Chapter 5

Analysis of Point Patterns

A primary goal in the analysis of point pattern data is to detect patterns.

Does the set of observed locations suggest clustering?

Equivalently, are certain areas associated with a greater likelihood of occurrence?

- This may indicate areas with greater risk of disease.

5.1 Types of Patterns

In public health, one pattern of particular interest is the presence of a tendency of locations to cluster together.

Events are **clustered** when they occur more frequently near one another than one would expect from a set of cases with no common causative factor (e.g., environmental exposure).

An **event** is an occurrence of interest (e.g., an incident of a disease).

An **event location** is where the event occurred.

Note:

- A **point** is any location in the study area where an event could occur.
- An **event location** is a particular location in the study area where an event *did* occur.

A **point pattern data set** consists of a collection of *observed event locations* and a *spatial domain of interest*.

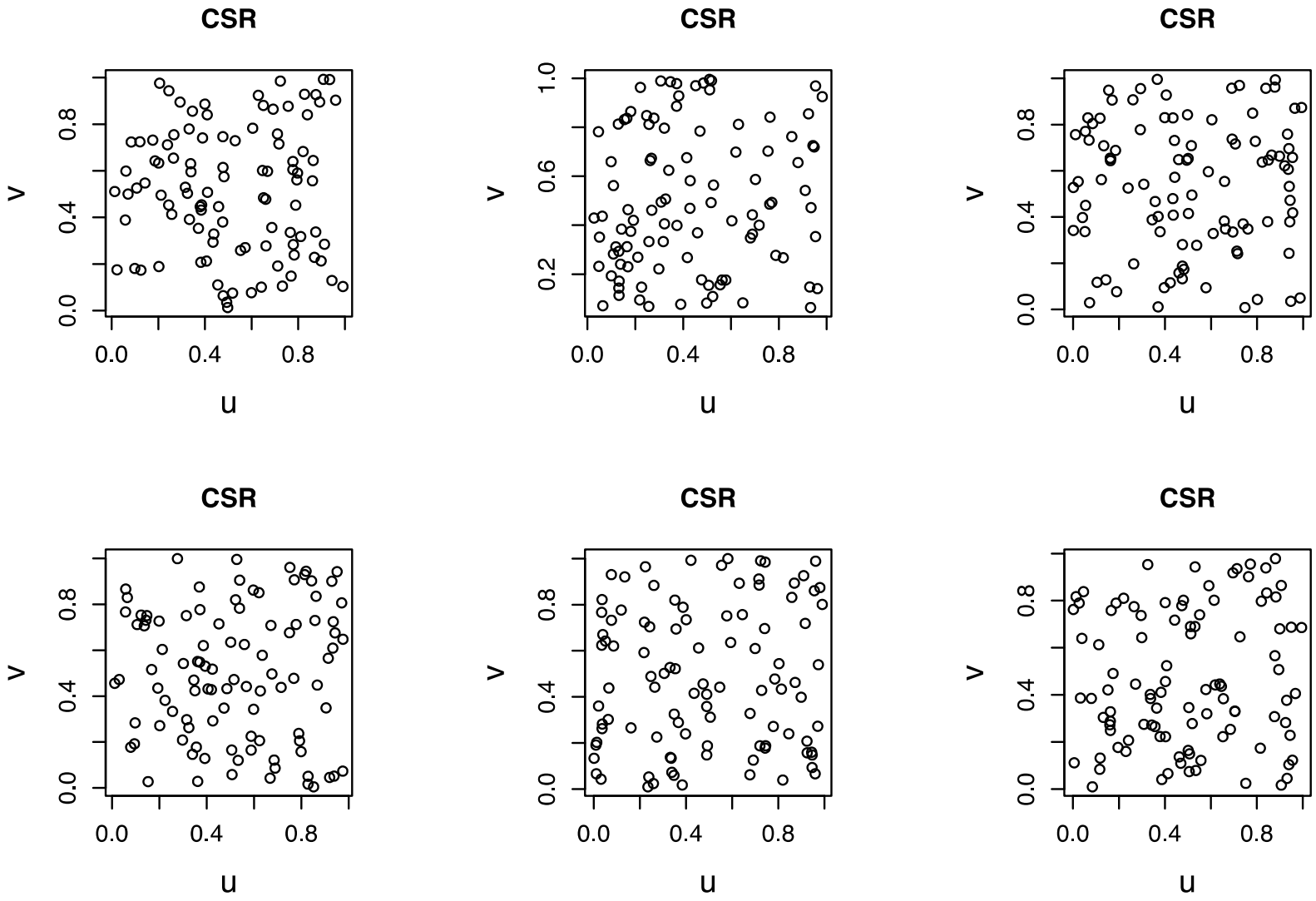
- The domain of interest is a very important aspect of the data set.
- Different choices for the domain can result in different inferences about the spatial data under consideration.
- The domain of interest is typically defined by the (thoughtful analyst), though there might be a natural domain of interest (like the borders of wildlife refuge).

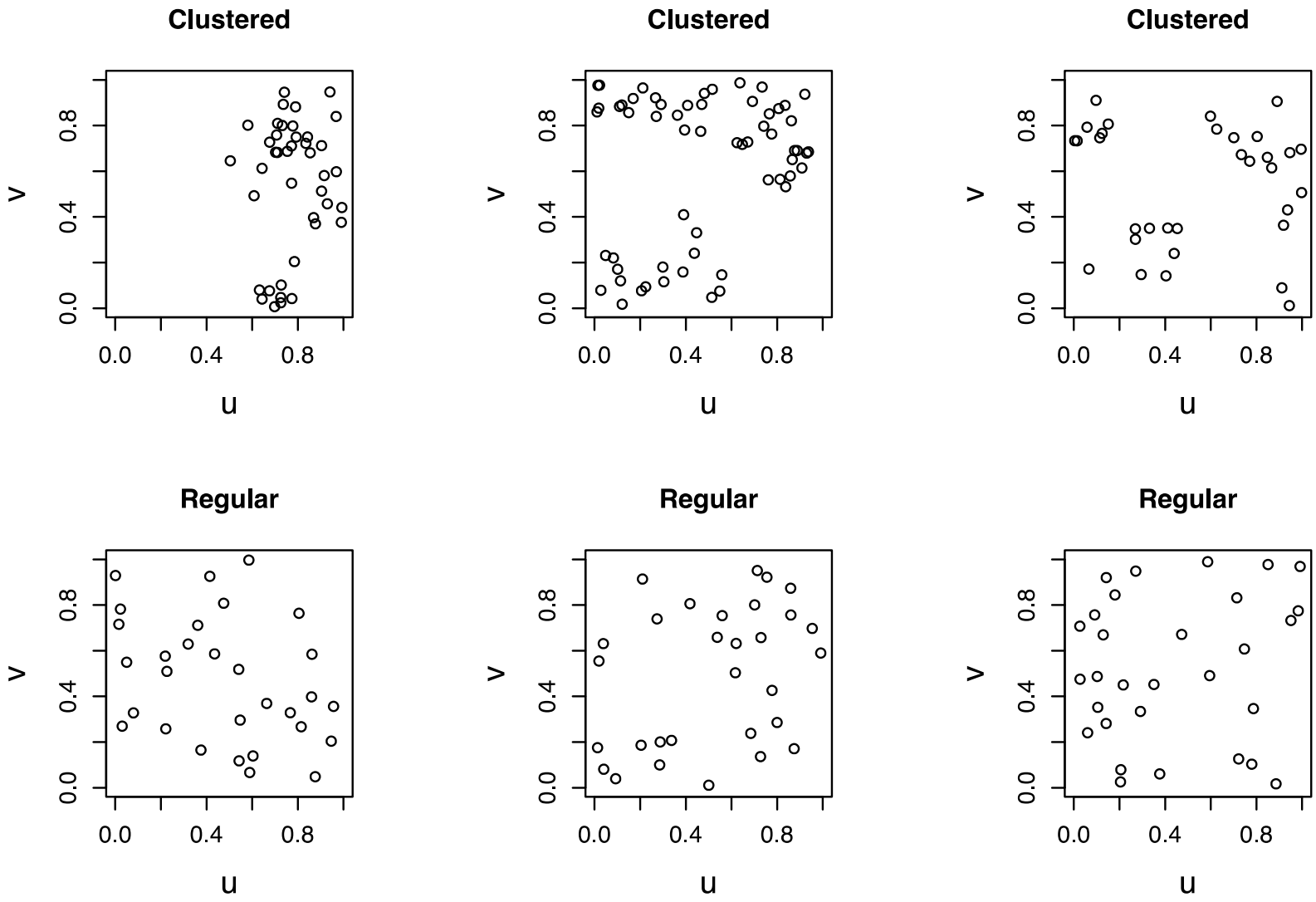
All the patterns we observe will be "random" in some sense, so we must be clear to define the types of "randomness" that are of interest.

