[redacted]

Stat M222 Final Project: Crime in Greater London

[redacted]

[redacted]

Introduction

Until humans manage to completely eliminate crime, it is something that society would like to manage and keep under control. Criminology is exactly the study of criminal behavior in general that has such an aim. There have been many studies that attempt to explain criminal behavior from various angles, including looking at the effects on the propensity to commit crime due to: gender, race or immigrant status, socioeconomic status, religion, and psychological traits (Lee 2009). Given all of these different perspectives, why not also use spatial-temporal data to study criminal activity and perhaps to also manage crime? This is the goal of this introductory analysis – to use spatial temporal data to see whether there might be any patterns to crime in the area of Greater London.

The rest of this paper will be as follows in our analysis of criminal activity in Greater London. We will first give an introduction to the data used as well as several caveats to this data. Next, we will introduce our methodology that we adopted for the sake of this analysis. We will then point to various diagnostic plots as an introduction to how the original data looks. Next, we will fit two different point process models to the data in an attempt to explain possible patterns to the crime data. Finally, we will summarize our findings and suggest further avenues we could explore about this data and topic in general for a future analysis. Overall, we find that there does seem to be clustering in a spatial-temporal sense of criminal activity in Greater London, but that the point process models we fit our data with do not seem to fit the data perfectly.

Introduction to the Data and its Caveats

The data we will work with in this analysis comes from open data about policing in England, Wales, and North Ireland. Specifically, it comes from the Data.Police.UK website that is operated by government of the United Kingdom as a part of an initiative to let individuals have easier access to government data. We choose to focus specifically on crime in the Greater London area, which encompasses the city of London at its center as well as its surrounding areas.

For more specifics on our data, we choose to only look at crimes perpetrated in March 2017 that disrupted the "public order." The Data.Police.UK website defines this as any "offenses which cause fear, alarm, or distress." (About Data.Police.UK) The reason that we chose to work with such a specific focus is because it seems to make sense to focus only on one type of crime as different types of crime may not necessarily be related to each other. Also, there seems to be quite an amount of data even in just looking at one month (there were 4019 recorded instances of crimes disrupting the "public order" that contained location data in Greater London in March 2017 alone). The data essentially contains only the locations of crimes, and the type of crime committed at each location.

There are several less than desirable features to the data we are working with, however. Due to privacy concerns, the data that the Data.Police.UK website provides is somewhat lacking in detail. Namely, while the data contains the month and year in which the incidents occurred, it does not even contain the date of occurrence, much less the time of day of occurrence. However, the data should come in chronological order, and for the purposes of this analysis we assume that the events are evenly spaced out in time. Another related drawback to the data we are using is that the locations of crime incidents are shifted slightly from their actual locations. This means that for each crime, the location reported will be something like the center of the nearest street, or the nearest public place such as a park or airport. There is no real way to deal with this as we do not have more detailed information, but this should not be a big issue as long as the shifts in locations are not too drastic, which the Data.Police.UK website seems to suggest. (About Data.Police.UK)

Methodology

For this analysis, we will adopt certain methodology. Here, instead of doing our analysis on the full list of 4019 crimes that disrupted the public order in Greater London during March 2017, we will instead take a sample of 200 of these crimes to conduct our analysis on. While this is not ideal, we believe

[redacted]

it should not matter too much as the distribution of the crimes spatially and temporally seem to be similar to the full data set (see Figures 1 and 2 in the appendix).

In fitting our two point processes models on our data, we will normalize the longitude and latitude coordinates of our data so that they both fall between [0, 1]. Basically, we ensure that the transformed locations of our data fall within the unit square. Also, as stated before, we assume that the crimes are evenly spaced out temporally, with 0 representing the time of the earliest crime and 1 representing the time of the latest crime.

We will try to fit two different point processes models on our sample – which are the Hawkes process and an inhomogeneous Poisson process.

In fitting a Hawkes process model on our data, we will use maximum-likelihood estimation to get estimates of the parameters. The exact form of the conditional intensity function under this model that we choose to fit on can be found on Note 1, which can be found next to Table 1 in the appendix. We will also conduct super-thinning based on this fitted model, and plot these super-thinned points as well as estimate the F, G, and J functions of the super-thinned points back-transformed to the original coordinate system on a custom spatial window defined by the actual borders of Greater London.

As for fitting an inhomogeneous Poisson process, we will use the Stoyan method to estimate the parameters of this model. The exact form of the conditional intensity function that we elected to fit on our data on can be found in Note 2, which can be found alongside Table 2 in the appendix. Like with the Hawkes model, for our fitted Poisson process we will also conduct super-thinning, and plot these super-thinned points as well as estimate the F, G, and J functions of the super-thinned points back-transformed to the original coordinate system on a custom spatial window defined by the actual borders of Greater London.

