RISK ANALYSIS: FREQUENTISTS AND BAYESIANS UNITE!

Jack P.C. Kleijnen
Tilburg University
Faculty of Economics and Business Administration
Tilburg, the Netherlands

ABSTRACT

Both deterministic and random simulation models may have unknown values for their parameters and input variables. Simulation analysts may therefore assume that these values have a statistical distribution (based on expert opinions or past data). This assumption results in so-called Risk Analysis or Uncertainty Analysis. Frequentists use crude Monte Carlo, Latin Hypercube Sampling, or some other method to sample values from this assumed input distribution; next, the simulation model transforms these input values into output values; repeating this sampling many times gives an Estimated Distribution Function for the outputs; this function provides an estimate of the probability of a specific 'disaster'. Software is abundant for this analysis. Bayesians use more specific input distributions; e.g., a conjugate prior distribution. They also obtain simulation data, but next they compute the posterior output distribution. Frequentists and Bayesians should unite to exchange experiences and theories.

1 Introduction

Before discussing Risk Analysis (RA)—also known as Uncertainty Analysis (UA)—I distinguish between two types of simulation models, namely *deterministic* and *random* (or stochastic) simulation models. By definition, random simulation uses Pseudo-Random Numbers (PRNs).

Deterministic simulation is often applied in *Computer Aided Engineering* (CAE) and *Computer Aided Design* (CAD). Examples are models of airplanes, automobiles, chemical processes, and computer chips. Recent surveys are Chen et al. (2003), Meckesheimer et al. (2001), Oden (2006), Post et al. (2004), Simpson et al. (2001), and Stinstra (2006).

Other types of deterministic models are used in Operations Research/Management Science (OR/MS); e.g., models of *routing protocols* in telematics and *project planning* in the 'Critical Path Method' (CPM) and 'Programme Evaluation and Review Technique' (PERT). These models usually assume known values for the components of the total routing or project respectively. A recent survey is Elmaghraby (1995).

Such a *deterministic* simulation model becomes a *random* simulation model if the exact values of the model's parameters or input variables are unknown so their values are sampled from a distribution function. This is called *RA* or *UA*; see Borgonovo and Peccati ((2006), Ridlehoover (2004), Scholtes (2004), and recent textbooks such as Evans and Olson (1998) and Vose (2000).

Note: The term 'Risk Analysis' is also used for the study of financial derivatives (such as American options); see the special track on this topic at recent Winter Simulation Conferences. There is some overlap between this type of analysis and the one I mean; e.g., both types may use Importance Sampling for rare events with major consequences; also see Helton et al. (2006b).

So, even a *deterministic* simulation model generates *random* output (also called response) if the model's input variables or parameters are sampled from a (prior) distribution because their values are uncertain. This uncertainty is called *subjective* or *epistemic*; see Helton et al. (2006b). There are various methods for obtaining subjective distributions based on *expert opinions*. Alternative representations of this uncertainty—such as fuzzy sets and evidence theory—are discussed in Helton et al. (2006a) and Helton et al. (2006b).

Random simulations (such as Discrete-Event Dynamic System or DEDS simulations) have *objective*, *aleatory*, or *inherent* uncertainty; see again Helton et al. (2006b). DEDS simulations represent real systems that without this inherent uncertainty would have a completely different character; e.g., a queueing model without uncertain arrival and service times becomes a scheduling model. Also see Cheng (2006b) and Zouaoui and Wilson (2004).

The parameters and input variables of a simulation model are called 'factors' in the statistical theory on *Design Of Experiments* (DOE). DOE is often used in simulation for

Sensitivity Analysis (SA). DOE gives 'better' estimators of factor effects (i.e., estimators with lower standard errors) than changing one factor at a time does. Moreover, DOE enables estimation of interactions among factors. See Kleijnen (2007).

Note: DOE in simulation is also called "Design and Analysis of Computer Experiments (DACE)' or 'Design and Analysis of Simulation Experiments (DASE)'; see Santner, Williams, and Notz (2003) and Kleijnen (2007) respectively.

