The Lawn Sprinkler Problem

– Matthew Grissom and Stanley Zalewski –

Can we use the spread of water droplets from a lawn sprinkler to find where the sprinkler is placed? More generally, can we find the origin point of a random spread of this type with the same techniques we use to find the center of a normal random spread?

We sought to find the answer to these questions by dropping coins on a grid to create a random spread. Then we attempted to find the origin.

Method:
We dropped a single coin 4 feet above the center of two gridlines 30 times, measuring the distance from each axis in centimeters. We repeated this process twice to get two separate random spreads. We then used 4 methods to find the origin point:

1. Range Method

   Taking the midpoints of the maximum and minimum values on each axis as the origin.

2. Median Method

   For n=30, we take the average of the 15th and 16th data points on each axis as the origin.

3. Mean Method

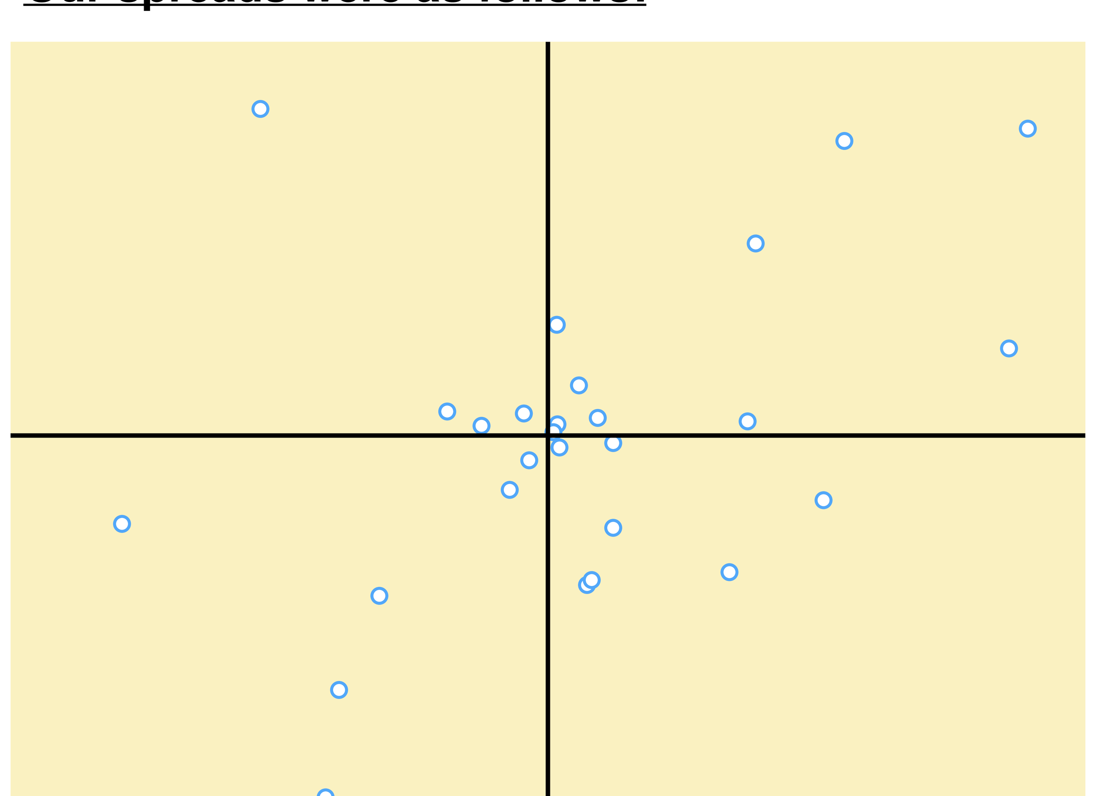
   Take the average of all 30 points on each axis as the origin.