Original Data and Diagnostic Plots

In Figure 1 of the appendix, we can see a plot of all of the incidents that disrupted public order in the Greater London area during March 2017. Based on this plot, there seem to be crimes over most of Greater London. However, there does seem to be more clustering towards the center of Greater London, as well as a temporal aspect to crimes committed (notice how the earlier crimes seem to be committed on the outer regions of Greater London, while the more recent crimes seem to be clustered towards the center). As for the sample of 200 crimes we will work with for the rest of this analysis, we can see similar spatial-temporal patterns as seen in Figure 2 of the appendix over all of the data.

Before fitting point process models, we can also look at the estimated F, G, and J functions of our sample of data. Based on Figure 3 in the appendix, the estimated F function seems to indicate more clustering of crimes at all distances than a stationary Poisson process would yield. The G function of Figure 4 in the appendix paints a similar story, except that for some reason at longer distances there seems to be inhibition rather than clustering. The estimated J function of figure 5 in the appendix combines the information yielded by the estimated F and G functions, seeming to indicate that there is clustering at lower distances between crimes but perhaps inhibition at higher distances.

<u>Fitting a Hawkes Process</u>

In Note 1 and Table 1 of the appendix, we see that according to the model fitted, we have an estimated background rate of $\mu = 36.185$ and a productivity of $\kappa = 0.605$. This means that if the fitted Hawkes model does truly describe the spatial-temporal patterns to our sample of crime in Greater London, we are looking at a sub-critical process in which one crime in Greater London on average generates 0.605 crimes near it, which would seem to indicate clustering. The standard parameters of all of the estimates seem to be reasonably low, which would seem to indicate this model is an ok fit.

After fitting this model, we can then conduct the process of super-thinning based on this fitted model and see if the resulting super-thinned points follow a stationary Poisson process (which it should, if the model fits the data well). As we can see in Figure 6 of the appendix, the super-thinned points seem to follow a stationary Poisson process, with no obvious gaps or clustering. Figure 7 of the appendix shows the super-thinned points back-transformed to their original coordinate system, plotted on a map of Greater London. Again, these super-thinned points seem to follow a stationary Poisson process. (The reason that not all of the map of Greater London is filled is most likely due to there being very little crime in those areas, and our sample probably not sampling those crimes. The super-thinned points under the fitted Poisson process will also exhibit this behavior.) As for the F, G, and J functions on the super-thinned points, the estimated functions should be very close to what the functions would be if under a stationary Poisson process. Unfortunately, as one can see in Figures 8, 9, and 10, this is not really true. The estimated F and G functions of our super-thinned points only seem to follow the Poisson curve at lower distance *r* values, while the estimated J function seems to indicate that there is still clustering among the super-thinned points. This is perhaps due what we saw with the original data – there seems to be clustering at shorter distances but inhibition at longer distances, but the form of the Hawkes model we attempted to fit does not really seem to account for this structure of data.

Fitting an Inhomogeneous Poisson Process

Note 2 and Table 2 in the appendix details the parameter estimates and their standard errors when we fit an inhomogeneous Poisson process of the form detailed in the same Note. One thing to immediately note is that the standard errors for all of the parameter estimates are higher than the absolute values of the parameter estimates themselves, suggesting not a very good fit. Here, even the super-thinned points plotted on the unit square in Figure 11 and the back-transformed super-thinned points plotted on the map of Greater London in Figure 12 don't seem to suggest a good fit (there are obvious gaps in the upper left-hand corner of the unit square and Greater London plots, hinting that our predictions of the conditional intensity in those areas are overestimates). The estimated F, G, and J function seen in Figures 13, 14, and 15 of the appendix all also seem to imply a less than optimal fit. Again, like in the fitted Hawkes process case, the estimated F and G functions of the super-thinned points seem to only follow the Poisson curve at short distances, and the J function still indicates clustering among the super-thinned points.

Conclusion and Further Extensions

Overall, we find that there seems to be some clustering in the spatial-temporal sense among incidences of crime in the Greater London area during March 2017, both by looking at the original data and through our fitted models. While we tried to fit a Hawkes process and an inhomogeneous Poisson process to our data, we found that the models only fit somewhat well to the data – these models seem to be able to only explain the clustering found at short distances but were not able to fully account for and explain what seems to be inhibition between crimes that take place at further distances.

There are definitely more avenues of exploration for this topic that we could potentially further explore in a future analysis. One obvious aspect to look at in the future, given more time, is to actually conduct an analysis on the full set of data instead of a sample of the data and see if our results end up being very different. We could also look at crimes disrupting the public order over several months rather than just for one month to see if the clustering pattern we saw here still remains and if our models fit better. Since we only fit two different point processes models to our data, we could also fit more different point process models (like Poisson processes of different forms, models that include covariates, etc.) and see if there are any that fit better, especially those that take into consideration the nature of our data seeming to exhibit clustering at close distance but inhibition at further distances. Of course, yet another perspective we could examine at is to see if the patterns we saw here with crimes disrupting the public order are still present with other types of crimes such as theft, drugs, arson, etc. Finally, though this is probably the least likely to happen (due to privacy concerns), we could try getting more precise data that include more details about each incident (that can be used as covariates) as well as more exact information on time and locations of crimes committed.