Complete spatial randomness (CSR) defines a situation where an event is equally likely to occur at any location within the study area, regardless of the locations of other events.

- Events follow a *uniform distribution* across the study area.
  - This is a statistical distribution—it does not mean equally spread across the spatial domain.
- Events are *independent* of one another.

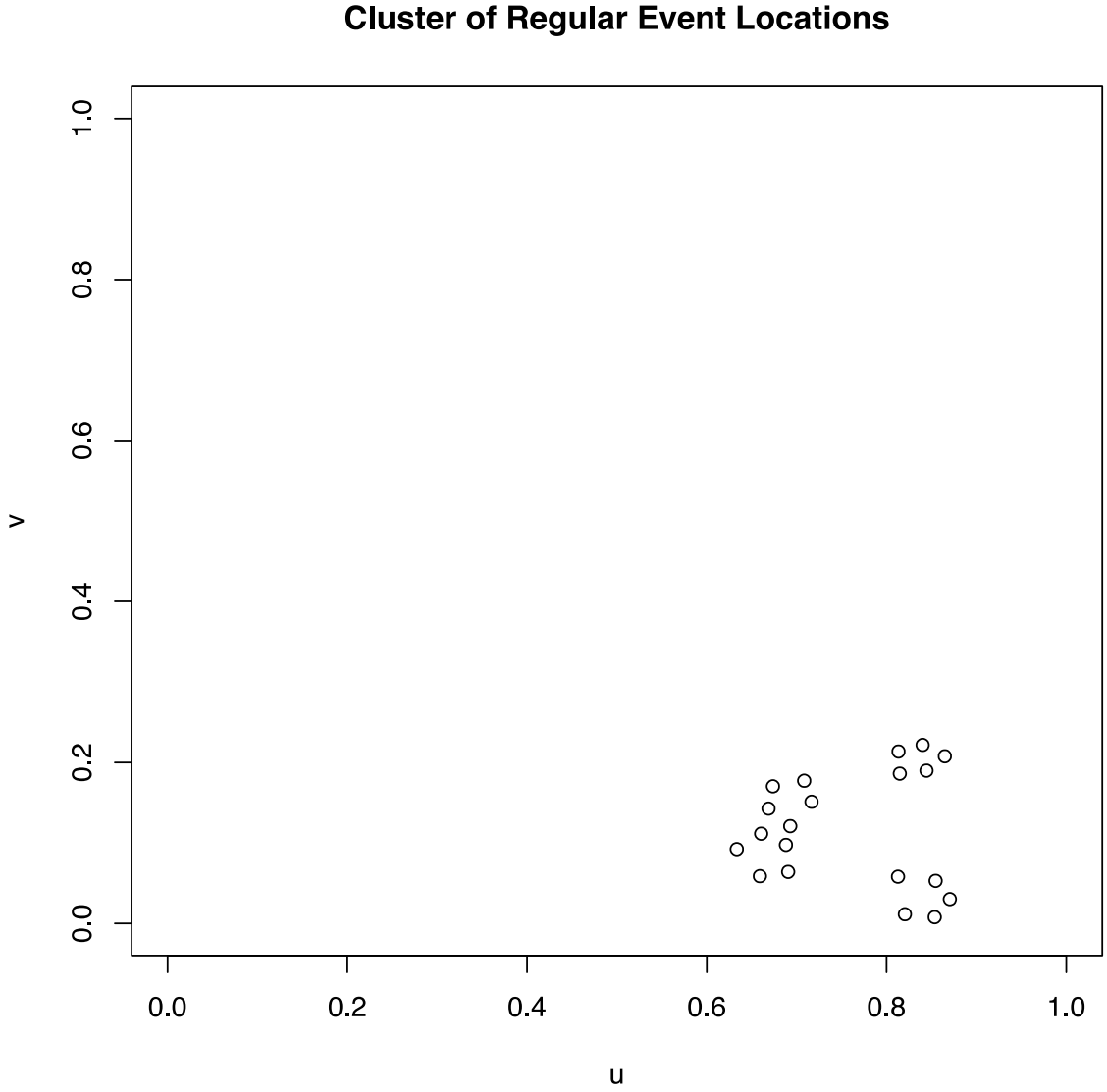
Consider the point patterns on the next few slides. The first six are examples of CSR, the next three show clustered point patterns, while the final three show regular point patterns.

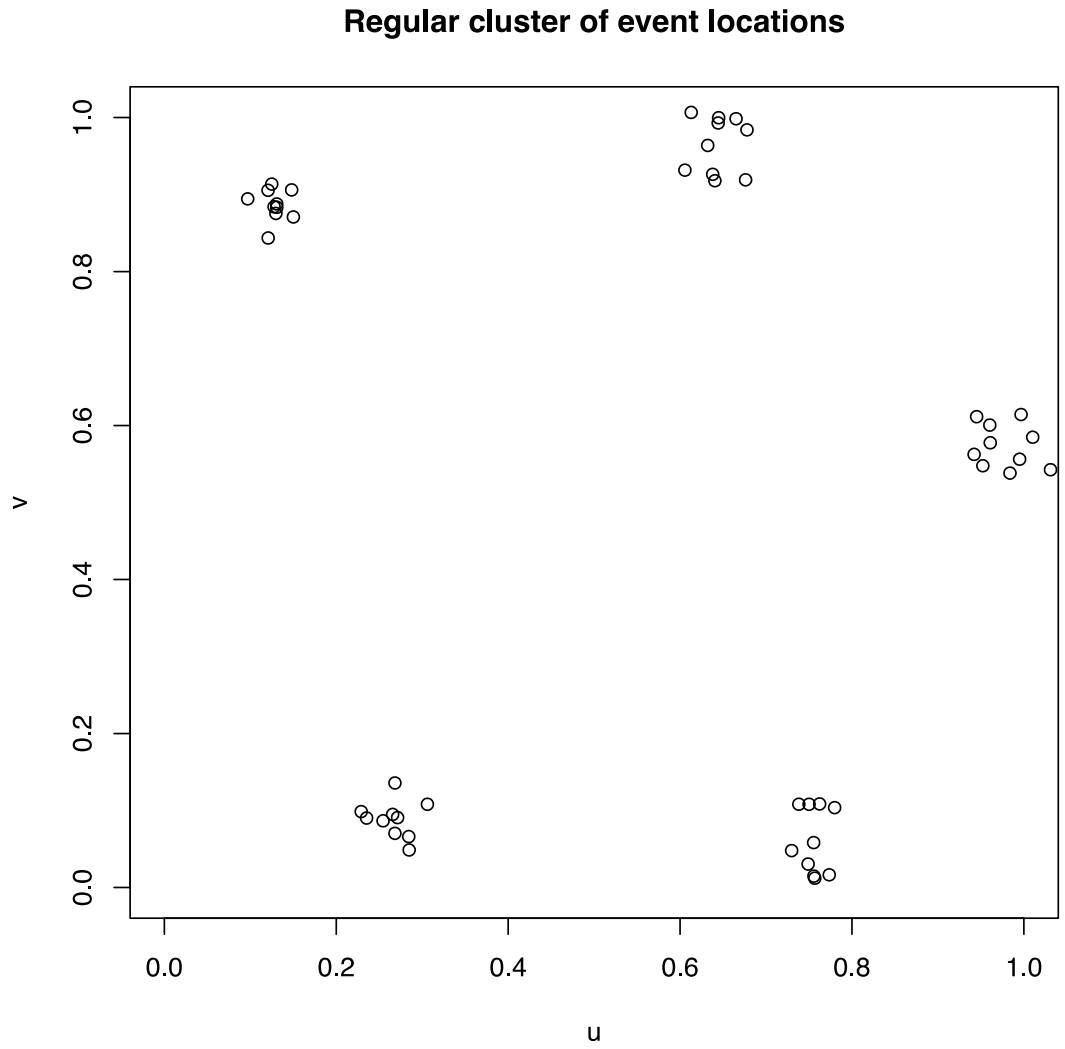

CSR
v
u
0.0
0.4
0.8
0.2
0.6
1.0


Clustered
Clustered
Clustered
Regular
Regular
Regular
v
u
0.0
0.4
0.8

Complete spatial randomness is the boundary condition between spatial processes that are more clustered than random and processes that are more regular than random.

- It may be difficult to visually distinguish between the three patterns.
- Observed patterns do not always fall neatly into one pattern.
    - We can have regular clusters of events locations.
    - We can have clusters of regular event locations.


