

## Imputation multiple de données manquantes par l'Analyse Statistique Implicative

Pablo Gregori<sup>1</sup>, Josep V. Felip-Bardoll<sup>2</sup> et Régis Gras<sup>3</sup>

<sup>1</sup>Departament de Matemàtiques  
Universitat Jaume I de Castellón  
E-12071 Castelló de la Plana (Spain)

<sup>2</sup>Centre d'Informàtica,  
Conselleria de Sanitat,  
Av. Campanar, 21  
E-46009 València (Spain)

<sup>3</sup>Ecole Polytechnique de l'Université de Nantes, Equipe Connaissance et Décision,  
Laboratoire d'Informatique de Nantes-Atlantique (LINA), UMR 6241,  
La Chantrerie – BP 50609, 44306 Nantes cedex

E-mail: [gregori@mat.uji.es](mailto:gregori@mat.uji.es) , [felip\\_jvibar@gva.es](mailto:felip_jvibar@gva.es) , [regisgra@club-internet.fr](mailto:regisgra@club-internet.fr)

**Résumé.** La plupart des études de recherche appliquée nécessitent des traitements sur des tableaux de données qui présentent fréquemment des données manquantes. Or la majorité des algorithmes statistiques ne travaillent que sur des tableaux de données complets. Donc chaque individu (file) se présentant avec une ou plusieurs lacunes doit être ignoré. L'imputation est une procédure qui consiste à mettre des valeurs dans les cases vides du tableau pour tirer parti des individus incomplets. Gras (2009) propose une méthode originale d'imputation basée sur l'Analyse Statistique Implicative (A.S.I.). Dans cette contribution, on analyse l'applicabilité de cette méthode et on propose l'utilisation de l'imputation multiple pour obtenir des estimations d'intensités d'implication d'un tableau de données obtenues d'une enquête sociologique.

**Abstract.** Most of applied research studies are based on datasheets presenting a fraction of missing data. Statistical algorithms can only be applied to complete datasheets, then every case (row) with one or more gaps must be fully discarded. Imputation of missing data is a step that allows one for taking profit of the information provided by those incomplete cases. Gras (2009) proposes an original data imputation method based on Statistical Implicative Analysis. In this contribution we analyze the applicability of this method and we propose the use multiple imputation, in order to provide estimations on the implication intensities of a data table issued from a sociological survey.