References

Ellis, Lee, Kevin M. Beaver, and John Paul. Wright. *Handbook of Crime Correlates*. Amsterdam: Elsevier/Academic, 2009. Print.

"About Data.police.uk." *Data.Police.UK*. Government of the United Kingdom, n.d. Web. 1 June 2017.

[redacted]

Appendix

Figure 1

Locations of all crime incidents disrupting the public order in March 2017

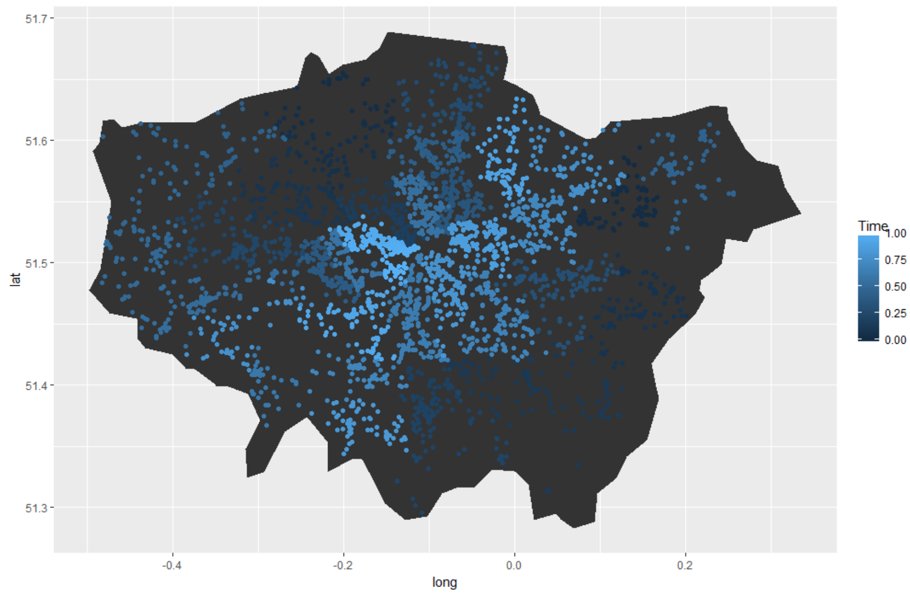


Figure 2

Locations of sample of 200 crime incidents disrupting the public order used for the analysis

[redacted]

Figure 3

Estimated F function on sample of 200 crime incidents used for the analysis

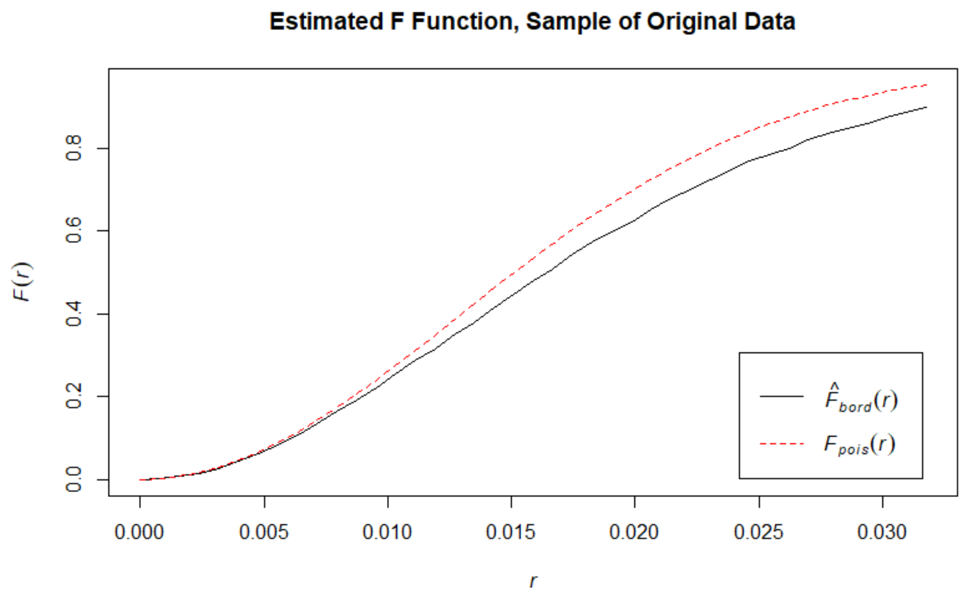
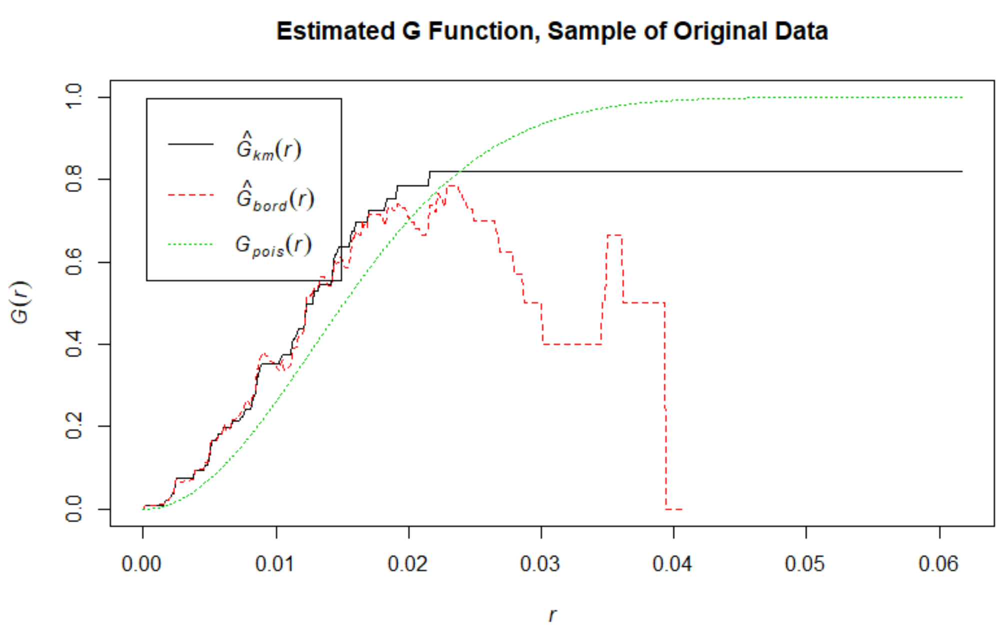


