Une règle d'exception en Analyse Statistique Implicative

Régis Gras *, Pascale Kuntz *, Einoshin Suzuki **

*Laboratoire d'Informatique de Nantes Atlantique FRE CNRS 2729
Equipe COD - Connaissances & Décision
Site Ecole Polytechnique de l'Université de Nantes
La Chantrerie BP 60601 44306 Nantes cedex
** Department of Informatics, ISEE, Kyushu University, Japan
regisgra@club-internet.fr , pascale.kuntz@polytech.univ-nantes.fr
suzuki@i.kyushu-u.ac.jp

Résumé. En fouille de règles, certaines situations exceptionnelles défient le bon sens. C'est le cas de la règle R : $a \to c$ et $b \to c$ et (a et b) $\to$ non c. Une telle règle, que nous étudions dans l'article, est appelée règle d'exception. A la suite des travaux précurseurs de E. Suzuki et Y. Kodratoff (1999), qui ont étudié un autre type de règle d'exception, nous cherchons ici à caractériser les conditions d'apparition de la règle R dans le cadre de l'Analyse Statistique Implicative.

1 Introduction

Depuis les travaux de Agrawal et al., (1993) les règles d'association ont été un modèle très utilisé pour extraire des tendances implicatives dans des bases de données. Rappelons que lorsqu'on dispose d'un ensemble E d'individus décrits par p variables $\{a, b, \ldots.\}$, qui peuvent être des conjonctions de variables atomiques et que l'on supposera ici binaires, une règle d'association $a \to b$ signifie que si a est vérifiée alors *généralement* b l'est également. Lorsque l'on extrait un ensemble de telles règles *partielles* d'association, il est pertinent de s'interroger sur les « relations » que ces règles entretiennent entre elles. Cette question a été abordée dans la littérature selon différents points de vue. Dans une optique de structuration de l'ensemble des règles, différentes méthodes de classification ont été proposées (e.g. Lent et al., 1997 ; Gras et Kuntz, 2005). Des représentations visuelles bien adaptées permettent également de mettre en évidence des dépendances entre les règles (e.g. Lehn, 2000 ou Couturier et Gras, 2005).

Si l'on étudie localement avec attention ces relations, on peut découvrir une situation qui défie l'intuition. Supposons que l'on ait, entre trois variables (par exemple, des attributs) a, b et c, conjonction de variables binaires dans l'étude présente et vérifiant $a \to c$ et $b \to c$. Dans des cas exceptionnels, on n'a pas (a et b) $\to c$, comme le bon sens nous le suggère, mais (a et b) $\to$ non c. Cette dernière règle sera appelée ici *règle d'exception.*

Remarquons que des travaux antérieurs (Suzuki et Kodratoff, 1999 ; Suzuki et Zytkow, 2005) considèrent comme situation d'exception la situation suivante :
$a \to c$ (dite règle de **sens commun**), non ($b \to c'$) (dite règle de **référence**) et (*a et b*) $\to c'$ (dite règle d'**exception**) où $c \neq c'$ et où a et b *sont respectivement des conjonctions(a =*