Patterns in Offender Distance Decay and the Geographic Profiling Problem.

Mike O'Leary

Towson University

2010 Fall Western Section Meeting
Los Angeles, CA
October 9-10, 2010

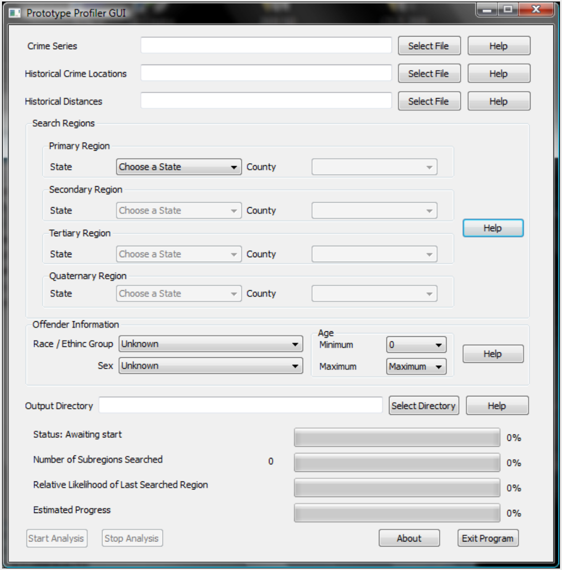
Acknowledgments

- This work has been supported by the National Institute of Justice, through grant 2009-SQ-B9-K014
- We thank the Baltimore County Police Department for providing the data used in this study
- We thank Phil Canter from the Baltimore County Police Department for his assistance.

The Geographic Profiling Problem

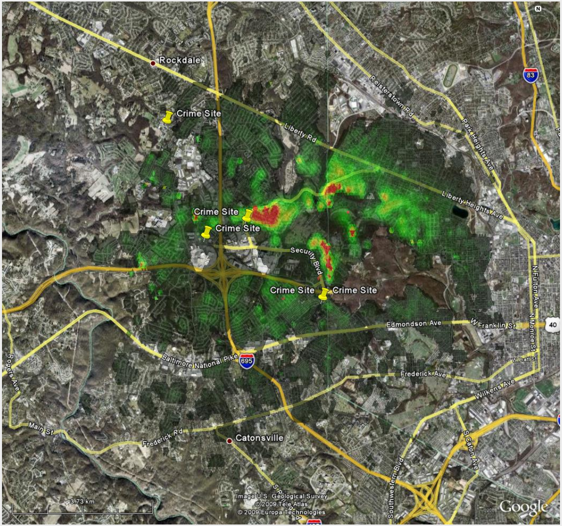
- The geographic profiling problem is to estimate the location of the home base of a serial criminal from the known locations of the elements of the offender's crimes.
  - The home base is also called the anchor point of the offender. It may be the offenders home, the home of a relative, a place of work, or even a favorite bar.
- We have developed a new tool for the geographic profiling problem.
  - It is free for download and use, and is entirely open source.
    - `http://pages.towson.edu/moleary/Profiler.html`
  - It is still in the prototype stage and is being evaluated by different police agencies across the country.

The Tool


Prototype Profiler GUI
Crime Series
Select File
Help
Historical Crime Locations
Select File
Help
Historical Distances
Select File
Help
Search Regions
Primary Region
State
Choose a State
County
Secondary Region
State
Choose a State
County
Tertiary Region
State
Choose a State
County
Quaternary Region
State
Choose a State
County
Help
Offender Information
Race / Ethinc Group
Unknown
Sex
Unknown
Age
Minimum
0
Maximum
Maximum
Help
Output Directory
Select Directory
Help
Status: Awaiting start
0%
Number of Subregions Searched
0
0%
Relative Likelihood of Last Searched Region
0%
Estimated Progress
0%
Start Analysis
Stop Analysis
About
Exit Program

