|  |  |
| --- | --- |
| Map  Description automatically generated  Modeling 911 calls in Montgomery, pennsylvania  Using a homogeneous and non-homogeneous Poisson process | By: Daniel Garcia  Dr. Olga Korosteleva |

Objective

The state of Pennsylvania ranks right in the middle in crime of all 50 U.S. states. Finding data on crime in Montgomery county, Pennsylvania inspired me to further analyze the emergency 911 calls and generate reliable information for use in staffing law enforcement, 911 call operators, law enforcement administrators, and management. This analysis aims to model 911 calls as a homogenous and non-homogenous Poisson process, where we can find probabilities of 911 calls given certain conditions and accurately staff to respond promptly.

Data

The data used was from Kaggle.com, an online community platform where users can interact with one another, share datasets, projects, and other collaborations.

The original dataset has 663,522 observations. After sub-setting by town, Lower Merion had 55,490 observations and East Greenville had 1,316 observations. The variables included latitude, longitude, description, zip code, title of 911 call, timestamp, township, address, and e. For the purposes of this study, I used only two variables: timestamp and township.

Overview of methods used

Homogenous:

The first thing I do after reading the data into RStudio is group the data by town to see which towns in Montgomery county occupy the most and least 911 calls. Lower Merion had the greatest number of 911 calls, and East Greenville had the 6th lowest 911 calls. I chose these two towns for comparison purposes. I then format the timestamp variable to only include the month, day, and year. This allows me to count the total 911 calls per day. Next, I find the average number of calls per day by taking the sum of the total calls and dividing by the total number of days, this is my lambda for a homogeneous Poisson process.

I then visualize the total number of calls per day by creating a histogram in ggplot. The outcome for Lower Merion showed that of a non-homogeneous Poisson process and showed a homogeneous Poisson process to be more appropriate for East Greenville. To get more numeric data I calculate wait times for both towns, then create simulations.

Non-homogeneous:

For a non-homogeneous process, I decided to format the timestamp to only include hours of the day and only include Lower Merion as the town. I grouped the data by hour—where 0 represents 12 AM and 23 represents 11 PM—then summed up all 911 calls, giving me the total 911 calls by hour of the day. I divided the number of hours in the day into 4 equal intervals and calculated the average calls per time interval. The output acted as my intensity rate function. Next, I integrated each average to derive the integrated intensity rate function and plot the intensity rate function and the integrated intensity rate function.

Next I define the integrated intensity function in RStudio and calculate 3 things: the expected number of calls between 10 AM and 1 PM, the probability there will be exactly 3 calls at or after 10 AM and before 2 PM, and the probability there will be more than 3 calls at or after 10 AM and before 2 PM.

Results

I found that the homogenous Poisson model was appropriate only for the town of East Greenville and the non-homogenous model was more appropriate for Lower Merion.

Predictions/Probabilities

In the homogenous Poisson model, we saw that average wait times until we get the 20th 911 call were 14 hours and 35 minutes and 11 days, 20 hours, and 52 minutes for Lower Merion and East Greenville, respectively. We also found that on average, Lower Merion can expect a 911 call every 44 minutes, where East Greenville can expect a 911 call every 14 hours and 15 minutes.

For the non-homogeneous Poisson model, we saw that the average number of 911 calls at or after 10 AM and before 2 PM was approximately 5.95. We also found that the probability there will be exactly 3 calls at or after 10 AM and before 2 PM was approximately 9.15%. And the probability there will be more than 3 calls at or after 10 AM and before 2 PM was approximately 84.68%.

Future Work

In the future, I hope to expand on my findings by analyzing more towns and comparing them to one another. I also would like to use natural language processing techniques to filter the description of the crimes and analyze specific types of crimes.

Appendix

**A) Code for Homogenous Poisson model on Lower Merion 911 calls**

**Graphical user interface, text, application, email

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated** **Graphical user interface, text, application, email

Description automatically generated** Graphical user interface, text, application

Description automatically generated Graphical user interface, text, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Chart, histogram

Description automatically generated Chart, scatter chart

Description automatically generated Graphical user interface, text, application, email, timeline

Description automatically generated Graphical user interface, text, application, email

Description automatically generated

**B) modeling 911 calls in East Greenville as a homogeneous Poisson process**

Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated Graphical user interface, application, table

Description automatically generated Graphical user interface, text

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Chart, histogram

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Chart, line chart

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**C) code for modeling 911 calls in Lower Merion as a non-homogeneous Poisson process**

Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated Graphical user interface

Description automatically generated Chart, line chart

Description automatically generated Graphical user interface, text, application, email

Description automatically generatedReferences

Chirico, M. (2020, July 29). Emergency - 911 calls. Kaggle. Retrieved November 30, 2022, from https://www.kaggle.com/datasets/mchirico/montcoalert