation on groups' in the combo box 'Switch for FAST'. Enter something as seed and a number of executions (e.g. 10 000).
2. Instead of generating the sample now, load the model first by clicking on 'Configure (Monte Carlo)' and then 'Select Model'.
3. Now click on 'Start (Monte Carlo)'. SIMLAB will generate the sample and run the model all together.
4. Run the SA: two pie charts will appear showing both the S_x and S_{Tx} estimated for the groups in Table 1.7. You can also look at the tabulated values. Try again using larger sample sizes.

1.5 The (less) simple correlated-input model

We have now reached a crucial point in our presentation. We have to abandon the last nicety of the portfolio model: the orthogonality (independence) of its input factors.[7]

We do this with little enthusiasm because the case of dependent factors introduces the following considerable complications.

1. Development (1.29) no longer holds, nor can any higher-order term be decomposed into terms of lower dimensionality, i.e. it is no longer true that

$$\frac{V(E(Y|C_s, P_s))}{V_y} = S^c_{C_s P_s} = S_{C_s} + S_{P_s} + S_{C_s P_s} \tag{1.38}$$

although the left-hand side of this equation can be computed, as we shall show. This also impacts on our capacity to treat factors into sets, unless the non-zero correlations stay confined within sets, and not across them.

2. The computational cost increases considerably, as the Monte Carlo tricks used for non-orthogonal input are not as efficient as those for the orthogonal one.

Assume a non-diagonal covariance structure **C** for our problem:

$$C = \begin{array}{c|cccccc} & P_s & P_t & P_j & C_s & C_t & C_j \\ P_s & 1 & & & & & \\ P_t & 0.3 & 1 & & & & \\ P_j & 0.3 & 0.3 & 1 & & & \\ C_s & . & . & . & 1 & & \\ C_t & . & . & . & -0.3 & 1 & \\ C_j & . & . & . & -0.3 & -0.3 & 1 \end{array} \tag{1.39}$$

We assume the hedges to be positively correlated among one another, as each hedge depends upon the behaviour of a given stock

[7] The most intuitive type of dependency among input factors is given by correlation. However, dependency is a more general concept than correlation, i.e. independency means orthogonality and also implies that the correlation is null, while the converse is not true, i.e. null correlation does not necessarily imply orthogonality (see, for example, Figure 6.6 and the comments to it). The equivalence between null correlation and independency holds for multivariate normal distributions.

Table 1.9 Estimated main effects and total effect indices for model (1.1) with correlated inputs (six factors).

Factor	S_x	S_{Tx}
P_s	0.58	0.35
P_t	0.48	0.21
P_j	0.36	0.085
C_s	0.01	0.075
C_t	0.00	0.045
C_j	0.00	0.02
Sum	1.44	0.785

price and we expect the market price dynamics of different stocks to be positively correlated. Furthermore, we made the assumption that the C_xs are negatively correlated, i.e. when purchasing more of a given hedge investors tends to reduce their expenditure on another item.

The marginal distributions are still given by (1.2), (1.23) above. The main effect coefficients are given in Table 1.9. We have also estimated the S_{Tx} indices as, for example, for P_s:

$$S_{TP_s} = \frac{E(V(Y|\mathbf{X}_{-P_s}))}{V(Y)} = 1 - \frac{V(E(Y|\mathbf{X}_{-P_s}))}{V(Y)} \qquad (1.40)$$

with a brute force method at large sample size. The calculation of total indices for correlated input is not implemented in SIMLAB.

We see that now the total effect terms can be smaller than the first-order terms. This should be intuitive in terms of bottom marginal variances. Remember that $E(V(Y|\mathbf{X}_{-i}))$ is the expected residual variance that one would end up with if all factors but X_i could be known or fixed. Even if factor X_i is still non-determined, all other factors have been fixed, and on average one would be left with a smaller variance, than one would get for the orthogonal case, due to the relation between the fixed factors and the unfixed one. The overall result for a non-additive model with non-orthogonal inputs will depend on the relative predominance of the interaction, pushing for $S_{Ti} > S_i$ as for the C_xs in Table 1.9, and dependency between input factors, pushing for $S_{Ti} < S_i$ as for the P_xs in Table 1.9. For all additive models with non-orthogonal

Table 1.10 Decomposition of $V(Y)$ and relative value of $V(E(Y|X_i))$, $E(V(Y|\mathbf{X}_{-i}))$ for two cases: (1) orthogonal input, all models and (2) non-orthogonal input, additive models. When the input is non-orthogonal and the model non-additive, $V(E(Y|X_i))$ can be higher or lower than $E(V(Y|\mathbf{X}_{-i}))$.