Cluster of Regular Event Locations
v
u
0.0
0.2
0.4
0.6
0.8
1.0


Regular cluster of event locations
v
u
0.0
0.2
0.4
0.6
0.8
1.0

The examples above indicate the importance of *spatial scale*.

Whether we classify event locations as regular, clustered, or random depends on the scale at which we are comparing the locations.

5.2 Spatial Point Processes

A **stochastic process** is a collection of random variables, say $\{X_1, X_2, \ldots, X_N\}$, described by a probability distribution.

- Often, these are measurements of the same variable at different times, locations, or on different people.
- E.g., heights of students in this class.

A **spatial point process** describes a stochastic process where each random variable represents the location of an event in space.

A **realization** of a spatial point process is a collection of locations generated under the spatial point process model (regardless of whether we know what that model is).

The previous graphics display realizations from various spatial point process models.

Spatial point processes are found in many fields (e.g., forestry, astronomy, cellular biology, etc.), but we will focus on applying spatial point process models to public health data.

5.2.1 Stationarity and Isotropy

A **stationary** process is invariant to translation within *d*-dimensional space.

- i.e., relationships between two events in a stationary process depend only on their relative positions, not on the event locations themselves.
- i.e., if we look in two similarly shaped subareas of the study area, the process will behave similarly in the two subareas.

An **isotropic** process is invariant to rotation about the origin.

- i.e., the process should behave similarly in all directions when a similar distance from the origin.

Neither stationarity nor isotropy imply the other property.

Stationarity and isotropy offer a form of "replication" within a data set.

- Typically, in statistics, we have independent and identically distributed data.
  - Each random variable has the same behavior.
  - We do not have this in spatial statistics.
- For a stationary process, two pairs of points in a realization that are separated by the same distance and relative direction should be subject to the same relatedness.

For a stationary and isotropic process, two pairs of points the same distance apart should be subject to similar properties.

Stationarity and/or isotropy are key ingredients to many of the inferences that we will make about point processes.

5.2.2 Spatial Point Processes and CSR

We next outline a class of spatial point processes that is equivalent to a CSR process.

Consider the family of stationary **homogeneous spatial Poisson point processes** defined by the following criteria:

1. The number of events, N, occurring within a finite region A is a random variable following a Poisson distribution with mean $\lambda|A|$ for some positive constant λ and $|A|$ denoting the area of A, i.e., $N \sim \text{Poisson}(\lambda|A|)$.
2. Given the total number of events N occurring within an area A, the locations of the N events represent an independent random sample of N locations, where each point is equally likely to be chosen as an event.

Criterion 2 represents the concept of CSR (events uniformly distributed over the study area).

Criterion 1 represents the stationarity of the process and introduces the intensity parameter λ, which represents the number of events expected per unit area.

The Poisson distribution allows the total number of events observed to vary from realization to realization, while maintaining a fixed (but unknown) expected number of events per unit area.

Note that for a homogeneous Poisson process, the intensity λ is the same in all parts of the domain. A sensible estimator of λ is $\hat{\lambda} = N/|A|$.

This estimator tends to work well, but other estimators might be better in certain situations (particularly sparse samples of the set of events observed).

An equivalent definition of the homogeneous Poisson process is given by Cressie (1993, p. 634):

1. The number of events in nonoverlapping regions are statistically independent.
2. For any region $A \subseteq D$ (D is the domain),
$$\lim_{|A|\to 0} \frac{\Pr[\text{exactly one event in } A]}{|A|} = \lambda > 0.$$
3. $$\lim_{|A|\to 0} \frac{\Pr[\text{two or more events in } A]}{|A|} = 0.$$

Component 1 is very important in the analysis of regional counts.

Component 2 implies that the probability of a single event in an increasingly small area A (adjusting for the the area of A) is constant.

Component 3 implies that no more than one event can occur at a point.

Since the intensity of events is constant at all locations in the study area, we say that the process is **homogeneous**.

The definition of a homogeneous Poisson process given by criteria 1 and 2 provides a two-stage approach for simulating realization of CSR in a study area D.

1. Generate a total number of points, $N(D)$, from a Poisson distribution with mean $\lambda|D|$ ($|D|$ denotes the area of D).
2. Place events within D according to a uniform distribution.
   a. If D is rectangular, we may generate u and v coordinates using uniform random number generators on the intervals corresponding to the width and height of D, respectively.
   b. If D is not rectangular, one option is to embed D within a larger rectangle R, and generate event locations uniformly in R until $N(D)$ events occur within D, and then use these $N(D)$ events as realizations.

5.2.3 Hypothesis Tests of CSR via Monte Carlo Methods

We detect clustering or regularity by detecting a departure from CSR.

We need methods that describe how much variation we expect under CSR and to tell us when an observed pattern of event locations appears to significantly differ from CSR.

Hypothesis tests and Monte Carlo simulation techniques provide versatile tools for such assessments.

Most hypothesis tests rely on asymptotic arguments (when $N \rightarrow \infty$) to derive the distribution of the test statistic under the null hypothesis.

- The domain is assumed to be a regular shape (e.g., a square or rectangle).
- The number of events $N \rightarrow \infty$.

These assumptions do not apply in many contexts, including ours, so we need a different approach.

Hypothesis tests are generally performed by comparing the observed value of a test statistic to its distribution under the null hypothesis.

Since we can easily generate data under the null hypothesis of CSR, we can approximate the distribution of our test statistic under CSR! (Even when our study area is oddly shaped, small sample size, etc.)

This suggests using a **Monte Carlo** (simulation-based) method of inference.

Procedure for Monte Carlo testing:

1. Calculate the test statistic based on the data observed. Denote this T_{obs}.
2. Generate a large number (N_{sim}) of data sets assuming CSR.
3. Calculate the test statistic for each of the N_{sim} simulated data sets.
4. Count the number of test statistics (the simulated and observed statistics) as large or larger than the observed test statistic. Denote this count ℓ.
5. The estimated p-value is

$$\widehat{\Pr}[\, T \geq T_{obs} \mid H_0 \text{ is true} \,] = \ell/(N_{sim} + 1).$$

We add one in the denominator since our estimate is based on $N_{sim} + 1$ values (the N_{sim} test statistics and the observed one).

Due to their reliance on simulated data, Monte Carlo methods are not precisely replicable (unless we set our number seed in R!).

- If each of us performed a test, we would get slightly different results.

The larger N_{sim} is, the more stable the conclusions will be.

- One can calculate the variability (a.k.a. Monte Carlo error) as a function of N_{sim}.

We can use Monte Carlo simulation in many contexts, not just when CSR is the null hypothesis.

5.2.4 Heterogeneous Poisson Processes

A Poisson process is homogeneous when the intensity, λ, is constant across the study area.

This is often too restrictive in the analysis of public health events since population is typically not distributed uniformly across space.

The **constant risk hypothesis (CRH)** is a competing model of "no clustering". Under CRH, each person has the same risk of disease during the observation period, regardless of location.

- Clusters of cases in high populations are likely to violate CSR but may not violate the CRH.
- We should expect more cases in high population areas.

The CRH requires a generalization of CSR where we define the intensity as a spatially varying function defined over our study area D (i.e., the intensity is a function $\lambda(s)$ of the spatial location $s \in D$).

The **heterogeneous Poisson process** is defined by the following criteria:

1. The number of events occurring within a finite region A is a random variable following a Poisson distribution with mean $\int_A \lambda(s)ds$.
2. Given the total number of events N occurring within an area A, the N events represent an independent sample of N locations, with the probability of sampling a particular point $\boldsymbol{s}$ proportional to $\lambda(s)$.

Translating this definition:

- The number of events in disjoint regions still follow independent Poisson distributions.
- Component 1 means that the number of events expected in an area A is related to the average intensity over the region.
- Component 2 means that the events are distributed more or less frequently related to the intensity $\lambda(s)$. We expect more events in areas with higher values of $\lambda(s)$.

A heterogeneous spatial intensity function necessarily means that the process is nonstationary since it is no longer translation invariant.

Heterogeneity results in an anisotropic process if the intensity function is anisotropic.

The intensity function $\lambda(s)$ is a first-order (mean or average) property of the random process describing the expected density of events in any locations of the region.

- How intense/dense do we expect the points to be at this (and immediately surrounding) locations?

Events will remain independent of one another, but clusters appear in areas of high intensity.

