Lab03: Spatial Structure and Processes

**Handed out:** Wednesday, March 4, 2020

**Return date:** Wednesday, March 25, 2019 (after Spring break)

**Grading:** The first installment of Lab03 counts 13 % towards your final grade.

**Objectives:** This lab explores spatial processes, spatial scale and basic spatial patterns with -functions.

**Format of answer:** Your answers (statistical figures and verbal description) should be submitted as ***hardcopy***. Add a running title with the following information: Lab03, your name and page numbers. You may use this document as template. Copy the requested statistical figures into your document. No trial and error answers are permitted. Label each answer properly with the bold task headings. You are expected to hand in professionally formatted answers: use a fixed pitch font, like **Courier New**, for any  code the use mathematical type-setting when equations are required. Copy and paste figures into your document. Make sure that each figure has a proper ***caption*** describing its content.

## Identification of Spatial Processes [2 points]

Study the measles in Cornwall example in document **CliffOrdChapt01Examples.pdf**, which can be found in the lecture notes.

Task 1: Answer the question “how do the different specifications of the spatial link matrices helped the authors to identify, which the prevailing spatial disease process at each week?”

The author observed the study area under 7 different spatial reference scale, generate plots of the dependent variable against time under each scale respectively. According to the results, you could find which reference scale is dominated at each stage.

## Spatial Scale [4 points]

The -script **Task2SpatialScale.r** shows you how to read point coordinates in a **dBase**-file into **spatstat** and how, for a given number of grid-cells how to calculate the *Variance-Mean-Ratio*. You will investigate in this task how an interpretation of a point pattern changes in dependence of the selected spatial scale.

Note: There is no need to show the map of the grid counts and the -test.

Task 2: For the point pattern in the file **HomoCSR.dbf** explore in dependence of the selected reference grid at either , , , or cells how the *Variance-Mean-Ratio* changes. Explain why at some spatial scales we observe a particular spatial dependence. [2 points]

Those points generated based on homogeneous surface, which means the expectation of the number of points is equivalent and disturbance follows the normal distribution. When the pixel size goes larger, the expectation of the number of points also increases. But the difference among disturbances barely rises. That is why VMR would decrease when the observed scale increase. When the cells size goes extreme large, we would find all of them have the same number of points, variance becomes 0, which indicates an extreme regular spaced pattern (perfectly spatial dependent).

|  |  |
| --- | --- |
| VMR 16: 1.094118 | VMR 8: 0.8888889 |
| VMR 4 : 0.3125 | VMR 2 : 0 |

Task 3: For the point pattern in the file **CSRClust.dbf** explore in dependence of the selected reference grid at either , , , or cells how the *Variance-Mean-Ratio* changes. Explain why at some spatial scales we observe a particular spatial dependence. [2 points]

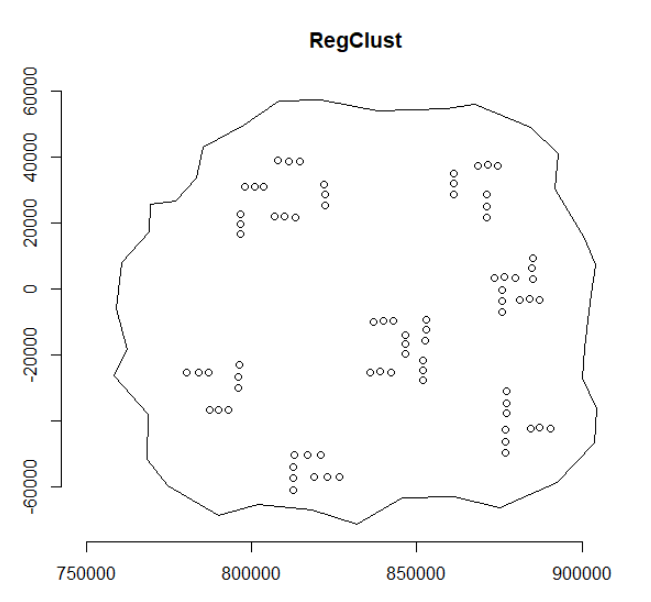
Those points generated based on clustered surface, which means the expectation of the number of points is not equivalent. That's why we find clusters when we have a high spatial resolution. Under this situation, VMR would larger than 1. When the pixel size goes coarser, the VMR value becomes fluctuated. Under a clustered distribution, it is hard to tell how the expectation and variance would change. it totally dependents on whether our reference boundary conflicts with the original cluster boundary.

