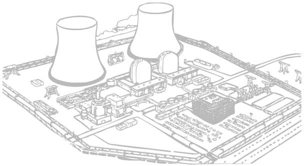
RISK ENGINEERING



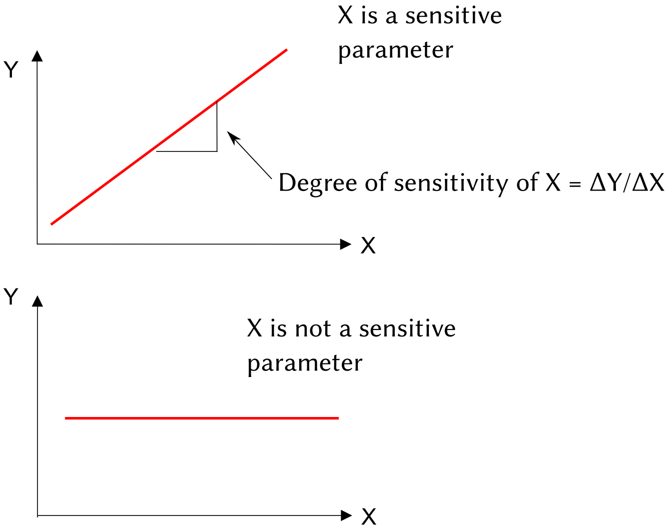
Sensitivity analysis for risk-related decision-making

Eric Marsden

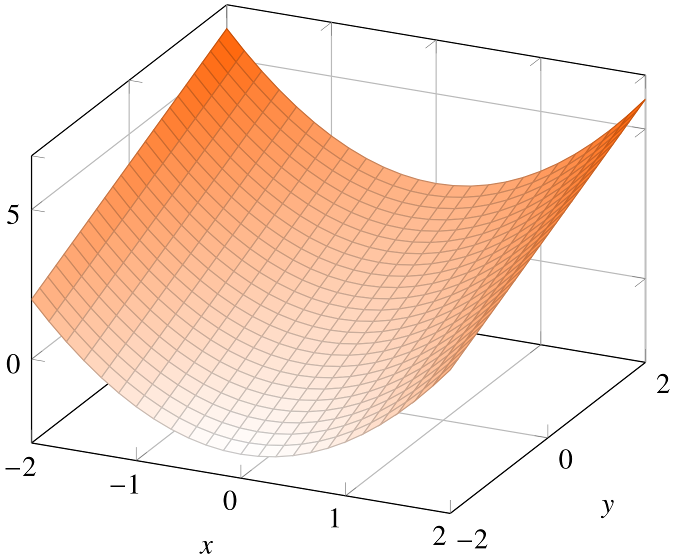
<eric.marsden@risk-engineering.org>

What are the key drivers of my modelling results?