Case (1) Orthogonal input factors, all models. For additive models the two rows are equal.	$V(E(Y\|X_i))$ top marginal. (or main effect) of X_i	$E(V(Y\|X_i))$ bottom marginal. (or total effect) of $\mathbf{X}_{-i}$
	$E(V(Y\|\mathbf{X}_{-i}))$ bottom marginal (or total effect) of X_i	$V(E(Y\|\mathbf{X}_{-i}))$ top marginal (or main effect) of $\mathbf{X}_{-i}$
Case (2) Non-orthogonal input factors, additive models only. If the dependency between inputs vanishes, the two rows become equal. For the case where X_i and $\mathbf{X}_{-i}$ are perfectly correlated both the $E(V(Y\|.))$ disappear and both the $V(E(Y\|.))$ become equal to $V(Y)$.	$V(E(Y\|X_i))$ top marginal of X_i	$E(V(Y\|X_i))$ bottom marginal of $\mathbf{X}_{-i}$
	$E(V(Y\|\mathbf{X}_{-i}))$ bottom marginal of X_i	$V(E(Y\|\mathbf{X}_{-i}))$ top marginal of $\mathbf{X}_{-i}$
	$V(Y)$ (Unconditional)	

input factors it will be $S_{Ti} \leq S_i$. In the absence of interactions (additive model) it will be $S_{Ti} = S_i$ for the orthogonal case. If, still with an additive model, we now start imposing a dependency among the input factors (e.g. adding a correlation structure), then S_{Ti} will start decreasing as $E(V(Y|\mathbf{X}_{-i}))$ will be lower because having conditioned on $\mathbf{X}_{-i}$ also limits the variation of X_i (Table 1.10).

We take factor P_s as an example for the discussion that follows. Given that for the correlated input case S_{TP_s} can no longer be thought of as the sum of all terms including factor P_s, what is the point of computing it? The answer lies in one of the possible uses that is made of sensitivity analysis: that of ascertaining if a given factor is so non-influential on the output (in terms of contribution to the output's variance as usual!) that we can fix it. We submit that if we want to fix a factor or group of factors, it is their S_{Tx} or $S_{T\mathbf{x}}$ respectively that we have to look at. Imagine that we want

to determine if the influence of P_s is zero. If P_s is totally non-influential, then surely

$$E(V(Y|\mathbf{X}_{-P_s})) = 0 \tag{1.41}$$

because fixing 'all but P_s' results in the inner variance over P_s being zero (under that hypothesis the variance of Y is driven only by non-P_s), and this remains zero if we take the average over all possible values of non-P_s. As a result, S_{TP_s} is zero if P_s is totally non-influential. It is easy to see that the condition $E(V(Y|\mathbf{X}_{-P_s})) = 0$ is necessary and sufficient for factor P_s to be non-influent, under any model or correlation/dependency structure among input factors.

1.6 Conclusions

This ends our analysis of the model in Equation (1.1). Although we haven't given the reader any of the computational strategies to compute the S_x, S_{Tx} measures, it is easy to understand how these can be computed in principle. After all, a variance is an integral. It should be clear that under the assumptions that:

1. the model is not so terribly expensive that one cannot afford Monte Carlo simulations, and
2. one has a scalar objective function Y and is happy with its variance being the descriptor of interest.

Then

1. variance based measures offer a coherent strategy for the decomposition of the variance of Y;
2. the strategy is agile in that the owner of the problem can decide if and how to group the factors for the analysis;
3. this strategy is model free, i.e. it also works for nasty, non-monotonic, non-additive models Y, and converge to easy-to-grasp statistics, such as the squared standardised regression coefficients β_x^2 for the linear model;

4. it remains meaningful for the case where the input factors are non-orthogonal;
5. it lends itself to intuitive interpretations of the analysis, such as that in terms of top and bottom marginal variances, in terms of prioritising resources to reduce the uncertainty of the most influential factors or in terms of fixing non-influential factors.