Key Issues to Consider Analyzing Point Pattern

- The location of each point is considered to be random
- Processes generating different random point patterns
  - Clustering (attraction) – random – regular (repulsion)
  - Parent-offspring processes
- Inhomogenous or homogenous underlying intensity within the study area
  - Lack of first order stationarity
  - Modeling first order effects before investigating second order effects
- Handling of edge effects and spatial scale.
- Do the random point locations have attached ***marks*** (point attributes, e.g., plant type or canopy diameter), which inform our understanding of the underlying spatial process?
- Which locational definition reflects the underlying data generating process?
  - Do we record where a crime happens or where the victims or offenders live?
  - Do we record the home, hospital or the work address of a person with a disease?
  - This two-reference points perspective may lead to a model of random lines connecting the two points.

Inhomogeneous means the expectation of all points are not a constant

Simple technique: Density based Quadrate Analysis

- Point pattern: Agglomerative advantages (clustering) versus spatial competition/repulsion (regular pattern maximizes the distance among points).

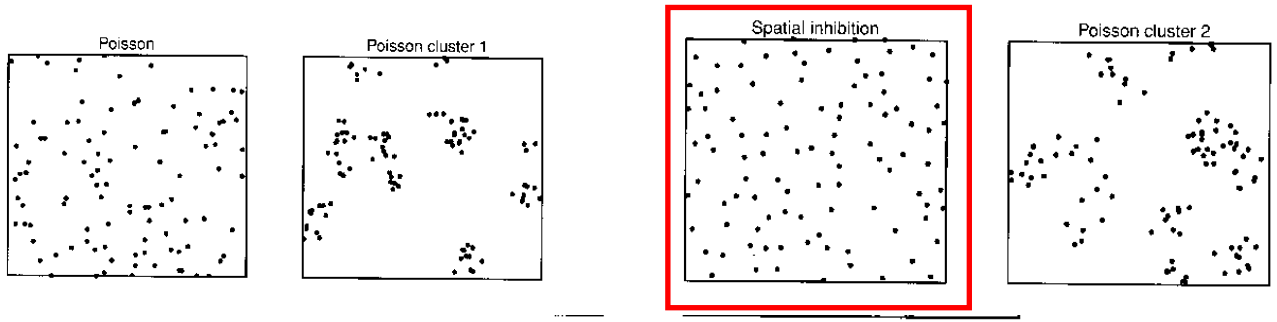


Figure 1.18 Different point pattern processes generating different types of random patterns: Poisson process (i.e. complete simple randomness); Poisson clusters (1 and 2 using different parameter values for spacing among clusters); spatial inhibition.

- *Assumption:* Constant first order effect (i.e., homogeneity), that is, the intensity of point process constant at each location
- *Underlying principle:* Overlay a grid with fixed grid cell size over the point pattern and count the number of points within each grid cell.
  Depending on the ***underlying data generating process*** the distribution of the point counts should follow a ***theoretical reference distribution***.
- *Problems with reference grid generation:* [a] boundary cells in irregular shaped study areas, [b] cell size and [c] shifts in the overlay grid pattern (spatial scale and aggregation problem)

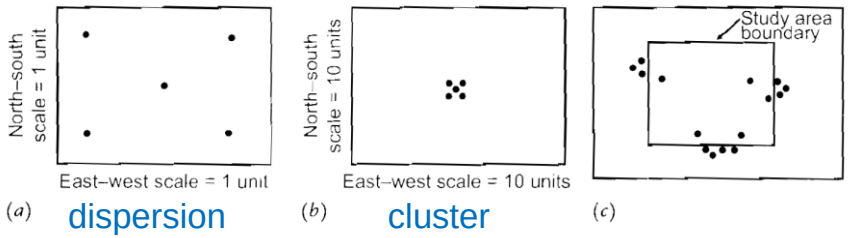


FIGURE 14-4. Study area boundaries.

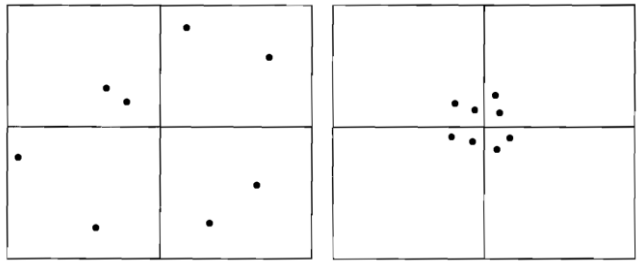


FIGURE 14-7. Uniform point patterns from quadrat analysis.