Sensitivity analysis: intuition

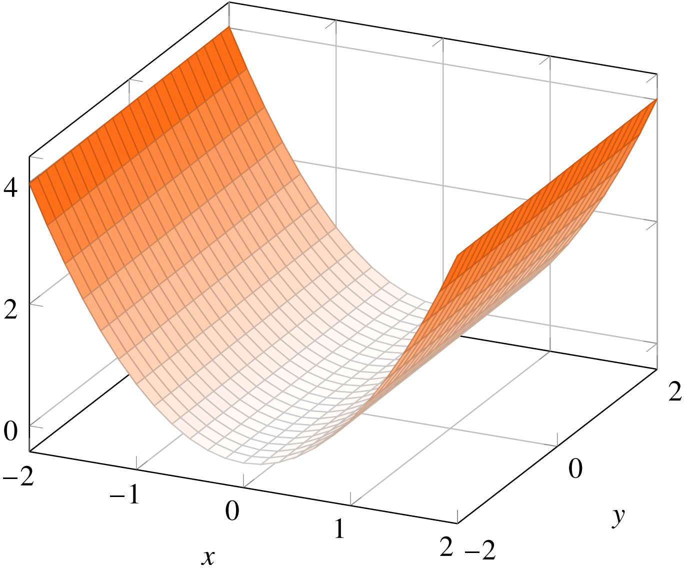


Sensitivity analysis: intuition



$f(x, y)$ is sensitive in x and in y

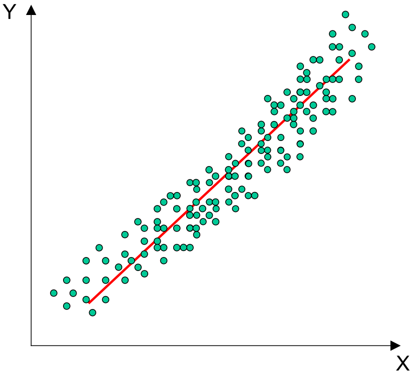
Sensitivity analysis: intuition



$f(x, y)$ is sensitive in x but not in y

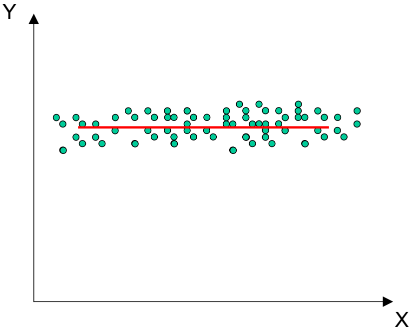
Sensitivity analysis: intuition

- ▷ Consider a case where $Y = f(x)$ and the function f is a “black box”
- ▷ We can sample Y for different values of X to reassemble the relationship
- ▷ Here: X is a sensitive parameter



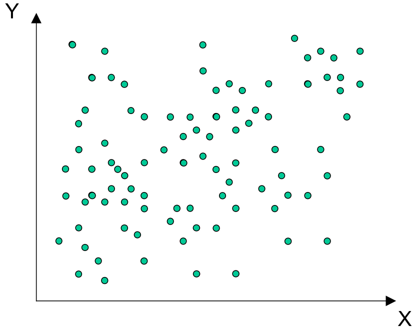
Sensitivity analysis: intuition

- Here, X is not sensitive
- Can be "seen" visually



Sensitivity analysis: intuition

- ▷ What can we say about the sensitivity of *X*?
- ▷ No graphical interpretation
- ▷ Consider also functions (or computer models, or spreadsheets) which have tens of inputs
  - you can't draw graphs in dimension 23! 
- ▷ We need a more sophisticated method than scatterplots...

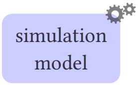


What is sensitivity analysis?

- The study of how the variation (uncertainty) in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the model inputs
- Answers the question *"What makes a difference in this decision problem?"*
- Can be used to determine whether further research is needed to reduce input uncertainty before making a decision
  - information is not free

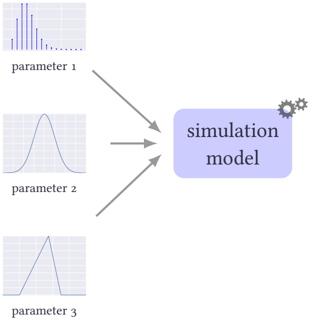
RISK ENGINEERING

Sensitivity and uncertainty analysis



Start with a simulation model that you want better to understand (often a computer model). It may be a "black box" (you don't know how it works internally).

Sensitivity and uncertainty analysis



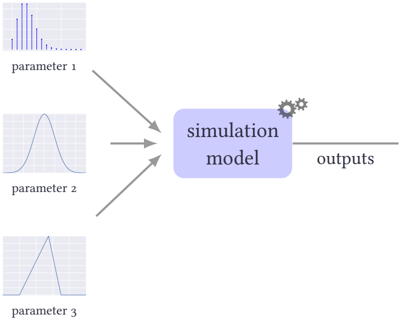
Start with a simulation model that you want better to understand (often a computer model). It may be a "black box" (you don't know how it works internally).

Define which uncertain input parameters you want to analyze.

Characterize their probability distributions.

