**GGRC32 – Essential Spatial Analysis**

**Fall Semester 2017**

**Lab 2: Point Pattern Analysis**

**By: Grisham Nathan**

**Due October 24th before class**

**Assignment #2 Questions**

***Please include both questions and answers. Only submit the Assignment 2 Questions and Answers. Make sure to type and show all of your calculations.***

**Part I: Nearest Neighbor Analysis**

Consider the point distribution on the NND tab in the accompanying spreadsheet. Without producing the distance matrix, fill in the NND column simply by observing the point distribution.

Q1) What is the average nearest neighbour distance for this dataset?

The average neighbour distance for this dataset is 0.1.

Q2) Assuming that the study area is a 2.5X3.2 rectangle, compute , , and .

=24/(2.5x3.2) =1/(2) = 0.26136/()

=24/8 = =0.26136/(6)

=3

Q3) Compute the Z score for this point pattern.

= (0.1-())/(0.26136/(6))

= -6.125503541

Researchers would like to know if this distribution can be considered uniform. Conduct the appropriate 1-sided hypothesis test. Set the significance level to 0.1.

Q4) State the null and alternative hypotheses

Q5) Calculate the critical Z value.

norm.s.inv(0.1)= -1.281551566

Q6) Compare the critical Z value to the compute Z score for this point pattern.



Q7) Make a conclusion regarding the spatial pattern.

* + Given our significance level of 0.1, is within the rejection zone of the distribution. As a result, we reject the null hypothesis that the data are randomly distributed.

Q8) Without doing the calculations, sketch and interpret G(d) and K(d) functions.

\*Sketched by manually plotting points of the G(d) and K(d) functions on excel .\*

From the flat line from 0.2 to 1, we can conclude that the 100% of the data is uniform with distance of separation among the points to be typically be less than 2. We know that all the points have Nearest Neighbour distances of 0.1, so that makes sense.

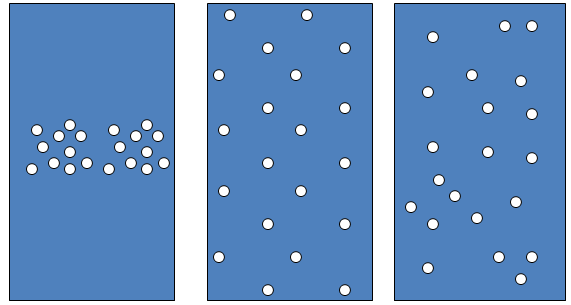
The average number of neighbours that have distances less than 1 are very low. However, the average number of neighbours that have distances less than 1 or less than a distance of bigger than 1 are very high.

Q9) How do these functions help you describe the spatial pattern over and above the simple testing performed using nearest neighbour distances?

The functions help me describe the spatial pattern over and above the simple testing performed using nearest neighbour distances by examining more complete descriptions of the spatial distribution of the data at multiple spatial scales.

**Part 2: Quadrat Analysis**

Q1) Identify if the patterns (A, B, and C) are random, clustered or uniform.



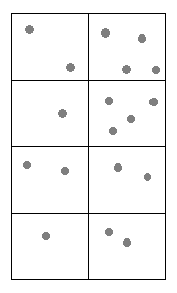
A B C

A: Clustered

B: Uniform

C: Clustered

**Quadrat analysis**: Perform a quadrat analysis on the image below. Calculate the variance-to-mean ratio (VMR) of the data, and indicate whether the data are random, uniform, or clustered based on the calculated VMR.



Q2) Complete the following table:

|  |  |
| --- | --- |
| **Quadrat #** | **Points per Quadrat** |
| 1 | 2 |
| 2 | 4 |
| 3 | 1 |
| 4 | 4 |
| 5 | 2 |
| 6 | 2 |
| 7 | 1 |
| 8 | 2 |

Q3) What is the mean of this dataset?

18/8

2.25

Q4) What is the variance (s2) of this dataset?

= (0.0625+3.0625+1.5625+3.0625+0.0625+0.0625+1.5625+0.0625)/(8-1)

=9.5/7

Q5) Using the mean and variance, calculate the VMR.

=(9.5/7)/2.25

=9.5/(7x2.25)

=9.5/15.75

=38/63 or about 0.6032

Q6) Given the VMR, what can we conclude about the point distribution?

Since VMR is less than 1, we can conclude that the point distribution is more uniform than random.

Q7) What is the χ2 statistic of the dataset?

=7 x (38/63)

=38/9 or about 4.22

Q8) Consider the formula for the degrees of freedom of the sampling distribution of VMR. How many degrees of freedom does this problem have?

This problem has 7 degrees of freedom.

Q9) If your alternative hypothesis had been that the point pattern was uniform (HA: VMR<1), would you reject the null hypothesis (H0: VMR=1) based on a 5% significance level?

pattern is random

pattern is uniform

= 2.167349919

> 🡪 Given our significance level of 10%, is not within the rejection zone of the distribution. As a result, I would not reject the null hypothesis based on a 10% significance level.

**Part 3: Function Analysis in ArcMap**

In this part of the assignment you will be using 2 point files and a boundary file for the City of Toronto. Add the 3 shapefiles into a new map in ArcMap.

Q1) Based on a visual analysis of the point files, would you guess that the Ambulance stations are Random, Uniform, or Clustered? How about the Bike Shop locations?

I would guess that the Ambulance stations are Random. As for the Bike Shop locations, I would guess they are Clustered.

We are going to be using the Multi-Distance Spatial Cluster Analysis (Ripley’s K Function) tool.

Compute a Ripley’s K Function analysis on the Ambulance stations.

* Set the ambulance stations as the input file
* Choose an appropriate output location
* Select 999 repetitions for the Monte Carlo Experiment
* Make sure you select the option to produce a graph of the outputs.

Q2) Paste the output graph into your answers. Interpret the output graph. Do Ambulances appear to be uniform, random, or clustered? At what spatial scales?



The output graph appears to show statistically significant dispersion of ambulances, between about 1800 meters and about 2900 meters, as the observed K function is below the confidence envelope. As a result, between about 1800 meters and 2900 meters, the ambulances display a more uniform distribution. Outside of 1800 meters and 2900 meters, the ambulances appear to be random as the Observed K function is within the confidence envelope.

Q3) What is the confidence level associated with your conclusions above?

Since 999 draws were created using the Monte Carlo Simulation, the confidence level associated with my conclusions above are 99.9%.

Compute a Ripley’s K Function analysis on the Bike shops.

* Set the bike shops as the input file
* Choose an appropriate output location
* Select 999 repetitions for the Monte Carlo Experiment
* Make sure you select the option to produce a graph of the outputs.

Q4) Paste the output graph into your answers. Interpret the output graph. Do bike shops appear to be uniform, random, or clustered? At what spatial scales?

**999 Repetitions**



The output graph appears to show statistically significant clustering of bike shops at all spatial scales, since the observed K function is above the confidence envelope at all distances.

Try re-running the analysis on bike shops. This time, use 9 repetitions, and select the Toronto boundary file as the user supplied study area.

**9 Repetitions, User supplied study area**



Q5) What did you notice about the run-time when using your custom study area? Thinking about the steps in a Monte Carlo analysis, why do you think the run-time changed so dramatically?

I noticed that it took longer to produce the K function analysis for custom study area than my first bicycle shops K function analysis. I think the run-time changed dramatically because even though we are calculating far less repetitions for the custom study area K function analysis, I supplied the Toronto boundary as my study area and so the Monte Carlo simulation has to take longer to simulate repetitions of random point patterns of equal size and study area than for my original K function analysis which just uses a minimum enclosing rectangle around the Toronto boundary.