Figure 4

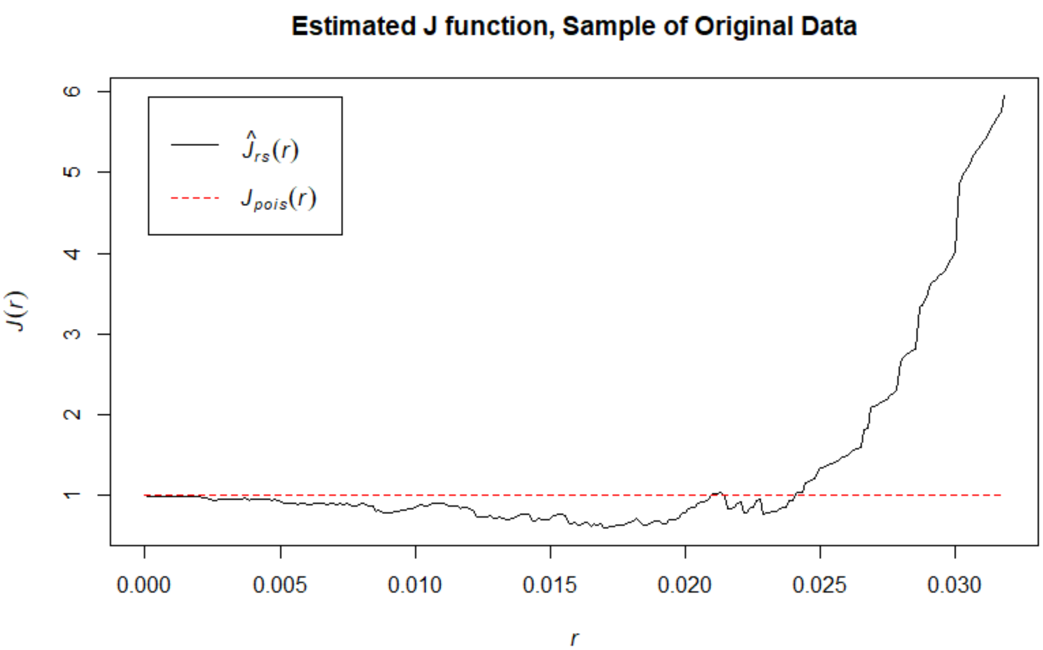
Estimated G function on sample of 200 crime incidents used for the analysis



[redacted]

Figure 5

Estimated J function on sample of 200 crime incidents used for the analysis



[redacted]

Note 1

Form of fitted conditional intensity under Hawkes process:

$$\lambda(t,x,y) = \mu\, \rho(x,y) + \kappa \sum_{\{t',x',y':\, t' < t\}} g(t-t_i)\, g(x-x_i, y-y_i)$$

Where:

- $\rho(x,y) = \frac{1}{X_1 Y_1}$
- $g(t) = \beta e^{-\beta t}$
- $g(x,y) = \frac{\alpha}{\pi}\, e^{-\alpha r^2}, x^2 + y^2 = r^2$
- *Over the space S = [0, X_1] x [0, Y_1] in time [0, 1]. Here, we set $X_1 = Y_1 = 1$, so over unit square.*
- *Parameters:* $\mu,\ \kappa,\ \alpha,\ \beta$

Table 1

Estimated parameters under the fitted Hawkes process model and their standard errors

Parameter	$\boldsymbol{\mu}$	$\boldsymbol{\kappa}$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$
Estimate	36.185	0.605	4.528	705.646
Std. Error	21.503	0.043	0.324	69.656

[redacted]

Figure 6

Super-thinned points based on fitted Hawkes process plotted on the unit square. Black dots are super-thinned points, while red dots are the original sample of 200 crime incidents. We used b = 50 here.

