Assessment\_4\_STA5001\_Michael\_Le\_21689299

2024-04-30

#Question 1. #Simulate a realization of the inhomogeneous Poisson process at [0,5]x[0,5] that exhibits all of the following properties:

#(a).points form one group;

#(b).the center of the group has a random location;

#(c).all the points of this group have an x coordinate that is within a distance of 1 from the x coordinate of the group’s center;

#y-coordinates of the points in this group uniformly take on all possible values within the range of [0,5].

#Repeat simulations twice and produce two plots of simulated points.

library(spatstat)

## Warning: package 'spatstat' was built under R version 4.3.3

## Loading required package: spatstat.data

## Warning: package 'spatstat.data' was built under R version 4.3.3

## Loading required package: spatstat.geom

## Warning: package 'spatstat.geom' was built under R version 4.3.3

## spatstat.geom 3.2-9

## Loading required package: spatstat.random

## Warning: package 'spatstat.random' was built under R version 4.3.3

## spatstat.random 3.2-3

## Loading required package: spatstat.explore

## Warning: package 'spatstat.explore' was built under R version 4.3.3

## Loading required package: nlme

## spatstat.explore 3.2-6

## Loading required package: spatstat.model

## Warning: package 'spatstat.model' was built under R version 4.3.3

## Loading required package: rpart

## spatstat.model 3.2-10

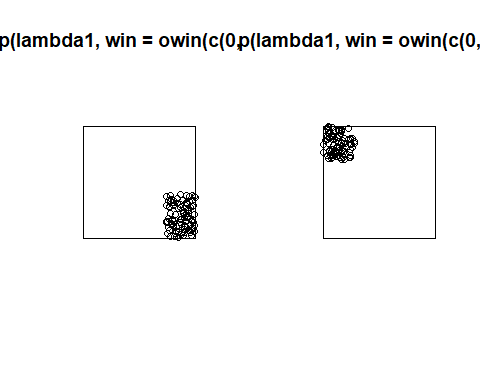
## Loading required package: spatstat.linnet

## Warning: package 'spatstat.linnet' was built under R version 4.3.3

## spatstat.linnet 3.1-4

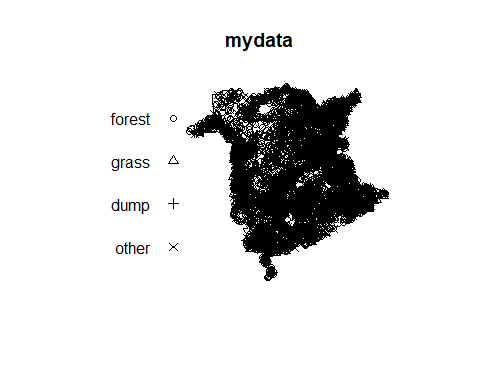
##   
## spatstat 3.0-7   
## For an introduction to spatstat, type 'beginner'

par(mfrow=c(1,2))  
  
  
#First simulation  
  
a<-runif(4,min = 0, max = 5)  
  
lambda1 <- function(x, y) {50\*as.numeric((abs(x -a[1])<1) & (abs(y -a[2])<1))}  
plot(rpoispp(lambda1, win=owin(c(0,5),c(0,5))))  
  
  
#Second simulation with two plots of simulated points  
a<-runif(4,min = 0, max = 5)  
  
lambda1 <- function(x, y) {50\*as.numeric((abs(x -a[1])<1) & (abs(y -a[2])<1))}  
plot(rpoispp(lambda1, win=owin(c(0,5),c(0,5))))



#Question 2.

mydata <- unmark(nbfires)  
marks(mydata) <- nbfires$marks$fire.type  
  
#Plot the data  
plot(mydata)