Adapted from figure by A. Saltelli

Sensitivity and uncertainty analysis

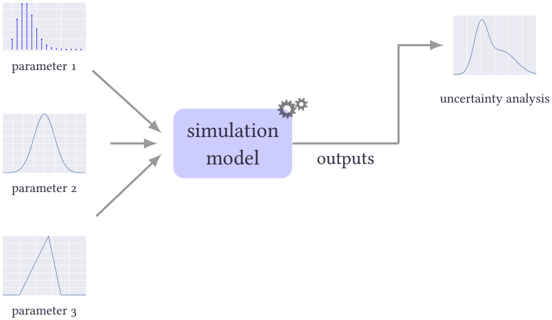


Propagate the input variability to the output variability by running the model a large number of times with inputs taken from the input probability distributions.

(This is a "Monte Carlo" or "stochastic simulation" method.)

Adapted from figure by A. Saltelli

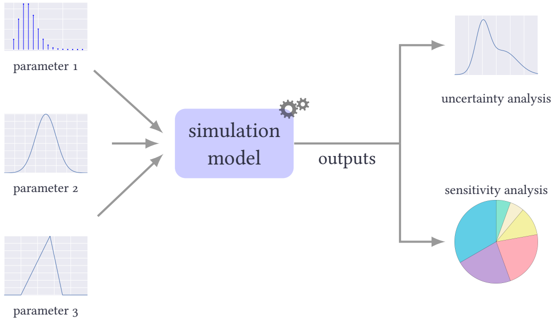
Sensitivity and uncertainty analysis



Uncertainty analysis: how does variability in the inputs propagate through the model to variability in the outputs?

Adapted from figure by A. Saltelli

Sensitivity and uncertainty analysis

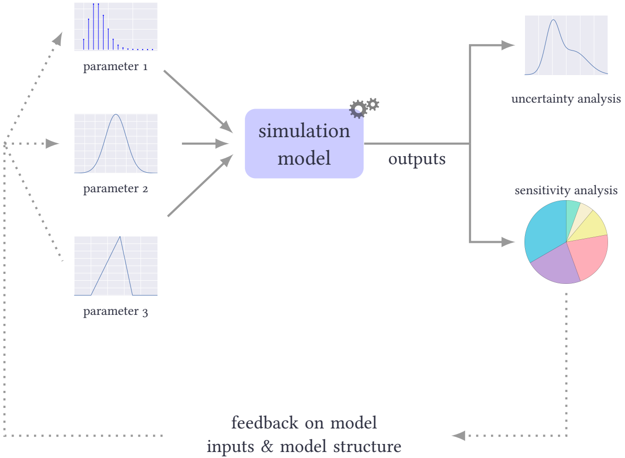


Uncertainty analysis: how does variability in the inputs propagate through the model to variability in the outputs?

Sensitivity analysis: what is the relative contribution of the variability in each of the inputs to the total output variability?

Adapted from figure by A. Saltelli

Sensitivity and uncertainty analysis



Insights gained from the sensitivity analysis may help to

- prioritize effort on reducing input uncertainties
- improve the simulation model

Adapted from figure by A. Saltelli

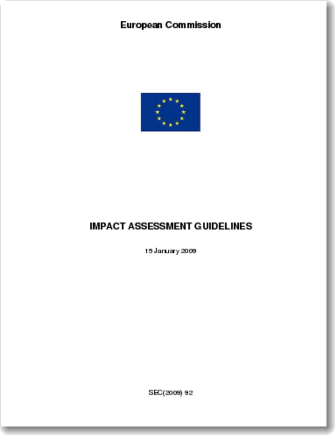
Applications of sensitivity analysis

- Risk communication
  - how much of my output uncertainty is *irreducible* (caused by aleatory uncertainty in input parameters)?
  - how much is *epistemic* (related to lack of knowledge, could be reduced with more research)?
- Optimize research investment to improve risk analysis
  - which uncertain input parameters contribute the most to model output uncertainty?
  - on which uncertain input parameters should I spend my research money to gain the biggest reduction in uncertainty?
- Model reduction
  - identify ineffective parameters
  - generate models with fewer parameters, but (almost) identical results (*metamodels* or *response surfaces*)

Application areas

The European Commission recommends sensitivity analysis in the context of its **impact assessment guidelines** (2009):

> *When the assumptions underlying the baseline scenario might vary as a result of external factors, you need to do a sensitivity analysis to assess whether the impacts of the policy options differ significantly for different values of the key variables.*