4. Rossmo's Formula

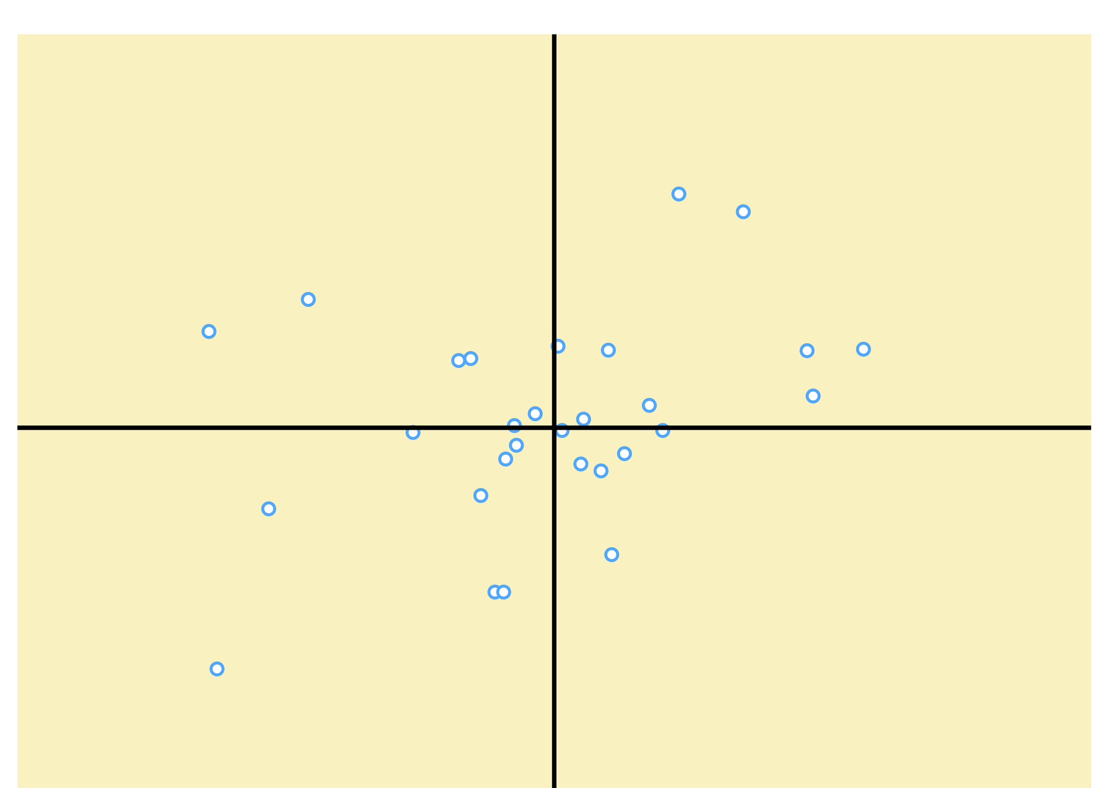
   Rossmo's formula uses the placement of points and idea of a 'buffer zone' to probabilistically model a location of origin. It was developed in the 1990s by criminologist Kim Rossmo to locate the residence of serial criminals based on their crime locations.

Our spreads were as follows:

Trial #1

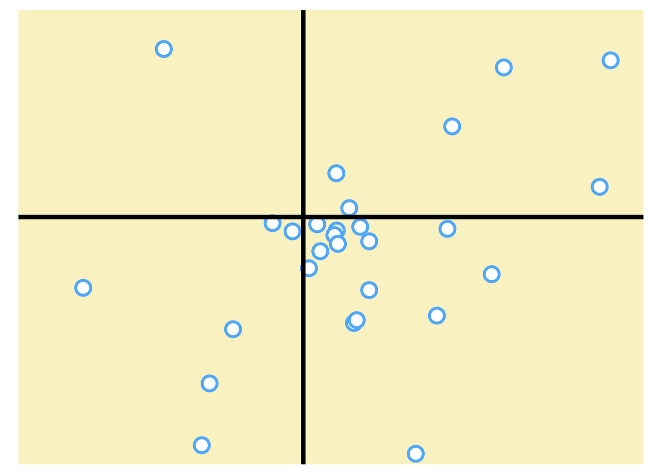


Trial #2

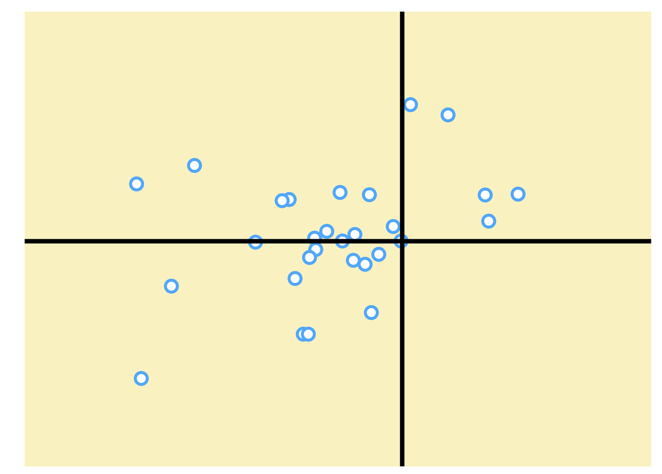


1. Range Method

Trial 1: (-3.525, 3.6)

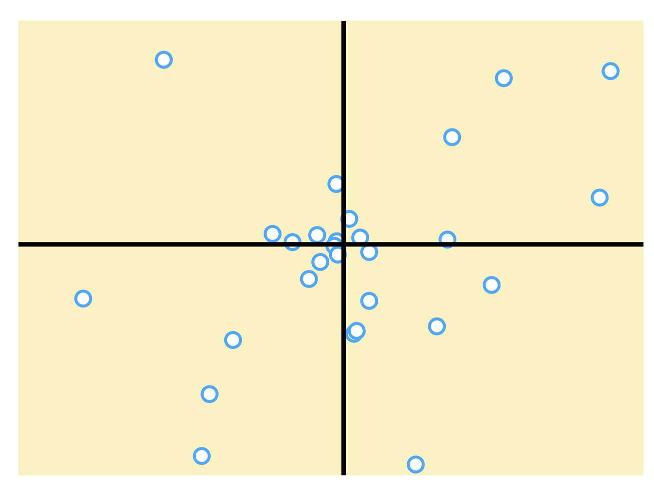


Trial 2: (8.225,-.35)

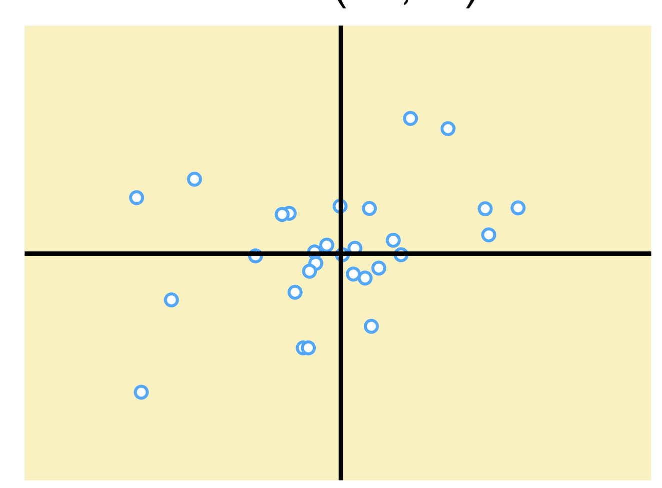


2. Median Method

Trial 1: (1.625, .675)