So it depends on the spatial unit

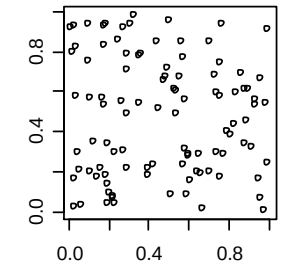
Reference distribution under the assumption of complete spatial randomness This is reference

- **Def. Complete spatial randomness**: Each location of the study area has equal probability of receiving an event and events are mutually independent, that is, the location of events do not influence each other.
- ***Simulation*** of a random point pattern under ***complete spatial randomness*** *(CSR)*: Select ***independently*** and ***randomly*** the latitude and the longitude of a point from a ***uniform distribution*** (Note: the spatial extent should match the bounding box of the study area)
- Under ***complete spatial randomness*** one can show that the number of points within a grid should theoretically follow a Poisson distribution (see derivation below):

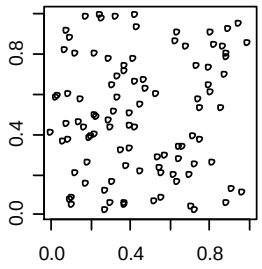
  $\Pr(X = x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$ with the constant ***intensity*** $\hat{\lambda} = \frac{\#\text{ of points in study region}}{\#\text{ of grid cells}}$ that we expect for each grid cell.
- Note: the expectation $\bar{x}$ and variance s^2 of the Poisson distribution are both equal to the intensity λ.
- This property leads to the variance-mean ratio test statistic $VMR = \frac{s^2}{\bar{x}}$ where the mean is $\bar{x} = \hat{\lambda}$ and the observed variance is $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$ with x_i the number of points in the i^{th} grid cell and n is the number of grid cells.
- Interpretation: [a] if the points are more clustered s^2 increases (a few cells with many points and many cells with a few points) and [b] if the points are regular spaced s^2 decreases because the number of points within a cell will be equal to the expected value.
- Thus $VMR > 1$ indicates spatial clustering and $VMR < 1$ points to regular spaced observations.

Evaluating the distribution of the VMR under the assumption of CSR by simulation experiments

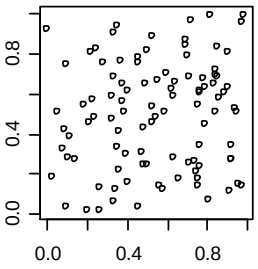
- Generate a set of CSR pattern in a simulation experiment.
- Calculate the *VMR* for each simulated pattern.
- Evaluate the distribution of the *VMR*. We expect that *VMR* varies around one.
    - Is the *VMR* of an observed empirical pattern in the lower tail then we can reject the zero hypothesis of CSR in favor for a more uniformly (disperse or spatial inhabitation) data generating process.
    - Is the *VMR* of an observed empirical pattern in the upper tail then we can reject the zero hypothesis of CSR in favor for a more clustered data generating process.
- Example with 200 simulations of 100 CSR points (see program **`SimulCSRQuad.r`**):



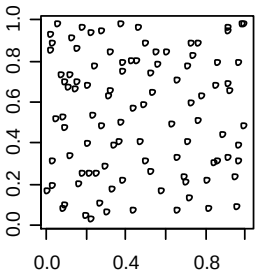
Simulated pattern 120 with VMR: 0.99



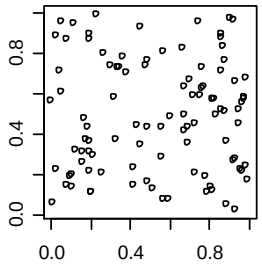
Simulated pattern 160 with VMR: 0.97



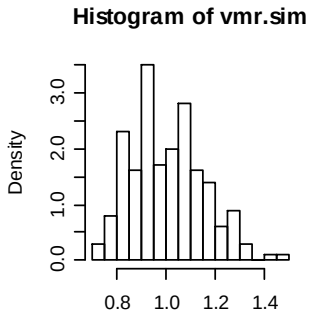
Simulated pattern 190 with VMR: 1.09



min VMR (disperse): 0.71



max VMR (clustered): 1.45



Distribution of VMR

Deriving the Poisson Distribution for CSR Point Patterns

- See also Chapter 5 in Burt, Barber and Rigby for the an empirical derivation of the Poisson distribution for the Binomial distribution
- Assume, without loss of generality, that we expect to observe one event in a unit square (i.e., $\lambda = 1$), but due to random variation the number of points may vary from cell to cell.
- Now let us split these unit squares into 4 equal sized sub-squares, then the expected value of observing an event in a quarter square becomes $\pi = \frac{1}{4}$.

