

Commentaires sur l'article de G. Celeux et C. Robert

I. C. Lerman

Université de Rennes 1, IRISA

L'article de G. Celeux et de C. Robert est intéressant à plus d'un titre. Sa conclusion dit tout et -sans doute par prudence- son contraire. Néanmoins, une tendance très marquée plaide en faveur des analyses "continues".

Mon opinion est tout autre. Je pense fondamentalement que -comme on dit en biologie moléculaire- la partie "codante" d'une variable continue est essentiellement discrète; et, cette discrétisation dépend du domaine d'application. Découvrir une "bonne" discrétisation de la variable, filtre de façon significative l'information nécessaire pour une "bonne" analyse des données. On peut alors tenir compte de façon logique et ordinale d'une structure sous-jacente à l'ensemble des modalités de la variable ainsi obtenue. Cependant, tout le problème est de pouvoir déterminer cette "bonne" discrétisation. À cet égard, il n'est pas évident que l'algorithme de Fisher, où d'ailleurs le nombre k de classes est fixé, va nous permettre de découvrir la "bonne" discrétisation, parce qu'il optimise le critère mathématique de l'inertie expliquée.

Nous avons à cet égard une certaine expérience en médecine et en traitement d'image (thèse de Nadia Ghazzali 1992). Pour ce qui est du premier domaine d'application, je voudrais citer un article de J. Y. Lafaye (1979) issu de son travail de thèse (de 3ème cycle), dont l'origine est l'analyse d'une idée brute que je lui avais soumise. Il est à cet égard tout à fait intéressant de considérer la thèse en médecine d'A.M. Kerjean (1978) que cite J. Y. Lafaye dans son article.

Bibliographie

- [1] Ghazzali N. (1992) *Comparaison et réduction d'arbres de classification, en relation avec des problèmes de quantification en imagerie numérique*. Thèse de doctorat de l'Université de Rennes I.
- [2] Kerjean A. M. (1978) *Tentative d'établissement de 100 typologies d'examen biologiques. Contribution à l'établissement du système "ADM"*. Thèse de doctorat d'État, Université de Rennes.
- [3] Lafaye J. Y. (1979) Une méthode de discrétisation de variables continues. *Revue de Statistique Appliquée* 27, 39-53.