Sample Results

- When the program runs, it produces an estimate for the offender's anchor point



Distance Decay

- Suppose that an offender with anchor point $\mathbf{z}$ chooses to offend at the location $\mathbf{x}$ according to the probability density $P(\mathbf{x}|\mathbf{z})$ (or more generally according to $P(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n|\mathbf{z})$.
    - We use a probability density because of our lack of knowledge about the behavior of the offender, not because the offender necessarily has a random component in the selection of target locations.
- Given a choice for P, Bayes theorem lets us estimate $P(\mathbf{z}|\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.
    - This gives us our search area for the offender's anchor point.
- Clearly $P(\mathbf{x}|\mathbf{z})$ has some dependence on the distance $d(\mathbf{x}, \mathbf{z})$ between the anchor point $\mathbf{z}$ and the crime site $\mathbf{x}$ in some distance metric d.
    - We call this factor the distance decay.

Distance Decay

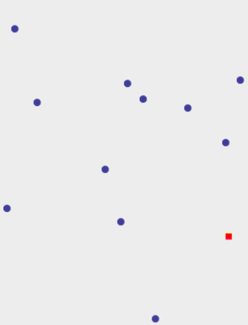
- There is no agreement as to the best mathematical form for distance decay.
- The three main software packages commonly used by police officers all use different distance decay functions:
  - Rossmo's approach is to use a piecewise rational function for distance decay, with a Manhattan metric.
  - Canter's approach is to use negative exponential functions combined with buffers and plateaus and the Euclidean metric.
  - Levine's approach it to let the analyst choose from a library of available distance decay functions and distance metrics.

Circle Theory

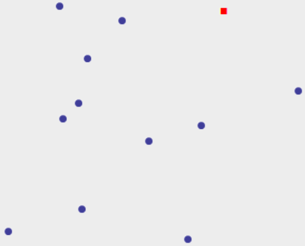
- Canters Circle hypotheses[1]: Given a series of crimes, construct the circle whose diameter is the segment connecting the two crimes that are farthest apart.
  - If the offender is a *marauder*, then their anchor point will lie in this circle.
  - If the offender is a *commuter*, then their anchor point will lie outside this circle.
    - Note that all of the crimes are not necessarily within the circle.
- For crimes like rape and arson, there is evidence that most offenders are marauders; for crimes like residential burglary the evidence shows a mixture of marauders and commuters.
- This is a binary approach- either someone is a commuter or they are a marauder.
  - This binary approach may not be suitable in many cases.

[1] Canter D. & Larkin, P. (1993). The environmental range of serial rapists. Journal of Environmental Psychology, 13, 63-69.

Which is the Commuter?



or



- Here the crime locations are blue points, and the offender's anchor point is a red square.

Commuters & Marauders

- We have created a different way to differentiate between commuters and marauders.
- Suppose that
  - The crimes are at $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.
  - The offender's anchor point is $\mathbf{z}$.
- For $1 \leqslant p < \infty$ define

$$\mu_p = \min_{\mathbf{y}} \left[\frac{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{y})^p}{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z})^p} \right]^{1/p}$$

- Note that $0 \leqslant \mu_p \leqslant 1$.
- Offenders with small μ_p correspond to μ_p-commuters, while offenders with large μ_p correspond to μ_p-marauders.

- If $p = 1$, then

$$\min_{\mathbf{y}} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{y}) = d(\mathbf{x}_i, \mathbf{y}_{\text{cmd}})$$

and

$$\mu_1 = \frac{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{y}_{\text{cmd}})}{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z})}$$

where $\mathbf{y}_{\text{cmd}}$ is the center of minimum distance of the crime series.

