

Détection d'agrégats pour données ponctuelles

Lionel Cucala, Université Toulouse I

Mots-clés : Processus ponctuels, détection d'agrégats, espacements.

Résumé :

Dans ce travail, nous nous intéressons à l'analyse de processus ponctuels, temporels ou spatiaux, et plus précisément à la recherche des zones de plus forte intensité des processus, appelées agrégats. Cette problématique apparaît notamment dans le domaine de l'épidémiologie lorsque l'on effectue une veille sanitaire.

Dans un premier temps, nous recensons les techniques de détection d'agrégats temporels, et notamment la plus populaire, la statistique de balayage (Nagarwalla, 1996). Nous montrons comment modifier cette statistique afin d'obtenir un test indépendant de toute hypothèse d'agrégation. La distribution de cette nouvelle statistique, appelée statistique des espacements anormaux, est débattue, et ses performances illustrées sur des jeux de données réels ou simulés.

Ensuite, nous montrons comment adapter cette statistique à des jeux de données spatiaux. Deux voies sont possibles. La première, identique à la généralisation de la statistique de balayage au cas spatial (Kulldorff, 1997), consiste à explorer une famille finie d'agrégats potentiels. La seconde s'appuie sur une transformation des données inspirée par Dematteï *et al.* (2005).

Bibliographie :

Dematteï, C., Molinari, N. and Daurès J.P. (2005). Arbitrarily shaped multiple spatial cluster detection for case event data. A paraître dans *Computational Statistics and Data Analysis*.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics. Theory and Methods*, **6**, 1481-1496.

Nagarwalla, N. (1996). A scan statistic with a variable window. *Statistics in Medicine*, **15**, 845-850.