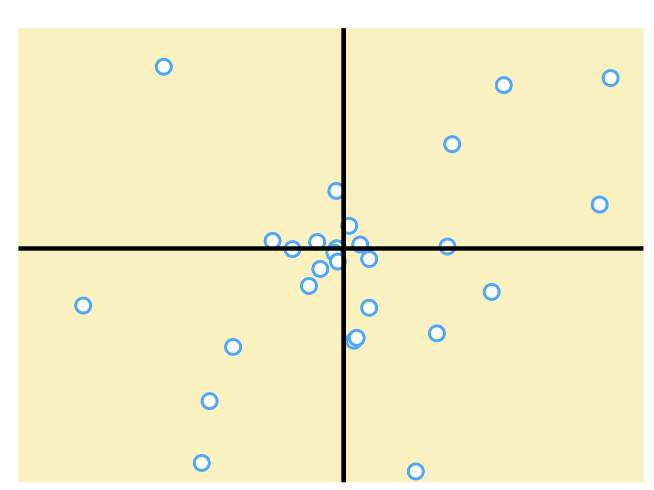


Trial 2: (.4, 0)

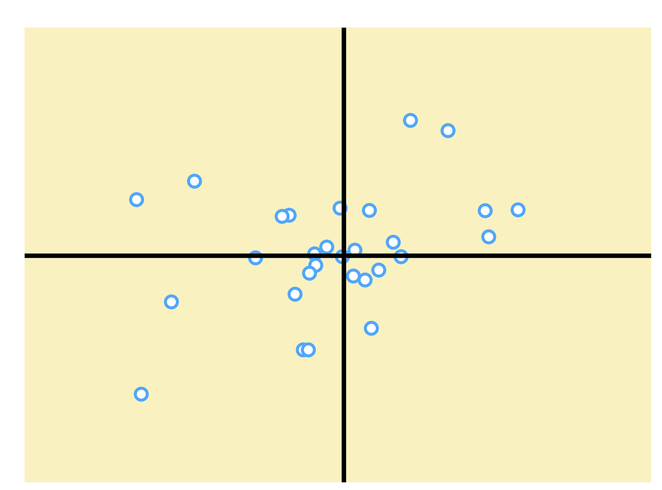


3. Mean Method

Trial 1: (1.673, 1.028)



Trial 2: (.808, .05)



4. Rossmo's Formula

Rossmo's formula is given as follows:

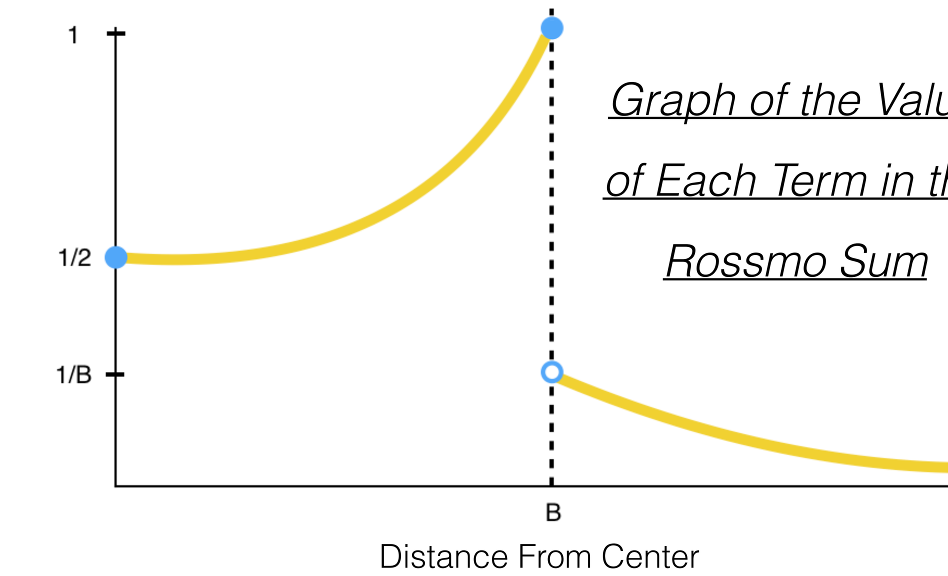
$$p_{i,j} = \sum_{n=1}^{crimes} \frac{\phi_{i,j}}{(|X_i - x_n| + |Y_j - y_n|)^f} + \frac{(1-\phi_{i,j})(B^{g-f})}{(2B - (|X_i - x_n| + |Y_j - y_n|))^g}$$

where $\phi_{i,j} = \begin{cases} 1, & (|X_i - x_n| + |Y_i - y_n|) > B \\ 0, & otherwise \end{cases}$

