Lab03: Spatial Structure and Processes

## Identification of Spatial Processes [2 points]

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 authors created seven different theoretical graphs to model the measles outbreak under various assumptions about how the disease spreads. They then measured the correspondence (as the z-score of the outbreak/no outbreak statistic) between each of those graphs and the actual spread data over the course of the outbreak to see what patterns were more accurate at different times. For example, graphs 2, 4, 5, and 6 were most accurate in the early stages, while graphs 1, 3, and 7 were most accurate after the peak of the outbreak. From this they concluded that diffusion based on population size/density was most important for explaining the early development of the outbreak while more local-scale spread was more relevant in the later stages.

## Spatial Scale [4 points]

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]

pat <- ppp(HomoCSR$x, HomoCSR$y, xrange=c(0,16), yrange=c(0,16))

> for (cells in c(16,8,4,2)) {

qc <- quadratcount(pat, nx=cells, ny=cells)

vqc <- as.vector(qc)

cat("VMR", cells, ": ",var(vqc)/mean(vqc), "\n", sep="")

}

VMR16: 1.094118

VMR8: 0.8888889

VMR4: 0.3125

VMR2: 0

The points seem to have been generated with a uniform intensity throughout the study area. Therefore, the cell point counts will tend to converge toward the mean as the number of cells decreases, causing the variance to decrease. In this particular case, the VMR is 0 with a 2x2 grid of cells because all four cells have exactly the same number of points. A large number of smaller cells results in a larger variance and therefore larger VMR because at that scale the random arrangement of points causes greater variability in point counts per cell.

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]

> pat2 <- ppp(CSRClust$x, CSRClust$y, xrange=c(0,16), yrange=c(0,16))

> for (cells in c(16,8,4,2)) {

qc <- quadratcount(pat2, nx=cells, ny=cells)

vqc <- as.vector(qc)

cat("VMR", cells, ": ",var(vqc)/mean(vqc), "\n", sep="")

}

VMR16: 2.917983

VMR8: 1.189673

VMR4: 1.348379

VMR2: 0.7963684

This pattern visually appears to be more clustered than the pattern in Task 2, and that is reflected in the greater VMR values at all four spatial scales. However, it follows the same pattern of decreasing VMR as the cell size increases because of the underlying uniform point generation intensity and the fact that larger cells are better able to capture similar numbers of points even when there is pronounced clustering at finer spatial scales.

## Systematic Spatial Point Pattern [4 points]

Task 4: Discuss the pattern in the shape-file **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]

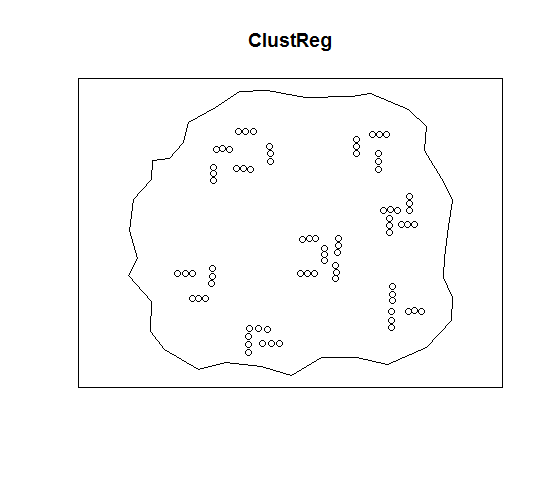
> clustRegPPP <- as.ppp(clustRegUTM)

> clustRegPPP <- clustRegPPP[win]

> clustRegPPP$marks <- NULL

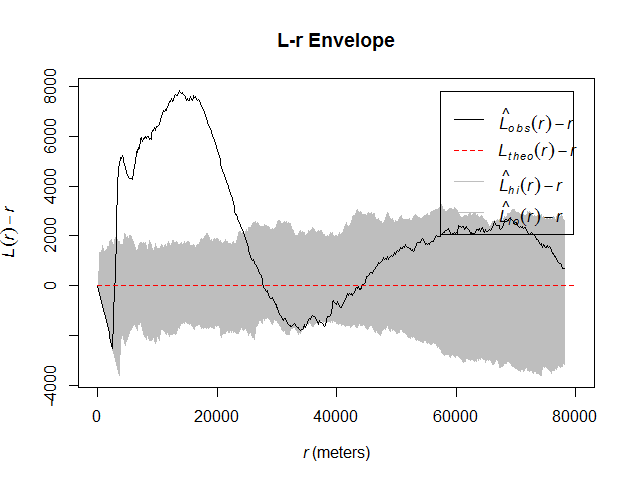
> plot(clustRegPPP, add=F, main="ClustReg")

> box()



> Lv <- envelope(clustRegPPP, fun=Lest, rmax=80000)

> plot(Lv, .-r ~ r, main="L-r Envelope")



The ClustReg point pattern is composed of several visually distinct clusters and each cluster contains 3–5 sub-clusters of three points arranged in a row. Based on the -function, the points are more dispersed than would be expected under CSR below a distance of about 3000 meters. Each point is separated from its closest neighbor(s) by a fairly regular distance of about 3000 m, so points are guaranteed to have 0 neighbors when the radius is less than 3000 m, in contrast to a CSR pattern which could produce points closer together than 3000 m. The points are more clustered than expected from around 3000 to 25000 m. The diameter of each cluster is roughly 25000 m, which explains why points have more than the expected number of neighbors within the 3000–25000 m range. Above 25000 m the -function indicates dispersion until about 45000 m which corresponds to the approximate distance between clusters, so above that range a given point is likely to pick up neighbors from other clusters. However, the only significant range is the clustering from 3000 to 25000 m where the -function is well outside the simulated confidence envelope.

