Expert Systems with Applications 37 (2010) 171–181

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A new data mining approach to estimate causal effects of policy interventions

F. Camillo *, Ida D'Attoma *

Department of Statistical Science, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy

ARTICLE INFO

Keywords:
Selection bias
Program evaluation
Data mining
Conditional space
Matrix decomposition

ABSTRACT

This paper presents a data driven approach that enables one to obtain a measure of comparability between-groups in the presence of observational data.

The main idea lies in the use of the general framework of conditional multiple correspondences analysis as a tool for investigating the dependence relationship between a set of observable categorical covariates X and an assignment-to-treatment indicator variable T, in order to obtain a global measure of comparability between-groups according to their dependence structure. Then, we propose a strategy that enables one to find treatment groups, directly comparable with respect to pre-treatment characteristics, on which estimate local causal effects.

1. Introduction

The availability of information concerning the processes aimed the monitoring of operational activities of bodies, institutions, private and public companies has over the last decades increased. This phenomenon led to the proliferation of semi-automatic control processes which largely rely on the advances made by information technology and on the development of statistical techniques that are peculiar of modern data mining.

On many different fields, new demands arise of an evaluation of the impacts that large-scale actions and policies generate on the various stake-holders, users or managers, involved in the production of relevant goods or services. Reference is made not only to the evaluation of marketing campaigns but also to the assessment of the impact of social or economic policies on the individual citizens or businesses. The modern dataflow within the organizations has turned the monitoring processes into a step of ordinary production process. The verification of the validity of the actions which are developed and implemented becomes part of tools which are available to private or public decision-makers.

The evaluation process is, by definition, a process which takes place following a large-scale action and provides a retrospective view of the events. It is clearly possible-but only in a pre-test setting-to validate a certain action or campaign ex-ante, though this validation cannot be made on large samples. For this reason, precisely because reference is made to social sciences, the pre-test is usually part of qualitative research.

The logic reference framework is one well known in literature as quasi-experiments or observational studies. Data analysis thus refers to the broader frame of causal inference which is useful to estimate causal effects. The closest approach to modern information treatment using information technology and statistics, i.e. data mining, is the one mainly developed by Rubin and known as Potential Outcome Framework (Rosenbaum & Rubin, 1983; Rubin, 1991, 2004, 2005, 2006). In this paper we will not dwell on discussing the cultural proximity between data mining and the potential outcome framework but it will clearly emerge that the similarity of purposes generates major synergies in operational applications. One of the operational aims of modern data mining as a statistical semi-industrial process linked to private and public production organizations is minimizing the researcher's subjective choices during data analysis strategies. For example, in the recent literature on data mining ample room is given to the concept whereby processes such as variable selection, model selection and measurement of model stability have to be dealt with in terms of IT implementation of strategies which are to the largest possible extent semi-automatic. Furthermore, modern data mining is increasingly faced with the need to estimate or allocate the result of more general predictive models to each individual (customers, citizens, points of sale, patients, etc.). As a consequence, the above models will have to be selected among those which best treat data at individual level and which are able to provide business and management intelligence systems with the results of the individual units which make up the reference databases. The latter concept also applies to the evaluation framework, where the effect of an action (a policy or a campaign) should be measured for each treated individual, subsequently releasing this information in the reference database at the level of the individual record of the estimated effect.

2. The Rubin starting point

In literature, there are different approaches to causal inference, but in this work, as starting point, we consider the Potential

* Corresponding authors. Tel.: +39 051 2098251.
E-mail addresses: furio.camillo@unibo.it (F. Camillo), Ida.dattoma2@unibo.it (I. D'Attoma).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.05.072

Outcome Framework, mainly due to the pioneering work of Rosenbaum and Rubin (1983); Rubin (1991, 2004, 2005, 2006), which is the existing approach more close to the data mining point of view.

First, we consider a teaching example to contextualize the fundamental problem tackled within the potential outcome framework.

Then, we will explain how our method tackles the same problem but according to a different point of view.

According to the potential outcome framework the causal effect estimation is considered a *missing data problem* due to the definition of the causal effect as the comparison of the potential outcomes that would have been observed for each unit under different treatment levels.

By considering, for example, a two-levels treatment, for each unit is possible to observe only one of the potential outcomes, because of the difference between the observed and observable.

Being the causal effect for a single individual not observed, the solution is to compute the average causal effect for the whole population or for some interesting sub-groups; but in the latter case, as we will show in the example, the estimation could be biased if the assignment mechanism is not random.

We report an example due to Rubin (2004) in which the author underlies the importance of the missing assignment mechanism in the estimation of the causal effect of interest.