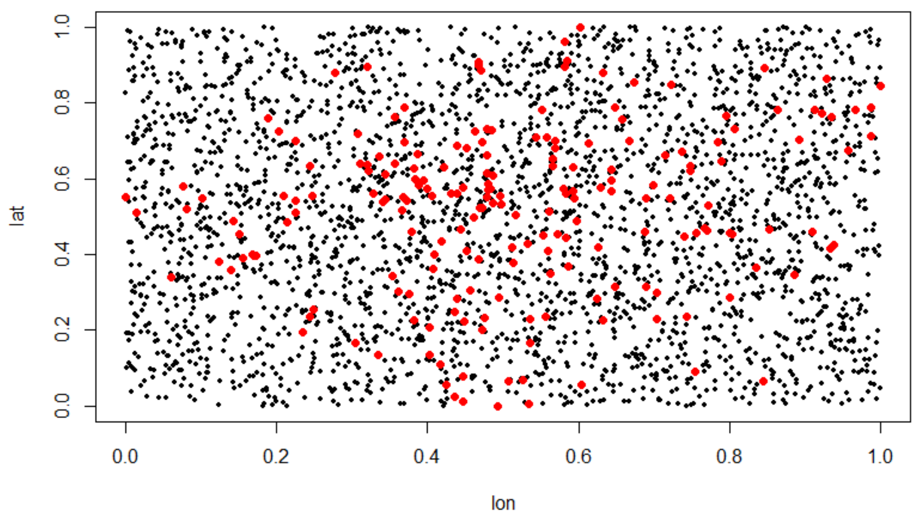
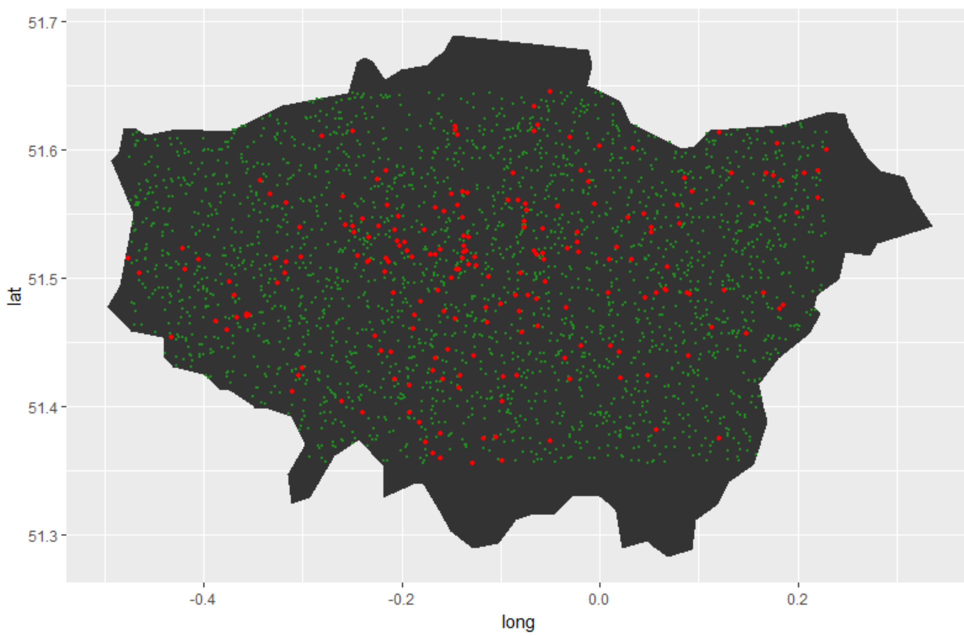


