Point Pattern Analysis: Basics

SERGIO REY

GPH 483/598
Geographic Information Analysis
School of Geographical Sciences
Arizona State University
Spring 2010

Outline

Outline

Outline

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Point Pattern Analysis Objectives

Goals

- Pattern detection
- Assessing the presence of *clustering*
- Identification of individual *clusters*

General Approaches

- Estimate intensity of the process
- Formulating an idealized model and investigating deviations from expectations
- Formulating a stochastic model and fitting it to the data

Outline

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region *A* and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region A and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region A and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region A and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region A and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Spatial Point Pattern

A set of events, irregularly distributed within a region A and presumed to have been generated by some form of stochastic mechanism.

Representation

$\{Y(A), A \subset \Re\}$, where $Y(A)$ is the number of events occurring in area A.

Events, points, locations

Event an occurrence of interest

Point any location in study area

Event location a particular point where an event occurs

Point Pattern Analysis Definitions

Region: A

- Most often planar (two-dimensional Euclidean space)
- One dimensional applications also possible
- Three-dimensional increasingly popular (space + time)
- Point processes on networks (non-planar)

Point Pattern Analysis Definitions

Region: A

- Most often planar (two-dimensional Euclidean space)
- One dimensional applications also possible
- Three-dimensional increasingly popular (space + time)
- Point processes on networks (non-planar)

Point Pattern Analysis Definitions

Region: A

- Most often planar (two-dimensional Euclidean space)
- One dimensional applications also possible
- Three-dimensional increasingly popular (space + time)
- Point processes on networks (non-planar)

Point Pattern Analysis Definitions

Region: A

- Most often planar (two-dimensional Euclidean space)
- One dimensional applications also possible
- Three-dimensional increasingly popular (space + time)
- Point processes on networks (non-planar)

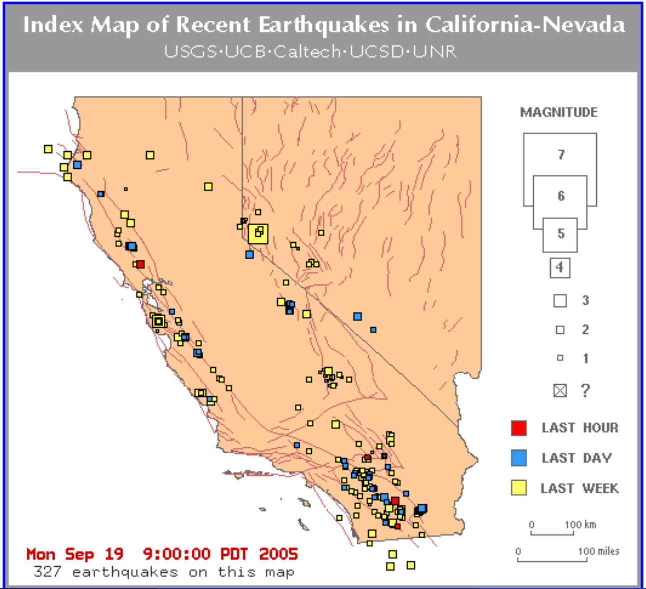
Point Pattern Analysis Definitions

Region: A

- Most often planar (two-dimensional Euclidean space)
- One dimensional applications also possible
- Three-dimensional increasingly popular (space + time)
- Point processes on networks (non-planar)

Space-Time Point Patterns

Space-Time Point Patterns



Point Patterns on Networks

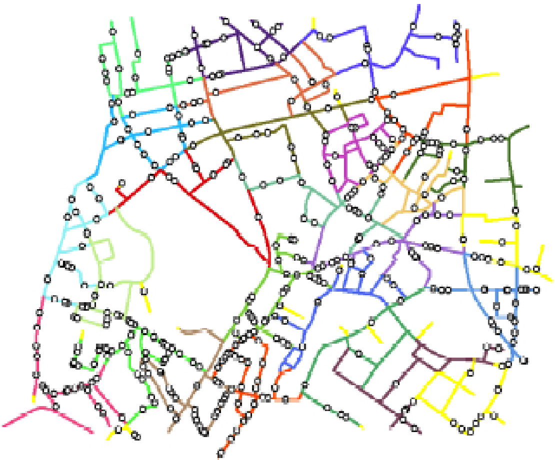


Figure 2: Retail stores assigned to the street network in Shibuya, Tokyo
(cells are indicated by different colors)

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Point Patterns

Unmarked Point Patterns

- Only location is recorded
- Attribute is binary (presence, absence)

Marked Point Patterns

- Location is recorded
- Non-binary stochastic attribute
- e.g., sales at a retail store, dbh of tree

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Realizations

Mapped Point Patterns

- *All* events are recorded and mapped
- Complete enumeration of events
- Full information on the realization from the process

Sampled Point Patterns

- *Sample* of events are recorded and mapped
- Complete enumeration of events impossible or intractable
- Partial information on the realization from the process
- Presence/“absence” data (ecology, forestry)

Outline

Mapped Patterns

not a sample

Selection Bias

- events are mapped
- but non-events are not

Points as Events

Mapped Patterns

not a sample

Selection Bias

- events are mapped
- but non-events are not

Points as Events

Mapped Patterns

not a sample

Selection Bias

- events are mapped
- but non-events are not

Points as Events

Mapped Patterns

not a sample

Selection Bias

- events are mapped
- but non-events are not

Research Questions

Location Only

are points randomly located or patterned

Location and Value

- marked point pattern
- is combination of location and value random or patterned

Both Cases

What is the Underlying Process?