|  |  |
| --- | --- |
| VMR 16 : 2.917983 | VMR 8 : 1.189673 |
| VMR 4 : 1.348379 | VMR 2 : 0.7963684 |

## Systematic Spatial Point Pattern [4 points]

Evaluate the -function and its confidence envelop for two systematic spatial patterns. Carefully interpret the -function in relation to the inter-event distance . The shape files, including a boundary file for the study area, can be found in the folder **SystematicPattern.zip**.

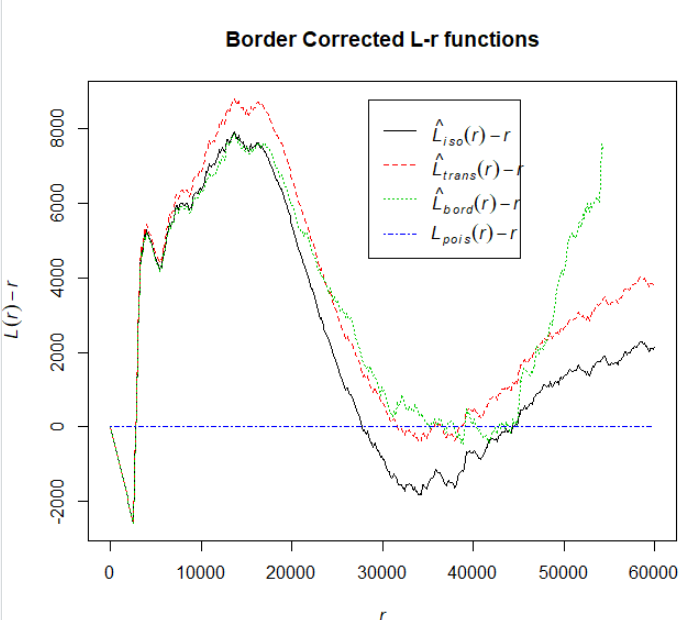
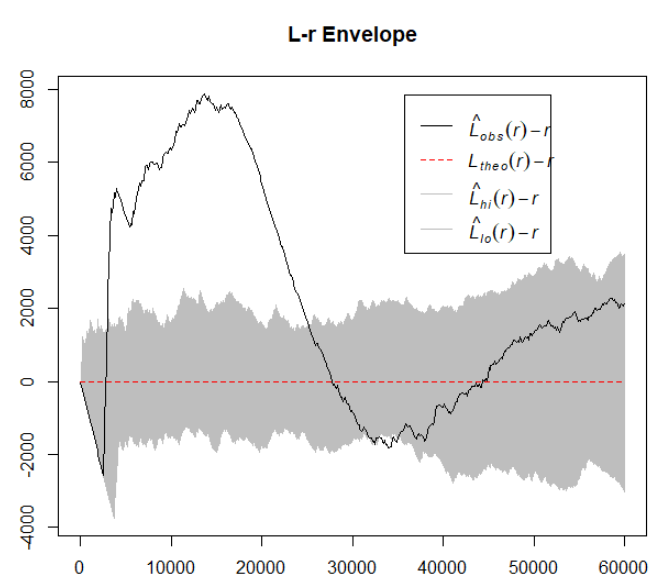
Task 4: Discuss the pattern in the shapefile **ClustReg.shp** and interpret its associate -function in relation to the inter-event distance . Also evaluate its significance in relation to with the confidence envelop simulated under CSR. [2 points]