Figure 7

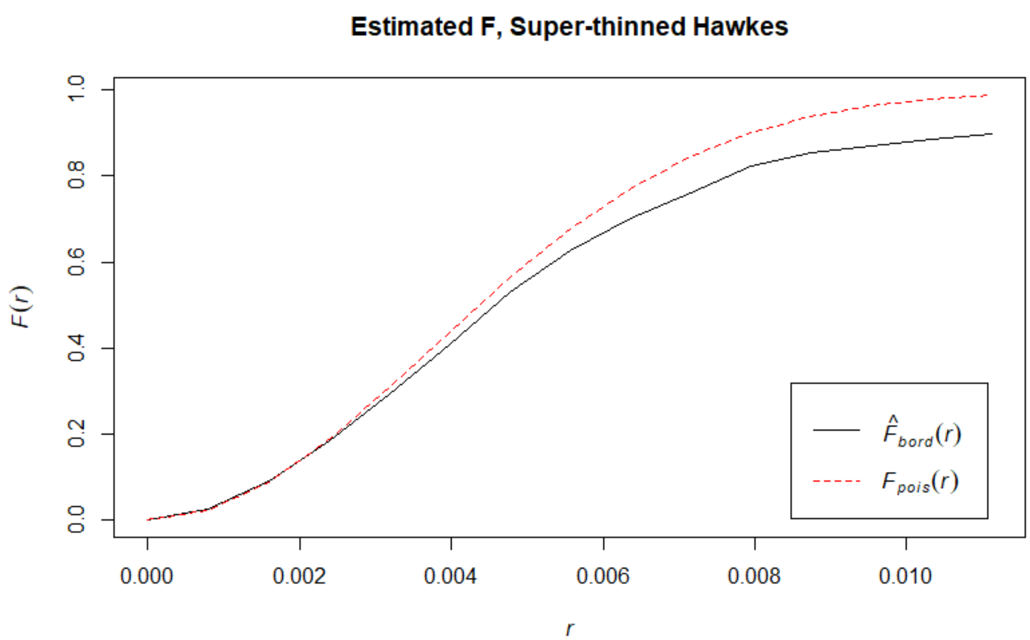
Back-transformed super-thinned points based on fitted Hawkes process plotted on a map of Greater London. Green dots are super-thinned points, while red dots are the original sample of 200 incidents.

[redacted]



<u>Figure 8</u>

Estimated F function of super-thinned points based on the fitted Hawkes process model



<u>Figure 9</u>

Estimated G function of super-thinned points based on the fitted Hawkes process model

[redacted]

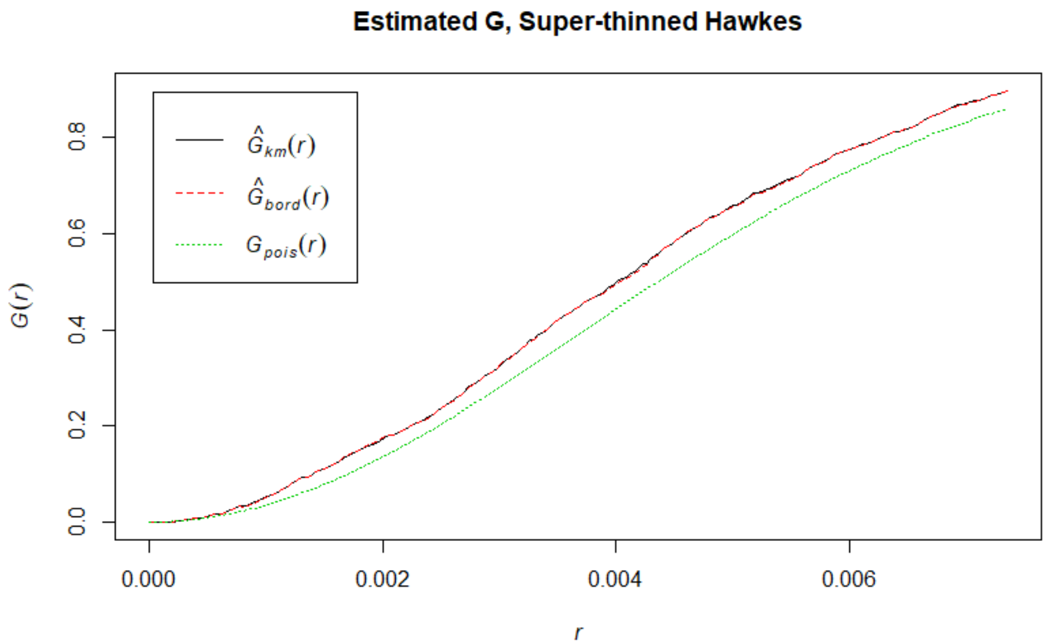
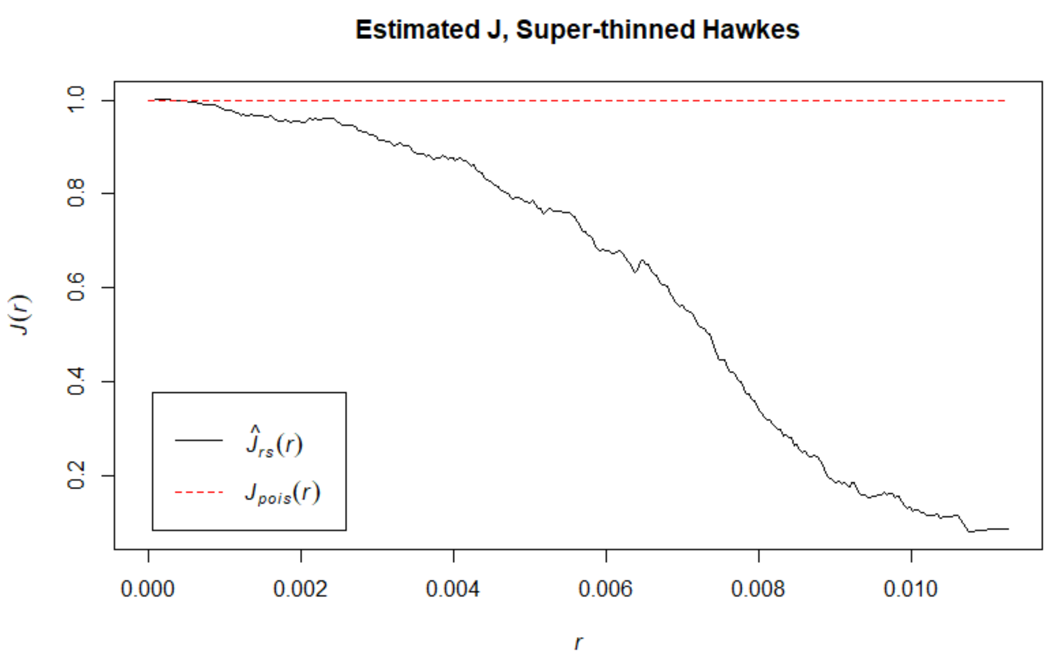