European Commission

IMPACT ASSESSMENT GUIDELINES

15 January 2009

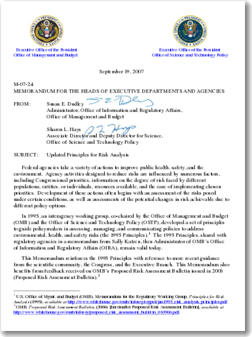
SEC(2009) 92

Source: EC guidance document at `ec.europa.eu/smart-regulation/impact/`

Application areas

Principles for Risk Analysis published by the US Office of Management and Budget:

> *Influential risk assessments should characterize uncertainty with a sensitivity analysis and, where feasible, through use of a numeric distribution.*
>
> *[…] Sensitivity analysis is particularly useful in pinpointing which assumptions are appropriate candidates for additional data collection to narrow the degree of uncertainty in the results. Sensitivity analysis is generally considered a minimum, necessary component of a quality risk assessment report.*



Executive Office of the President
Office of Management and Budget

Executive Office of the President
Office of Science and Technology Policy

September 19, 2007

M-07-24
MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Susan E. Dudley
Administrator, Office of Information and Regulatory Affairs,
Office of Management and Budget

Sharon L. Hays
Associate Director and Deputy Director for Science,
Office of Science and Technology Policy

SUBJECT: Updated Principles for Risk Analysis

Federal agencies take a variety of actions to improve public health, safety, and the environment. Agency activities designed to reduce risks are influenced by numerous factors, including Congressional priorities, information on the degree of risk faced by different populations, entities, or individuals, resources available, and the ease of implementing chosen priorities. Development of these actions often begins with an assessment of the risks posed under certain conditions, as well as assessments of the potential changes in risk achievable due to different policy options.

In 1995, an interagency working group, co-chaired by the Office of Management and Budget (OMB) and the Office of Science and Technology Policy (OSTP), developed a set of principles to guide policymakers in assessing, managing, and communicating policies to address environmental, health, and safety risks (the 1995 Principles).[1] The 1995 Principles, shared with regulatory agencies in a memorandum from Sally Katzen, then Administrator of OMB's Office of Information and Regulatory Affairs (OIRA), remain valid today.

This Memorandum reinforces the 1995 Principles with reference to more recent guidance from the scientific community, the Congress, and the Executive Branch. This Memorandum also benefits from feedback received on OMB's Proposed Risk Assessment Bulletin issued in 2006 (Proposed Risk Assessment Bulletin).[2]

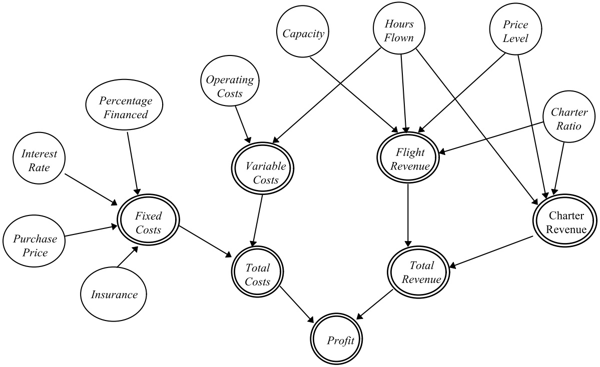
[1] U.S. Office of Mgmt. and Budget (OMB), Memorandum for the Regulatory Working Group, *Principles for Risk Analysis* (1995), *available at* http://www.whitehouse.gov/omb/inforeg/regpol/jan1995_risk_analysis_principles.pdf.
[2] OMB, *Proposed Risk Assessment Bulletin*, (2006) [hereinafter Proposed Risk Assessment Bulletin], *available at* http://www.whitehouse.gov/omb/inforeg/proposed_risk_assessment_bulletin_010906.pdf.

Source: *Updated principles for risk analysis*, US OMB, `whitehouse.gov/sites/default/files/omb/assets/regulatory_matters_pdf/m07-24.pdf`

Sensitivity analysis: the process

1. Specify the **objective of your analysis**
   - Example: *"Which variables have the most impact on the level of risk?"*
2. Build a model which is suitable for automated numerical analysis
   - Example: spreadsheet with inputs in certain cells and the output of interest in another cell
   - Example: computer code that can be run in batch mode
3. Select a sensitivity analysis method
   - most appropriate method will depend on your objectives, the time available for the analysis, the execution cost of the model
4. Run the analysis
5. Present results to decision-makers

Step 2: build a model which can be run by a computer



A useful starting point is to build an *influence diagram* of the relevant variables.