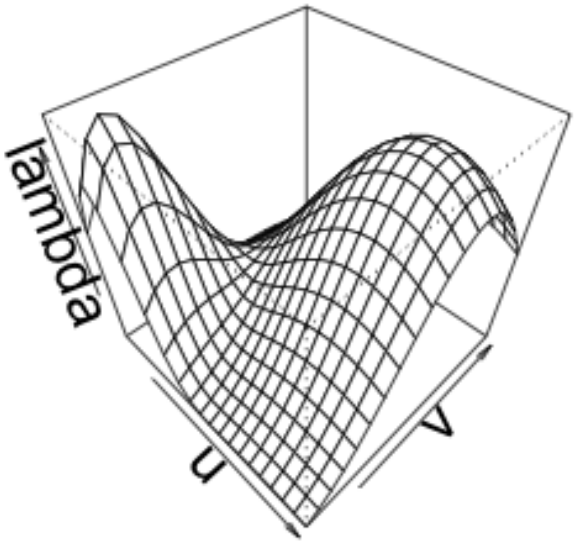
Under a heterogeneous Poisson process, clusters occur due to heterogeneities in the intensity function.

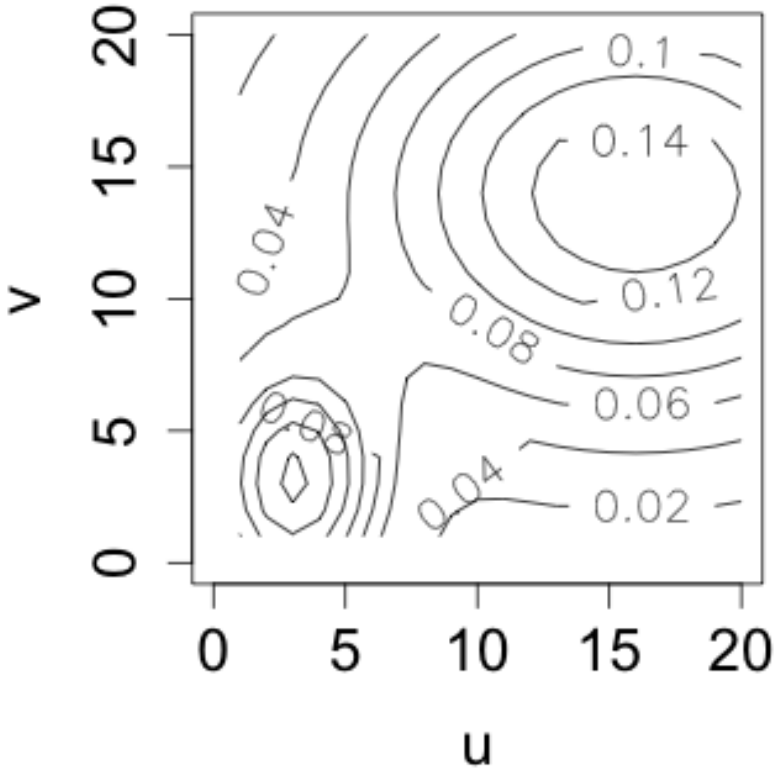
Under the constant risk hypothesis, our concept of what constitutes a cluster depends on the size of the population at risk.

- Our model of no clustering should depend on the observed population density across our study area.
- As population sizes increases, so should our expected number of cases under a model of constant individual-level risk of disease at all locations.

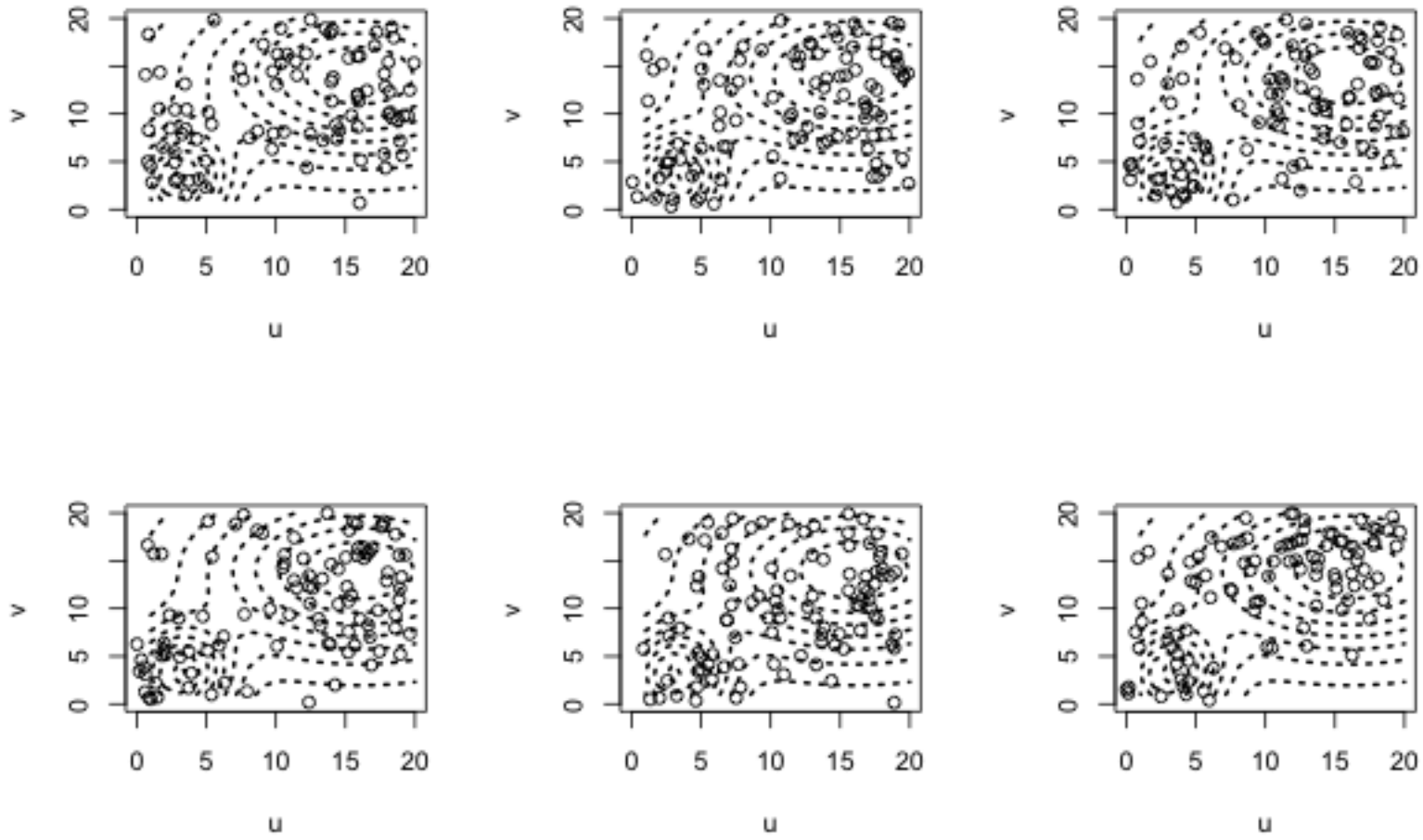
The heterogeneous Poisson process offers a convenient null model (no clustering) for our tests of disease cluster that allows for geographic variations in population size.

The following intensity function exhibits two areas of high intensity – one at (3,3) and another at (16,4).


lambda
u
v


0.1
0.14
0.12
0.08
0.06
0.04
0.02
0.04
0.08
v
u
0
5
10
15
20

Six realizations from the above heterogeneous process:



Notice the lack of events between the modes of the intensity function and the greater number of events close to the modes.