Commuters & Marauders

- If $p = 2$, and d is Euclidean distance, then

$$\min_{\mathbf{y}} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{y})^2 = d(\mathbf{x}_i, \mathbf{y}_{\text{centroid}})^2$$

and

$$\mu_2 = \sqrt{\frac{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{y}_{\text{centroid}})^2}{\sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z}_{\text{anchor}})^2}}$$

where

$$\mathbf{y}_{\text{centroid}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i$$

is the centroid of the crime series

Commuters & Marauders

- If $p = \infty$, we define

$$\mu_\infty = \min_{\mathbf{y}} \frac{\max_{1 \leqslant i \leqslant n} d(\mathbf{x}_i, \mathbf{y})}{\max_{1 \leqslant i \leqslant n} d(\mathbf{x}_i, \mathbf{z})}.$$

- Suppose that:
  - d comes from Euclidean distance
  - C is the circle whose diameter is the segment between the two crime sites farthest apart.
  - All of the crimes lie in the circle C
  - $\mathbf{p}$ is the center and r is the radius of the circle C

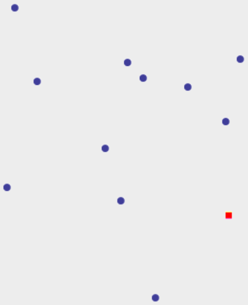
  then

$$\min_{\mathbf{y}} \max_{1 \leqslant i \leqslant n} d(\mathbf{x}_i, \mathbf{y}) = r$$

  and

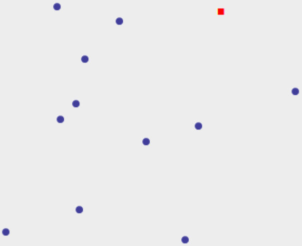
$$\arg\min_{\mathbf{y}} \max_{1 \leqslant i \leqslant n} d(\mathbf{x}_i, \mathbf{y}) = \mathbf{p}.$$

Which is the Commuter?



$\mu_2 = 0.58$
(Canter Marauder)

or



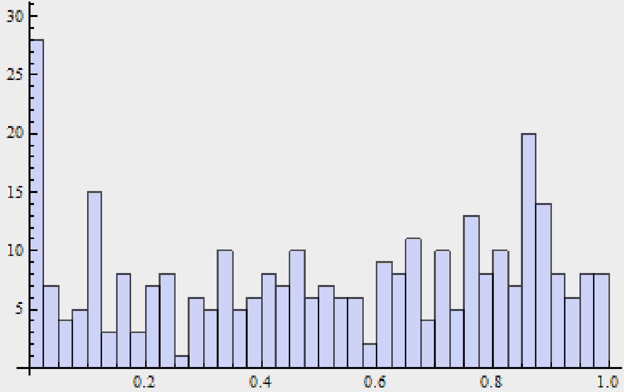
$\mu_2 = 0.56$
(Canter Commuter)

Data

- We have data for residential burglaries in Baltimore County
  - 5863 solved offenses from 1990-2008
  - We have 324 crime series with at least four crimes
    - A series is a set of crimes for which the Age, Sex, Race, DOB and home location of the offender agree.
  - The average number of elements in a series is 8.1, the largest series has 54 elements.
- We have data for non-residential burglaries in Baltimore County
  - 2643 solved offenses from 1990-2008
  - We have 167 crime series with at least three crimes.
  - The average number of elements in a series is 7.87, the largest series has 111 elements.
- We have data for bank robberies in Baltimore County
  - 602 solved offenses from 1993-2009.
  - We have 70 crime series with at least three crimes.
  - The average number of elements in a series is 4.51, the largest series has 15 elements.

Commuters & Marauders

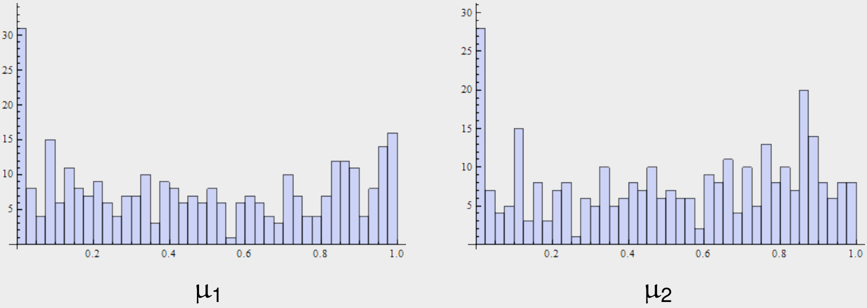
- What is the distribution of μ_2 commuters and marauders for residential burglary?