Use this diagram to build your model (*e.g.* profit = revenue - expenses), or check that all relevant variables are integrated in your existing numerical model.

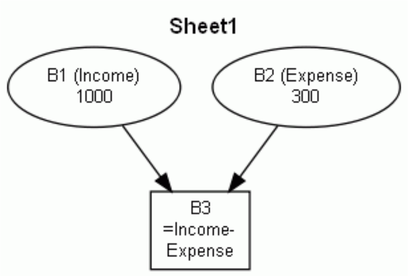
It should be possible to run the model in an automated way (called “batch mode”).

Figure from Clemen, R. T., *Making Hard Decisions: An Introduction to Decision Analysis*, 1996

Aside: extracting a model from a spreadsheet

In some cases, the model you want to analyze will already be implemented in a spreadsheet such as Microsoft Excel.

The free Trace plugin for Excel can extract a dependency graphs that depict relationships between cells in your spreadsheet. This can help to build a model that is easy to run in batch mode.



Trace tool can be downloaded for free from `christopherteh.com/trace/`

Step 3: sensitivity analysis methods

1. Basic approach: tornado diagram
2. Screening methods
3. OAT "one at a time" methods
4. Local sensitivity analysis
5. Global sensitivity analysis

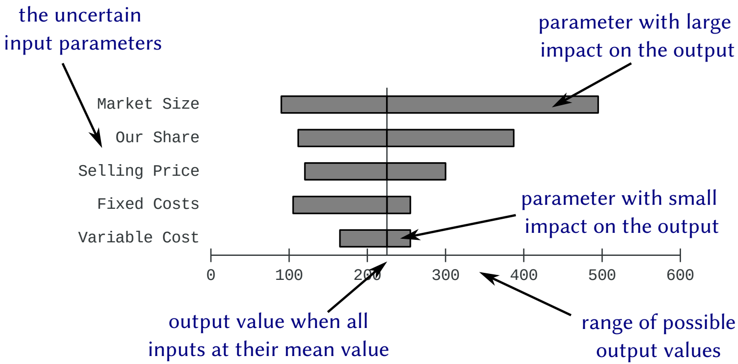
increasing sophistication

more information available

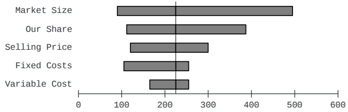
RISK ENGINEERING

Basic SA: tornado diagram

- **Tornado diagram**: way of presenting basic sensitivity information
  - mostly used for project risk management or net present value estimates
  - sometimes called "what-if" analysis



How to produce a tornado diagram

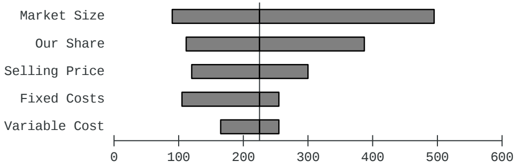


- ▷ Determine lower bound, upper bound and best estimate of each uncertain input parameter
  - = 10%, 90% and 50% quantiles of parameter's probability distribution
- ▷ For each uncertain parameter, calculate model output for lower and upper bounds, while taking best estimate for all other uncertain parameters
- ▷ Draw a horizontal bar for each uncertain parameter between value for lower bound and value for upper bound
- ▷ Vertical order of uncertain parameters given by width of the bar
  - parameters which lead to large output "spread" (have more impact) at top
- ▷ Draw a vertical line at position of the expected value
  - calculated using best estimate for each uncertain parameter