Research Questions

Location Only

are points randomly located or patterned

Location and Value

- marked point pattern
- is combination of location and value random or patterned

Both Cases

What is the Underlying Process?

Research Questions

Location Only

are points randomly located or patterned

Location and Value

- marked point pattern
- is combination of location and value random or patterned

Both Cases

What is the Underlying Process?

Research Questions

Location Only

are points randomly located or patterned

Location and Value

- marked point pattern
- is combination of location and value random or patterned

Both Cases

What is the Underlying Process?

Research Questions

Location Only

are points randomly located or patterned

Location and Value

- marked point pattern
- is combination of location and value random or patterned

Both Cases

What is the Underlying Process?

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Points on a Plane

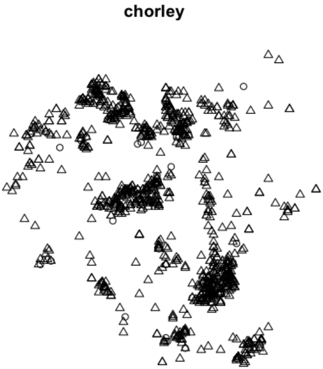
Classic Point Pattern Analysis

- points on an isotropic plane
- no effect of translation and rotation
- classic examples: tree seedlings, rocks, etc

Distance

- no directional effects
- no translational effects
- straight line distance only

Events: Point Map



Points in Context

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Intensity

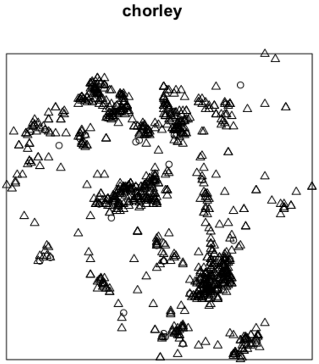
First Moment

- number of points N, area of study $|A|$
- intensity: $\lambda = N/|A|$
- area depends on bounds, often arbitrary

Artificial Boundaries

- bounding box (rectangle, square)
- other (city boundary)

Bounding Box



District Boundary

Convex Hull

Tightest fit

various algorithms

Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1-m/N)}]$
- m: number of vertices of convex hull

Convex Hull

Tightest fit

various algorithms

Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1-m/N)}]$
- m: number of vertices of convex hull

Convex Hull

Tightest fit

various algorithms

Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1 - m/N)}]$
- m: number of vertices of convex hull

Convex Hull

Tightest fit

various algorithms

Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1-m/N)}]$
- m: number of vertices of convex hull

Convex Hull

Tightest fit

various algorithms

Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1-m/N)}]$
- m: number of vertices of convex hull

Convex Hull

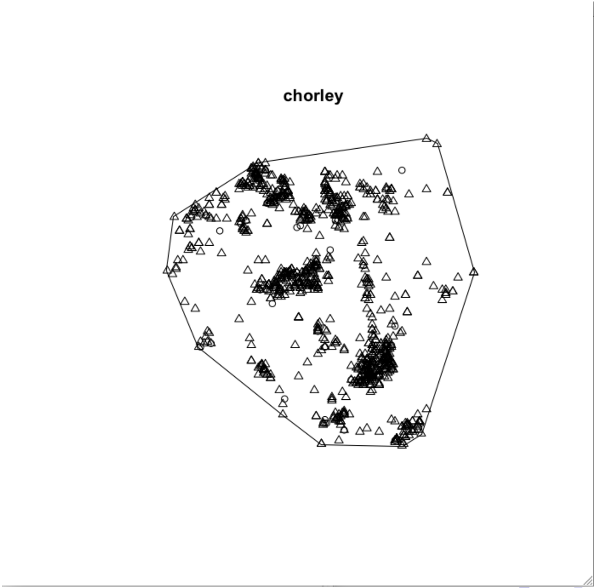
Tightest fit

various algorithms

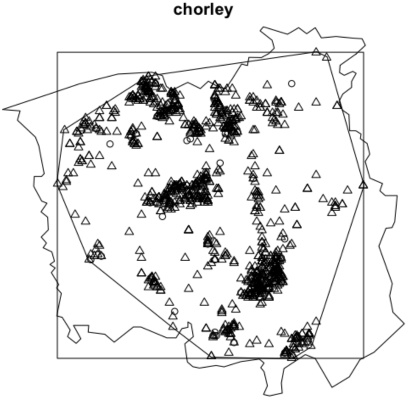
Rescaled Convex Hull (Ripley-Rasson)

- adjust to properly reflect spatial domain of point process
- use centroid of convex hull
- rescale by $1/[\sqrt{(1-m/N)}]$
- m: number of vertices of convex hull

