IVEN VAN MECHELEN

Prediction of a dichotomous criterion variable by means of a logical combination of dichotomous predictors

Mathématiques et sciences humaines, tome 102 (1988), p. 47-54
<http://www.numdam.org/item?id=MSH_1988__102__47_0>

Math. Inf. Sci. hum. (26^e année, n°102, *1988*, p.47-54).

PREDICTION OF A DICHOTOMOUS CRITERION VARIABLE BY MEANS OF A LOGICAL COMBINATION OF DICHOTOMOUS PREDICTORS

Iven VAN MECHELEN [1]

INTRODUCTION

Logical relations between dichotomous variables are of interest in several areas of human sciences. A number of methods have been developed already to detect such relations. For example Lerman, Gras and Rostam (1981) proposed a technique to discover all significant pairwise implications within a set of dichotomous variables. Another example is Van Buggenhaut's (1987) approach to reveal collections of more complex logical relations.

A particular case where logical relations are of use is the situation in which one wants to predict a given criterion by means of a logical combination of predictor variables. In this contribution we propose a method to find, for a given dichotomous criterion C and a given set of n dichotomous potential predictor variables $(P_1,\ldots,P_n)$, a logical

[1] Université de Leuven, Tiensestraat 102, B 3000 Leuven, Belgium.
The author is research assistant of The National Fund for Scientific Research (Belgium).

combination of predictors that is equivalent with the criterion. Four types of logical combination rules will be considered:

1) ordinary disjunction: The criterion event is predicted to take place iff at least one out of a subset consisting of k selected predictor variables has value one. In other words, the criterion event is predicted to take place iff at least one of a set of k each sufficient conditions has been fulfilled. More formally:

$$(P_{i_1} \text{ or } P_{i_2} \text{ or } \ldots \text{ or } P_{i_k}) \Longleftrightarrow C \quad \text{with } \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\}$$

2) generalized disjunction: Both a subset of predictors and a subset of negated predictors are disjunctively combined. More formally:

$$[P_{i_1} \text{ or } \ldots \text{ or } P_{i_k} \text{ or } (\text{not } P_{j_1}) \text{ or } \ldots \text{ or } (\text{not } P_{j_{k'}})] \Longleftrightarrow C$$

$$\text{with } \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\} \text{ and } \{j_1,\ldots,j_{k'}\} \subset \{1,\ldots,n\}$$

3) ordinary conjunction: The criterion phenomenon is predicted to take place if all selected predictors have value one. Stated in other words: the criterion event is predicted to take place iff a set of singly necessary and jointly sufficient conditions have been fulfilled simultaneously. Or more formally:

$$(P_{i_1} \text{ and } P_{i_2} \text{ and } \ldots \text{ and } P_{i_k}) \Longleftrightarrow C \quad \text{with } \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\}$$

4) generalized conjunction: this is a conjunctive combination of both predictors and negated predictors. In other words: the criterion event is predicted to take place iff both a set of necessary conditions and a set of exclusion conditions have been fulfilled. More formally:

$$[P_{i_1} \text{ and } \ldots \text{ and } P_{i_k} \text{ and } (\text{not } P_{j_1}) \text{ and } \ldots \text{ and } (\text{not } P_{j_{k'}})] \Longleftrightarrow C$$

$$\text{with } \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\} \text{ and } \{j_1,\ldots,j_{k'}\} \subset \{1,\ldots,n\}$$

DISCREPANCIES

The equivalence of a criterion with a logical combination of predictors can be read off from a truth table like Table 1. For the example in Table 1 holds that the criterion C is equivalent with the generalized conjunctive combination P_1 and (not P_2).