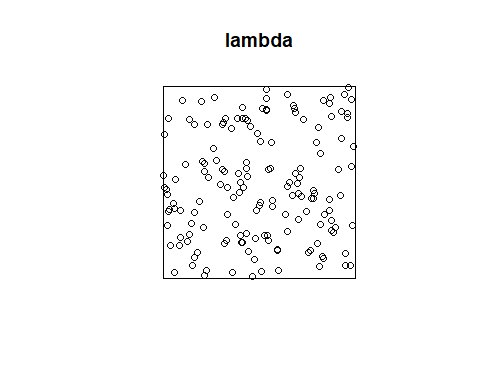


#Assume stationary. What is the value of the estimated constant intensity?   
summary(mydata)$intensity

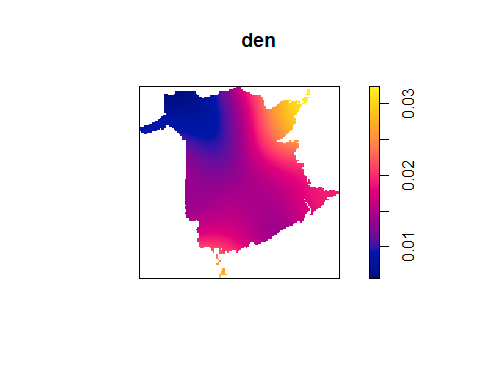
## [1] 0.01572195

#0.01572195

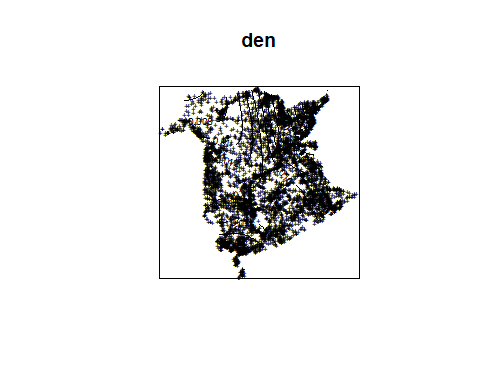
#Simulate and plot a Poisson process with the estimated constant intensity multiplied by 10000.  
lambda<-rpoispp(10000\*summary(mydata)$intensity)  
plot(lambda)



#Estimate and plot non-constant intensity.   
  
den<-density(mydata)  
plot(den)



#Plot the data points and a contour plot for the estimated non-constant intensity in the same figure  
contour(den)   
points(mydata, cex = 0.5, pch = "+")

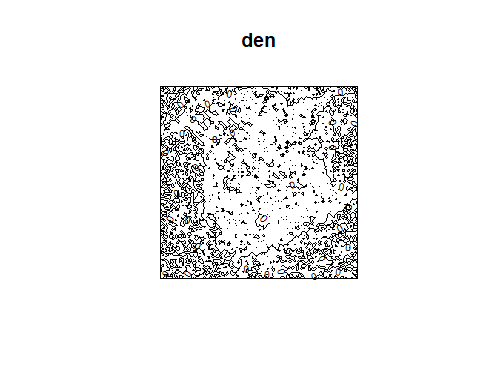


#Alternative solution,  
x<-mydata$x  
y<-mydata$y  
myppp<- ppp(mydata$x,mydata$y,c(13.19402,988.46830),c(5.489932,956.133395))

## Warning: 2 points were rejected as lying outside the specified window

## Warning: data contain duplicated points

den <- density(myppp, sigma = 1)  
contour(den)



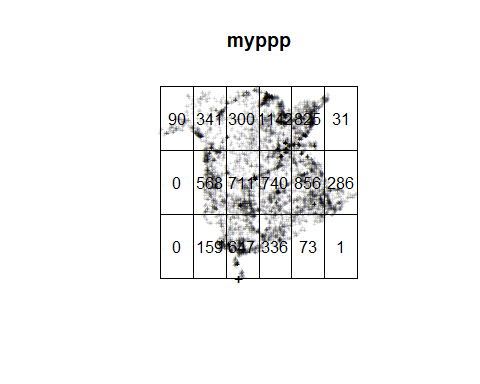
#Use quadrat counts to investigate the intensity. Plot the data points and quadrat counts in the same figure.   
  
quadratcount(myppp, nx = 6, ny = 3)