- There does not appear to be a sharp distinction between commuters and marauders in this data

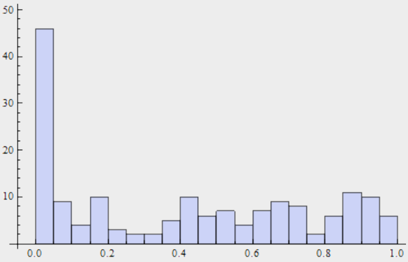
Commuters & Marauders

- There is little difference between μ_1 and μ_2 for residential burglary

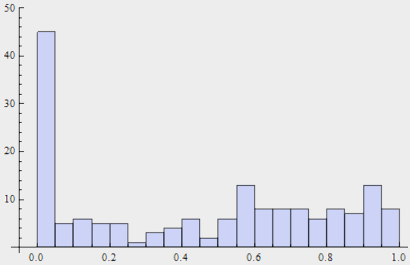


Commuters & Marauders

- Non-residential burglary shows a decided preference for commuters.



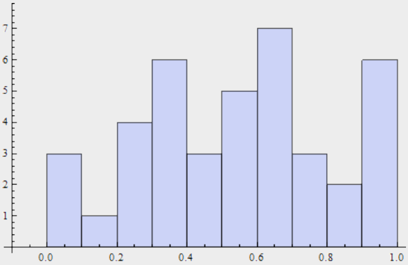
μ_1



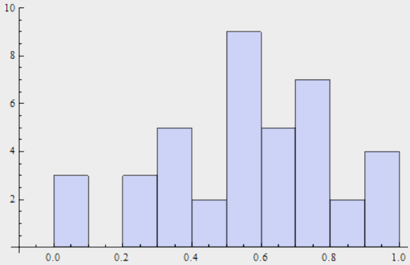
μ_2

Commuters & Marauders

- Bank robberies show a slight preference towards marauders.



μ_1



μ_2

Distance Decay

- The distance decay patterns of offenders are of fundamental importance in the geographic profiling problem.
- Though we have data for the distance from the offenders home to the offense site for a large number of solved crimes, we cannot directly use this information to draw inferences about the behavior of any individual offender.
  - To do so is to commit the *ecological fallacy*.
- There are two sources of variation- the variation within each individual, and the variation between individuals.
  - If all of the individuals behaved in the same fashion, then the aggregate data can be used to draw inference about the (common) underlying behavior.

Distance Decay

- If the only quantity that varies between offenders is the average offense distance, then the resulting scaled distances should exhibit the same behavior regardless of the offender.
  - In particular, this will allow us to aggregate the data across offenders and draw valid inference about the (assumed) universal behavior.
- For each serial offender with crime sites $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ and home $\mathbf{z}$, estimate the average offense distance α by

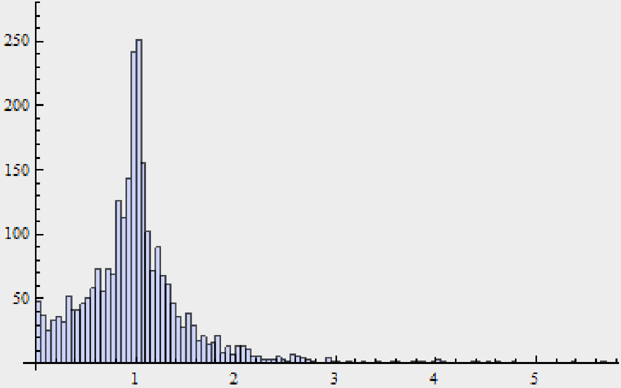
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z})$$

and now consider the set of scaled distances

$$\rho_i = \frac{d(\mathbf{x}_i, \mathbf{z})}{\hat{\alpha}}$$

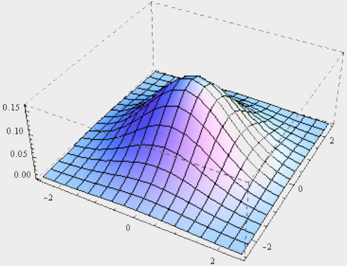
Distance Decay

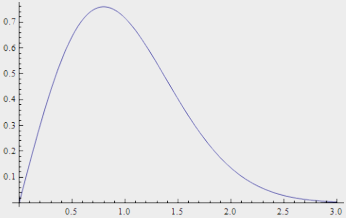
- What do we obtain when we graph not offense distance, but scaled distance?



