Available online at www.sciencedirect.com
SCIENCE DIRECT
Reliability Engineering and System Safety 91 (2006) 1135–1141
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

A new estimator for sensitivity analysis of model output: An application to the e-business readiness composite indicator

Stefano Tarantola*, Michela Nardo, Michaela Saisana, Debora Gatelli

European Commission[1], Joint Research Centre, Institute for the Protection and Security of the Citizen, Applied Statistics Group, TP 361, Via E. Fermi, 1, Ispra (VA), 21020, Italy

Available online 19 January 2006

Abstract

In this paper we propose and test a generalisation of the method originally proposed by Sobol', and recently extended by Saltelli, to estimate the first-order and total effect sensitivity indices. Exploiting the symmetries and the dualities of the formulas, we obtain additional estimates of first-order and total indices at no extra computational cost. We test the technique on a case study involving the construction of a composite indicator of e-business readiness, which is part of the initiative "e-Readiness of European enterprises" of the European Commission "e-Europe 2005" action plan. The method is used to assess the contribution of uncertainties in (a) the weights of the component indicators and (b) the imputation of missing data on the composite indicator values for several European countries.

Keywords: e-business survey; Robustness analysis; Uncertainty analysis; Sensitivity indices; Weights; Imputation

1. Introduction

In this paper we propose a methodology for sensitivity analysis, which is a generalisation of that proposed in [1] (a review is also offered in [2]) at no extra cost for the analysis. We then test it on a study for the construction of a composite indicator of e-business readiness. Composite indicators are aggregate measures that are calculated as weighted combinations of selected sub-indicators via underlying models of the policy domains of interest. Much discussion surrounds the legitimacy of such indicators. Composites are increasingly used by media and policy makers to communicate information on the situation of countries or regions in various policy fields such as environment, economy or technological development (reviews in [3,4]). Opponents lament that composites are mixes of dubious interpretation yet expensive to obtain. Nevertheless, organisations such as the UN, the OECD and the European Commission make a growing use of such measures. In particular the OECD and the joint research centre (JRC) have recently undertaken the joint preparation of a handbook of good practices of composite indicators building [5].

The e-business readiness composite indicator is aimed at measuring the progress of European enterprises towards a more extensive take up and use of digital technologies. We focus our analysis on the weighting scheme used to aggregate sub-indicators, and on the sensitivity of the composite indicator to different weighting schemes and to incomplete data.

As far as weighting is concerned, JRC suggested and deployed a participatory technique, called "budget allocation", which allows an expert to express his/her opinion upon the relative importance of sub-indicators (see Section 4).

The issue of sensitivity is crucial for the assessment of composites. The communication from the European Commission on structural indicators [6] recognises the importance to assess the sensitivity of the message conveyed by composites with respect to the weights employed. Here we consider an additional source of uncertainty in the evaluation of the composite indicator, the uncertainty due to missing data.

*Corresponding author. Tel.: +39 0332 789928; fax: +39 0332 785733.
E-mail address: stefano.tarantola@jrc.it (S. Tarantola).

[1]The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

0951-8320/$ - see front matter
doi:10.1016/j.ress.2005.11.048

As we shall see in Section 3 we use a *Multiple Imputation* technique (based on Markov Chain Monte Carlo algorithms, henceforth MCMC) for the treatment of missing data. This method is appealing in that it provides confidence bounds for the imputed data [7,8]. Imputed data are, indeed, estimated values. Different imputed data may result in significantly different composite indicator values. Thus, the effect of the imputation on the resulting composite indicator must be acknowledged using both uncertainty and sensitivity analysis.

2. The e-readiness composite indicator

The e-Europe 2005 Action Plan [9] calls for a benchmarking of the target that "by 2005, Europe should have (…) a dynamic e-business environment". It also sets out a number of policy indicators to monitor progress in the implementation of the Action Plan.

One of these benchmarking indicators is the composite indicator on e-business readiness. This is made of two core groups (see Table 1): (a) *Adoption of ICT* by business, and (b) *Use of ICT* by business; each group is composed by six sub-indicators.

The composite indicator, Y_c, for a given country c, is a weighted sum of k sub-indicators X_{ic} (6 available for *Adoption* and 6 for *Use of ICT*) and k weights w_i: $Y_c = \sum_{i=1}^{k} X_{ic} \cdot w_i$. The analysis is conducted using the dataset for the year 2003 (data coverage greater than 92%, see Table 2). Therefore, the first step in our analysis is filling up missing data.

3. Multiple imputation of missing data

3.1. Methodology

The application of the MCMC-based technique for multiple imputation [7,8] permits to substitute the missing values with plausible quasi-random draws from their conditional distribution given the observed data. A set of M (in this case $M = 20$) estimates or imputed values for each of the 7 missing data is generated. The standard deviation of the M values reflects the a priori uncertainty over the true value of each missing number and the mean over the M estimates is used as imputed value (highlighted in grey in Table 2) to complete the dataset. Table 3 reports mean and standard deviation of the imputed values for Germany and the UK. The dataset in Table 2 is the starting point for the calculation of the composite indicator.

4. Selection of weights

A rather common way to assign weights is to involve experts' opinion. In the *budget allocation* method [10], each expert is given a "budget" of 100 points, and is asked to distribute the budget over the sub-indicators by allotting more points to those indicators which are felt as more important. For each sub-indicator, the average weight across the experts is used in the aggregation procedure.

Fourteen experts were involved in the exercise. The sets of weights obtained, for each core group, are given in Table 4. This table shows that experts' opinion varies to a large extent: for instance, indicator a4 (percentage of persons employed using computers in their normal week, at least once a week) ranges from weight 0 (assigned by the Finnish expert) to 15 (assigned by the Portuguese expert).

Contrarily to the common use of average weights, where the information from the single expert vanishes, we believe it is important to retain the identity of the experts and acknowledge, in our model of composite indicator, the uncertainty due to expert selection.

5. Uncertainty analysis

Given the variability of expert selection, and the uncertainty coming from the imputation of the missing data, the composite indicator for the different countries is

Table 1
List of sub-indicators for thee-business readiness composite indicator

Adoption of ICT by business	
a1	% of enterprises that use Internet
a2	% of enterprises that have a Website/home page
a3	% of enterprises that use at least two security facilities at the time of the survey
a4	% of total number of persons employed using computers in their normal work routine (at least once a week)
a5	% of enterprises having a broadband connection to the Internet
a6	% of enterprises with a LAN and using an Intranet or Extranet
Use of ICT by business	
b1	% of enterprises that have purchased products/services' via the internet, electronic data interchange or any other computer-mediated network where these are >1% of total purchases
b2	% of enterprises that have received orders via the internet, electronic data interchange or any other computer-mediated network where these are >1% of total turnover
b3	% of enterprises whose IT systems for managing orders or purchases are linked automatically with other internal IT systems
b4	% enterprises whose IT systems are linked automatically to IT systems of suppliers or customers outside their enterprise group
b5	% of enterprises with Internet access using the internet for banking and financial services
b6	% of enterprises that have sold products to other enterprises via a presence on specialised Internet market places