RA and SA answer different questions. *SA* answers the question: 'Which are the most important factors in the simulation model of a given real system?'. *RA* answers the question: 'What is the probability of a given event happening (e.g., a nuclear or a financial disaster)?' Kleijnen and Helton (1999) discuss a nuclear waste case-study; also see Helton et al. (2006b). A financial example is the estimation of the 5% quantile of the Net Present Value (NPV) distribution in Fu, Glover, and April (2005). Food safety risks (e.g., foot and mouth disease, terrorist food poisoning, natural disasters such as extreme weather) are discussed in Bourlakis and Weightman (2004). I further discuss the similarities and dissimilarities between SA and RA in Kleijnen (1994) and Kleijnen (1997); also see Oakley and O'Hagan ((2004). SA may help identify those RA inputs that have distributions that are important so they require further refinement; see Helton et al. (2006b).

In Section 2, I discuss frequentist RA in some detail. In Section 3, I briefly discuss Bayesian approaches. In Section 4, I propose that frequentists and Bayesians unite. In Section 5, I summarize some conclusions. A list with many references enables further study of RA.

2 Frequentist Risk Analysis

RA may use the *Monte Carlo* method, which is defined as the method that uses PRNs (in general, PRNs are used to simulate a DEDS or solve a multiple integral; such an integral may arise in physics, mathematical statistics, etc.). More specifically, RA may proceed as follows.

1. RA samples a combination of factor values—also called a 'scenario'—from the joint distribution of possible factor values. (If the factors are assumed to be independent, then this joint distribution is simply the product of the marginal distributions; a simple alternative assumes a multivariate Gaussian distribution; in case of a nonnormal joint distribution, Spearman's correlation coefficient may be used for Latin Hypercube Sampling (LHS) and crude Monte Carlo sampling;see Helton et al. (2006b).)
2. RA uses this combination as input into the (either deterministic or random) simulation model of the real system. RA uses the given simulation model to transform this input into output, which is also called 'propagation of uncertainty'.
3. RA repeats Steps 1 and 2 a number of times (say, 100 or 1000 times) to obtain an Estimated Distribution Function (EDF) of the response of interest.
4. RA uses the EDF resulting from Step 3 to estimate the 'disaster' probability.

Personally I was involved in the following three *applications*, using both SA and RA.

- Kleijnen and Helton (1999) tries to identify the important factors in a large-scale simulation model that was developed at Sandia National Laboratories in Albuquerque, New Mexico (NM). This simulation model is mainly deterministic (because it represents physical processes), but some aleatory randomness is also modeled (namely, human drilling intrusions into the Waste Isolation Pilot Plant, WIPP). The simulation estimates the probability of leakage from the WIPP near Carlsbad, NM. Several performance measures (over a planning horizon of 10,000 years) are considered, in order to obtain permission for building the WIPP. The number of factor combinations is one hundred (the simulation experiment uses LHS, to sample these combinations).
- Van Groenendaal and Kleijnen (2002) performs an environmental investment analysis for the deterministic simulation model of a biogass plant in China. This article applies the same RA and SA as Kleijnen and Helton (1999) does.
- Kleijnen and Gaury (2003) performs a RA using an academic random simulation model of a production line, to estimate the probability of a managerial disaster—for different production pull control systems including a Kanban system.

Note: Robbins, Medeiros, and Dum (2006) investigates the consequences of uncertain arrival rates in call centers, but does not use a formal RA framework.

If the simulation model is *expensive* (i.e. it requires much computer time per run), then RA may sample not the simulation model but its *metamodel* approximation. For example, Giunta et al. (2006) uses crude Monte Carlo, LHS, and orthogonal arrays to sample from specific metamodel types, namely Kriging models and Multivariate Adaptive Regression Splines (MARS). It turns out that the true mean output can be better estimated through sampling many 'cheap' observations from the metamodel; this metamodel is estimated from relatively few observations of the expensive simulation (because that publication estimates an expected value, it does not perform what I call RA).The use of Kriging models and Bayesian Risk Analysis is also briefly discussed in Ŝevĉíková , Raftery, and Waddell (2007).

Popular *software* for RA is 'Crystal Ball' and '@Risk', which are add-ins to Microsoft's Excel spreadsheet software; see the software review in Holland (2005). This software enables crude Monte Carlo and LHS. LHS is also possible through the MATLAB Statistics toolbox subroutine 'lhs' (see Huang et al. 2006), Sandia's DAKOTA software (see Giunta et al. 2006 and http://endo.sandia.gov/DAKOTA), and the European Commission's Joint Research Center (JRC) SIMLAB software (see Saltelli et al. 2004) on http://simlab.jrc.cec.eu.int/.