Convex Hull



Multiple Boundaries



Intensity Estimates

	Area	Intensity
	km^2	$cases/km^2$
District Boundary	315.155	3.29
Bounding Box	310.951	3.33
Convex Hull	229.421	4.52

N=1036

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Points on a Network

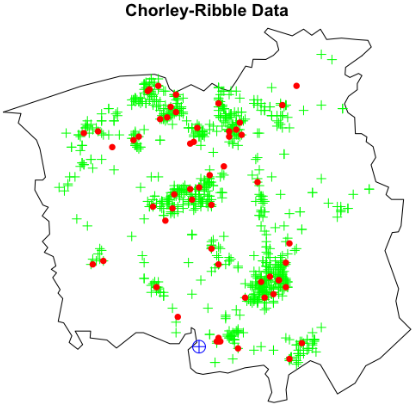
Realistic Location

- addresses
- remove impossible locations (lakes)

Network Distance

- shortest path along network
- Manhattan block distance
- distance vs. travel time or cost

Case-Control Design: Lancashire Cancer



Marked Point Patterns

Both Location and Value

- Patterns in the Location
- Value Associated with Location

Marked Point Patterns

Both Location and Value

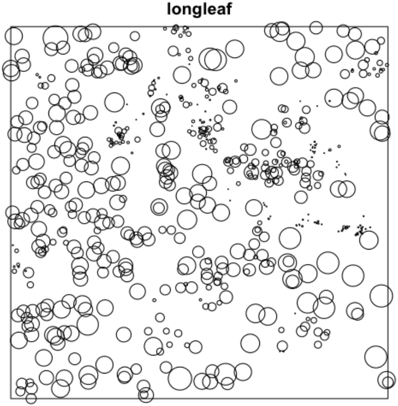
- Patterns in the Location
- Value Associated with Location

Marked Point Patterns

Both Location and Value

- Patterns in the Location
- Value Associated with Location

Marked Point Pattern: Longleaf Pine



Multi-Type Patterns

Multiple Categories

- Patterns in Single Category
- Association between Patterns in Other Categories
- Repulsion or Attraction

Multi-Type Patterns

Multiple Categories

- Patterns in Single Category
- Association between Patterns in Other Categories
- Repulsion or Attraction

Multi-Type Patterns

Multiple Categories

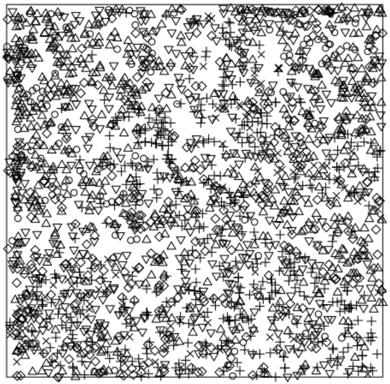
- Patterns in Single Category
- Association between Patterns in Other Categories
- Repulsion or Attraction

Multi-Type Patterns

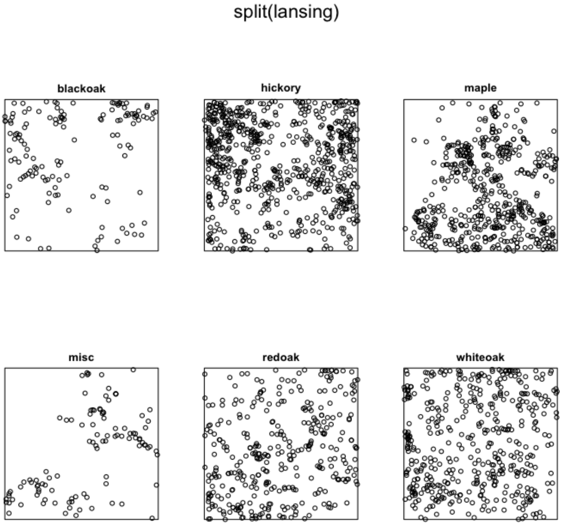
Multiple Categories

- Patterns in Single Category
- Association between Patterns in Other Categories
- Repulsion or Attraction

Multi-Type Pattern: Lansing Woods



Multi-Type Pattern: Lansing Woods



Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

Event Counts

- points aggregated by areal unit
- spatially extensive variable

Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

Areal Aggregation

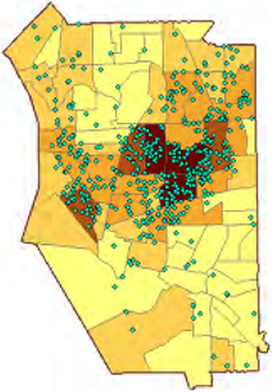
Event Counts

- points aggregated by areal unit
- spatially extensive variable

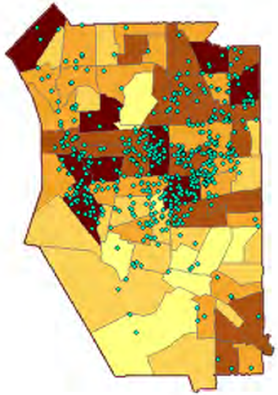
Rates

- events / population at risk
- non-homogeneous population at risk
- risk = probability of an event
- rate is an estimate of underlying risk