Tornado diagram example: profit calculation

Profit = (SellingPrice - VariableCost) × MarketSize × MarketShare - FixedCosts

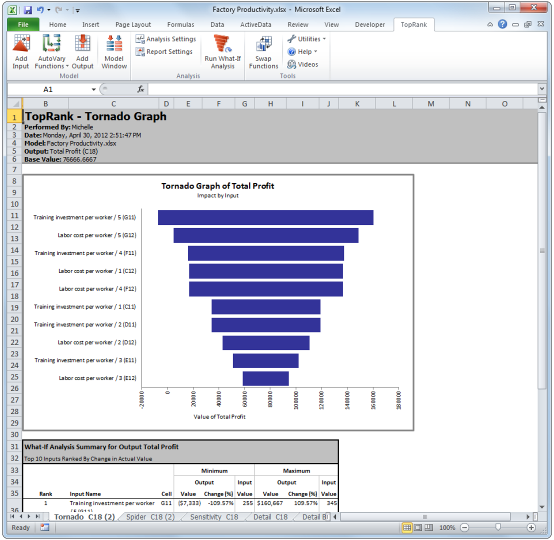
Parameter	Lower	Expected	Upper
Selling price	140	175	200
Market size	8	12	20
Our share	0.18	0.25	0.35
Variable cost	30	40	60
Fixed costs	150	180	300



Interpretation: given these assumptions, the market size parameter has most influence on profitability.

Plot generated with free Excel plugin by home.uchicago.edu/ rmyerson/addins.htm

Relevant commercial tools



Example tools with Excel integration:

- ▷ Palisade TopRank®
- ▷ Oracle Crystal Ball®

Typically quite expensive...

Screening methods

- "Screening" is a preliminary phase that is useful when the model has a large number of parameters
  - allows the identification, with a limited number of calculations, of those parameters that generate significant variability in the model's output
- Simple "OAT" (one at a time) screening method: change one factor at a time and look at effect on output
  - while keeping other factors at their nominal value

☺ Intuitive approach, can be undertaken by hand

☺ If model fails, you know which factor is responsible for the failure

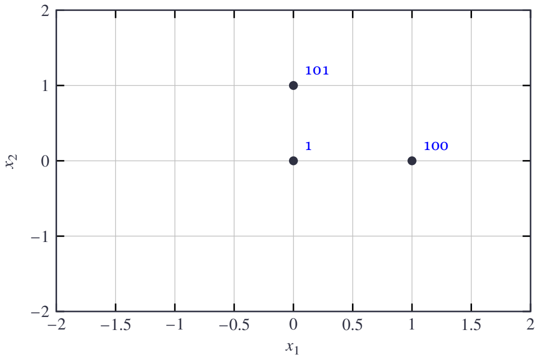
☹ Approach does not fully explore input space, since simultaneous variations of input variables are not studied

☹ Cannot detect the presence of interactions between input variables

OAT screening method: example

Real modelling situations have many more than two variables, so output cannot be plotted

Rosenbrock function: $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ over [-2, 2]²




```
> def rosenbrock(x1, x2):
    return 100*(x2-x1**2)**2 + (1-x1)**2
> rosenbrock(0, 0)
1
> rosenbrock(1, 0)
100
> rosenbrock(0, 1)
101
> rosenbrock(1, 1)
0
> rosenbrock(-1, -1)
404
```

Both x_1 and x_2 seem to be sensitive variables in this example

RISK ENGINEERING

The Elementary Effects screening method

- The *elementary effect* for the i-th input variable at $\mathrm{x} \in [0, 1]^k$ is the first difference approximation to the derivative of $f(\cdot)$ at x:

$$EE_i(\mathrm{x}) = \frac{f(\mathbf{x} + \Delta \mathbf{e}_i) - f(\mathbf{x})}{\Delta}$$

  where $\mathbf{e}_i$ is the unit vector in the direction of the i-th axis

- Intuition: it's the **slope of the secant line** parallel to the input axis

- Average $EE_i(\mathbf{x})$ for various points $\mathbf{x}$ in the input domain to obtain a measure of the relative influence of each factor

$$\mu_i = \frac{1}{r} \sum_{j=1}^{r} \left| EE_i(x_j) \right|$$

Elementary effects method: example

▷ Consider $y(\mathbf{x}) = 1.0 + 1.5x_2 + 1.5x_3 + 0.6x_4 + 1.7x_4^2 + 0.7x_5 + 0.8x_6 + 0.5(x_5 x_6)$ where