  Its complement of ***not*** observing an event is $\pi^C = 1 - \frac{1}{4} = \frac{3}{4}$
- Furthermore, let us assume that at most one event (or none) can be observed in a quarter-square, i.e., the event is rare.
- Then there are in total 5 possible outcomes:
  [1] no quarter square has an event
  [2] one quarter square has an event, the remaining ones have none
  [3] two quarter squares have each an event, the remaining ones have none
  [4] three quarter squares have each an event, the last one has none
  [5] four quarter squares have each an event
- For instance, for two events (any two quarter-squares receiving each one event) we have the 6 following possible samples (sampling without replacement irrespectively of the order in which the two events have

been drawn) with their associated probabilities:

$$\Pr[1,1,0,0] = \pi \cdot \pi \cdot (1-\pi) \cdot (1-\pi) = \pi^2 \cdot (1-\pi)^{4-2}$$
$$\Pr[1,0,1,0] = \pi \cdot (1-\pi) \cdot \pi \cdot (1-\pi) = \pi^2 \cdot (1-\pi)^{4-2}$$
$$\Pr[1,0,0,1] = \pi \cdot (1-\pi) \cdot (1-\pi) \cdot \pi = \pi^2 \cdot (1-\pi)^{4-2}$$
$$\Pr[0,1,1,0] = (1-\pi) \cdot \pi \cdot \pi \cdot (1-\pi) = \pi^2 \cdot (1-\pi)^{4-2}$$
$$\Pr[0,1,0,1] = (1-\pi) \cdot \pi \cdot (1-\pi) \cdot \pi = \pi^2 \cdot (1-\pi)^{4-2}$$
$$\Pr[0,0,1,1] = (1-\pi) \cdot (1-\pi) \cdot \pi \cdot \pi = \pi^2 \cdot (1-\pi)^{4-2}$$

binomial distribution

- Pooling these probabilities together we get the associate binomial probability for of 2 events out of 4

$$\Pr(\#\text{points} = 2) = \underbrace{\frac{4!}{2! \cdot (4-2)!}}_{=6} \cdot \pi^2 \cdot (1-\pi)^{(4-2)} = 0.2109 \text{ with } \pi = \tfrac{1}{4}$$

- Let us continue splitting the study area up further:

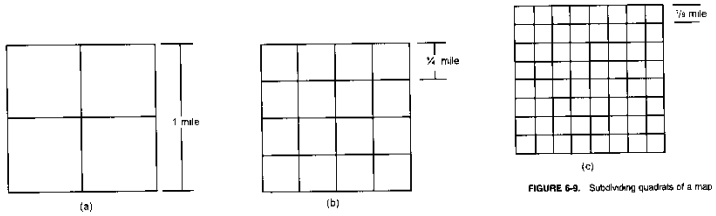


FIGURE 6-9. Subdividing quadrats of a map

- If we split the unit square into 16 equal sized sub-squares then the probability of observing 2 events becomes

$$\Pr(2) = \frac{16!}{2! \cdot (16-2)!} \cdot \pi^2 \cdot (1-\pi)^{(16-2)} = 0.1899 \text{ with } \pi = \tfrac{1}{16}$$

- If we split the unit square further into 64 equal sized sub-squares then the probability of observing 2 events becomes

$$\Pr(2) = \frac{64!}{2! \cdot (64-2)!} \cdot \pi^2 \cdot (1-\pi)^{(64-2)} = 0.1854 \text{ with } \pi = \tfrac{1}{64}$$

- Ultimately, if we split the unit square into *n* equal sized sub-squares then the probability of observing 2 events becomes

$$\Pr(2) = \frac{n!}{2! \cdot (n-2)!} \cdot \left(\frac{1}{n}\right)^2 \cdot \left(\frac{n-1}{n}\right)^{(n-2)} \text{ with } \pi = \tfrac{1}{n}$$