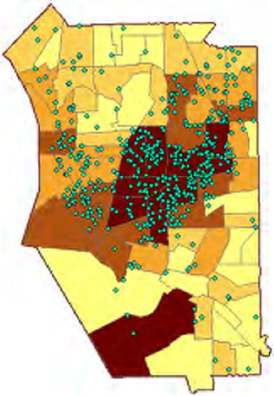
Homicide Counts by Census Tracts



Population Count by Census Tracts



Homicide Rate by Census Tracts



Outline

Central Tendency

Purpose

- Provide a "center point"
- Similar to first moment of a distribution
- "Representative point"

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a “center point”
- Similar to first moment of a distribution
- “Representative point”

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a "center point"
- Similar to first moment of a distribution
- "Representative point"

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a "center point"
- Similar to first moment of a distribution
- "Representative point"

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a "center point"
- Similar to first moment of a distribution
- "Representative point"

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a “center point”
- Similar to first moment of a distribution
- “Representative point”

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a “center point”
- Similar to first moment of a distribution
- “Representative point”

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a “center point”
- Similar to first moment of a distribution
- “Representative point”

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Central Tendency

Purpose

- Provide a "center point"
- Similar to first moment of a distribution
- "Representative point"

Measures

- Mean Center
- Weighted Mean Center
- Median Center
- Center of Minimum Distance

Example Data

i	x_i	y_i	w_i
1	20	40	10
2	30	60	20
3	34	52	10
4	40	40	20
5	44	42	10
6	48	62	80
7	50	10	10
8	60	50	90
9	90	90	100

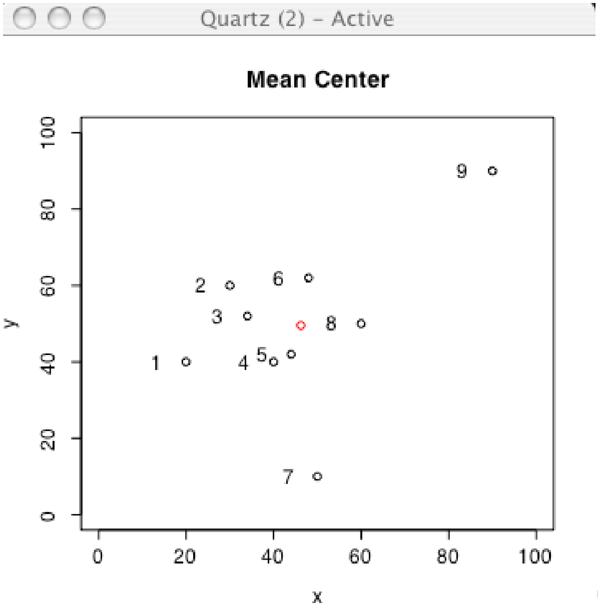
Mean Center

(x_m, y_m)

$$x_m = 1/n \sum_{i=1}^{n} x_i \tag{1}$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \tag{2}$$

Mean Center



Weighted Mean center

(x_m, y_m)

$$x_m = 1/n \sum_{i=1}^{n} x_i \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (3)$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (4)$$

w_i weight

- Marked point patterns
- Continuous mark
- Not categorical mark

Weighted Mean center

(x_m, y_m)

$$x_m = 1/n \sum_{i=1}^{n} x_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{3}$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

w_i weight

- Marked point patterns
- Continuous mark
- Not categorical mark

Weighted Mean center

(x_m, y_m)

$$x_m = 1/n \sum_{i=1}^{n} x_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{3}$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

w_i weight

- Marked point patterns
- Continuous mark
- Not categorical mark

Weighted Mean center

(x_m, y_m)

$$x_m = 1/n \sum_{i=1}^{n} x_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{3}$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \frac{w_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

w_i weight

- Marked point patterns
- Continuous mark
- Not categorical mark

Weighted Mean center

(x_m, y_m)

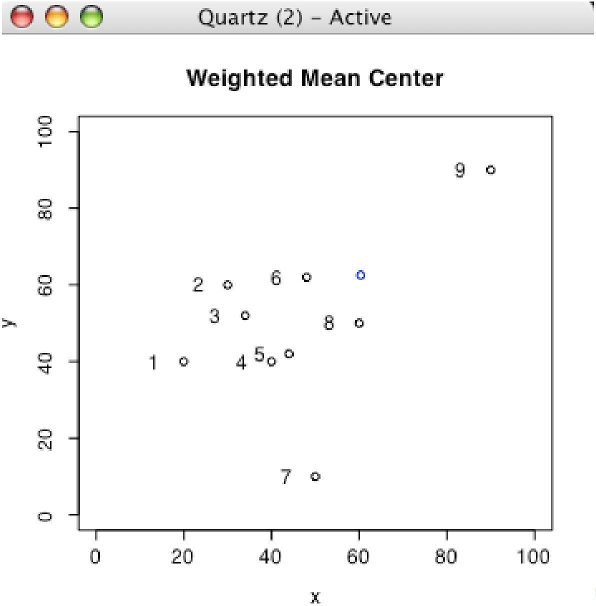
$$x_m = 1/n \sum_{i=1}^{n} x_i \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (3)$$

$$y_m = 1/n \sum_{i=1}^{n} y_i \frac{w_i}{\sum_{i=1}^{n} w_i} \quad (4)$$

w_i weight

- Marked point patterns
- Continuous mark
- Not categorical mark

Weighted Mean Center



Median Center

Definition(s)

English Statistics The intersection of two orthogonal axes, each which has an equal number of points on either side.