Figure 10

Estimated J function of super-thinned points based on the fitted Hawkes process model



[redacted]

Note 2

Form of fitted conditional intensity under inhomogeneous Poisson process:

$$\lambda(t,x,y) = \beta_1 + \beta_2 x + \beta_3 y$$

- *Over the space S = [0, 1] x [0, 1] in time [0, 1].*
- *Parameters:* $\beta_1, \ \beta_2, \ \beta_3$

Table 2

Estimated parameters under the fitted inhomogeneous Poisson process model and their standard errors

Parameter	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$
Estimate	26.158	-27.489	27.122
Std. Error	39.124	49.346	42.060

[redacted]

Figure 11

Super-thinned points based on inhomogeneous Poisson process plotted on the unit square. Black dots are super-thinned points, while red dots are the original sample of 200 crime incidents. We used b = 50 here.

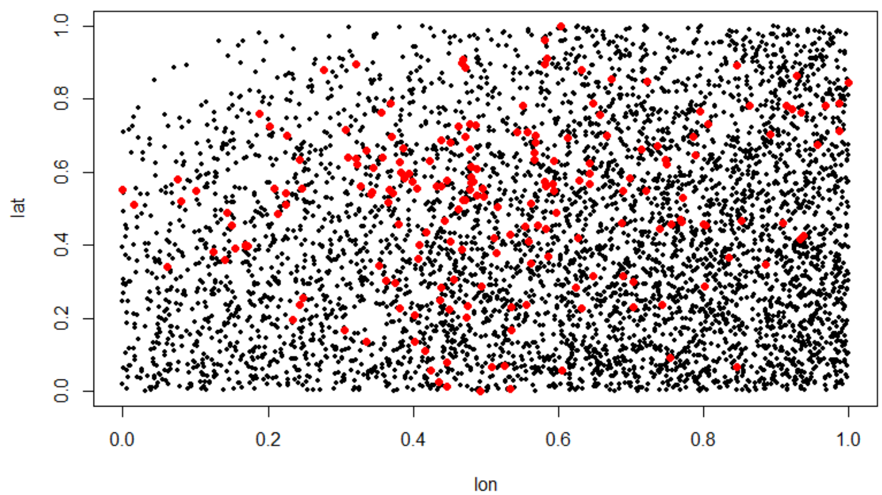


Figure 12:

Back-transformed super-thinned points based on fitted inhomogeneous Poisson process plotted on a map of Greater London. Green dots are super-thinned points, while red dots are the original sample of 200 crime incidents.

[redacted]

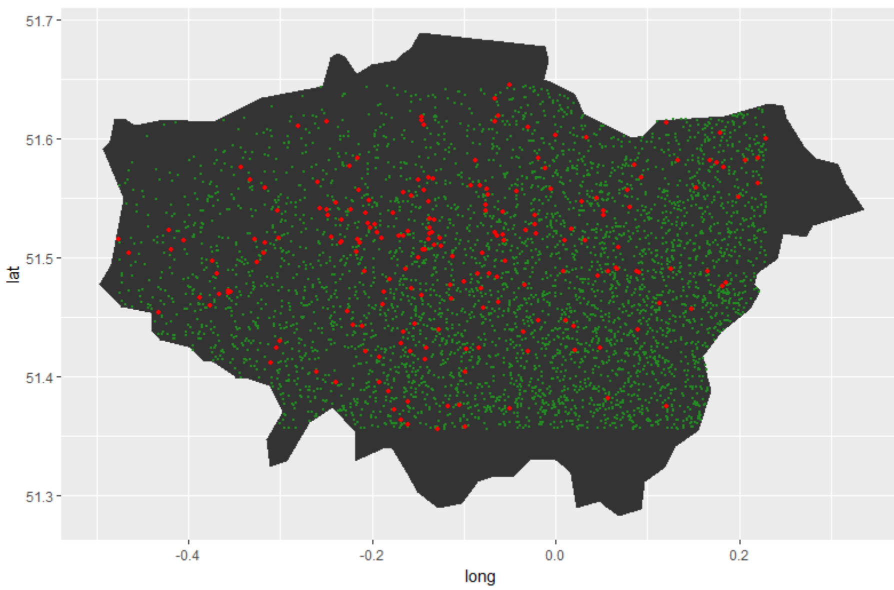


Figure 13

Estimated F function of super-thinned points based on the fitted inhomogeneous Poisson process model