- Summing up: This is how to deduce the equation of Poisson distribution

n	4	16	64	256	1024
BinomDist(2|n, $\frac{1}{n}$ *)*	0.21094	0.18991	0.18539	0.18430	0.18403

- As we make the sub-divisions ever finer then *n* goes to infinity and we get

$$\Pr(X = x) = \lim_{n\to\infty}\left\{\frac{n!}{x! \cdot (n-x)!} \cdot \left(\frac{1}{n}\right)^x \cdot \left(\frac{n-1}{n}\right)^{(n-x)}\right\} = \frac{\exp(-\lambda) \cdot \lambda^x}{x!}$$

For $X = 2$ we get $\Pr(X = 2) = 0.1839$ with $\lambda = n \cdot \pi$ and $\pi = \frac{\lambda}{n}$ if we assume that the probability of observing one point in the original unit square was one.

Derivation of the Asymptotic Poisson Distribution (not test relevant)

- To evaluate the limit and come up with the Poisson distribution takes some basic algebra:

- Set $\pi = \frac{\lambda}{n}$ then

$$\Pr\left(X = x \mid n, \frac{\lambda}{n}\right) = \frac{n!}{(n-x)!\cdot x!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1 - \frac{\lambda}{n}\right)^{(n-x)}$$

$$= \frac{n!}{(n-x)!} \cdot \frac{1}{n^x} \cdot \left(1 - \frac{\lambda}{n}\right)^{-x} \cdot \frac{\lambda^x}{x!} \cdot \left(1 - \frac{\lambda}{n}\right)^n$$

$$= n \cdot (n-1) \cdots (n-x+1) \cdot \frac{1}{n^x} \cdot \left(1 - \frac{\lambda}{n}\right)^{-x} \cdot \frac{\lambda^x}{x!} \cdot \left(1 - \frac{\lambda}{n}\right)^n$$

Taking the limit we get the Poisson distribution

$$\underbrace{\lim_{n\to\infty} \frac{n \cdot (n-1) \cdots (n-x+1)}{n^x}}_{=1} \cdot \underbrace{\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-x}}_{=1} \cdot \frac{\lambda^x}{x!} \cdot \underbrace{\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n}_{=\exp(-\lambda)}$$

- Key assumptions are
  - The number of the events in two mutually exclusive units is independent (comes from multiplicative expression of the binomial distribution)
  - The probability of the occurrence of an event in a small unit is small and proportionally to the size of the unit, i.e., the event is rare (i.e., $\pi = \frac{\lambda}{n}$)
  - The probability of two or more events in just *one* small unit is near zero
- Show the Poisson distribution with the code: **SIMULPOINTPATBARPLOT.R**

The Moments of the Poisson Distribution (not test relevant)

- The expected value of the Poisson distribution is

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{\exp(-\lambda) \cdot \lambda^x}{x!} = \lambda \text{ where } x \in \{0,1,2,\ldots,\infty\}$$

Proof: $\sum_{x=0}^{\infty} x \cdot \frac{\exp(-\lambda) \cdot \lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{\exp(-\lambda) \cdot \lambda^{x-1} \cdot \lambda}{(x-1)!} = \lambda \cdot \underbrace{\sum_{x=1}^{\infty} \frac{\exp(-\lambda) \cdot \lambda^{x-1}}{(x-1)!}}_{=1}$

because the sum of the probabilities over the support is 1.

- The variance of a Poisson distribution is $Var(X) = \sum_{x=0}^{\infty} \left(x - E(X)\right)^2 \cdot \frac{\exp(-\lambda) \cdot \lambda^x}{x!} = \lambda$

Descriptive Frey Plot

- Concept of the Fry plot to identify clustering, CSR, and regular patterns

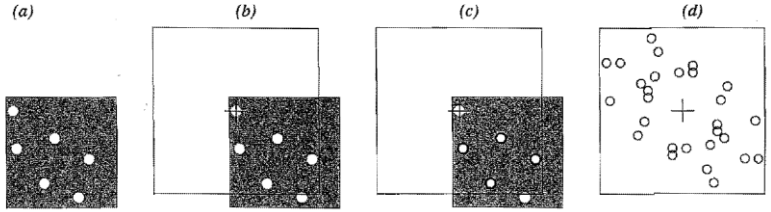


Figure 7.3. *Stages in forming the Fry plot.* (a): *data point pattern, printed on paper.* (b): *transparency superimposed on point pattern so that centre of transparency (+) lies above the first data point.* (c): *other data points are copied (○) onto the transparency.* (d): *final result.*

In mathematical terms the Fry plot is a scatterplot of the vector differences $x_j - x_i$ between all pairs of distinct points in the pattern.