Table 1
Truth table for hypothetical predictors and criterion

P_1	P_2	P_1 and (not P_2)	C	[P_1 and (not P_2)] $\leftrightarrow$ C
1	0	1	1	1
1	0	1	1	1
1	1	0	0	1
0	0	0	0	1

However, in practice a condition of perfect logical equivalence is seldom met. There are two main reasons for this: (1) the values of predictor and criterion variables can be disturbed by error (e.g. due to various measurement problems) and (2) our models are simplifications of reality; other predictors than the ones considered can play a role and the correct logical combination rule can be more complex than the rule under consideration.

In view of this problem, it is not feasible to evaluate the truth value of empirical equivalence relations in an all-or-none fashion. It makes more sense to accept less than perfect relations and to evaluate relations in terms of their number of discrepancies (i.e. the number of data points that contradict the relation).

In the case of a logical combination of predictors and a criterion there are two types of discrepancies: (1) "false positives" where the combination of predictors has a value of one and the criterion a value of zero, and (2) "false negatives" with the combination of predictors not showing the criterion value of one.

The equivalence relation between the combination of predictors and the criterion consists of two implication relations on which the two types of discrepancies have a differential impact: "False positives" only contradict the implication from the combination of predictors to the criterion, "false negatives" only contradict the inverse implication. Therefore, the two implication relations can be differentially loaded with error. Otherwise, it can be useful to further examine this by computing for each implication relation an index of implicative strength (Lerman, Gras & Rostam, 1981).

ALGORITHM

The aim of our analysis is to find for a given set of predictors, a given criterion and a given combination rule, a logical combination of the predictors that is optimally equivalent with the criterion, that is, a combination with the minimal number of discrepancies.

Probably the only method to detect optimal combinations is the enumerative one. However, even with a small number of predictors this method is not workable. Therefore, a heuristic is proposed here that leads to satisfactory, yet possibly suboptimal solutions. This heuristic is based on Boolean regression. Such a regression consists in a stepwise construction of a Boolean sum of predictors; in each step the predictor that maximally reduces the number of residual discrepancies is added to the Boolean sum of the previous step. -Up till now, such a Boolean regression was mostly not used in its own right, although it is part of the procedures in Boolean factor analysis (Mickey, Mundle & Engelman, 1983) and in hierarchical classes analysis (De Boeck & Rosenberg, in press).-

For the case of an ordinary disjunctive combination rule it is evident that ordinary Boolean regression of the criterion upon the predictors will lead to the optimal combination looked for. For the generalized disjunction, a Boolean regression of the criterion upon both the set of predictors and their negations will do the job.

For the conjunctive combination rule one has to regress the negation of the criterion upon the negation of the predictors. This leads to an optimal disjunctive combination:

$$[(\text{not } P_{i_1}) \text{ or } \ldots \text{ or } (\text{not } P_{i_k})] \Longleftrightarrow (\text{not } C)$$

After negation of both sides this appears to be equivalent with the conjunctive combination looked for:

$$(P_{i_1} \text{ and } \ldots \text{ and } P_{i_k}) \Longleftrightarrow C$$

Finally, for the generalized conjunctive rule the optimization problem is solved by means of a Boolean regression of the negation of the criterion upon both the set of predictors and their negations.

FURTHER OPTIONS

During the Boolean regression the total number of discrepancies between the combination of predictors and the criterion is minimized in a stepwise procedure. As standard option, false positives and false negatives are given an equal weight in this total number of discrepancies. However, the practical importance of the two types of discrepancies might be different depending on the situation. For example, in a situation of personnel selection where predictors are scores on admission tests and the criterion is success in the job, it might be less important to reject candidates that would turn out to be successful (false positives), than to enroll candidates that will fail (false negatives). In such cases it is useful to minimize during the regression a weighted sum of discrepancies with a differential weighting for the two types of discrepancies. For example, a higher weight for the false positives will lead to less false positives and consequently to a raising of the positive predictive power of the final logical combination. (Analogous reasoning holds for the false negatives and the negative predictive power.)