Note: LHS has no strict mathematical relationship between the number of input combinations (say) n and the number of factors k, whereas in classic designs there is such a relationship; e.g., a two-level fractional factorial design has $n = 2^{k-p}$ with $0 < p < k$; see Kleijnen (2007). For details on LHS and related designs (such as maximin designs and orthogonal arrays) I refer to Helton et al. (2006b), Huang et al. (2006), Jin, Chen, and Sudjianto (2005), Rafajlowicz and Schwabe (2006), Santner, Williams, and Notz (2003), Stinstra(2006), and Yeh, Li, and Sudjianto ((2000).

Note: Helton, Davis, and Johnson (2005) and Helton et al. (2006b) partition their LHS design with sample size $n = 300$ into three subsamples of equal size (namely, 100), to test the 'stability' of RA and SA results. As an alternative, I would suggest *bootstrapping*; i.e., resampling (without replacement) the original 300 observations. A complication, however, is that these 300 observations are not strictly independent in LHS.

3 Bayesian Risk Analysis

Frequentist RA and Bayesian RA resemble each other, since both assume that the parameters of the simulation model are unknown and have specific distributions. However, the Bayesian paradigm is as follows.

1. Bayesians select the prior distributions in a more formal way; e.g., they select a so-called *conjugate prior*.
2. They obtain simulation I/O data (like the frequentists do).
3. They calibrate the metamodel's parameters; i.e., they compute the *posterior* distribution (or likelihood), using the well-known Bayes theorem.

Recent references with many additional references, are Bayarri et al. (2005), Cheng (2006a), Hankin (2005), Ng and Chick (2006), Oakley and O'Hagan (2004), Rajagopal and del Castillo (2005), and Zouaoui and Wilson (2004); also see Kleijnen (2001).

Bayesian model averaging formally accounts—not only for the uncertainty of the input parameters—but also for the uncertainty in the form of the (simulation) model itself; see Chick (2006), Rajagopal and del Castillo (2005), and Zouaoui and Wilson (2004). Also see *Bayesian melding* in Ŝevĉíková , Raftery, and Waddell (2007).

Because I am not a Bayesian, I refrain from further discussion of the Bayesian approach to RA.

4 Frequentists and Bayesians unite!

Sample size determination in Bayesian RA is the focus of Ng and Chick (2006); i.e. that publication focuses on the allocation of the limited sampling budget to the various input parameters that can be better estimated when additional data are collected; also see Chick (2006) and Zouaoui and Wilson (2004). For a classic, frequentist approach to sample size determination I again refer to Helton et al. (2006b).

In general, I think that the Bayesian approach is very interesting—especially from an academic point of view. Practically speaking, however, frequentist RA has been applied many more times (see, e.g., the applications at Sandia). My conclusion is: frequentists and Bayesians should exchange experiences, theories, etc.

5 Conclusions

A fact is that in both deterministic and random simulation models, the exact values of their parameters and input variables are not known. Simulationists may therefore assume that these values have a statistical distribution; i.e., they may apply so-called RA. Frequentists often use either crude Monte Carlo or more refined LHS to sample values from such an input distribution, and feed those values into their simulation model to obtain the EDF of the output, which provides an estimate of the 'disaster' probability that is of interest to the clients. Bayesians, however, use more specific input distributions; e.g., a conjugate prior distribution. They also obtain simulation data, but next they compute the posterior output distribution. Frequentists and Bayesians should unite to exchange experiences and theories—which fits the purpose of this Workshop!