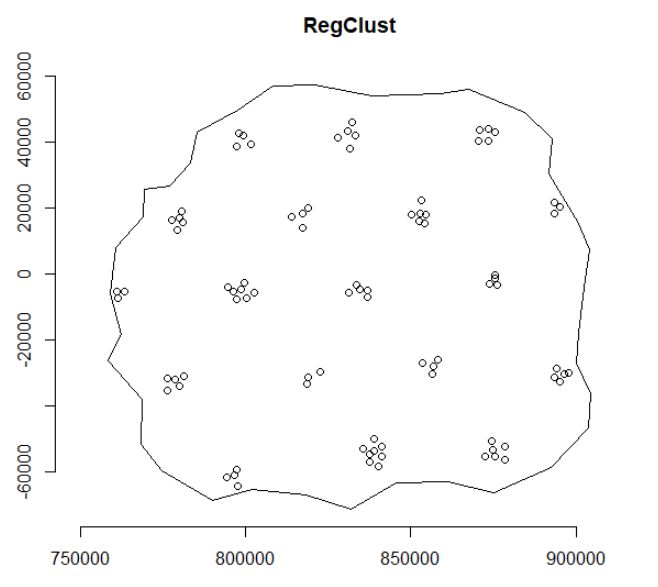
A quick view of the data, we may find some obvious clusters in the study area.

In the left figure, the blue-dashed line gives the theoretical Poisson distribution. The black line is the L-r function. When the reference radius size gets very smaller (below 3000), the backline is below the blue line, which indicates regular patterns are observed. As the r-value increased, the backline becomes higher than the blue line, under this situation, indicates that there is clustering present in the data.

The right figure almost the same with the left one, but use a data generating process to get multiple simulated patterns. The dark shaded area indicates the 95% confidence interval of fluctuation of l-r value departs from the theoretical Poisson distribution. When the dark line below or beyond that dark shade area, indicates we observed a significant regular or cluster pattern in that distance.