Task 5: Discuss the pattern in the shape-file **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]

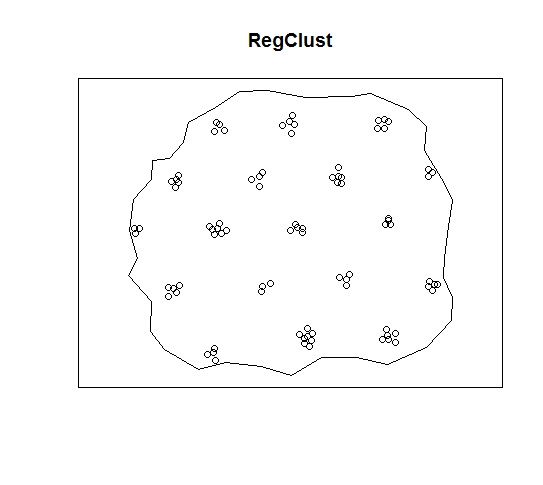
> regClustPPP <- as.ppp(regClustUTM)

> regClustPPP <- regClustPPP[win]

> regClustPPP$marks <- NULL

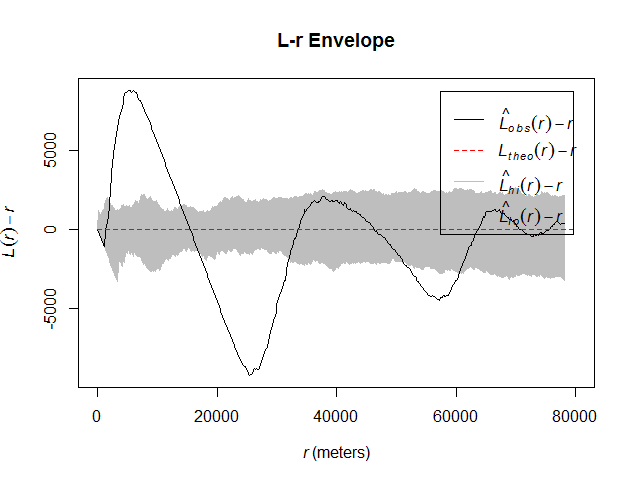
> plot(regClustPPP, add=F, main="RegClust")

> box()



> Lv <- envelope(regClustPPP, fun=Lest, rmax=80000)

> plot(Lv, .-r ~ r, main="L-r Envelope")



RegClust is somewhat similar to ClustReg in that there are multiple visually distinct clusters of points, although RegClust does not appear to have any sub-clusters or arrangement of points into lines with regular spacing. The two closest points are about 1100 m apart, so the -function indicates dispersion below that radius because a CSR pattern could potentially have closer points. Above 1100 m the -function indicates significant clustering until a switch to significant dispersion above about 16000 m. 16000 m is roughly half of the distance between clusters, which explains why the -function indicates clustering above approximately 33000 m since at that radius a given cell will tend to count neighbors from other clusters. As the radius increases, the -function continues to fluctuate between dispersion and clustering with decreasing amplitude, probably because of the very regular arrangement of clusters in a near-perfect hexagonal pattern.

## Heterogeneous Poisson Process [3 points]

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

N <- 100 # Number of points to generate

pts <- data.frame("x"=numeric(N), "y"=numeric(N))

i <- 1

while(i <= N) {

# 1. Generate candidate point

pntX <- runif(1)

pntY <- runif(1)

intensity <- pntX + pntY

normIntensity <- intensity / 2

# 2. Generate uniform random number

rand <- runif(1)

# 3. Keep or reject candidate point

if(rand <= normIntensity) {

pts$x[i] <- pntX

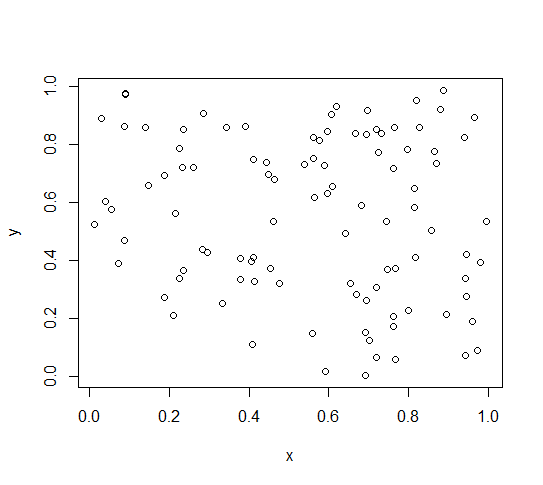
pts$y[i] <- pntY

i <- i + 1

}

}

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



It is a random point pattern but with a greater density of points toward the upper right corner compared to the lower left corner because of the intensity gradient used to decide whether to keep or reject uniform-random points.