American The center of minimum travel.

Median Center

Definition(s)

English Statistics: The intersection of two orthogonal axes, each which has an equal number of points on either side.

American: The center of minimum travel.

Median Center

Definition(s)

English Statistics The intersection of two orthogonal axes, each which has an equal number of points on either side.

American The center of minimum travel.

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \tag{5}$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \quad (5)$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \qquad (5)$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \qquad (5)$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \tag{5}$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \quad (5)$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \quad (5)$$

Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

English Median Center

Manhattan Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} |x_i - x_m| + |y_i - y_m| \tag{5}$$

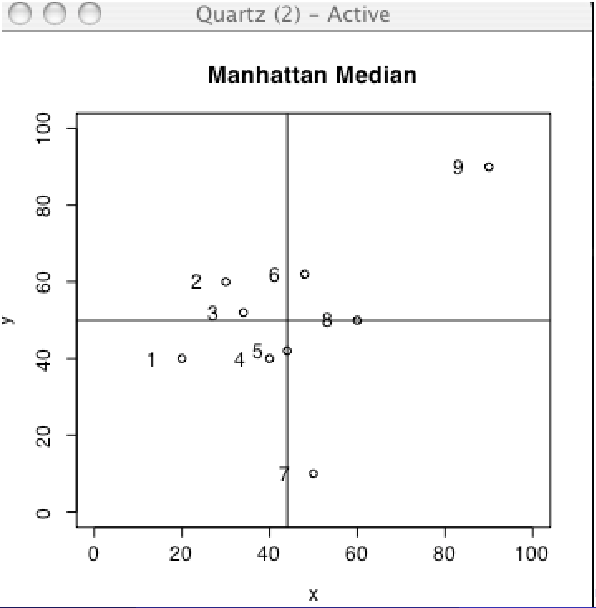
Advantages

- Can be found very quickly
- No calculations are typically required (other than intersection)

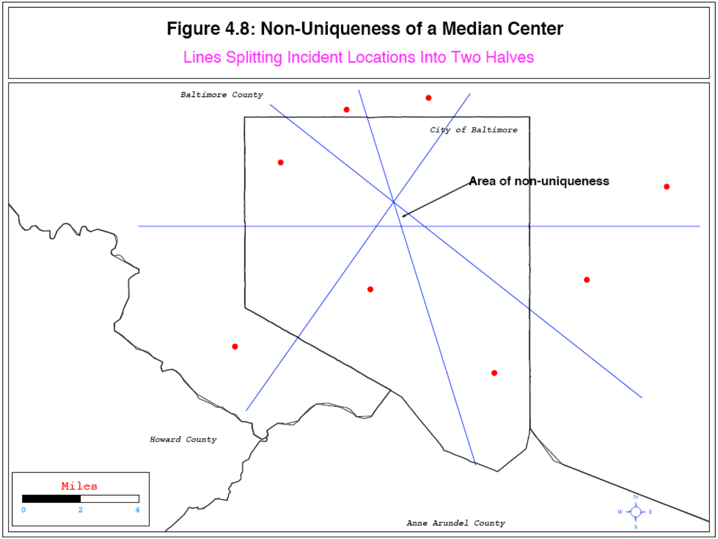
Disadvantage

- Never unique with even n
- Always unique with odd n
- Not unique under axis rotation

Manhattan Median



Non-Uniqueness



Center of Minimum Travel

Euclidean Median

The location from which the sum of the Euclidean distances to all points in a distribution is a minimum.

Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \tag{6}$$

Weighted Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \frac{w_i}{\sum_{i=1}^{n} w_i} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \tag{7}$$

Center of Minimum Travel

Euclidean Median

The location from which the sum of the Euclidean distances to all points in a distribution is a minimum.

Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \quad (6)$$

Weighted Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \frac{w_i}{\sum_{i=1}^{n} w_i} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \quad (7)$$

Center of Minimum Travel

Euclidean Median

The location from which the sum of the Euclidean distances to all points in a distribution is a minimum.

Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \tag{6}$$

Weighted Euclidean Median

$$Min\ f(x_m, y_m) = \sum_{i=1}^{n} \frac{w_i}{\sum_{i=1}^{n} w_i} \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \tag{7}$$

Euclidean Median

Weber Problem

Find the optimal location for a factory: one that minimizes transport costs between sources of raw materials and delivery to the market.

