Assignment 14

Overview

* All calculations must be done using R

Resources

The SpatialEpi package includes a data set of cases of lung cancer in Pennsylvania in 2002. You can load this package and data set:

**library(SpatialEpi)**  
**data(pennLC)**

You can extract a SpatialPolygons layer from this data set:

**penn.map <- pennLC$spatial.polygon  
penn.map <- latlong2grid(penn.map)  
plot(penn.map)**

Note that the structure of this data set is more complex than others. “pennLC” is a “list” of data frames, and pennLC$data is a data frame containing population and case counts of lung cancer for a total of 16 strata levels per county (strata are all combinations of gender; 4 age classes; and race (white and non-white)).

For this problem set, we will ignore strata by focusing on the total population and total number of cases in each county. To get the total population in each county, we sum across strata:

**population <- tapply(pennLC$data$population, pennLC$data$county, sum)**

Similarly, to get the total number of cases of lung cancer in each county, we sum across strata:

**cases <- tapply(pennLC$data$cases, pennLC$data$county, sum)**

Next, we calculate the statewide incidence rate for lung cancer, *r*,  by dividing the total number of cases of lung cancer in the state by the total population in the state:

**r <- sum(cases) / sum(population)**

For each county, we can compute the expected number of lung cancer cases by multiplying the population of each county by the statewide incidence rate:

**expected <- r \* population**

Finally, we compute the *standardized mortality ratio* (SMR) by dividing the number of cases in each county by the expected number of cases in each county. If the rate of lung cancer in a county is equal to the statewide incidence rate, the SMR will be 1. If a county has *more*lung cancer cases than expected, the SMR will be greater than 1. If a county has *fewer* lung cancer cases than expected, the SMR will be less than 1.

**SMR <- as.numeric(cases / expected)**

Questions

1. Create a color patch map where counties with SMR < 0.9 are in blue, and counties with SMR > 1.1 are in red. Describe any hotspots and coldspots of lung cancer.  (3 points)
2. Calculate local Moran’s I on the SMR using queen’s case adjacency. Do you identify any counties with significant spatial autocorrelation with neighboring counties? (4 points)
3. Calculate the Getis-Ord statistic on SMR using queen’s case adjacency. Do you identify any significant hotspots or coldspots in SMR of lung cancer? (4 points)