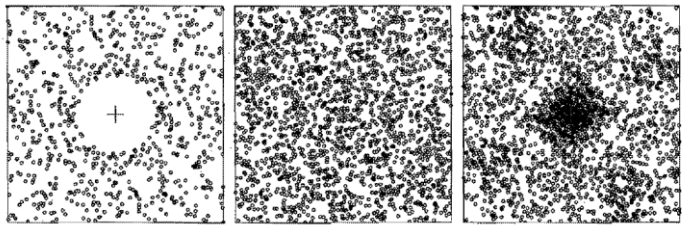


Figure 7.4. *Fry plots for the three patterns in Figure 7.1.*

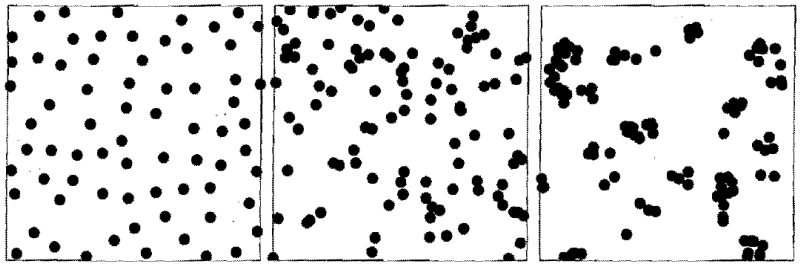


Figure 7.1. *Classical trichotomy between regular* (Left), *independent* (Middle), *and clustered* (Right) *point patterns. All three patterns are in the unit square.*

Grab every points into the center of sheet, do not record the center point but others,
then do this for all points in the data set

Notation of Point Processes

Symbol	Meaning
$\mathbf{s}_i$	Coordinates of point i: $\mathbf{s}_i = (x_{i1}, x_{i2})^T$
R	Study region of any shape
$R = \lvert \mathrm{R} \rvert$	Area of the study region R
A	Sub-set of the study region $\mathrm{A} \subseteq \mathrm{R}$
$Y(\mathrm{A})$	Number of events in sub-region A
$E[Y(\mathrm{A})]$	Expected number of events in sub-region A
$Cov[Y(\mathrm{A}_i), Y(\mathrm{A}_j)]$	Covariance between the number of events in sub-regions A_i and A_j where A_i and A_j are centered around $\mathbf{s}_i$ and $\mathbf{s}_j$, respectively

- The intensity at a point $\mathbf{s}$ is derived by shrinking the area $d\mathrm{A}$ into the point $\mathbf{s}$

$$\lambda(\mathbf{s}) = \lim_{d\mathrm{A} \to \mathbf{s}} \left\{ \frac{E[Y(d\mathrm{A})]}{d\mathrm{A}} \right\}$$

- If $\lambda(\mathbf{s})$ is constant for all points $\mathbf{s}$ then the point pattern process is first order stationary.
- Concept of independence: The number of events in one area does not influence the number of events in another areas.

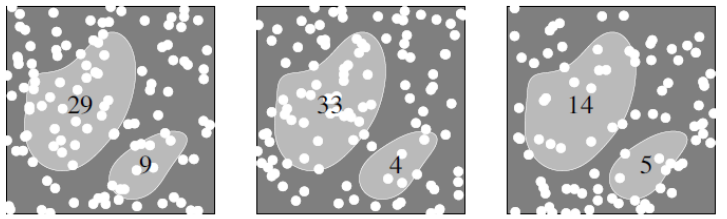


Figure 5.7. *Concept of independence between outcomes in disjoint regions (light grey shading).*

- And the second order intensity is

$$\gamma(\mathbf{s}_i,\mathbf{s}_j) = \lim_{dA_i \to \mathbf{s}_i, dA_j \to \mathbf{s}_j} \left\{ \frac{E\left[Y(dA_i) \cdot Y(dA_j)\right]}{dA_i \cdot dA_j} \right\}$$

ideally equal to 0

> 0 means attraction

- If the second order intensity depends only on the distance $h = d(\mathbf{s}_i,\mathbf{s}_j)$ between any two points and not on the direction then the point pattern process is called ***isotropic***, that is $\gamma(\mathbf{s}_i,\mathbf{s}_j) = \gamma(h)$.