Solutions

- No closed form solution
- First iterative alogrithm: Kuhn and Kuenne (1962)
- Important for more general location allocation problems

Euclidean Median

Weber Problem

Find the optimal location for a factory: one that minimizes transport costs between sources of raw materials and delivery to the market.

Solutions

- No closed form solution
- First iterative alogrithm: Kuhn and Kuenne (1962)
- Important for more general location allocation problems

Euclidean Median

Weber Problem

Find the optimal location for a factory: one that minimizes transport costs between sources of raw materials and delivery to the market.

Solutions

- No closed form solution
- First iterative alogrithm: Kuhn and Kuenne (1962)
- Important for more general location allocation problems

Euclidean Median

Weber Problem

Find the optimal location for a factory: one that minimizes transport costs between sources of raw materials and delivery to the market.

Solutions

- No closed form solution
- First iterative alogrithm: Kuhn and Kuenne (1962)
- Important for more general location allocation problems

Euclidean Median

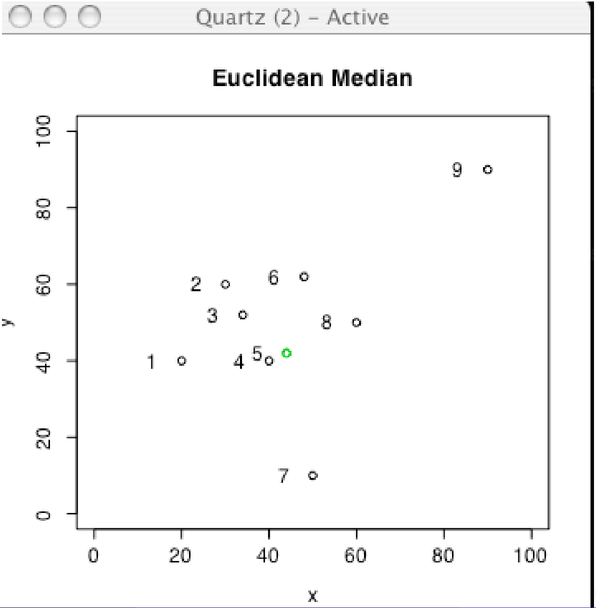
Weber Problem

Find the optimal location for a factory: one that minimizes transport costs between sources of raw materials and delivery to the market.

Solutions

- No closed form solution
- First iterative alogrithm: Kuhn and Kuenne (1962)
- Important for more general location allocation problems

Euclidean Median



Outline

Dispersion and Orientation

Measures

- Standard Distance
- Major/minor axes
- Standard Deviational Ellipse

Dispersion and Orientation

Measures

- Standard Distance
- Major/minor axes
- Standard Deviational Ellipse

Dispersion and Orientation

Measures

- Standard Distance
- Major/minor axes
- Standard Deviational Ellipse

Dispersion and Orientation

Measures

- Standard Distance
- Major/minor axes
- Standard Deviational Ellipse

Standard Distance

Euclidean Based

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_m)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_m)^2}{n}} \quad (8)$$

Uses

- Similar to standard deviation
- Combine with Mean Center for "outlier detection"
- Sensitive to extreme values

Standard Distance

Euclidean Based

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_m)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_m)^2}{n}} \tag{8}$$

Uses

- Similar to standard deviation
- Combine with Mean Center for "outlier detection"
- Sensitive to extreme values

Standard Distance

Euclidean Based

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_m)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_m)^2}{n}} \tag{8}$$

Uses

- Similar to standard deviation
- Combine with Mean Center for “outlier detection”
- Sensitive to extreme values

Standard Distance

Euclidean Based

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_m)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_m)^2}{n}} \tag{8}$$

Uses

- Similar to standard deviation
- Combine with Mean Center for “outlier detection”
- Sensitive to extreme values

Standard Distance

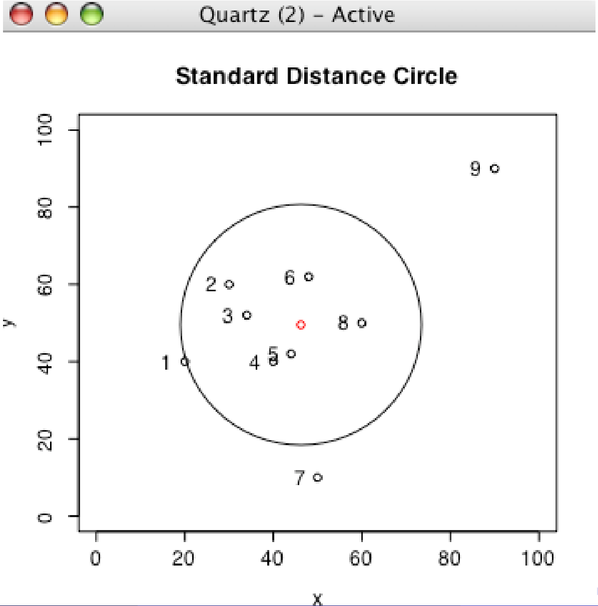
Euclidean Based

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_m)^2}{n} + \frac{\sum_{i=1}^{n}(y_i - y_m)^2}{n}} \qquad (8)$$