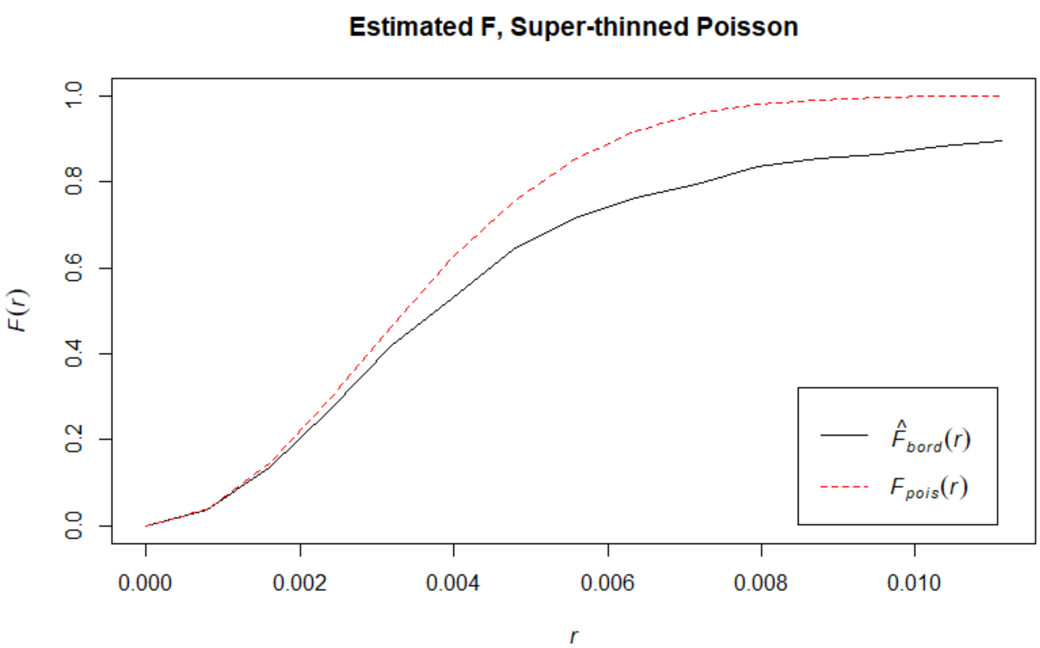


Figure 14

Estimated G function of super-thinned points based on the fitted inhomogeneous Poisson process model

[redacted]

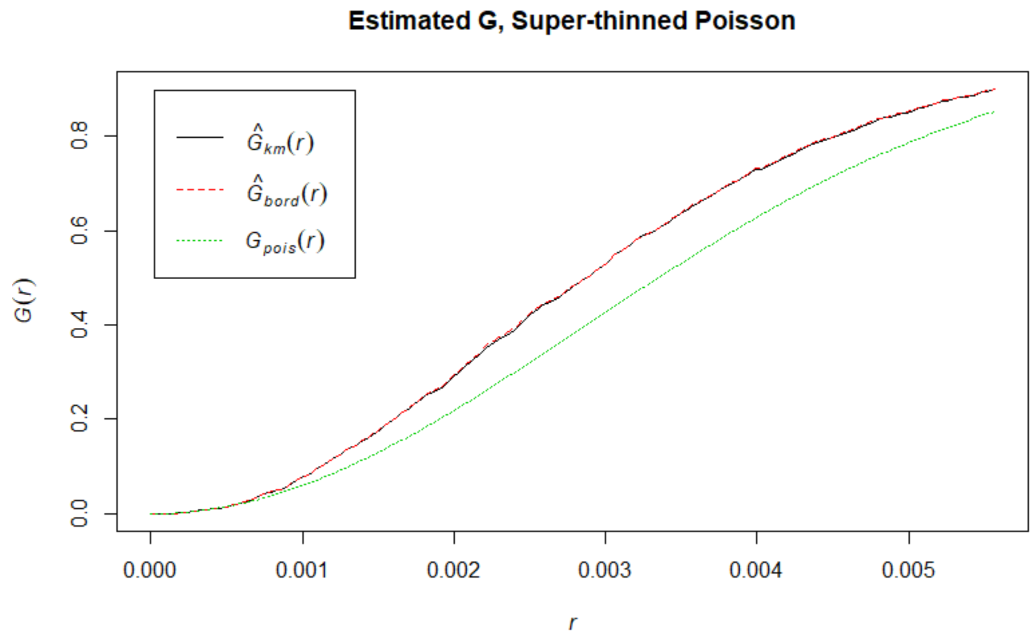


Figure 15

Estimated J function of super-thinned points based on the fitted inhomogeneous Poisson process model