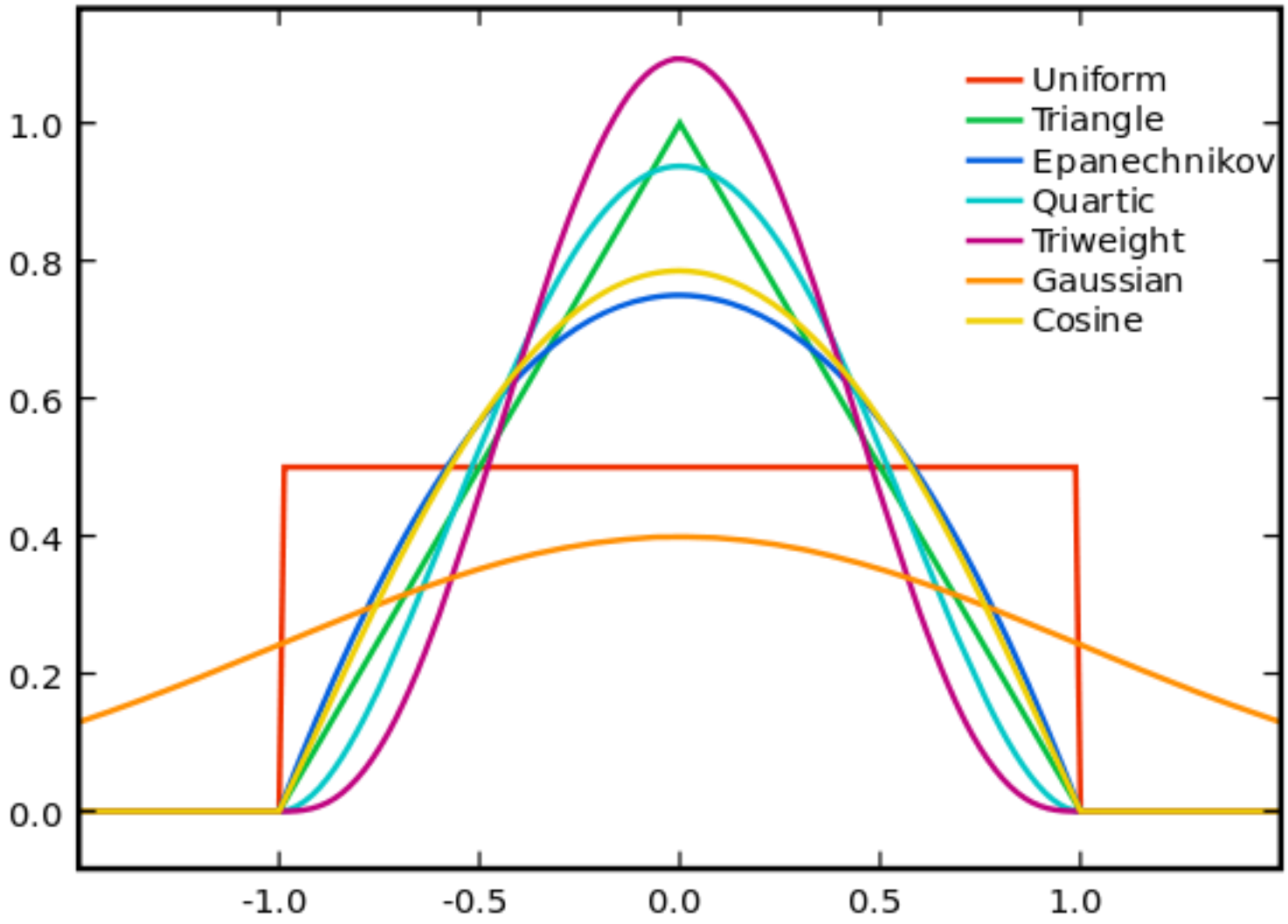
The heterogeneous Poisson process provides a convenient null model (of no clustering) for our tests of clustering that allows for geographic variations in population size.

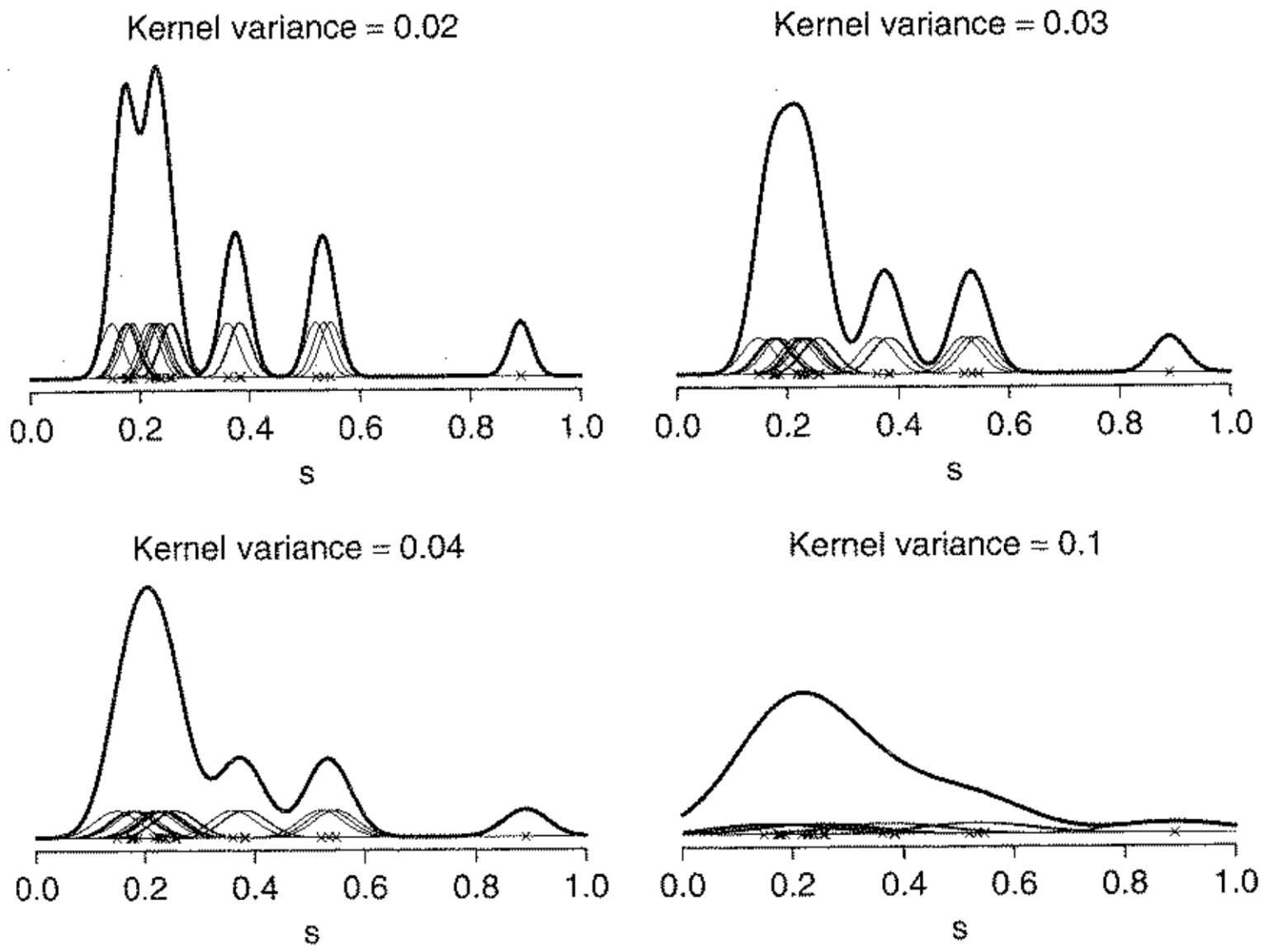
5.2.5 Estimating Intensity Functions

A contour plot of the intensity function for a heterogeneous Poisson process indicates areas with higher and lower probabilities of an event occurring.

Kernel density estimation is often used to estimate the intensity function $\lambda(s)$ of a point process.

- Data are smoothed using a weighted average of nearby points using a weighting mechanism known as a kernel.
- Some common one-dimensional kernel functions are the uniform, triangle, quartic, triweight, and Gaussian kernels.
- These different kernel functions all assign weights slightly differently.
- **Bandwidth** refers to the spread or variance of the kernel.


Uniform
Triangle
Epanechnikov
Quartic
Triweight
Gaussian
Cosine
1.0
0.8
0.6
0.4
0.2
0.0
-1.0
-0.5
0.0
0.5
1.0


Kernel variance = 0.02
0.0
0.2
0.4
0.6
0.8
1.0
s
Kernel variance = 0.03
0.0
0.2
0.4
0.6
0.8
1.0
s
Kernel variance = 0.04
0.0
0.2
0.4
0.6
0.8
1.0
s
Kernel variance = 0.1
0.0
0.2
0.4
0.6
0.8
1.0
s

To estimate the shape of the intensity function $\lambda(s)$ underlying the distribution of events, we plot the curve defined by summing the heights of all kernels at any point along the s-axis.

The kernel variance reflects the square of the **bandwidth** (the sd) or extent of the influence of each data and governs the overall smoothness of the intensity estimate.

- The larger the bandwidth, the smoother the intensity estimate.

Small detail: kernel estimation methods often focus on estimating a probability density function $f(s)$ rather than an intensity function $\lambda(s)$.

For a spatial point process:

- A density function defines the probability of observing an event at location s.
- An intensity function defines the expected number of events per unit area at location s.
- These functions differ only by a constant of proportionality, so they provide you with the same information. The relative patterns are the same.
- Dividing $\lambda(s)$ by its integral over D will return the density function $f(s)$.

The definition of a kernel density estimate in one dimension based on observations $u_1, u_2, \ldots, u_N$ is

$$\tilde{f}(u) = \frac{1}{Nb} \sum_{i=1}^{N} \text{kern}\left(\frac{u - u_i}{b}\right),$$

where u is a location in one-dimensional space, kern$(\cdot)$ is a kernel function, and b is the bandwidth.