Uses

- Similar to standard deviation
- Combine with Mean Center for “outlier detection”
- Sensitive to extreme values

Standard Distance Circle



Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Relative to Standard Distance

- Measures dispersion
- Sensitive to *shape* of distribution
- Measures dispersion in two dimensions

Components

- Angle of rotation
- Dispersion along major axis
- Dispersion along minor axis

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Standard Deviational Ellipse

Major, minor axes

- Major axis defines the direction of maximum spread in the distribution
- Minor axis is orthogonal to major axis
- Minor axis defines the direction of minimum spread

Steps

1. Determine rotation angle of Y-axis
2. Calculate standard deviations for transposed axes
3. Determine length of axes
4. Determine area of the ellipse

Rotation Angle Θ

$$\Theta = ARCTAN\left\{\left(\sum_i (x_i - \bar{x})^2 - \sum_i (y_i - \bar{y})^2\right) + \left[\left(\sum_i (x_i - \bar{x})^2 - \sum_i (y_i - \bar{y})^2\right)^2 + 4\left(\sum_i (x_i - \bar{x})(y_i - \bar{y})\right)^2\right]^{1/2}\right\} / 2\sum_i (x_i - \bar{x})(y_i - \bar{y})$$

Standard Deviations On Transposed Axes

S_x

$$S_x = \sqrt{2\frac{(\sum_{i=1}^{n}(x_i - \bar{x})Cos(\Theta) - \sum_{i=1}^{n}(y_i - \bar{y})Sin(\Theta))^2}{n-2}} \quad (9)$$

S_y

$$S_y = \sqrt{2\frac{(\sum_{i=1}^{n}(x_i - \bar{x})Sin(\Theta) - \sum_{i=1}^{n}(y_i - \bar{y})Cos(\Theta))^2}{n-2}} \quad (10)$$

Standard Deviations On Transposed Axes

S_x

$$S_x = \sqrt{2\frac{(\sum_{i=1}^{n}(x_i - \bar{x})Cos(\Theta) - \sum_{i=1}^{n}(y_i - \bar{y})Sin(\Theta))^2}{n-2}} \tag{9}$$

S_y

$$S_y = \sqrt{2\frac{(\sum_{i=1}^{n}(x_i - \bar{x})Sin(\Theta) - \sum_{i=1}^{n}(y_i - \bar{y})Cos(\Theta))^2}{n-2}} \tag{10}$$

Ellipse Axes

Lengths

$$L_x = 2S_x \tag{11}$$

$$L_y = 2S_y \tag{12}$$

Mid Point

Mean Center of Point Pattern (x_m, y_m)

Area

$$A = \pi S_x S_y \tag{13}$$

Ellipse Axes

Lengths

$$L_x = 2S_x \tag{11}$$

$$L_y = 2S_y \tag{12}$$

Mid Point

Mean Center of Point Pattern (x_m, y_m)

Area

$$A = \pi S_x S_y \tag{13}$$

Ellipse Axes

Lengths

$$L_x = 2S_x \tag{11}$$

$$L_y = 2S_y \tag{12}$$

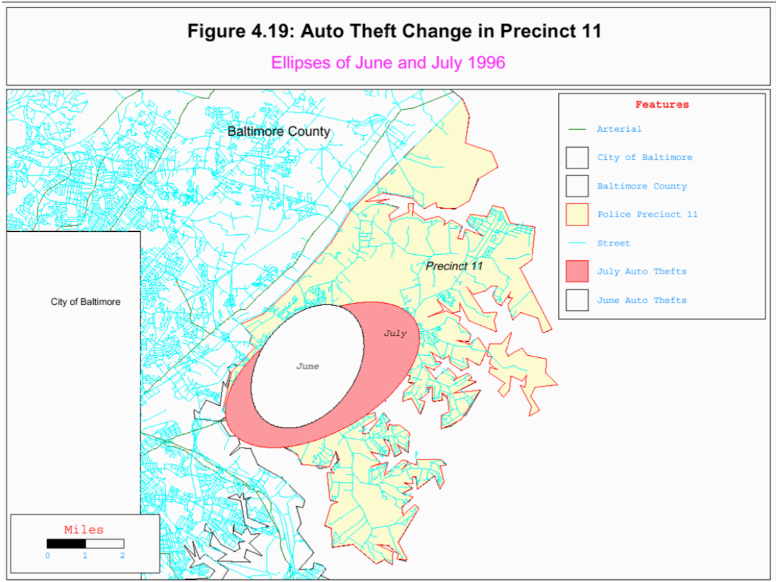
Mid Point

Mean Center of Point Pattern (x_m, y_m)