Logically speaking a differential weighting of the two types of discrepancies amounts to a differential weighting of the two implications that are contained in the equivalence relation between the combination of predictors and the criterion. To give a higher weight to false positives, means that one predominantly looks at the implication from the combination of predictors to the criterion; by giving a higher weight to false negatives one can scrutinize the inverse implication.

ILLUSTRATION

A case study was set up in order to illustrate the proposed logical combination method. This study concerns the prediction of political preferences of a subject by means of a logical combination of perceived attributes of the politicians.

The subject for this study was a 25 year old male graduate student. First he was asked to list some important Belgian political figures he was more or less familiar with. Nineteen politicians were chosen: Annemie Neyts, Guy Verhofstadt, Paul Staes, Ludo Dierickx, Frank Swaelen, Leo

Tindemans, Marc Eyskens, Jean-Luc Dehaene, Jean Gol, Gérard Deprez, Wilfried Martens, Charles-Ferdinand Nothomb, José Happart, Jaak Gabriëls, Karel Dillen, Kris Merckx, Karel Van Miert, Willy Claes, Louis Tobback. Then the subject was asked to describe each of these politicians by means of one or more typical attributes. The twelve obtained attributes were: arrogant, realistic, idealistic, patronizing, diligent, eloquent, slimy, sly, diplomatic, willing to listen, experienced, has a quality program. Next, a matrix of politicians by attributes was constructed. The subject was asked to fill in this matrix by assigning to cell (i,j) of the matrix a value of 1 if the i-th politician had the j-th attribute, and a value of 0 otherwise. Finally he was asked to tell about each candidate whether he (she) was acceptable to him as member of a government. Of the nineteen politicians six were judged acceptable. The acceptability was used as a criterion, that we have tried to predict by means of the twelve attributes.

The results of the logical combination analysis show that, for the ordinary as well as for the generalized disjunctive combination rule, one single attribute "diligent" appears as optimal predictive combination (with 21% discrepancies). "Diligent" was set forth by the subject as "gives me the impression to work hard". The positive predictive power is 60% and the negative predictive power 100%. Therefore, "diligence" is a necessary, but not sufficient characteristic for a politician to be preferred by our subject. The negative predictive power of 100% shows that a further disjunctive extension is not possible. On the other hand, with the simple as well as with the generalized conjunctive combination rule, the conjunction of "diligent" and "diplomatic" appears as the optimal combination (with 10.5% discrepancies). The positive predictive power of this combination is 100% and the negative predictive power 86.7%. So "diligent <u>and</u> diplomatic" is a sufficient characteristic of acceptable politicians. Although this combination is not strictly necessary, overall it appears to be the best predictive combination. With regard to the interpretation the further description of "diplomatic" as "flexible in solving conflicts" is useful. It must be concluded that our subject accepts a politician in a government function if he (she) not only works hard but is also skill ful in steering clear of the rocks that, particularly in Belgian politics, are pretty numerous!

BIBLIOGRAPHY

(1) De Boeck, P., Rosenberg, S., "Hierarchical Classes analysis: Model and data analysis", *Psychometrika*, (in press)

(2) Lerman, I.C., Gras, R., Rostam, H., "Elaboration et évaluation d'un indice d'implication pour des données binaires", *Math. Sci. hum.*, 74, (1981), 5-35.

(3) Mickey, M. R., Mundle, P., Engelman, R., "Boolean factor analysis", in W.J. Dixon (Ed.), *BMDP statistical software*, Berkeley, University of California Press, 1983, 538-545.

(4) Van Buggenhaut, J., "Questionnaires booléens: schémas d'implications et degrés de cohésion", *Math. Sci. hum.*, 98, (1987), 9-20.

NOTE

The computer program that was used in the case study is called "Combination Rule Analysis". It was written in Fortran 77. A Microsoft Fortran version for IBM compatible Personal Computers, and a short manual may be obtained from Iven Van Mechelen, Department of Psychology, Tiensestraat 102, B 3000 Leuven, Belgium.