Calculation of Distance:
Rossmo's Formula uses a specialized distance called Manhattan distance, which is based on traveling through the gridded streets of Manhattan.

The Value of the Sum:
Unlike a traditional probability, the resulting value is not normalized on the interval [0,1]. Instead, we take a high value to be equivalent to a high probability.

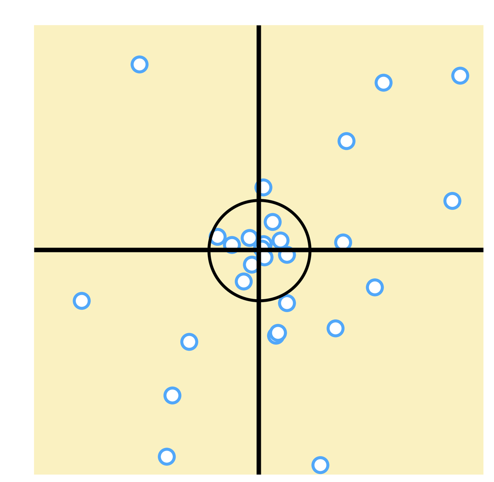
Selecting a Buffer Zone:
We selected a buffer zone that accounted for a third of our data points. As you can see below, this mostly accounted for the middle cluster of points for each trial.



Graph of the Value of Each Term in the Rossmo Sum

Ideally, all points land on the center, giving a value of .5n, however this is not realistic. We expect a small cluster of points inside the buffer zone and more outside, hence our center is the maximized value of the sum with an upper bound of 15.

Trial 1: B=9



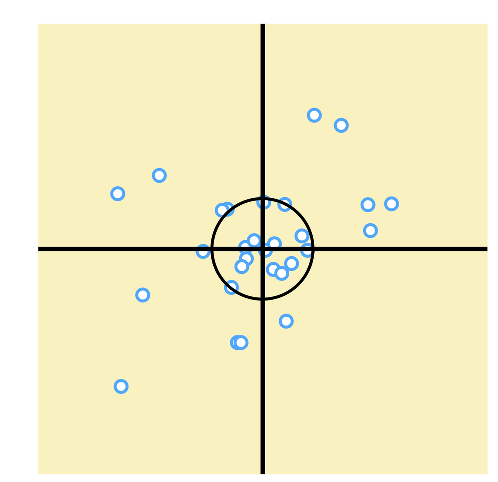
(0,0): Ω=7.6605

(10,10): Ω=1.227

P_1 = 7.66/15 = 51%

P_2 = 1.227/15 = 8%

Trial 2: B=9



(0,0): Ω=9.1438

(10,10): Ω=2.6127

P_1 = 9.1438/15 = 61%

P_2 = 2.6127/15 = 17%

The Problem with Buffer Zones:
Initially, we set buffer zone values of B_1=36.7552 and B_2=25.2717. However, upon calculating probabilities, we found that our Ω values for (0,0) and (10,10) were nearly identical. This is because too wide of a buffer zone captures too many data points no matter its placement. In application, a buffer zone is previously defined as the expected neighborhood of a serial criminal

Conclusions

1. Contrary to our expectations of traditional statistical methods, median predicted center slightly more effectively than mean.
2. As expected, range was the least effective method for predicting origin.
3. In Rossmo's formula, a large buffer zone clearly demonstrated a high probability for the origin point, but lacked specificity of location for said point.
4. With a realistic buffer value, Rossmo's formula was as effective at measuring center as the traditional methods.