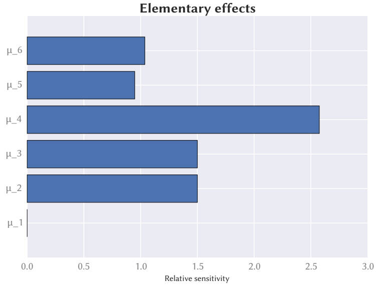
- $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$
- $0 \leq x_1, x_2, x_4, x_5, x_6 \leq 1$
- $0 \leq x_3 \leq 5$

▷ Note:

- $y(\cdot)$ is functionally independent of x_1
- $y(\cdot)$ is linear in x_2 and x_3 and non-linear in x_4
- $y(\cdot)$ contains an interaction in x_5 and x_6

▷ Sensitivity results:

- $\mu_1 = 0$ as expected
- influence of x_4 is highest
- influence of x_2 and x_3 is equal, as expected



Download full details as a Python notebook at `risk-engineering.org`

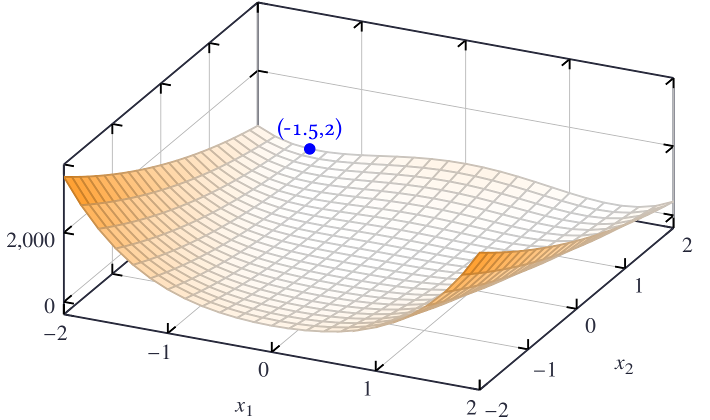
Local sensitivity analysis methods

- Local sensitivity analysis: investigation of response stability over a **small region of inputs**
- Local sensitivity with respect to a factor is just the partial derivative wrt that factor, evaluated at that location
- Simple example: Rosenbrock function
  - $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x_1, x_2 \in [-2, 2]$
  - $\frac{\partial f}{\partial x_1} = -400x_1(-x_1^2 + x_2) + 2x_1 - 2$
  - $\frac{\partial f}{\partial x_2} = -200x_1^2 + 200x_2$

RISK ENGINEERING

Rosenbrock example

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x_1, x_2 \in [-2, 2]$$



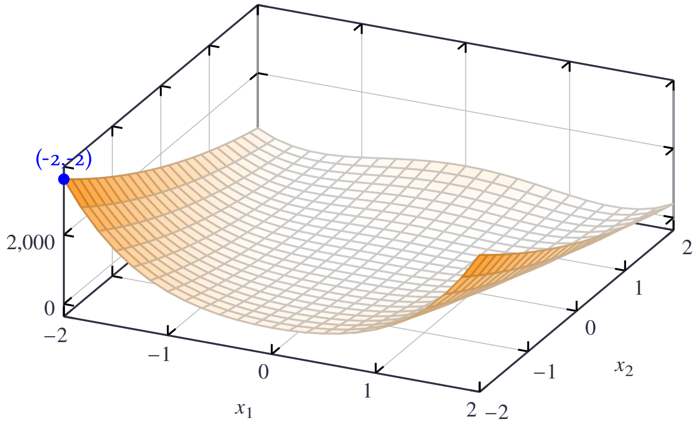
$$\frac{\partial f}{\partial x_1} = -155$$