## x  
## y [13.2,176) [176,338) [338,501) [501,663) [663,826) [826,988]  
## [639,956] 90 341 300 1142 825 31  
## [322,639) 0 568 711 740 856 286  
## [5.49,322) 0 159 647 336 73 1

Q <- quadratcount(myppp, nx = 6, ny = 3)  
plot(myppp, cex = 0.5, pch = "+")

## Warning in plot.ppp(myppp, cex = 0.5, pch = "+"): 2 illegal points also plotted

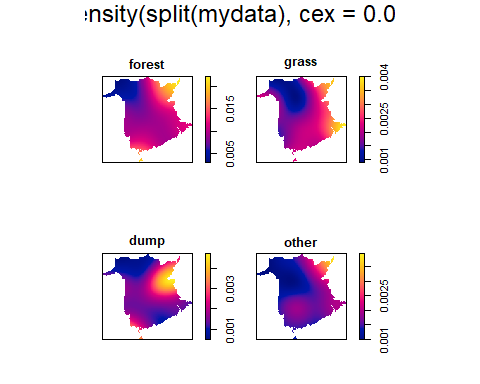
plot(Q, add = TRUE, cex = 1)



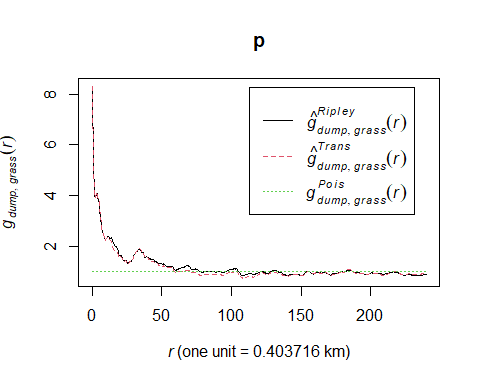
#Elaborate on the conclusions drawn from the previous questions:

#Explanation: #After computing all the results from above, #we find out the intensity of the estimated constant density is 0.01572195. After plotting ipp, where the intensity can vary over space as a function of the position. Conclude that the intensity may be non-stationary. Which corresponds intensity measure values in each subregion.

#Separate the data into the sub-patterns of points by types  
#and plot their intensities.  
mydata <- unmark(nbfires)  
marks(mydata) <- nbfires$marks$fire.type  
plot(density(split(mydata),cex = 0.05))

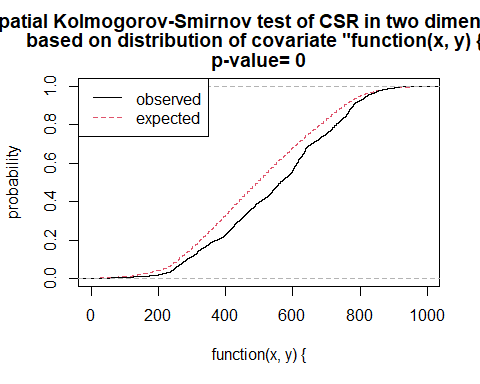


#Plot the cross-type pair correlation function for "dump" and "grass" marks. Interpret the plot.  
  
p <- pcfcross(mydata, "dump", "grass")  
plot(p)



#The plot suggest that there is a inhibition between dump and grass at all scales except very small distances.

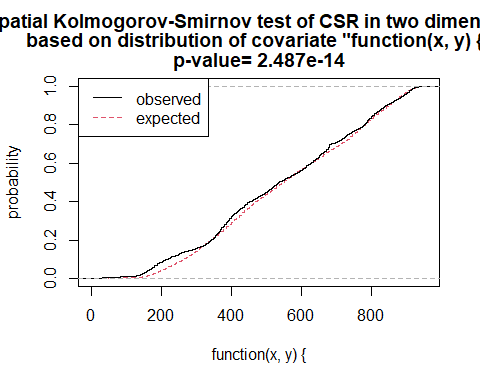
#Remove the marks from the point pattern and perform the spatial Kolmogorov-Smirnov test for the uniform distribution of the x coordinate.  
  