REFERENCES

Bayarri, M.J. et al. (2005), A framework for validation of computer models. Working Paper, National Institute of Statistical Sciences (NISS), Research Triangle Park, North Carolina

Borgonovo, E. and L. Peccati (2006), Uncertainty and global sensitivity analysis in the evaluation of investment projects. *International Journal of Production Economics*, 104, no. 1, pp. 62-73

Bourlakis, M.A. and P.W.H. Weightman, editors (2004), *Food supply chain management*. Blackwell Publishing, Oxford

Chen, V.C.P., K.-L. Tsui, R.R. Barton and J.K. Allen (2003), A review of design and modeling in computer experiments. In: *Handbook of Statistics; Volume 22*, edited by R. Khattree and C.R. Rao, Elsevier, Amsterdam, pp. 231-261

Cheng, R.C.H. (2006a), Resampling methods. *Handbooks in Operations Research and Management Science, Volume 13*, edited by S.G. Henderson and B.L. Nelson, pp. 415-453

Cheng, R.C.H. (2006), Validating and comparing simulation models using resampling. *Journal of Simulation*, in press

Chick, S.E. (2006), Subjective probability and Bayesian methodology. *Handbooks in Operations Research and Management Science, Volume 13*, edited by S.G. Henderson and B.L. Nelson, Elsevier/North Holland, pp. 225-257

Elmaghraby, S.E. (1995), Activity nets: a guided tour through some recent developments. *European Journal of Operational Research*, 82, no. 3, pp. 383-408

Evans, J.R. and D.L. Olson (1998), *Introduction to simulation and risk analysis*. Prentice-Hall, Upper Saddle River, New Jersey

Fu, M.C., F.W. Glover, and J. April (2005), Simulation optimization: a review, new developments, and applications. *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 83-95

Giunta, A.A., McFarland, J. M., Swiler, L.P., and Eldred, M.S. (2006), The promise and peril of uncertainty quantification using response surface approximations, *Structure and Infrastructure Engineering*, 2, nos. 3 - 4, pp. 175-189

Hankin, R.K.S. (2005), Introducing BACCO, an R bundle for Bayesian analysis of computer code output. *Journal of Statistical Software*, 14, no. 16, pp. 1-21

Helton, J.C , F.J. Davis, and J.D. Johnson (2005), A comparison of uncertainty and sensitivity results obtained with random and Latin hypercube sampling. *Reliability Engineering and Systems Safety*, 89, pp. 305-330

Helton, J.C , J.D. Johnson, W.D. Oberkampf, and C.J. Sallaberry (2006a), Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. *Reliability Engineering and Systems Safety*, 91, pp. 1414-1434

Helton, J.C , J.D. Johnson, C.J. Sallaberry, and C.B. Storlie (2006b), Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and Systems Safety*, 91, pp. 1175-1209

Holland, W. (2005), Crystal Ball v7.01 Professional. *OR/MS Today*, 32, no. 2, pp. 54-57

Huang, D., T. T. Allen, W. Notz, and N. Zheng (2006), Global optimization of stochastic black-box systems via sequential Kriging meta-models, *The Journal of Global Optimization*, 34, 441-466

Jin, R, W. Chen, and A. Sudjianto (2005), An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134, no. 1, pp. 268-287

Kleijnen, J.P.C. (1994), Sensitivity analysis versus uncertainty analysis: when to use what? *Predictability and nonlinear modelling in natural sciences and economics,* edited by J. Grasman and G. van Straten, Kluwer, Dordrecht, the Netherlands, pp. 322-333

Kleijnen, J.P.C. (1997), Sensitivity analysis and related analyses: a review of some statistical techniques. *Journal Statistical Computation and Simulation*, 57, nos. 1-4, pp. 111-142

Kleijnen, J.P.C. (2001), Comments on M.C. Kennedy & A. O'Hagan's Bayesian calibration of computer models, *Journal Royal Statistical Society, Series B*, 63, Part 3, 2001, pp. 458-459

Kleijnen, J.P.C. (2007), *DASE: Design and analysis of simulation experiments.* Springer Science + Business Media

Kleijnen, J.P.C. and E.G.A. Gaury (2003), Short-term robustness of production-management systems: a case study. *European Journal of Operational Research*, 148, no. 2, pp. 452-465

Kleijnen, J.P.C. and J. Helton (1999), Statistical analyses of scatter plots to identify important factors in large-scale simulations,1: review and comparison of techniques. *Reliability Engineering and Systems Safety*, 65, no. 2, pp. 147-185

McKay, M.D., R.J. Beckman, and W.J. Conover (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21, no. 2, pp. 239-245 (reprinted in *Technometrics*, 42, no. 1, 2000, pp. 55-61)