Measures for second order relationships among points

The distance matrix

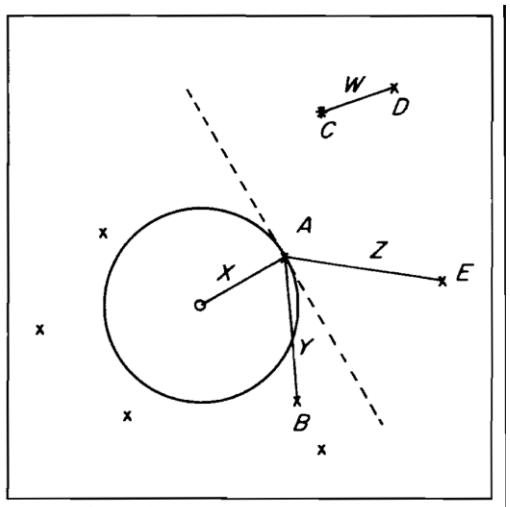
- The distance can also be written in matrix notation: Let $\mathbf{u} \equiv \mathbf{s}_i - \mathbf{s}_j$ with $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \equiv \begin{pmatrix} x_{i1} - x_{j1} \\ x_{i2} - x_{j2} \end{pmatrix}$, then the distance becomes $d(\mathbf{s}_i,\mathbf{s}_j) = \sqrt{\mathbf{u}^T \cdot \mathbf{u}} = \sqrt{(u_1,u_2) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}} = \sqrt{u_1 \cdot u_1 + u_2 \cdot u_2}$
- The ***second order*** properties are related to the pairwise distances among the events. Depending on which information from the distance matrix we are using we get different measures for second order properties.

- A distance matrix for a set of *n* events is given by

$$\mathbf{D} = \begin{pmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & 0 & d_{23} & \cdots & d_{2n} \\ d_{31} & d_{32} & 0 & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \cdots & 0 \end{pmatrix}$$

D is symmetric with zeros on the diagonal.

- Two different distance matrices are used in point pattern analysis:
  - *Event-event* distance W among the *n* events.
  - *Point-event* distance X between a randomly selected point and an event. In total there are *m* randomly selected points



Key * A randomly chosen plant
x Other plants
o A randomly chosen point
A, B, C, D, E. Plant locations
W, X, Y, Z. Nearest-neighbour distances

Figure 1.23 Various nearest-neighbour measurements

The area of circle is the area of the center point

Relationship between Nearest Neighbor Distance and Quadrat Counts under CSR

- CSR means that events follow a homogenous Poisson process. Homogenous means that the intensity λ (the first order effect) is constant over the study area $\mathbb{R}$ and the process is stationary with respect to the expectation.
- It further assumes that $Y(\mathbb{A}_i)$ and $Y(\mathbb{A}_j)$ are independent.

- Then the number of events $Y = y$ in a sub-area A follows a Poisson distribution

  $f_{Y(\mathrm{A})}(y) = \frac{(\lambda \cdot A)^y}{y!} \exp(-\lambda \cdot A)$ where $\lambda \cdot A$ is the expected number of events in area A and A is the size of the area.

- Under complete spatial randomness the theoretical distribution of nearest neighbor $d_{i,nearest}$ becomes the exponential distribution with

  $F(d_{i,nearest}) = 1 - \exp(-\lambda \cdot \underbrace{\pi \cdot d_{i,nearest}^2}_{\text{area of circle with radius } d_{i,nearest}})$

Statistics based on Nearest Neighbor Distances

- Minimum distance between event *i* and its neighboring events

  $w_i = \min\{d_{i1}, \ldots, d_{i,(i-1)}, d_{i,(i+1)}, \ldots, d_{in}\}$ because it measures data in finer scale, so we do not need homogeneous assumption

- Minimum distance between random point *k* and its closest neighboring events

  $x_i = \min\{d_{k1}, \ldots, d_{ki}, \ldots, d_{kn}\}$

- The nearest distances are ***relatively robust*** against particular ***first order stationarity violations***.
  Example: the nearest inter-event distance does not vary between two point clusters.

