Exercise 1 – Exploring a spatial point pattern

In the years 1974-1983 locations of 974 cases of lung and larynx cancer in Chorley-Ribble, Lancashire, were recorded in the dataset `southlancs`.

(i) Load the `R` package `splancs` and the dataset `southlancs`. The first two columns contain the grid eastings and northings of the cancer cases. The third column contains an indicator, if a lung or a larynx cancer case was observed (lung=0, larynx=1).

(ii) Extract the rows containing cases of larynx cancer from the dataset and plot their locations.

(iii) Estimate the density of the larynx cancer risk surface using a kernel estimator with different appropriate bandwidths (e.g. 1000, 2000). You can use function `kernel2d()` (for more information on this function call the `R` help using `?kernel2d`); the polygon of the study area is contained in the object `southlancs.bdy`.

(iv) Plot the estimated kernel densities using the functions `contour()` or `image()`.

(v) Add the location of an old incinerator (object `old.incinerator`) to the plot, which is suspected to elevate the risk of cancer. What is problematic about this plot?

Exercise 2 – Distance to a waste site, Poisson regression analysis (Ahrens et al.,1999)

In the years 1978-1982 all leukemia cases in upstate New York, which is divided in 790 population districts, were recorded. For each of the 790 districts the location of its centroid, the total population and the number of leukemia cases are available. It is of interest, if the proximity to the Monarch Chemicals waste site, which is in the south of the study region, elevates the risk of leukemia.

(i) Read the dataset `newyork.txt` into an `R` object `newyork` using `read.table()`. The four columns of the dataset contain x- and y-coordinates of the district centroids, total population and number of leukemia cases in each district (`names(newyork)`). Create a plot of the locations of the district centroids and add the Monarch Chemicals waste site, located at $(-0.14, -67.19)$.

(ii) Install the `R` package `fields` and use the function `rdist()` to calculate the Euclidean distance between each district centroid and the waste site.
(*) If you want, you can try to write your own `R` function, which calculates the Euclidean distance between two arbitrary locations and store the Euclidean distance between each district centroid and the waste site in a vector.

(iii) Perform a Poisson regression analysis including the inverse distance between each district centroid and the waste site as covariate. You can use the function `glm()` with option `family="poisson"`; include the `log()` total population of each district as an offset in the model. Interpret the output of the model.

Ahrens et al. (1999): Leukemia clusters in upstate New York: How adding covariates changes the story, *Environmetrics*, **12**, pp. 659-672.

Exercise 3 (*) Mapping shapefiles in `R`

It is possible to read polygons stored in a shapefile into `R` and map them. Here is a small example for a map of 185 Swiss regions.

- Load the library `maptools` and read in the shapefile `switzerland` using `readShapeSpatial()`; assign the `BEZIRKSNR` as `IDvar` for each polygon. Plot the map of the 185 Swiss regions using `plot()`.
- Assgin a blue color to all Swiss lakes. The `IDvar`'s of the lakes are 100, 200, 300, 500, 1000, 1400, 1700, 1900, 2000, 2100, 2200, 2300 and 2400. Plot the map again with a blue color for all lakes.