Meckesheimer, M., R.R. Barton, T.W. Simpson, and A.J. Booker (2001), Computationally inexpensive metamodel assessment strategies. *AIAA Journal*, 40. no. 10, pp. 2053-2060

Ng, S.H. and S.E. Chick (2006), Reducing parameter uncertainty for stochastic systems. *ACM Transactions on Modeling and Computer Simulation*, pp. 1-24

Oakley, J. and A. O'Hagan (2004), Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal Royal Statistical Society, Series B*, 66, part 3, pp. 751-769

Oden, J.T., Chair (2006), *Revolutionizing engineering science through simulation.* National Science Foundation (NSF), Blue Ribbon Panel on Simulation-Based Engineering Science

Post, J., G. Klaseboer, E.D. Stinstra, and J. Huétink (2004), A DACE study on a three stage metal forming process made of Sandvik Nonoflex, *Proceedings of Numiform 2004*, Ohio, pp. 475-480

Rafajlowicz, E. and R. Schwabe (2006), Halton and Hammersley sequences in multivariate nonparametric regression. *Statistics & Probability Letters*, 76, pp. 803-812

Rajagopal, R. and E. del Castillo (2005), Model-robust process optimization using Bayesian model averaging. *Technometrics*, 47, no. 2, pp. 152-163

Ridlehoover, J. (2004), Applying Monte Carlo simulation and risk analysis to the facility location problem. *The Engineering Economist*, 49, no. 3, pp. 237-252

Robbins, T.R., D. J. Medeiros, and P. Dum (2006), Evaluating arrival rate uncertainty in Call Centers. *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 2180-2187

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto (2004), *Sensitivity analysis in practice; a guide to assessing scientific models.* Wiley, Chichester, England

Santner, T.J., B.J. Williams, and W.I. Notz (2003), *The design and analysis of computer experiments*. Springer-Verlag, New York

Scholtes, S. (2004), Insight 2.0. *OR/MS Today*, 31, no. 1, pp. 42-45

Ševčíková, H., A.E. Raftery, and P.A. Waddell (2007), Assessing uncertainty in urban simulations using Bayesian melding *Transportation Research Part B: Methodological,* 41, pp. 652-669

Simpson, T.W., J. Peplinski, P.N. Koch, and J.K. Allen (2001), Metamodels for computer-based engineering design: survey and recommendation. *Engineering with Computers*, 17, no. 2, pp. 129-150

Stinstra, E.D. (2006), *The meta-model approach for simulation-based design optimization*. Doctoral dissertation, Tilburg University, Tilburg, The Netherlands

Van Groenendaal, W.J.H. and J.P.C. Kleijnen (2002), Deterministic versus stochastic sensitivity analysis in investment problems: an environmental case study. *European Journal of Operational Research*, 141, no. 1, pp. 8-20

Vose, D. (2000), *Risk analysis; a quantitative guide; second edition*. Wiley, Chichester, United Kingdom

Yeh, K., W. Li, and A. Sudjianto (2000), Algorithmic construction of optimal symmetric Latin Hypercube designs, *Journal of Statistical Planning and Inference*, 90, pp. 145-159

Zouaoui, F. and J.R. Wilson (2004), Accounting for input-model and input-parameter uncertainties in simulation. *IIE Transactions***,** 36, pp. 1135-1151

AUTHOR BIOGRAPHY

JACK P.C. KLEIJNEN is professor of ‘simulation and information systems’ at Tilburg University. He is a member of the Department of Information Systems and Management and the Operations Research Group of the Center for Economic Research (CentER). He also teaches at the Eindhoven University of Technology, in the Postgraduate International Program in Logistics Management Systems. He is an ‘external fellow’ of the Mansholt Graduate School of Social Sciences of Wageningen University. His research concerns the statistical design and analysis of simulation experiments, information systems, and supply chains. He has been a consultant for several organizations in the USA and Europe, and serves on many international editorial boards and scientific committees. He spent several years in the USA, at universities and private companies. He received a number of international awards, including the INFORMS Simulation Society’s ‘Lifetime Professional Achievement Award (LPAA)’ of 2005. More information can be found on http://center.uvt.nl/staff/kleijnen/ .