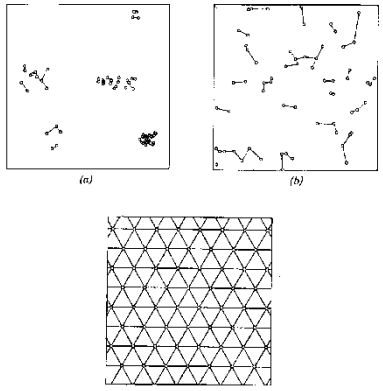
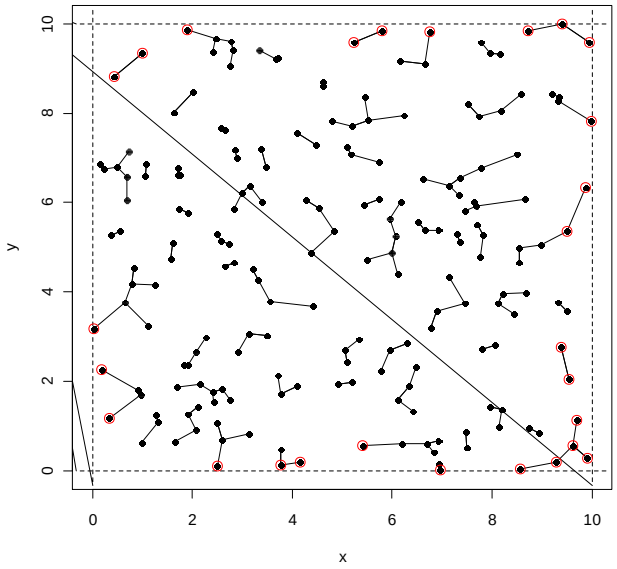


FIGURE 5.14 Nearest-neighbor statistics, R, for patterns of points on maps. (*a*) Points grouped into five clusters, $R = 0.34$. (*b*) Points scattered at random. $R = 0.91$. (*c*) Points arranged on a hexagonal net, $R = 2.15$. Point density, λ, is the same for all patterns. Adapted from Olea (1984).



All the red points are more closer to the boundary than any other points in the study area

Edge Corrections:

- Key terms:
  - b_i is the distance of event $w(s_i)$ or random point $x(s_i)$ to the ***nearest boundary point*** (such as the convex hull)
  - w_i is the distance of the event $w(s_i)$ to the ***nearest event***, whereas, x_i is the distance of the random point $x(s_i)$ to the ***nearest event***.
  - The indicator functions $I(\ )$ are 1 if its argument is true and otherwise 0:
    - $I(w_i \leq w) = 1$ if the shortest inter-event distance of event $w(s_i)$ is less than the distance w.
    - $I(w_i \leq w < b_i) = 1$ if the shortest inter-event distance of event $w(s_i)$ is less than w ***and*** the w is less than the distance b_i of event $w(s_i)$ to the boundary.

- $I(w < b_i) = 1$ if the distance w is less than the distance b_i of event $w(s_i)$ to the boundary.

- Use **`BoundCorrectionGFunction.r`** to show edge correction at different distances w: Shrinking area interpretation of boundary condition for events $w(s_i)$:

- Cumulative probability distributions:
  - Empirical cumulative probability distribution for event-event minimum distance

$$\hat{G}(w) = \frac{\sum_{i=1}^{n} I(w_i \le w)}{n} \text{ or with edge correction } \hat{G}(w) = \frac{\sum_{i=1}^{n} I(w_i \le w < b_i)}{\sum_{i=1}^{n} I(w < b_i)}$$

only consider those points

  - Empirical cumulative probability distribution for point-event minimum distance

$$\hat{F}(x) = \frac{\sum_{i=1}^{m} I(x_i \le x)}{m} \text{ or with edge correction } \hat{F}(x) = \frac{\sum_{i=1}^{m} I(x_i \le x < b_i)}{\sum_{i=1}^{m} I(x_i < b_i)}$$

- Behavior of distribution function $\hat{G}(w)$ for clustering (inter-event attraction) (see plot volcanic craters)