Distance Decay

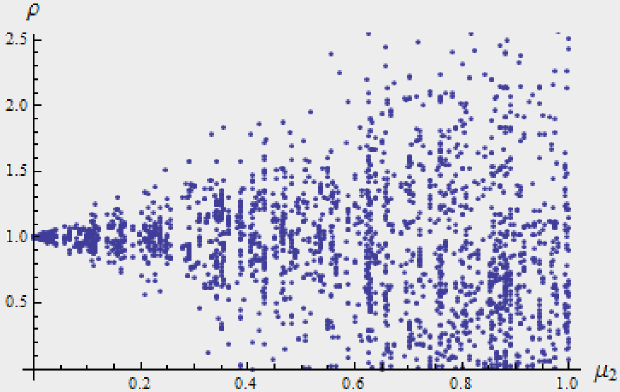
- When considering distance, it is important to realize that it is a *derived* quantity.
  - Offenders do not select a distance- they select a target.
- For example, if the offender selects a target from a two-dimensional normal distribution; then the distribution of distances is a Rayleigh distribution.





Distance Decay

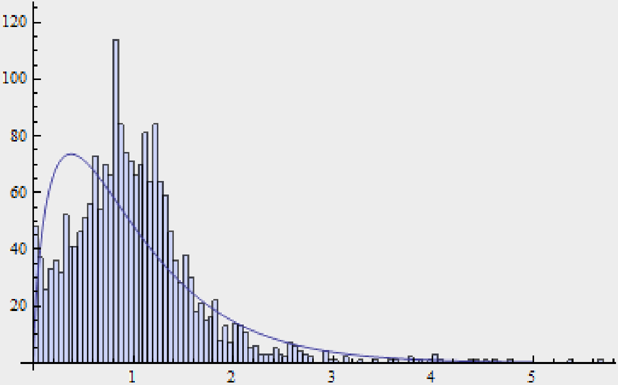
- It is useful to look at the dependence of μ_2 versus ρ for our residential burglars



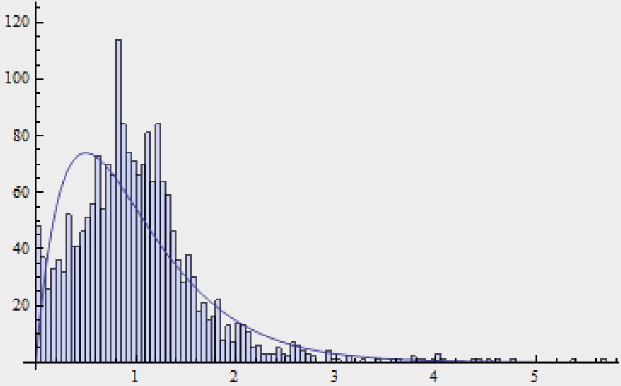
- Commuters ($\mu_2 \approx 0$) exhibit very different behavior than marauders ($\mu_2 \approx 1$).
- Focus our attention only on non-commuters- say $\mu_2 \geqslant 0.25$.