  
KS <- cdf.test(mydata, function(x, y) {x})  
plot(KS)



pval <- KS$p.value  
pval

## [1] 0

#Perform the spatial Kolmogorov-Smirnov test for the uniform distribution of the y coordinate.  
  
KS <- cdf.test(mydata, function(x, y) {y})  
plot(KS)



pval <- KS$p.value  
pval

## [1] 2.4869e-14

#Explanation: For the homogenous Poisson process assuming the intensity is constant, #The p-value for the x-value is 0. Thus, the test does reject the #hypothesis of CSR.The plot demonstrates functions for the observation and expected distribution functions that confirms it.Similarly for the y-value, the p-value is 3.22e-14. We reject the hypothesis of CSR as well, we can conclude that this pattern is completely spatially random.

#Question 3a.

library(spatstat)   
library(lubridate)

## Warning: package 'lubridate' was built under R version 4.3.3

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

library(stpp)

## Warning: package 'stpp' was built under R version 4.3.3

## Loading required package: rpanel

## Warning: package 'rpanel' was built under R version 4.3.3

## Loading required package: tcltk

## Package `rpanel', version 1.1-5: type help(rpanel) for summary information

## Loading required package: splancs

## Warning: package 'splancs' was built under R version 4.3.3

## Loading required package: sp

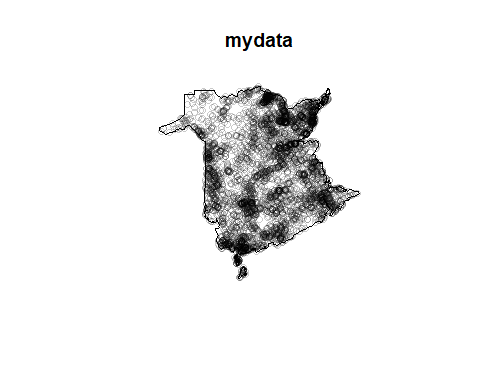
## Warning: package 'sp' was built under R version 4.3.3

##   
## Spatial Point Pattern Analysis Code in S-Plus  
##   
## Version 2 - Spatial and Space-Time analysis

data(nbfires)  
str(nbfires)

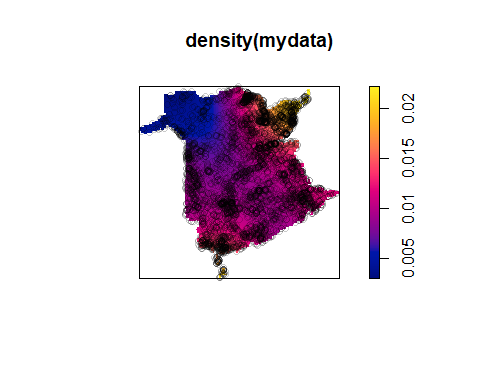
## List of 6  
## $ window :List of 5  
## ..$ type : chr "polygonal"  
## ..$ xrange: num [1:2] 0 1000  
## ..$ yrange: num [1:2] 0 959  
## ..$ bdry :List of 6  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:500] 412 415 415 416 419 ...  
## .. .. ..$ y: num [1:500] 123 124 124 123 122 ...  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:54] 835 834 836 837 838 ...  
## .. .. ..$ y: num [1:54] 910 912 913 914 916 ...  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:92] 823 826 828 828 833 ...  
## .. .. ..$ y: num [1:92] 865 867 869 869 874 ...  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:79] 385 384 384 385 387 ...  
## .. .. ..$ y: num [1:79] 90.4 91.7 91.8 93.7 95.9 ...  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:66] 395 395 396 397 395 ...  
## .. .. ..$ y: num [1:66] 64.8 64.8 65.9 67.9 70.3 ...  
## .. ..$ :List of 2  
## .. .. ..$ x: num [1:80] 403 404 407 409 411 ...  
## .. .. ..$ y: num [1:80] 0.417 3.753 6.103 9.212 13.578 ...  
## ..$ units :List of 3  
## .. ..$ singular : chr "km"  
## .. ..$ plural : chr "km"  
## .. ..$ multiplier: num 0.404  
## .. ..- attr(\*, "class")= chr "unitname"  
## ..- attr(\*, "class")= chr "owin"  
## $ n : int 7108  
## $ x : num [1:7108] 762 654 633 598 639 ...  
## $ y : num [1:7108] 778 671 625 648 611 ...  
## $ markformat: chr "dataframe"  
## $ marks :'data.frame': 7108 obs. of 9 variables:  
## ..$ year : Factor w/ 16 levels "1987","1989",..: 13 13 13 13 13 13 13 13 13 13 ...  
## ..$ fire.type : Factor w/ 4 levels "forest","grass",..: 1 1 2 1 2 1 1 1 2 1 ...  
## ..$ dis.date : POSIXct[1:7108], format: "2000-05-04 19:15:00" "2000-04-18 14:00:00" ...  
## ..$ dis.julian: num [1:7108] 125 109 124 124 110 ...  
## ..$ out.date : POSIXct[1:7108], format: "2000-05-05 12:00:00" "2000-04-18 19:00:00" ...  
## ..$ out.julian: num [1:7108] 125 109 124 124 110 ...  
## ..$ cause : Factor w/ 9 levels "unknown","rrds",..: 8 3 8 1 1 8 6 6 6 8 ...  
## ..$ ign.src : Factor w/ 16 levels "unknown","cigs",..: 4 2 3 1 1 5 1 1 1 6 ...  
## ..$ fnl.size : num [1:7108] 1.6 7 1 0.1 0.5 2 0.5 1.5 0.5 1.2 ...  
## - attr(\*, "class")= chr "ppp"