This idea can be extended to two-dimensions in more than one way. One possibility is the **product kernel**, which is defined as

$$\tilde{f}(u,v) = \frac{1}{N b_u b_v} \sum_{i=1}^{N} \text{kern}\left(\frac{u-u_i}{b_u}\right) \text{kern}\left(\frac{v-v_i}{b_v}\right),$$

where b_u and b_v are the bandwidths in the u- and v-directions.

- The product kernel assumes no interaction or dependence between the u and v coordinates and event locations.

Two items must be specified to do kernel estimation:

1. The kernel function.
2. The bandwidth.

Common kernel functions include the uniform, triangle, quartic (biweight), triweight, and Gaussian kernels. (Table 5.1)

- Most kernels will produce similar results.
- When computational efficiency is important, it is better to choose a kernel that has finite support.
  - Finite support means that only observations within a certain distance are used in the estimate.
  - The Gaussian does not have finite support, but the others mentioned above do.

Table 5.1 Common One-Dimensional Kernel Functions Centered at the ith Observation u_i[a]

Kernel	Kern(u^*)
Uniform	$\frac{1}{2b} I\left(\lvert u^*\rvert \leq 1\right)$
Triangle	$\frac{1}{b(2-b)}\left(1-\lvert u^*\rvert\right) I\left(\lvert u^*\rvert \leq 1\right)$
Quartic (biweight)	$\frac{15}{16b}\left[1-(u^*)^2\right]^2 I\left(\lvert u^*\rvert \leq 1\right)$
Triweight	$\frac{35}{32b}\left[1-(u^*)^2\right]^3 I\left(\lvert u^*\rvert \leq 1\right)$
Gaussian	$\frac{1}{\sqrt{2\pi b}} \exp\left[-\frac{1}{2}(u^*)^2\right]$

[a] u^* represents the difference between location u and u_i divided by the bandwidth b [*i.e.*, $(u-u_i)/b$] (see text). The function I(expression) is the indicator function taking the value 1 if the expression is true and 0 otherwise.

Bailey and Gattrell define another two-dimensional kernel function that extents the quartic kernel:

$$\text{kern}(\boldsymbol{s}) = \begin{cases} \frac{3}{\pi}(1 - \boldsymbol{s}'\boldsymbol{s})^2 & \boldsymbol{s}'\boldsymbol{s} \leq 1 \\ 0 & \text{otherwise,} \end{cases}$$

where $\boldsymbol{s}'$ is the transpose of $\boldsymbol{s}$.

The estimated intensity is then

$$\sum_{\|s - s_i\| \leq b} \frac{3}{\pi b^2}\left(1 - \frac{\|s - s_i\|^2}{b^2}\right)^2.$$

The bandwidth of the kernel estimator has a profound impact on the results.

- The theoretical optimum (the bandwidth minimizing the mean integrated squared error between the true density and the estimated density) cannot be found.
  - You need to know the true density to find it—but if we already knew the true density, there is no point in estimating it.
- The choice of bandwidth depends on the purpose of the estimation.
- Small bandwidths show more local behavior, while large bandwidths provide smoother estimates.

- Bandwidths are often chosen to minimize the Asymptotic Mean Integrated Square Error (AMISE), a measure of how closely the estimated density resembles the true density, on average, as the sample size increases.
- Scott suggested using the bandwidth (in the u-direction) $\hat{b}_u = \hat{\sigma}_u N^{-1/(dim+4)}$, where $\hat{\sigma}_u$ is the sample standard deviation of the u-coordinates, N is the number of events in the study area, and dim is the dimensions of the study area (typically two).
- You can do the same thing for the bandwidth in the v-direction.

Early Medieval Graves Sites

Alt and Vach (1991) describe an archaeological investigation of an early medieval burial ground in Neresheim Baen-Wurttemberg, Germany. The anthropologists and archaeologists involved wonder if this particular culture tended to place grave sites according to family units. The archaeologists consider 152 grave sites and use inherited features in the teeth of excavated skeletons (namely, missing or reduced wisdom teeth) to mark a subset of 31 graves. We will analyze 143 of these 152 grave sites, which includes 30 of the original 31 grave sites.

The research question is whether the spatial pattern of the 30 graves with affected teeth differs from the pattern of the 113 unaffected graves.

5.3 K Function

A **first-order** property of a random variable relates to its mean. A **second-order** property relates to its variance.

The intensity function of a point process informs us of the mean or **first-order** properties of the point process.

A **second-order** property of a spatial process informs us of the relative position of events or the interrelationship between events. E.g., how often do events occur within a given distance of other events. Second-order properties of spatial point processes allow the analyst to summarize spatial dependence between events over a wide range of possible spatial scales.

The **K function** is the most common form of analysis for spatial point processes.

The K function for any positive distance (spatial lag) h is defined as

$$K(h) = \frac{E[\text{\# of events within } h \text{ of a randomly chosen event}]}{\lambda}.$$

The numerator does not include the randomly chosen event in the count.

The K function assumes that the process is stationary and isotropic since the intensity is constant and the distance in the numerator is in any direction.

- Analogues of the K function for nonstationary and anisotropic processes are available, but we will not discuss them.
- Ripley (1977) showed that specifying $K(h)$ for all h is equivalent to specifying $\text{var}(N(A))$, the variance of the number of events occurring in a subregion A of the study region. So, the K function really does describe the second-order properties of the spatial point process.

What do we expect $K(h)$ to be under CSR?

Recall that the area of a circle with radius r is πr^2.

Also, for a homogeneous spatial Poisson point process, the expected number of events within a finite region A is $\lambda|A|$, where $|A|$ is the area of A.

Then under CSR the numerator of $K(h)$ is $\lambda\pi h^2$ and we expect $K(h) = \pi h^2$.

- This is what we expect under CSR.

If the process is clustered relative to CSR, we expect more poin events within distance h of a randomly chosen event location, so the numerator of $K(h)$ becomes larger and $K(h) > \pi h^2$.

If the process is more regular than CSR, we expect fewer event locations within distance h of a randomly chosen event location, so the numerator of $K(h)$ becomes smaller and $K(h) < \pi h^2$.

We will estimate $K(h)$ for our data and compare it to what we expect under CSR to determine whether our data is CSR or something else.

- If our $\widehat{K}(h)$ is substantially smaller than πh^2, we think the process is regular at that scale (relative to CSR).
- If our $\widehat{K}(h)$ is substantially larger than πh^2, we think the process is clustered at that scale (relative to CSR).

5.3.1 Estimating the K Function

We estimate the K function by replacing the expectation in the definition of K with a sample average. The basic estimator of $K(h)$ is

$$\widehat{K}(h) = \hat{\lambda}^{-1} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \delta(d(i,j) < h),$$

where $d(i,j)$ denotes the distance between events i and j, and $\delta()$ is 1 if the statement in () is true, and 0 otherwise. $\hat{\lambda} =$ (# events in A)/$|A|$, and is calculated under the assumption of stationarity and homogeneity.

What is $\widehat{K}(h)$ doing in its numerator? It is averaging (hence the $1/N$) the number of events within distance h of each event in the data set.

This estimator will undercount the number of events within h of an event that is located near the boundary of the study area (the radius of the circle around the event is only partially in the study area).

An edge-corrected estimator for $K(h)$ is

$$\widehat{K}_{ec}(h) = \hat{\lambda}^{-1} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_{ij}^{-1} \delta(d(i,j) < h),$$

where w_{ij} is the proportion of the circumference of the circle centered at event i with radius $d(i,j)$ that lies within the study area.

- This is easiest to calculate for rectangular or circular study areas since w_{ij} depends on the shape of the study area.
- w_{ij} is not necessarily equal to w_{ji}
- $w_{ij} = 1$ if the distance between events i and j is less than the distance of event i to the boundary of the study area.

5.3.2 Diagnostic Plots Based on the K Function

To assess the spatial pattern of our data, we want to compare our estimated K function to what we expected under CSR.

The visual comparison is made easier if we consider a transformation of $K(h)$ and $\widehat{K}_{ec}(h)$.

Specifically, since we expect $K(h) = \pi h^2$ under CSR, we expect that $\left(\frac{K(h)}{\pi}\right)^{\frac{1}{2}} - h = 0$.

Similarly, if we define $\widehat{L}(h) = \sqrt{\widehat{K}_{ec}(h)/\pi}$, then under CSR we expect $\widehat{L}(h) - h = 0$.

Thus, departures of $\hat{L}(h) - h$ from zero provide evidence of departures from CSR at that spatial scale.

Values of $\hat{L}(h) - h$ that are more than zero suggest greater clustering than CSR and values below zero suggest the pattern is more regular than CSR.