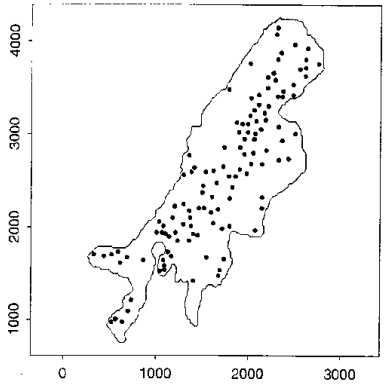


Fig. 3.1 Volcanic craters in west Uganda

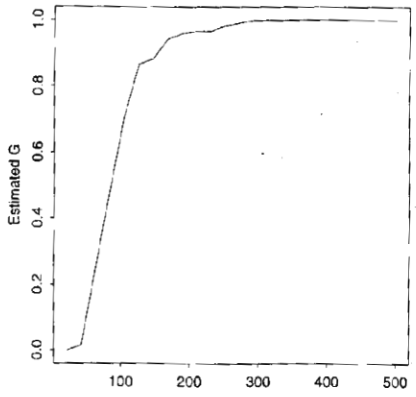


Fig. 3.6 Nearest neighbour distribution function for volcanic craters

F functions show the opposite way from G functions

Those two methods could use for estimating if there have any clusters. (Gest and F functions)

- Discuss expected behavior of distribution function $\hat{G}(w)$ for regularity (dispersion or inter-event repulsion) or clustering (inter-event attractions)
- $\hat{F}(x)$ measures the distribution of distances from arbitrary point locations to their nearest events. It is called the ***empty space function***, because it is a measure of average space left between events. It is a continuous distribution function.
  => The average empty space distance will be different from the inter-events distances.
- The distributions are
  - $\hat{F}(x) < F_{pois}(x)$ for clustered patterns
  - $\hat{F}(x) \cong F_{pois}(x)$ for random patterns
  - $\hat{F}(x) > F_{pois}(x)$ for regular patterns

This is shrink speed, for clustered pattern, the buffer of points are overlap. so, one we change the radius, they would shrinking slowerly.

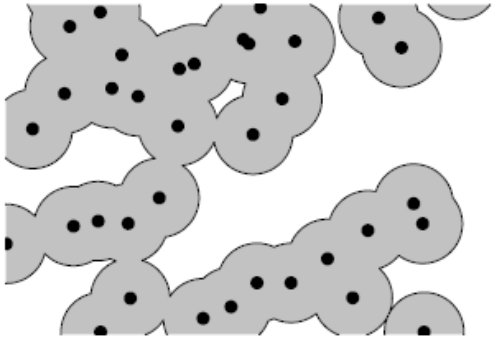


Figure 8.9. *Dilation set* $\mathbf{X}_{\oplus r}$ *(grey shading) of a point process* $\mathbf{X}$ *(points shown as* $\bullet$*). The empty-space probability* $F(r)$ *is the average fraction of area occupied by* $\mathbf{X}_{\oplus r}$.

- The theoretical cumulative distribution under CSR is the exponential distribution:

$$F = G = 1 - \exp\left(-\lambda \cdot \underbrace{\pi \times r^2}_{\text{area of a disk}}\right)$$

  with r either w for events or x for random points. This equation does not account for edge-effects.
- Whether the departure of $\hat{G}$ from G (or $\hat{F}$ from F) is significant must be evaluated by simulation, because it depends on the shape of the study area and the number of points.
- Discussion of nearest neighbor approach:
  - Uses selected information from the distance matrix, i.e., only the shortest distance of each point to its neighboring point

- Robust against moderate deviations from first order stationarity as long as we can assume that the intensity $\lambda(\mathbf{s})$ does not change substantially over short distances

- Demonstrate *G* and *F* function with the program **`G-F-Functions.r`**

Modes of Edge Corrections:

- A guard buffer of a given distance can be formed extending inwards form the edge of the study region *R* Events within the guard buffer do not function as pivot during the analysis, however, events outside the guard buffer can refer to theses excluded events.
  The guard buffer reduces the number of observations used and thus this method works with less information.

Figure 7.15. *Border method of edge correction for the K-function. When estimating $K(r)$, distances d_{ij} are measured only from points x_i lying at least r units away from the boundary. These are the points falling inside $W_{\ominus r}$ (shaded region).*

- Isotropic Correction

Figure 7.17. *Calculation of sampling bias for the isotropic correction. The contribution from a pair of data points x_i, x_j is determined by drawing a circle through x_j centred at x_i. The fraction p of the circle's* perimeter *which falls inside the observation window (grey line) is measured. The edge correction weight for this pair of points is* $1/p$.