Distance Decay

- If we assume that offenders follow the kinetic model of Mohler and Short, then the distribution of scaled distances should follow a modified Bessel function K_0 with mean 1:

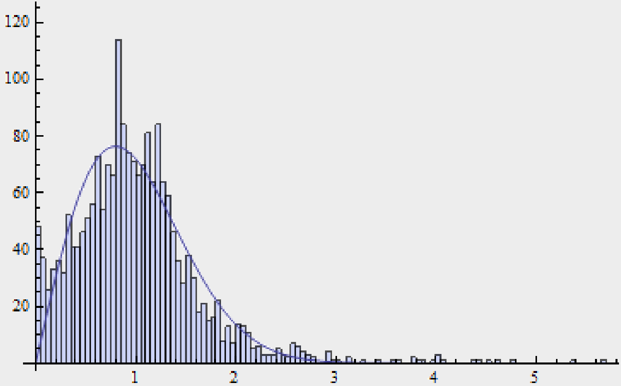


- If instead we assume that the offenders follow a two-dimensional negative exponential, then the distribution of scaled distances follows a Gamma (shape parameter 2, mean 1):



Distance Decay

- If we assume that each offender chooses targets from a two-dimensional normal distribution with their own average offense distance, then the distribution of scaled distances should follow a Rayleigh distribution with mean 1:

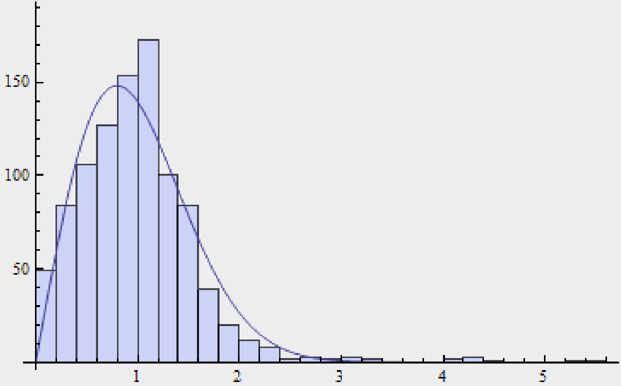


Distance Decay

- It is important to know that no parameters were used to generate this fit.
- If the offender selects a crime site according to a bivariate normal distribution with average offense distance α, (where α may be different for different offenders) then the scaled distances must follow a Rayleigh distribution with average 1.

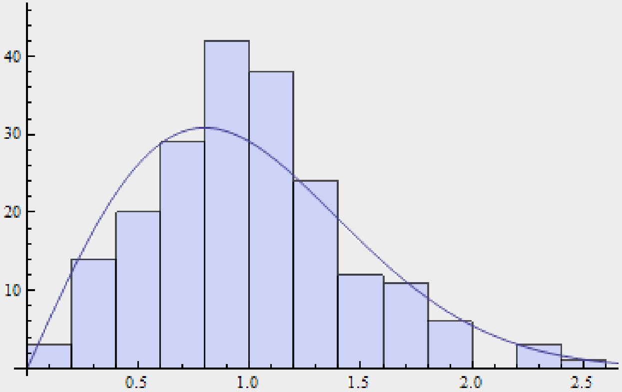
Distance Decay

- The agreement with the Rayleigh distribution does not appear to be happenstance. Here is what occurs for non-residential burglaries with $\mu_2 \geqslant 0.25$



Distance Decay

- Here is what occurs for bank robberies with $\mu_2 \geqslant 0.25$



Distance Decay

- It is possible that these fits are caused by something peculiar to the geography of Baltimore County.
- However, we are not the first to examine scaled distances.
  - Warren, Reboussin, Hazelwood, Cummings, Gibbs, and Trumbetta (1998). *Crime Scene and Distance Correlates of Serial Rape*, Journal of Quantitative Criminology **14** (1998), no. 1, 3559.
  - In that paper, they graphed scaled distances for serial rape:

Distance Decay

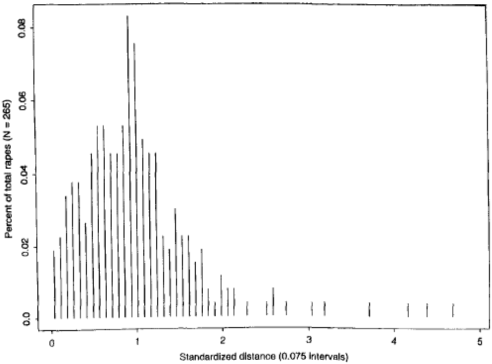
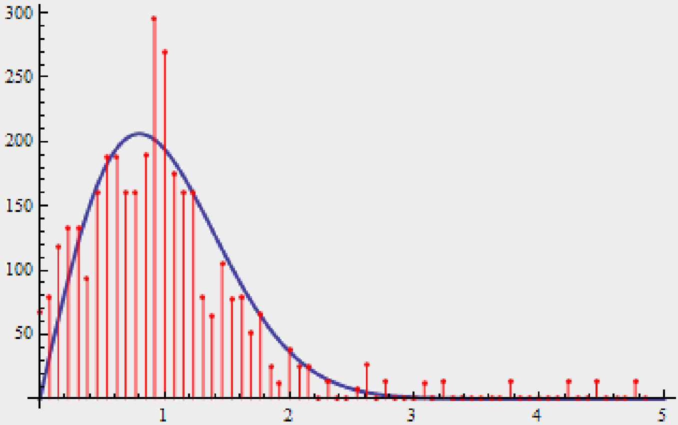


Fig. 2. Proportion of rapes by standardized distance from residence to rape location. Cases with five or more rapes.

two reasons. First, the nonrepresentative nature of the data diminishes the meaningfulness of significance levels. Second, the applied purpose of the paper heightened the need to present the data in a visually clear and practically interpretable form. Distance was found to vary with the demographic characteristics of the offender as well as certain "signature" and "modus

Distance Decay

- Our Rayleigh distribution with mean 1 appears to fit this data as well:

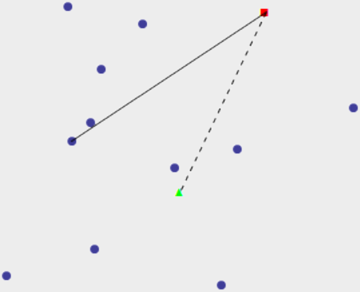


Distance Decay- Caveats

- It is important to note that, though compelling, these graphs do not provide justification that offenders follow a bivariate normal distribution.
  - Agreement is necessary, but not sufficient for this conclusion.
  - There are other two dimensional distributions whose distribution of distances also is Rayleigh.
- We still do not understand the situation yet with commuters.
  - The Warren *et. al.* data is for serial rape, which is known to be well approximated by circle theory- suggesting that this data set may be weighted away from commuters, which our theory does not yet handle.
- Though not presented, we obtained a similar degree of fit using μ_1 instead of μ_2 to characterize commuters and marauders.

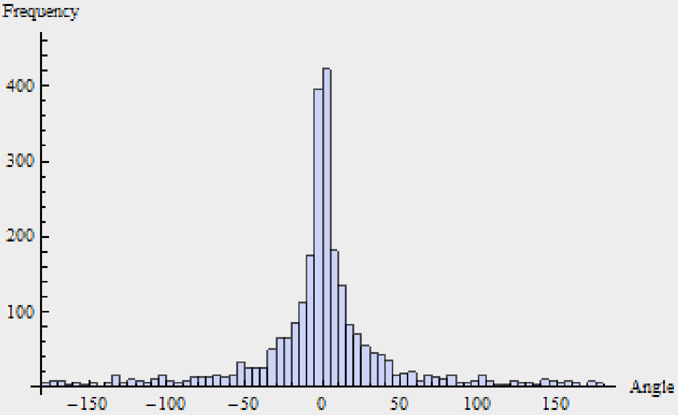
Angular Dependence

- If our idea that the underlying distribution is bivariate normal is correct, then there should be no angular dependence in the results.
- To measure angles, let the blue dots represent crime locations, the red square the anchor point, and the green triangle the centroid of the crime series.
- Then measure the angle between the ray from the anchor point to the crime site and the ray from the anchor point to the centroid.