We will refer to the plot of $\hat{L}(h) - h$ versus h as an $\hat{L}$ plot.

To be able to make inference, we still need to be able to assess the amount of uncertainty in our estimate of $L(h)$.

5.3.3 Monte Carlo Assessments of CSR Based on the K Function

We will use Monte Carlo simulation to assess the uncertainty in the estimate of the K function.

- Generate many realizations of spatial point processes under CSR. Calculate the function $\widehat{K}_{ec}(h)$ for each one.
- Calculate **envelopes** around $\hat{L}(h) - h$ specifying the amount of uncertainty.
  - The envelopes can be lines connecting the minimum and maximum values of $\hat{L}(h) - h$ across simulations.
  - If you have enough simulations, you can construct envelopes defining certain percentiles (e.g., the 5^{th} and 95^{th} percentiles) of the simulated $\hat{L}(h) - h$ at each h.

The envelopes help us assess whether there is a departure from CSR at the various spatial scales.

If want to perform a hypothesis test, we can conduct a Monte Carlo test with test statistic $T = \hat{L}(h) - h$ for a value of h specified prior to analysis.

- This is a test of CSR at a specific spatial scale.

We could also consider a Monte Carlo test using the test statistic

$$T = \max_{0 \leq h \leq h^*} \hat{L}(h) - h\,,$$

where h^* is a maximum spatial scale to consider.

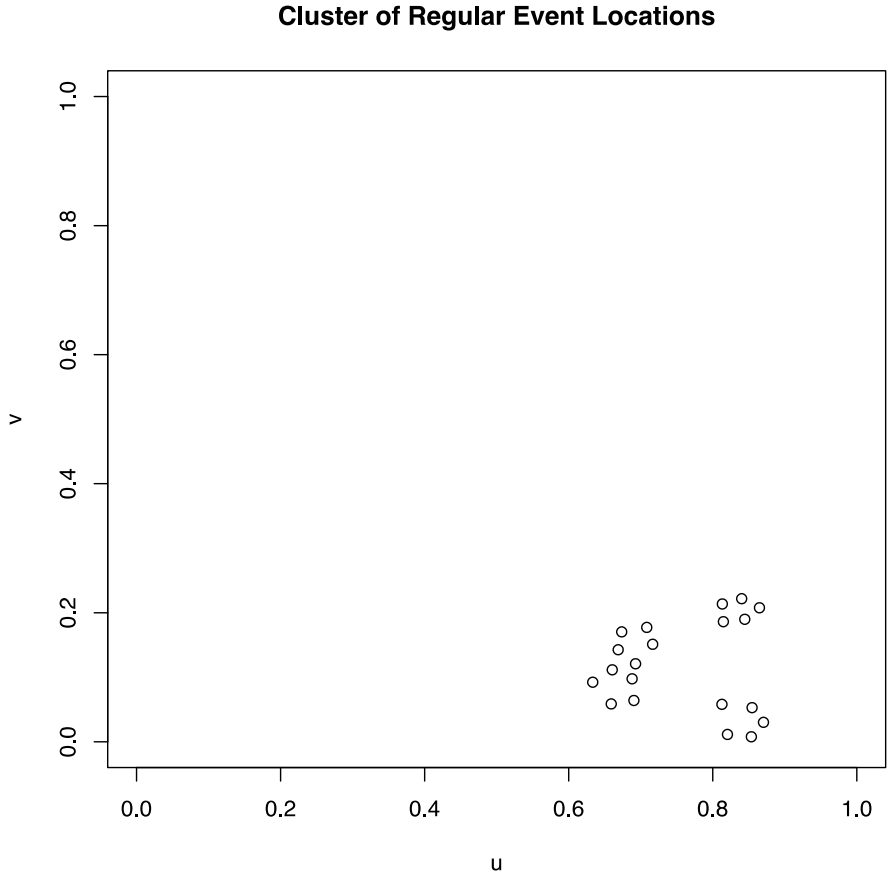
- This is a test of whether the event locations are clustered at some spatial scale.

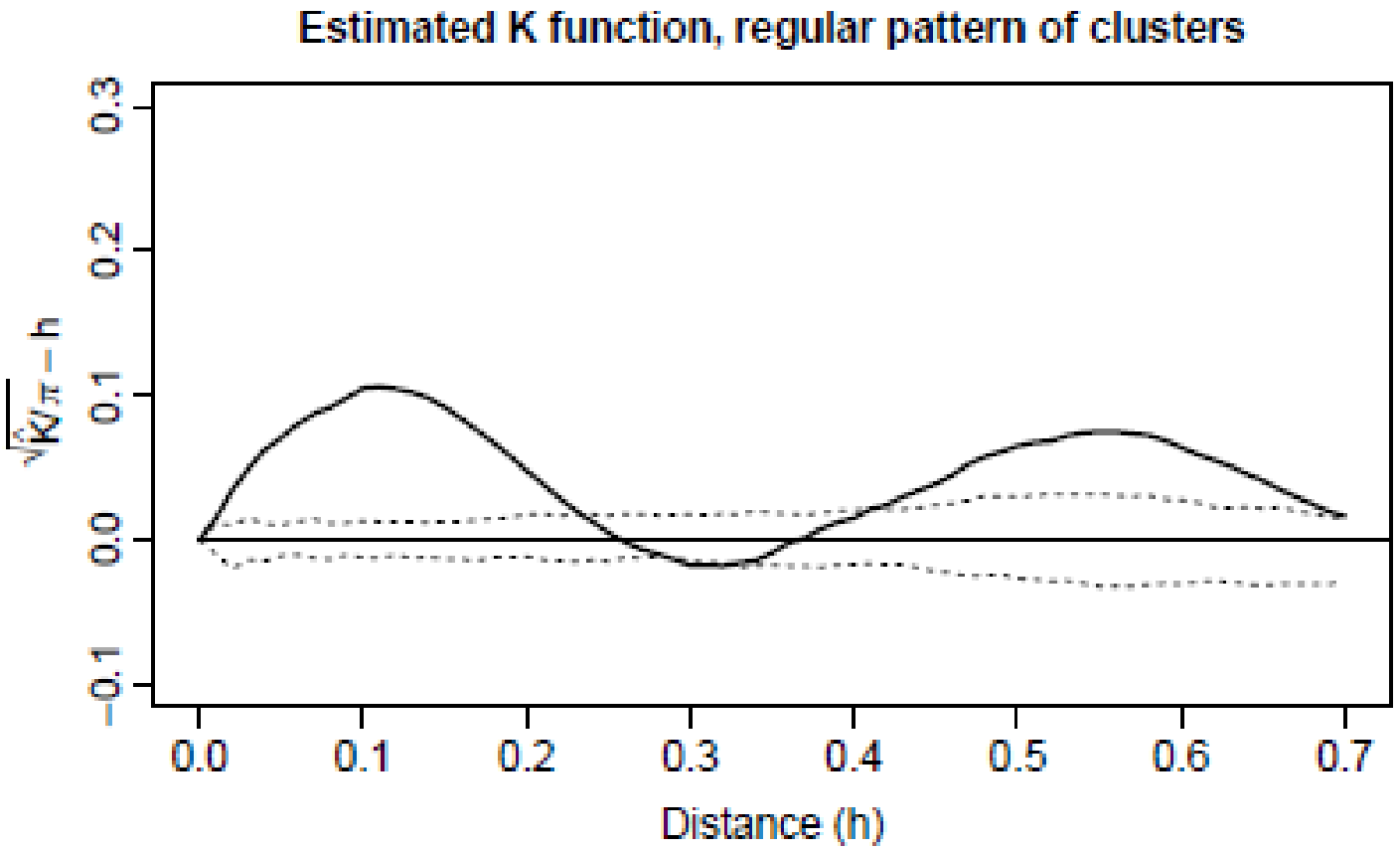
Alternatively, we could use

$$\max_{0 \leq h \leq h^*} |\hat{L}(h) - h\,|$$

to test for a departure from CSR at some spatial scale.

Let's construct an $\hat{L}$ plot for the point pattern observed below, along with envelopes of the min and max $\hat{L}(h) - h$ value across 500 simulated point processes under CSR.

