Assa: algorithms for stochastic sensitivity analysis

Michiel J. W. Jansen

Biometris, Wageningen University and Research Centre, P.O. Box 100, 6700 AC Wageningen, The Netherlands
E-mail: michiel.jansen@wur.nl

Abstract: ASSA is a public-domain open-source library of algorithms for stochastic sensitivity analysis in ANSI C. It should serve as a documented collection of basic and more sophisticated algorithms in that field. Its open character ought to advance the applicability, the quality and a restrained kind of completeness of the collection.

Keywords: software, Monte Carlo, latin hypercube, analysis of variance, regression-based, regression-free

1. INTRODUCTION

Various software products exist for stochastic sensitivity analysis (SSA; the adjective 'stochastic' is added to distinguish the subject from deterministic sensitivity analysis). Saltelli, Chan and Scott [1] contains an overview of software available in the year 2000. The software packages mentioned there are closed in the sense that you can hardly change or add components. The section in [1] on generic algorithms is still far from complete. Thus, there does not seem to exist a fairly complete, coherent, and documented collection of algorithms for SSA in a basic programming language like C or Fortran. The ASSA project has the purpose to begin filling this gap. The collection is available in the public domain, in such a form that everyone can use the software freely. It is hoped that users will suggest improvements or additions. The long-term goal is a collection of documented algorithms in the spirit of the famous series of Numerical Recipes [2], but with a slightly different legal status.

Model builders should be enabled to incorporate the algorithms into their own software, for instance in order to accompany model statements with an indication of inaccuracy due to input uncertainty. Another application is inclusion of SSA algorithms into generic frameworks for building, coupling and analysing models.

The language used is ANSI C, written in such a style that translation into another basic programming language should pose no serious problems.

At present, ASSA consists mainly of conventional algorithms for sensitivity analysis. Apart from auxiliary routines, the algorithms can be divided into algorithms for constructing model input samples and algorithms for analysing the corresponding model output samples. All sensitivity analyses are variance-based.

The long-term goals are: a gradually improving and extending collection of basic and advanced algorithms for SSA, leading to a moderate form of standardisation; uniform description of these algorithms via C-programs; a form of publication inviting comments and additions, while enabling flexible use of the algorithms.

Sensitivity Analysis of Model Output
Kenneth M. Hanson and François M. Hemez, eds.
Los Alamos National Laboratory, 2005; http://library.lanl.gov/

2. OVERVIEW OF ALGORITHMS

2.1. Input generators

Random generators are used to construct a sample from the distribution of unknown parameters or other model inputs. In the present version, the statistical properties of the inputs can be described in terms of their grade correlation (often loosely called rank correlation). Each of the individual variables is defined through the type and parameters of its distribution.

There are two basic random generators: uniform(0,1) and multinormal(μ, Σ). For the rest, the drawing of random samples is done in two steps.

The first step draws a sample from the k-dimensional unit-hypercube. Each of the k variables thus sampled is more or less randomly and more or less uniformly distributed over the interval (0,1). Some examples: independent; dependent with given grade correlation; latin hypercube [3]; latin hypercube with forced rank correlation [4]; and by way of example a systematic sample constructed from a saturated main-effect design.

The next step transforms these (0,1) variables into variables with the required distribution. The distributions currently available are: uniform, triangular, normal, log-normal, beta and gamma. Auxiliary routines are supplied to derive the standard parameters of distributions from information about means and variances, or the quantiles.

2.2. Analysis

In ASSA's present version, all sensitivity analyses are variance-based, i.e. they perform some kind of analysis of variance on the model output. During the 1990's there seems to have grown consensus that this form of sensitivity analysis is quite adequate for most purposes. The algorithms provide the possibility to estimate the variance contributions of *groups* of inputs, which often facilitates the interpretation of the results, especially when variables belonging to different groups are stochastically independent. There is an algorithm for the most common form of sensitivity analysis: the one based on linear regression. A spline-regression-based analysis is on the list of desiderata. All ingredients are present for a Sobol' type sensitivity analysis [5]. An algorithm for winding stairs analysis [6] is under construction. For the time being, only two simple test functions are included, whose sensitivity properties can be calculated analytically.

2.3. Auxiliary routines

ASSA contains routines to summarise the statistical properties of a sample: variance matrix, correlation matrix, mean, variance, median and rank-correlation. There is an algorithm to check if a symmetric matrix is positive definite. Graphical routines are not included.

3. SUPPORTING SOFTWARE REQUIRED

The current version of ASSA frequently uses algorithms from Numerical Recipes in C (NRC; Press et al., 1992). Thus, one may only use this version of ASSA in applications where one is entitled to use the algorithms from Numerical Recipes. The NRC procedures used for ASSA serve mainly to allocate and free memory space for vectors and matrices, to generate uniform random numbers, and to calculate special functions relating to probability distributions.

4. AVAILABILITY AND LEGAL MATTERS

ASSA's manual and source code can be downloaded free of charge from the website of NPB, the Dutch Nature Policy Assessment Office: www.natuurplanbureau.nl. Once there, click "publicaties" in the left column; next click "werkdocumenten" in the left column; then click "2004" in the central block; and finally download the ASSA manual and source code from the displayed list of documents. A link to that website indicated will be offered to the SAMO site http://sensitivity-analysis.jrc.cec.eu.int.

Permission to use, copy and distribute this software and its documentation for any purpose is granted without fee, provided that the entire package is kept together and that this permission and disclaimer notice appears in all copies. NPB, Biometris, Plant Research International and Wageningen UR make no warranty of any kind, expressed or implied, including without limitation any warranties of merchantability and/or fitness for a particular purpose. NPB, Biometris, Plant Research International and Wageningen UR do not assume any liability for the use of this software. In no event will NPB, Biometris, Plant Research International or Wageningen UR be liable for any additional damages, including any lost profits, lost savings, or other incidental or consequential damages arising from the use of or inability to use, this software and its accompanying documentation, even if NPB, Biometris, Plant Research International or Wageningen UR has been advised of the possibility of such damages.

ACKNOWLEDGEMENTS

The research enabling the construction of ASSA has been performed in the framework of strategic research of Biometris and in various consultation projects. The actual writing up was supported by NPB, the Dutch Nature Policy Assessment Office. The stimulating conversations with Harm Houweling of that office constituted vital support. I am grateful to Jacques Withagen of Biometris for code-checking and comparison of ASSA results with results from other software.

REFERENCES

1. A. Saltelli, K. Chan and E. M. Scott. *Sensitivity analysis*. Wiley, Chichester, 2000.
2. W. H. Press, S. A. Teukolsky, W. T. Vettering and B. P. Flannery. *Numerical recipes in C: the art of scientific computing, second edition*. Cambridge University, Cambridge, 1992.
3. M. D. McKay, R. J. Beckman and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21, 239-245, 1979.
4. R. L. Iman and W. J. Conover. A distribution-free approach to inducing rank correlation among input variables, *Communications in Statistics*. Part B, Simulation and Computation, vol 11, 311-334, 1982.
5. I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models, *Matematicheskoe Modelirovanie* 2, 112-118, 1990, (in Russian). Translated in: *Mathematical Modelling and Computational Experiments*, vol 1, 407- 414, 1993.
6. M. J. W. Jansen, W. A. H. Rossing and R. A. Daamen. Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In: J. Grasman and G. Van Straten (eds.), *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, p334-343, Kluwer, Dordrecht, 1994.