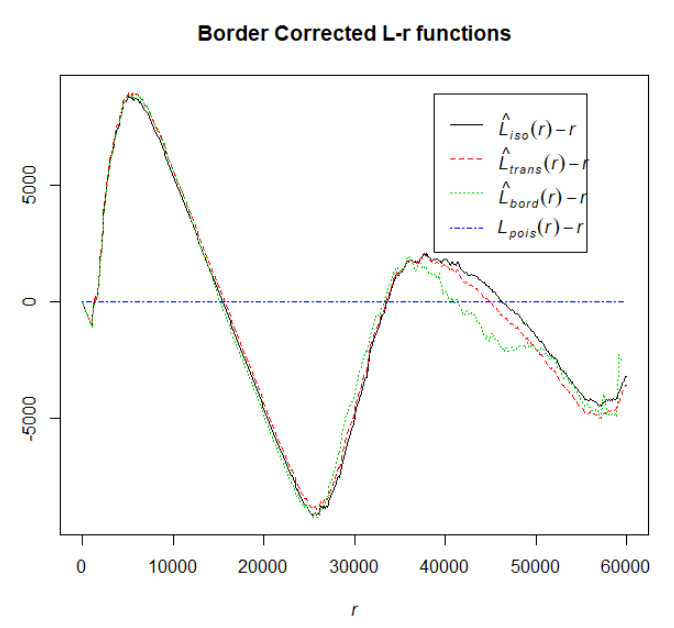
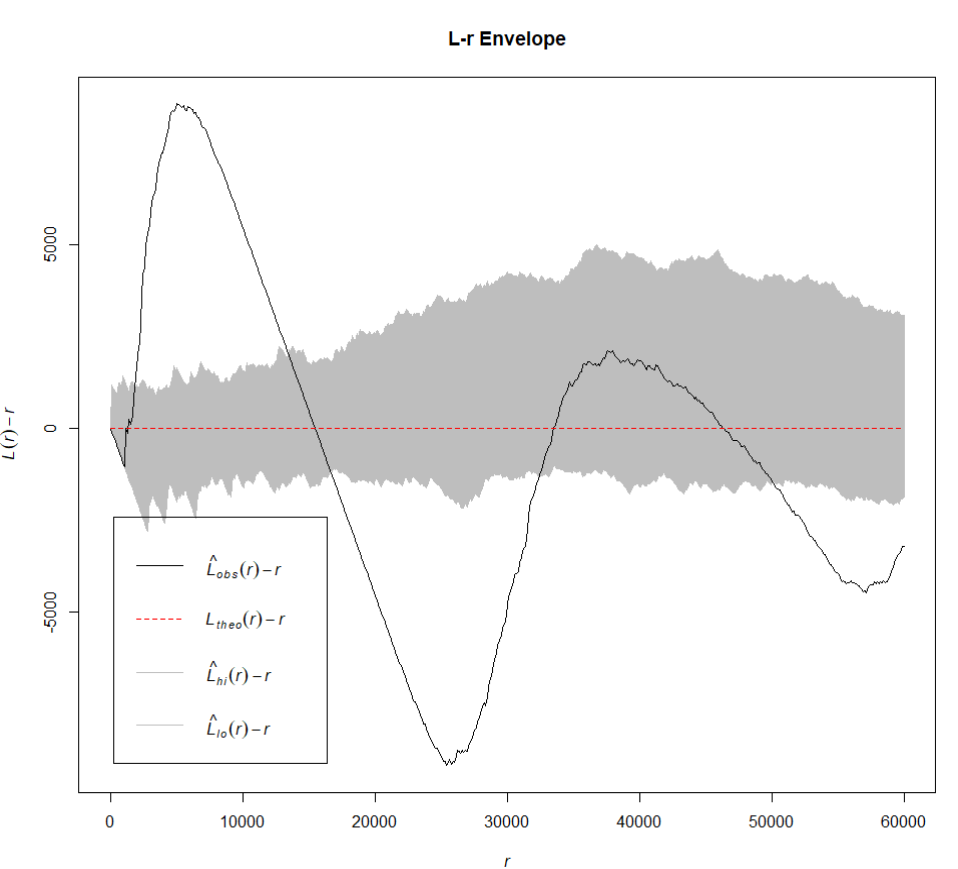
 

Task 5: Discuss the pattern in the shapefile **RegClust.shp** and interpret its associate -function in relation to the inter-event distance . Also evaluate its significance in relation to with the confidence envelop simulated under CSR. [2 points]



A quick view of the data, we may also find some obvious clusters but seems evenly distributed in the study area.

Same to the previous one, the blue-dashed line gives the theoretical Poisson distribution. When the reference radius size smaller than 1500, the black line is above the blue line, which means clustered patterns are observed. But only the radius between approximately 3000 to 12000, the black line is above the dark shade area, which means those clusters are significant. Same principle for the part below the black line.

## Heterogeneous Poisson Process [3 points]

Study exercise 5.6 the document **WallerChapt05Exercises.pdf**. Write an -script without the help of the **spatstat** library that simulates an inhomogeneous CSR process of 100 random points in a square 0 to 1 by 0 to 1 window. The underlying intensity surface is defined by , that is in the lower left corner and in the upper right corner of the square window.

Task 6: Show your -script. [2 points]

## Intensity function

lambda <- function(x,y) return((x+y)/2) # scale to probability

## define matrix of sample points

ppts <- matrix(NA, nrow=100, ncol=2)

icount <- 1 # counter of accepted sample points

while (icount <= 100){

ppt <- runif(2,min=0, max=1) # generate proposal point

if (runif(1,min=0, max=1) <= lambda(ppt[1],ppt[2])){ # if TRUE => accept point

ppts[icount,] <- ppt

icount <- icount+1

} # end::if

} #end::while

Task 7: Show a plot of your resulting point pattern and interpret it. [2 points]

From both the generation algorithm of the points and results, we could find that the right upper corner has a higher intensity than the lower-left corner. It means they are not spatially random.

