Available online at www.sciencedirect.com

ScienceDirect

Procedia Social and Behavioral Sciences 2 (2010) 7629–7630

Procedia Social and Behavioral Sciences

Sixth International Conference on Sensitivity Analysis of Model Output

Sampling plans for the estimation of moment-independent importance measures

W. Castaings[a,c], E. Borgonovo[b], S. Tarantola[c]1*

[a]*Institut de Mécanique des Fluides de Toulouse, Université de Toulouse, Allée Camille Soula, F-31400 Toulouse, France*
[b]*Dept. of Decision Sciences and ELEUSI, Bocconi University, Via Roentgen 1, 20135, Milano, Italy.*
[c]*Joint Research Center of the European Commission, Via Enrico Fermi 1, 21027, Ispra (VA), Italy*

Abstract

The numerical and computational aspects underlying the approximation of moment independent sensitivity measures are discussed. Sampling plans based on column substitution and column permutations are evaluated and compared for both analytical test cases and a practical application. The influence of the sampling strategy (simple random, latin hypercube or quasi-random sequences) is investigated.

Keywords: Type your keywords here, separated by semicolons ;

1. Main text

Probabilistic sensitivity analysis methods are based on the design of computer experiments sampling the space of uncertain inputs in order to characterise the underlying input-output mapping. The analysis of the spatial dispersion around the expectation of the variable of interest is natural given the common gaussianity assumption underlying most of our analysis techniques. However, variance is not the only nor the best measure of uncertainty. For example, variance is not adapted when the model output is characterized by a heavy-tail or a multi-modal probability distribution function. It does not reflect the decision-maker's state of belief and basing the analysis on this moment can result in misleading conclusions. The objective of the present contribution is to propose and compare different sampling strategies for the estimation of the moment-independent sensitivity measures proposed by *Borgonovo* (2006).

Let us denote $X=(X_1,X_2,...,X_k)$ the vector of model inputs and Y the output of interest which are treated as random variables characterised by probability distribution functions $f_X(x)$ and $f_Y(y)$. The importance of X_i is defined as:

$$\delta_i = E_{X_i}\left[\; _i(X_i) \right] \tag{1}$$

where

* Corresponding author. Tel.: +33 (0)561-285-809
E-mail address:william.castaings@imft.fr.

doi:10.1016/j.sbspro.2010.05.151

$$s_i(X_i) = \int \left| f_Y(y) - f_{Y|X_i}(y) \right| dy \tag{2}$$

The approximation of δ_i, the normalized expected shift in the probability distribution of Y produced by fixing X_i, involves the resolution of a multi-dimensional integral using Monte Carlo methods. The sampling strategy adopted in order to generate the conditional input samples is the key aspect of the calculation method.

Substituted columns sampling plans and permuted columns sampling plans (*Morris et al*, 2006) are proposed and discussed in the context of moment independent sensitivity analysis. Using Latin Hypercube sampling and quasi-Monte Carlo sequences, the previously mentioned sampling plans are evaluated for analytical test cases derived by *Borgonovo et al* (2010). In order to assess the accuracy and robustness of the different approaches, numerical estimates and calculated for 50 replicates using an increasing base sample size. Additional experiments were carried out for the LevelE model, a model simulating the radiological dose to humans over geological time scales due to the underground migration of radionuclides from a nuclear waste disposal site.

The results show the superiority of Quasi-Random sequences over Simple Random or Latin Hypercube sampling (similar outcome was obtained for variance-based sensitivity analysis by Kucherenko (2007)). Moreover, using substituted column sampling plans, performances can be improved by reducing the number of sample points characterising the external integration loop. Although a slight positive bias was observed for non-influent model inputs, the use of permuted column sampling plans is the most efficient calculation method.

2. References

Borgonovo E., 2006: “Measuring uncertainty importance: Investigation and comparison of alternative approaches”. *Risk Analysis*, 26(5), pp. 1349–1361.

Borgonovo E., Castaings, W., Tarantola, S., 2010: “Moment Independent Uncertainty Importance: New Results and Analytical Test Cases” *work in progress*

Kucherenko S., 2007: “Application of global sensitivity indices for measuring the effectiveness of quasi-monte carlo methods and parameter estimation”. In: *Proceedings of the Fifth International Conference on Sensitivity Analysis of Model Output (SAMO 2007)*, pp. 35-36.

Morris M.D., Moore L. M. and McKay, M.D., 2006: “Sampling plans based on balanced incomplete block designs for evaluating the importance of computer model inputs”. *Journal of Statistical Planning and Inference*, 136(9), pp 3203–3220.