$$\frac{\partial f}{\partial x_2} = -50$$

Local sensitivity is low

Rosenbrock example

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x_1, x_2 \in [-2, 2]$$



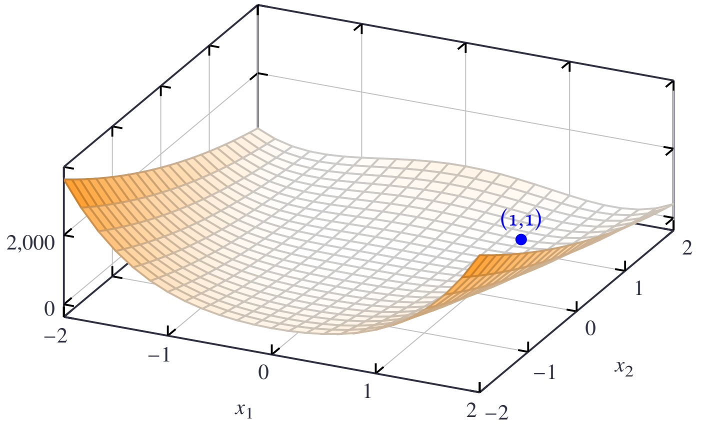
$$\frac{\partial f}{\partial x_1} = \text{-4806}$$

$$\frac{\partial f}{\partial x_2} = \text{-1200}$$

Local sensitivity is high

Rosenbrock example

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x_1, x_2 \in [-2, 2]$$



$$\frac{\partial f}{\partial x_1} = 0$$

$$\frac{\partial f}{\partial x_2} = 0$$

Local sensitivity is zero

Local sensitivity analysis methods

- The calculation of partial derivatives can be automated for software packages using *automatic differentiation* methods
  - the source code must be available
  - → `autodiff.org`
- This method does not allow you to detect **interaction effects** between the input variables
- Method cannot handle correlated inputs
- Widely used for optimization of scientific software

Global sensitivity analysis methods

- Examine effect of changes to all input variables *simultaneously*
  - over the **entire input space** you are interested in (for example the uncertainty distribution of each variable)
- Methods based on analysis of variance (often using Monte Carlo methods)
- Typical methods that are implemented in software packages:
  - Fourier Analysis Sensitivity Test (FAST), based on a multi-dimensional Fourier transform
  - the method of Sobol'
- In general, this is the most relevant method for risk analysis purposes
  - allows the analysis of **interactions** between input variables

RISK ENGINEERING

Sensitivity indices

- The *sensitivity index* of a parameter **quantifies its impact on output uncertainty**
  - measures the part of output variance which can be attributed to variability in the parameter
- Properties:
  - $S_i \in [0,1]$
  - $\sum_i S_i = 1$
- First-order index: $S_j = \frac{\mathrm{Var}(\mathbb{E}[z|x_j])}{\mathrm{Var}(z)}$
  - measures "main effect"
- Total effect index: $T_j = \frac{\mathbb{E}[\mathrm{Var}(z|x_j)]}{\mathrm{Var}(z)}$
  - measures residual variability due to interactions between x_i and other parameters

Estimating sensitivity indices using SciPy and SALib

```
import numpy
from SALib.sample import saltelli
from SALib.analyze import sobol

def rosenbrock(x1, x2):
    return 100 * (x2 - x1**2)**2 + (1 - x1)**2

problem = {
    "num_vars": 2,
    "names": ["x1", "x2"],
    "bounds": [[-2, 2], [-2, 2]]
}
sample = saltelli.sample(problem, N, calc_second_order=True)
Y = numpy.empty([sample.shape[0]])
for i in range(len(Y)):
    x = sample[i]
    Y[i] = rosenbrock(x[0], x[1])
Si = sobol.analyze(problem, Y, calc_second_order=True)
```

Download full details as a Python notebook at `risk-engineering.org`

The SALib python library is free software available from `github.com/SALib/SALib`

Further reading

- Appendix on sensitivity analysis of the EPA guidance on risk assessment, `epa.gov/oswer/riskassessment/rags3adt/pdf/appendixa.pdf`
- OpenTURNS software platform for uncertainty analysis (free software), `openturns.org`
- Dakota software platform for uncertainty analysis (free software), `dakota.sandia.gov`
- Case study: *Using @RISK for a Drug Development Decision: a Classroom Example*, from `palisade.com/cases/ISU_Pharma.asp`

Feedback welcome!

Was some of the content unclear? Which parts were most useful to you? Your comments to `feedback@risk-engineering.org` (email) or `@LearnRiskEng` (Twitter) will help us to improve these materials. Thanks!

`@LearnRiskEng`

`fb.me/RiskEngineering`

For more free content on risk engineering, visit `risk-engineering.org`