Area

$$A = \pi S_x S_y \tag{13}$$

Standard Deviational Ellipse



Outline

Shape Analysis of Point Patterns

Geometry

- Bounding Box
- Convex Hulls

Shape Analysis of Point Patterns

Geometry

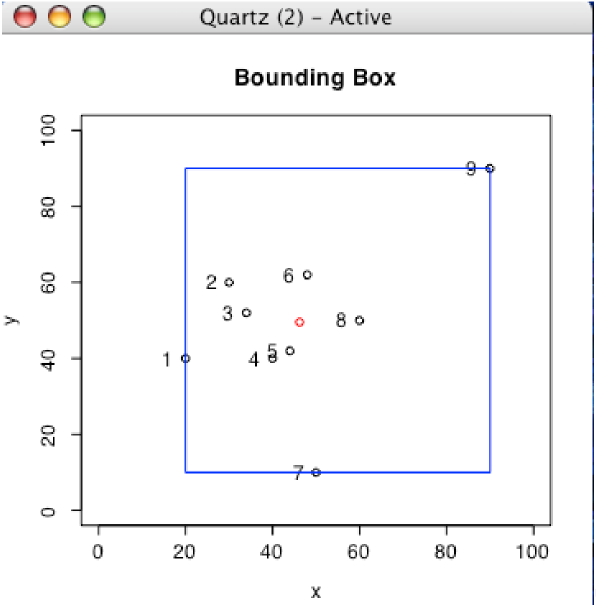
- Bounding Box
- Convex Hulls

Shape Analysis of Point Patterns

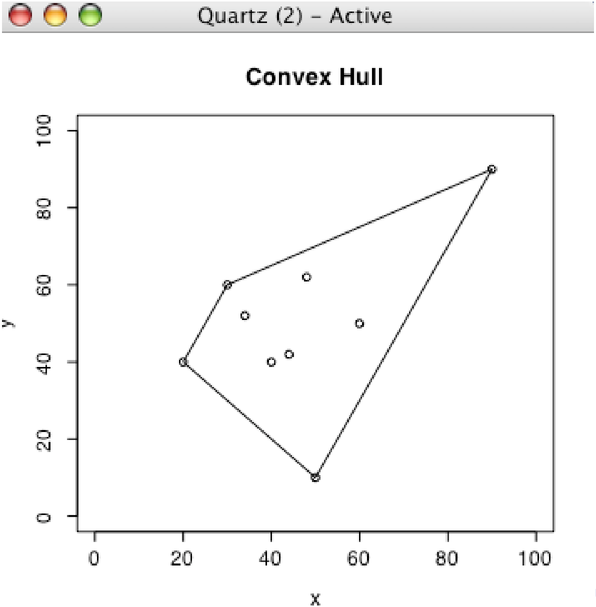
Geometry

- Bounding Box
- Convex Hulls

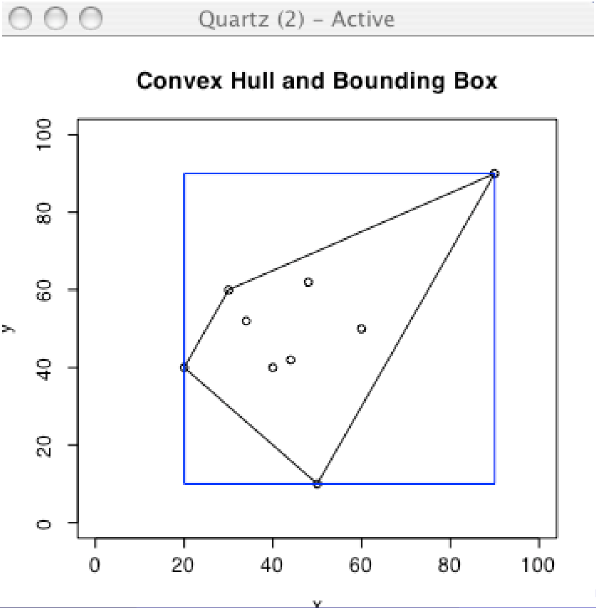
Bounding Box



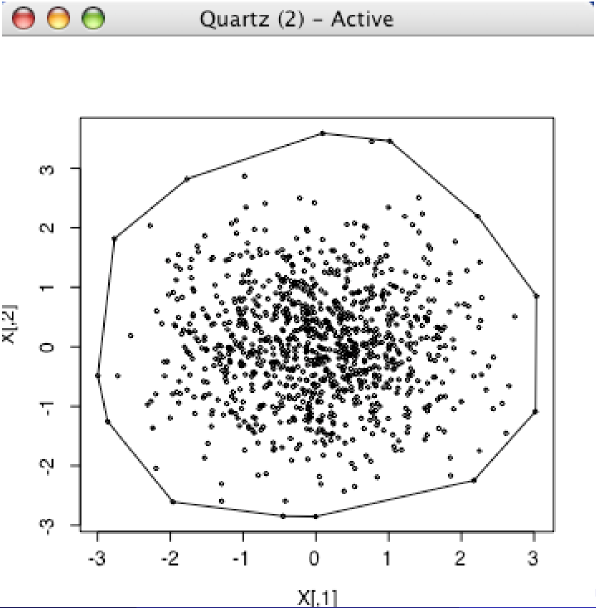
Convex Hull



Convex Hull and Bounding Box



Convex Hull (Large n)



Nested Convex Hulls