mydata <- nbfires[nbfires$marks$fire.type =="forest"]  
mydata <- unmark(mydata)  
  
  
#mydata <- unmark(nbfires)  
#marks(mydata) <- nbfires$marks$fire.type  
  
#Plot the data  
plot(mydata)

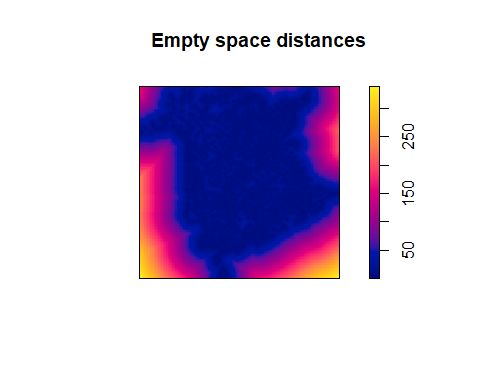


#Explanation: The dataset gives 4627 positions containing forest type fire #in a 1000 by 958.9142 meter rectangular sampling region in the Brunswick forest #fires that is occuring for the years between 1987 and 2003. Where many occured from while there is a less on the top-left corner.

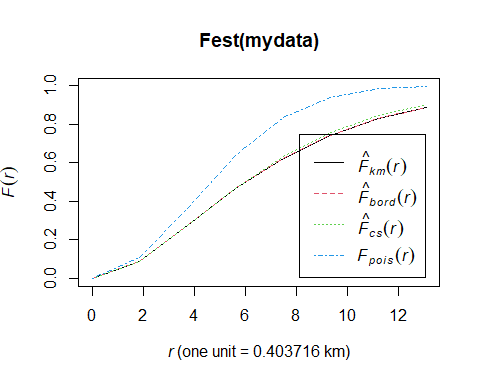
#Produce density plots.  
  
plot(density(mydata))  
plot(mydata,add=TRUE,cex=1)



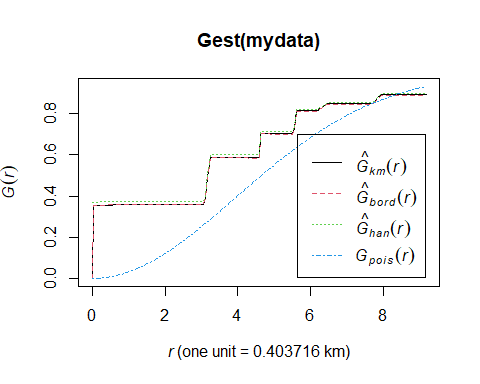
#Plot the image of empty space distances for the locations of fires.  
  
