

Larynx cancer data

Description

The data consist of incidence of larynx cancer diagnosed during 10 years (1982-1991) in the districts of **Mersey** and **West Lancashire** in the north-east of England (larynx-data.txt).

They have detected 876 cases in 144 electoral districts. The study variable is the Standardized morbidity rate, SMR. The SMR is calculated as the ratio between the number of observed cases (O) and the number that would be expected in the general population (E). The interpretation of the coefficient is: if the SMR is equal to 1, the number of cancers detected is equal to that expected, higher than 1 there is a higher number of cancers than expected (excess of risk), and lower than 1 there is a lower number of cancers than expected.

Complementary material:

Map of north-east of England in format Arview software (note to import the map to R package you need that the three files are in the same directory):

NWengland.shp

NWengland.dbf

NWengland.shx

Perform an exploratory analysis to study the possible spatial correlation of the larynx incidence (Standardized Morbidity Rate SMR)

- 1) Build a spatial polygon data frame for the England larynx cancer data.
- 2) Calculate the adjacency matrix, considering neighbors, those regions that share geographic limits and the spatial weights, the standardised to sum unity row.
 - a. Identify the regions with higher and lower number of neighbours
- 3) Calculate the Moran's I and the Geary's C indicators and check the independence of the process.
- 4) Calculate the local Moran indicators and plot the results. Identify clusters of regions and hotspots.