

空間スキャン統計量を用いた複数クラスタの検出と評価

名古屋大学医学系研究科
Loughborough University

高橋邦彦
島津秀康

疾病の発生状況などを含め、我々の健康に関連するイベントの地域的または時間的な分布や傾向を調査する空間疫学研究において、そのイベントの発生リスクが特定の場所・時間に集積しているかどうか、疾病の集積性を検討することがある。そのための分析方法として、対象空間内における集積（クラスタ）の有無を統計的仮説検定に基づいて評価する疾病集積性の検定（disease clustering test）がいくつか提案されている。特に空間内の各点で観測されたデータからクラスタの位置（地域，時間）を同定し、その統計的有意性を評価する cluster detection test (CDT) として、空間スキャン統計量を用いた circular scan statistic (Kulldorff, 1997) や flexible scan statistic (Tango and Takahashi, 2005) などは、ソフトウェア SaTScan, FleXScan とともに、空間疫学研究の実践において幅広く利用されている。これらの統計量は、対象空間内に単一のクラスタが存在することを仮定した尤度比統計量に基づいて導出されており、原則として一つのクラスタを most likely cluster (MLC) として評価し検出するものになっている。しかし実際には空間内に複数のクラスタが存在する場合も多く、そのため従来法では、単一クラスタの同定を再帰的に適用しながら、尤度比統計量に基づいて、互いに排反な複数クラスタを secondary cluster として順次同定することが行われている。その際の複数クラスタの有意性評価としては、各々が MLC として存在した場合を仮定したもとの、クラスタごとに評価が行われている。

最近我々は、一般化線形モデルによる複数クラスタのモデル化を行い、情報量基準の視点から適切なクラスタ数を選択する複数クラスタの同定法と、それら複数クラスタを全体として統計的有意性の評価を行う新たな複数クラスタ検出の枠組みを提案し、時間集積性への適用を行った (Takahashi and Shimadzu, 2018)。この方法は単一クラスタ検出のためのスキャン統計量を含有する形で導出され、それを一般化して議論できるものとなっており、従来の secondary cluster による検出法よりも精度よく複数クラスタを検出できることが示された。本講演では、これを平面上の集積性検定に拡張した複数クラスタの検出法を検討し、実例への適用といくつかのシミュレーションを通して、その検出の性能について評価を行う。

参考文献

- Kulldorff M (1997). A spatial scan statistic. *Communications in Statistics-Theory and Methods* 26:1481-1496.
- Tango T, Takahashi K (2005). A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 4:11.
- Takahashi K, Shimadzu H (2018). Multiple-cluster detection test for purely temporal disease clustering: Integration of scan statistics and generalized linear models. *PLoS ONE* 13(11): E0207821.