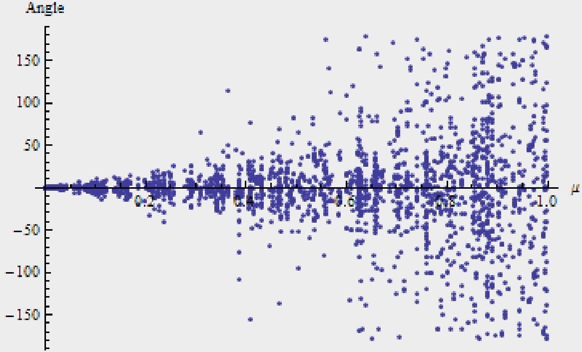
Angular Dependence

- The residential burglary data shows a striking relationship- nearly all of the crime sites lie in the same direction as the centroid.



Angular Dependence

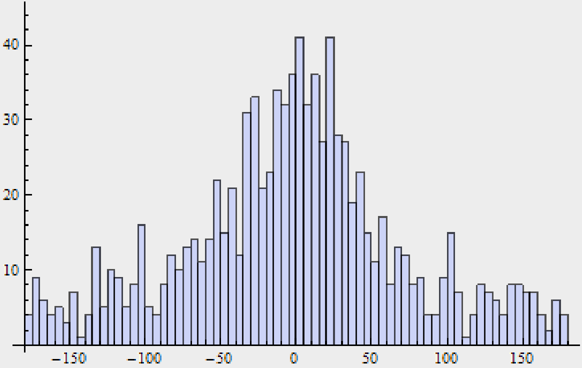
- We can again examine the angular variation as μ_2 varies.



- Even for relatively large values of μ_2, the data is clustered near the zero angle.

Angular Dependence

- The strong central peak reamains, even if we restrict our attention to series with $\mu_2 > 0.7$:



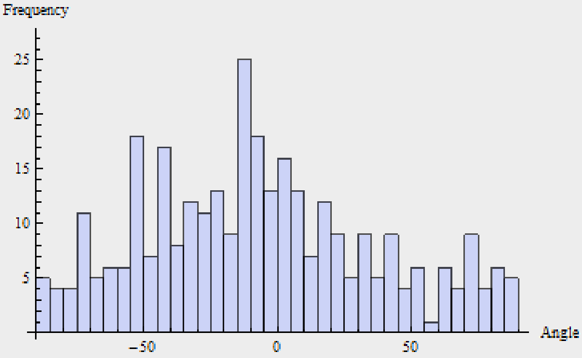
- Note the dramatic changes in the vertical scale between these images!

Angular Dependence

- Clearly there is a strong relationship between the directions the offender took to the different crime sites.
- Moreover, this relationship appears to be strong whether the offender is a commuter or a marauder.
- This suggests that weak information about direction would be more valuable than strong information about distance if one wanted to reduce the area necessary to search for the offender.

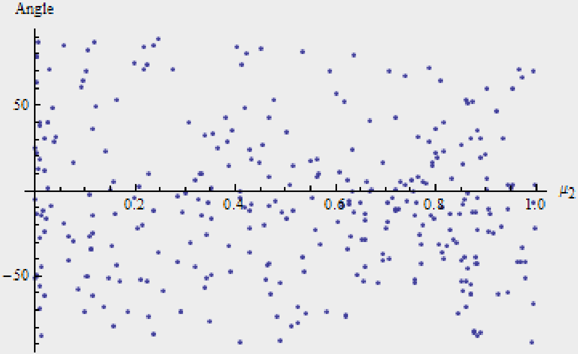
Angular Dependence

- One approach would be to calculate the principal component axes for the set of crime sites.
- It does not appear that there is a strong relationship between the principal axis and the direction to the offender's anchor point in this data.



Angular Dependence

- Moreover, this lack of a relationship persists whether or not the offender is a commuter or a marauder.



- Clearly much more work needs to be done to understand this phenomenon.

Questions?

Mike O'Leary
Department of Mathematics
Towson University
moleary@towson.edu

`http://pages.towson.edu/moleary/Profiler.html`