emp <- distmap(mydata)  
plot(emp, main = "Empty space distances")



#Compute and plot the F function for the locations of forest fires. Explain what kind of pattern for this data the function suggests.  
plot(Fest(mydata))

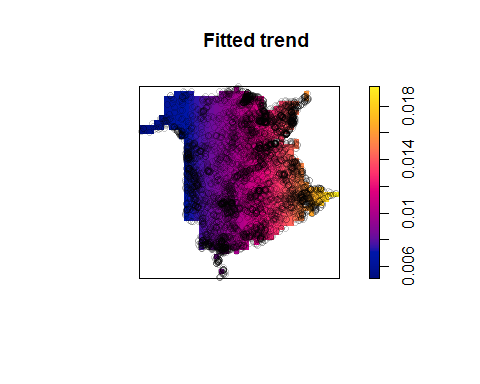


#Compute and plot the G function for the locations of forest fires. Explain what kind of pattern for this data the function suggests.   
plot(Gest(mydata))



#Both functions suggest clustering behaviour as the theoretical curve G is under empirical estimators, similarly for F.

#Fit an in-homogeneous Poisson model with an intensity that is a linear function of x and y coordinates. Plot the corresponding trend and data locations in the same image.   
fit1<-ppm(mydata,~x)   
plot(fit1, how = "image", se = FALSE)



#Question 3b.

mytime <- nbfires[nbfires$marks$fire.type =="forest"]$marks$dis.date  
time1 <- round\_date(ymd\_hms("1987-01-01 00:00:00"), unit = "hour")  
Date\_h <- round\_date(ymd\_hms(mytime), unit = "hour")

## Warning: 1 failed to parse.

#Calculate a vector that represents the differences between the incident times and the initial time moment. Display the first few values of this vector.   
hours\_diff <- as.numeric(difftime(Date\_h, time1, units = "hours"))   
head(hours\_diff)

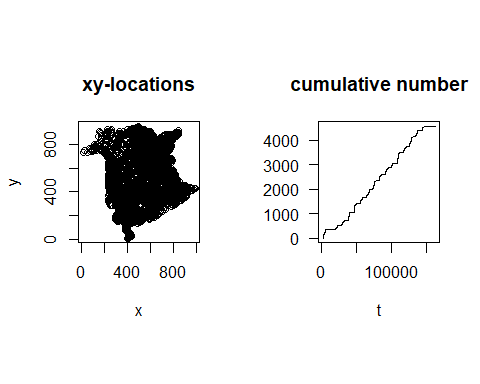
## [1] 116947 116558 116920 116292 116342 116342

#Create a data frame, using the vector of time differences and the x and y coordinates extracted from mydata. Remove incomplete cases from it.   
df1 <- data.frame( x = mydata$x, y = mydata$y, t = hours\_diff )   
df1 <- df1[complete.cases(df1), ]

#Create an 'stpp' object using this data frame and generate a static image that shows the spatial locations and a plot with a cumulative number of cases over time.   
X1 <- as.3dpoints(df1)   
str(X1)

## 'stpp' num [1:4558, 1:3] 274 239 678 269 324 ...  
## - attr(\*, "dimnames")=List of 2  
## ..$ : NULL  
## ..$ : chr [1:3] "x" "y" "t"

plot(X1)



#Run animation and show the obtained final plot  
#NOTE: break the runtime, if case the loop is not finite.  
dev.new()  
animation(X1, runtime = 20)   
dev.off()

## png   
## 2

#Explaination: This patterns shows stationary behavior. Less points occurred at the top-left sub-regions and the appears most frequently in central and lower sub-regions. Overall time rate of events looks constant, but at smaller time resolution it seems exhibit some periodic behaviour

NOTE: [Assignment-4 Solution (pdf) - CliffsNotes](https://www.cliffsnotes.com/study-notes/7138731) (Use this as reference point)