Christophe A.N. Biscio

The accumulative persistence function, a useful functional summary statistic for topological data analysis

Joint with Jesper Møller

Recently, many techniques have been developed in persistent homology, a branch of topological data analysis, with the purpose to recover topological information of a compact manifold from points randomly sampled on it. Examples of such information are the number of connected components, holes, and tunnels of the compact manifolds. However, persistent homology techniques have been very limited applied in spatial statistics although they provide information which is inaccessible by moment based functional summary statistics such as the K-function.

In this talk, we first give a brief introduction on persistent homology, focusing on how to obtain and interpret topological information from a point pattern. Second, to summarise this information we introduce a new functional summary statistic called the accumulative persistence function (APF) and possessing several attractive properties:

- It is a one-dimensional function easier to handle than the two-dimensional functions usually considered in persistence homology;
- for example, confidence regions are easier to plot and interpret for a one-dimensional function than for a two-dimensional function;
- the APF is a natural way of constructing a monotonic function, and this will ease the proof of e.g. convergence theorems.

Third, we apply the APF and the useful concept of extreme rank envelopes to a make a goodness-of-fit test for a point process model, and discuss the associated confidence region for the APF. In a simulation study we test complete spatial randomness in different cases of alternative models and in particular when methods based on moments properties failed as for instance extreme rank envelope methods based on the K function. Finally, we briefly discuss other possible applications of using the APF, including functional boxplots, the confidence region for the mean of APFs, the two sample problem, clustering, and supervised classification.