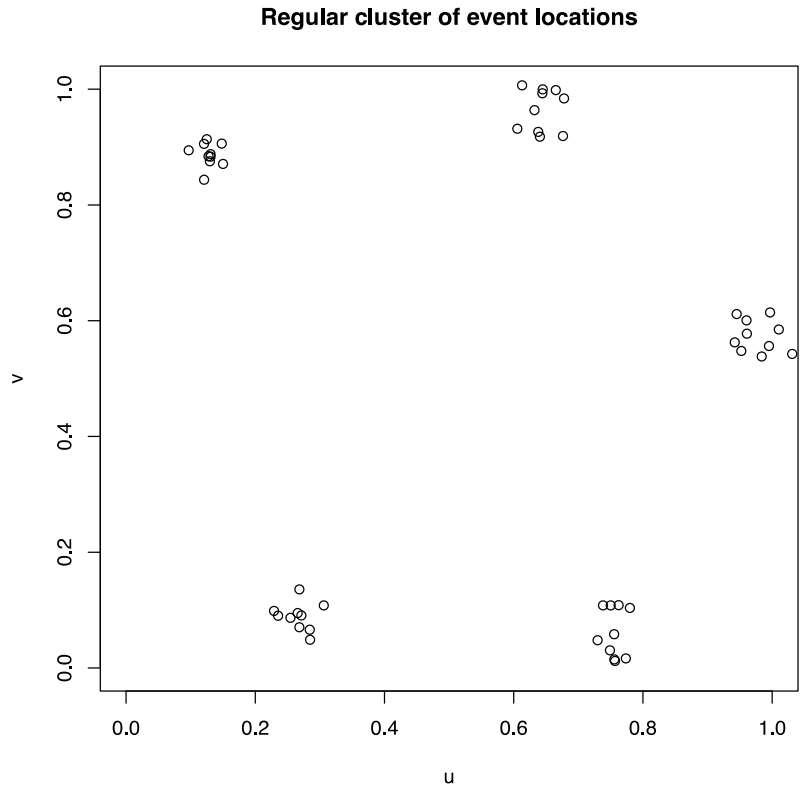


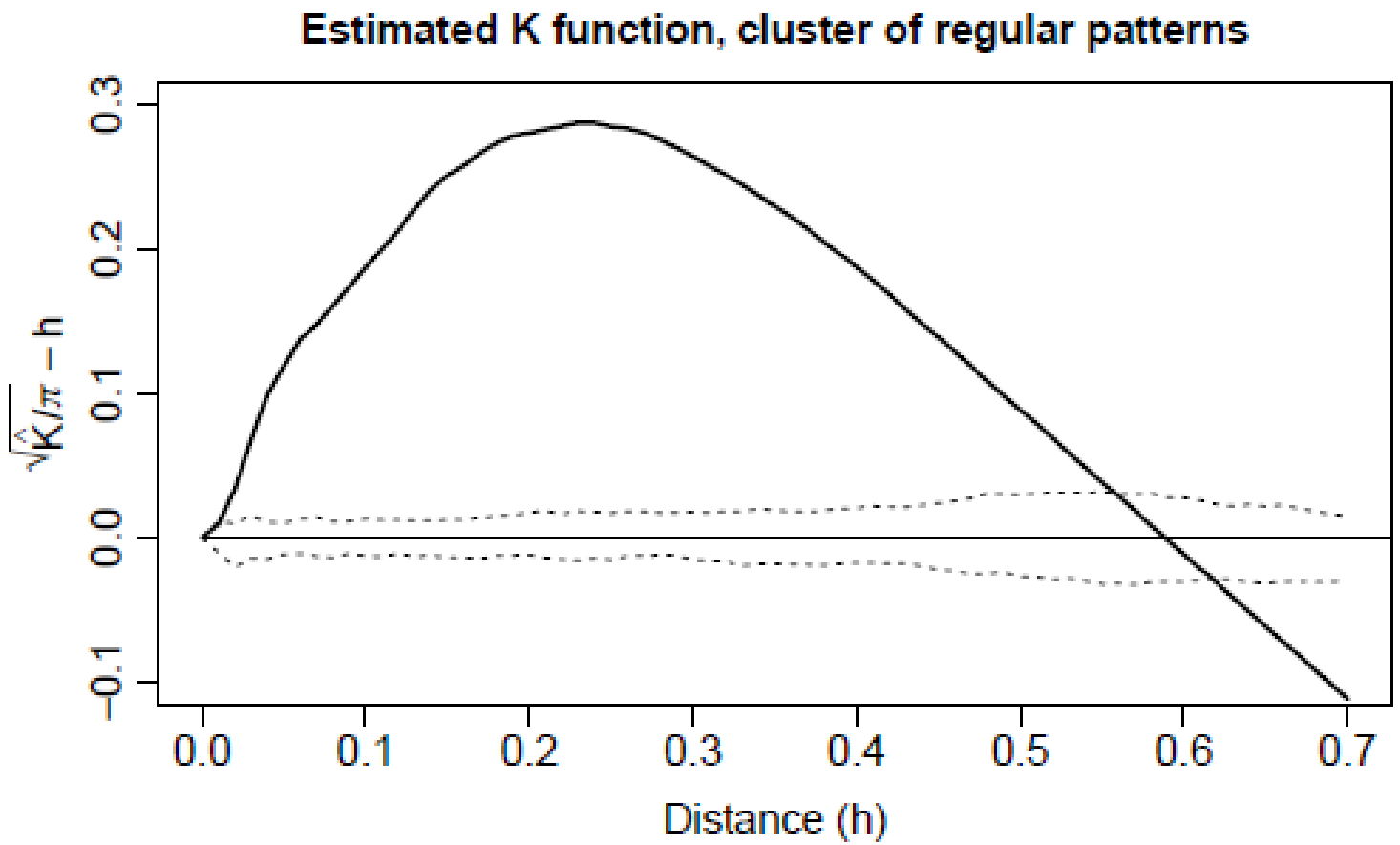

Estimated K function, regular pattern of clusters
$\sqrt{\hat{K}/\pi} - h$
0.3
0.2
0.1
0.0
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Distance (h)

We have evidence of clustering, relative to CSR, for $h \leq 0.2$ and $h \geq 0.4$ since $\hat{L}(h) - h$ is above the envelopes in these ranges.

- We observed more pairs of events at these distances than what we would expect under CSR.
- The first bump corresponds to events in a single cluster.
- The second bump corresponds to pairs of events in different clusters.
- This plot suggests several tight, distinct clusters, with regularity between clusters.

Let's construct an $\hat{L}$ plot for the point pattern observed below, including envelopes of the min and max $\hat{L}(h) - h$ value across 500 simulated point processes under CSR.




Estimated K function, cluster of regular patterns
$\sqrt{\hat{K}/\pi} - h$
0.3
0.2
0.1
0.0
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Distance (h)

We have evidence of clustering, relative to CSR, for $h \leq 0.55$ and regularity for $h \geq 0.6$ (relative to CSR).

- The regular pattern occurs only at the smallest scales, so it isn't really captured in our plot.
- The events are close enough that for the most part, it simply looks like clustering.

Early medieval grave sites continued:

We continue our analysis of this data by considering the second-order properties of the data set.

5.3.4 Roles of First- and Second-Order Properties

First-order (intensity function) and second-order (K function) analysis provide complementary insight about our point patterns.

- Second-order analysis provides insights into the global aspects of the point pattern (patterns of clustering and regularity).
- First-order analysis provides us local insight into where the patterns differ.

Two processes can have the same intensity and K function yet be two very different processes.

Intensities and K-functions provide partial insight into the behavior of the probabilistic mechanisms generating our point processes.

5.4 Other Spatial Point Processes

Homogeneous and heterogeneous Poisson point processes are only two of many different types of point processes.

The **Poisson cluster process** defines a point process where every event is part of a cluster.

- A set of parent locations is generated (usually as a Poisson process).
- Next, child locations are generated relative to each parent location.
    - The number of children usually follows a Poisson process.
    - The child locations follow a spatial probability density function, typically peaked at the parent locations.

- Only the child locations are observed.
- Some correction for edge effects must be made.
- The Neyman-Scott processes are a subclass of the Poisson cluster process.

Contagion/Inhibition Processes focus on directly modeling interevent interactions such that the occurrence of an event raises (or lowers) the probability of observing subsequent events nearby.

- Contagion models are particularly appropriate for modeling the spread of infection diseases.
- Inhibition models are appropriate when each event precludes the occurrence of other events in a nearby area (e.g., the territory of an animal).
- A wide variety of models are used to model inhibition or contagion processes.

Cox processes assume that both the events AND the intensity function are drawn from a probability distribution.

- They are said to be doubly stochastic because the random location of events depends on a random process itself.
- E.g., a process is CSR but the intensity function is drawn from a probability distribution.
- **Log Gaussian Cox processes** are a subclass of Cox processes that allow for complex modeling of the intensity function based on covariates and spatial correlation.

5.4.4 Distinguishing Processes

It is difficult to distinguish the different types of spatial processes based on one realization of events.

- Multiple realizations may allow us to do this.
- Cox processes are Neyman-Scott processes under certain conditions (the converse is much more difficult to prove).