Consider the case in which four patients get assigned to one of two medical operations (A and B). If a unit gets assigned to A then the treatment's level equals 1; and 0 viceversa. The outcome (Y) is defined as the number of years lived after the operation. Denote with $Y(1)$ the potential outcome under the operation A, and with $Y(0)$ the potential outcome under the operation B.

The individual causal effect of interest, $\tau_i = Y(1) - Y(0)$ is defined as the difference between the number of years lived after operation A minus the number of years lived after operation B. Unfortunately, the quantity τ_i could never be directly observed because, after the experiment, we can observe only one of the two potential outcomes.

In the second and third columns of Table 1 cells in bold indicate the observed potential outcomes, while the remaining cells indicate what we cannot directly observe, that is the counterfactual.

If we consider the mean of each potential outcome based on the available information, then it seems that treatment A is better than B, given that patients under operation A will live one year more than patients under operation B.

The achieved conclusion could be wrong, because by considering the counterfactual, the average of the individual causal effect favors the operation A, giving an average benefit of three years.

The misunderstanding is due to the fact that the assignment mechanism is not random, and each potential outcome could be related to a different population.

The use of the potential outcomes for the definition of the causal effects comes before Rubin. In fact, as reported in Rubin (2005), Neyman formalized the potential outcomes for defining individual causal effect already in 1923 within the context of randomized experiment.

Two fundamental aspects make the potential outcome framework innovative and promising with respect to both the definition and measurement of causal effects in observational studies.

First, potential outcomes and covariates are defined as scientific entities (Rubin, 2006), no matter which design researchers use for the analysis: experimental designs, observational studies, or other methods.

Second, a formal probabilistic model is defined: the model that takes into account the selection mechanism, the process that created missing and observed potential outcomes.

Table 2 (Rubin, 2005) formally describes the potential outcome notation. Units (n) denote physical objects. Covariates (X) are variables whose values exist before the unit gets assigned to treatment. In this sense, covariates values are not influenced by the assignment-to-treatment. The observed potential outcome denotes the outcome corresponding to the treatment received.

The framework formally described above requires the SUTVA assumption (Stable Unit Treatment Value Assumption) as labeled in Rubin (1980). The SUTVA assumption implies no interference between units, that is the assignment of unit i to one treatment's level and the potential outcomes must be not dependent on which treatment's level unit j got assigned.

By considering that an assignment mechanism can be thought of as a special type of missing data mechanism that creates missing potential outcomes (Rubin, 1976, 1978), the assignment mechanism is then explicitly formulated. The assignment mechanism explains the reasons of missing and observed potential outcomes, by taking into account a probabilistic model for T (the treatment) given the value of X, $Y(0)$ and $Y(1)$ labeled in Rubin (2006) *the science*.

Given the model for the assignment mechanism:

$$Pr(T|X, Y(1), Y(0)) \tag{1}$$

to understand the importance of the potential outcome framework, it is needed to consider, for example, randomized experiments and some of their important properties for what concerns the assignment mechanism. Within randomized experiments, the probability that a unit gets assigned to a treatment's level is known and independent on potential outcomes.

Further, experimental designs are defined as:

1. Unconfounded
$$Pr(T_i|X_i, Y_i(0), Y_i(1)) = Pr(T_i|X_i) \tag{2}$$
2. Probabilistic
$$0 < e_i < 1 \tag{3}$$
where $e_i \equiv Pr(T_i = 1|X_i)$

If the assignment mechanism is probabilistic and unconfounded, then the assignment mechanism is strongly ignorable.

If the strong ignorability property holds, then the average causal effect estimate will be unbiased and:

$$E[Y_i(0)|T_i = 0] = E[Y_i(0)|T_i = 1] = E[Y_i|T_i = 0] \tag{4}$$

Table 1
An example adapted from Rubin.

Unit	$Y(1)$	$Y(0)$	$\tau_i = Y_i(1) - Y_i(0)$	T
1	1	**6**	−5	0
2	3	**12**	−9	0
3	**9**	8	1	1
4	**11**	10	1	1
Mean	6	9	−3	
Observed mean	10	9	1	

Table 2
The potential outcome notation.

Units	X	Potential outcomes $Y(1)$	Potential outcomes $Y(0)$	Individual causal effects	Summary causal effects
1	X_1	Y_1^1	Y_1^0	$Y_1^1 - Y_1^0$	Y_i^1 vs Y_i^0 for a common set of units
i	X_i	Y_i^1	Y_i^0	$Y_i^1 - Y_i^0$	
n	X_n	Y_n^1	Y_n^0	$Y_n^1 - Y_n^0$	