A short introduction to models and inference for spatial point processes.

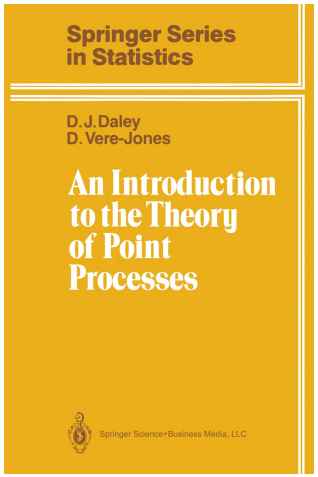
Frédéric Lavancier

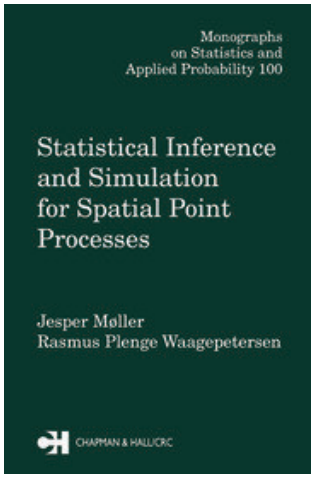
1. Examples of points patterns

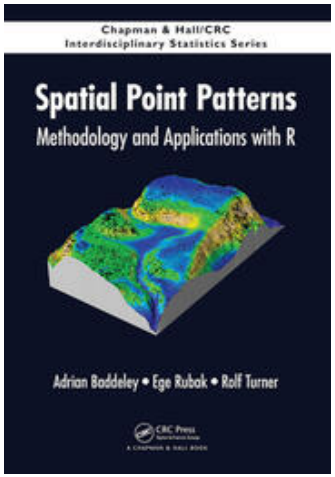
2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

References





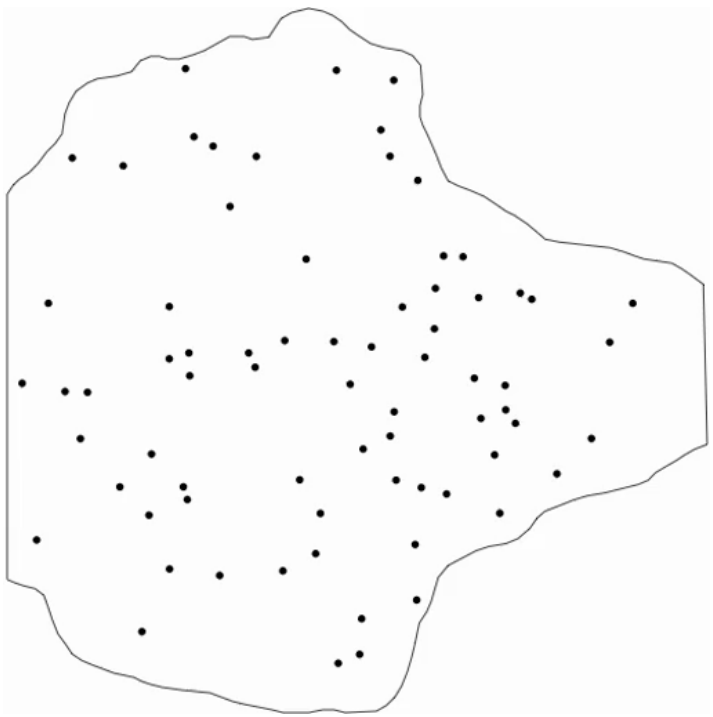


- D.J Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes.* Springer Series in Statistics, 1998.
- J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton, 2004.
- A. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns : Methodology and Applications with R.* CRC Press, 2015.

All illustrations have been implemented with the **spatstat library** in **R**. Almost all presented datasets also come from this library.

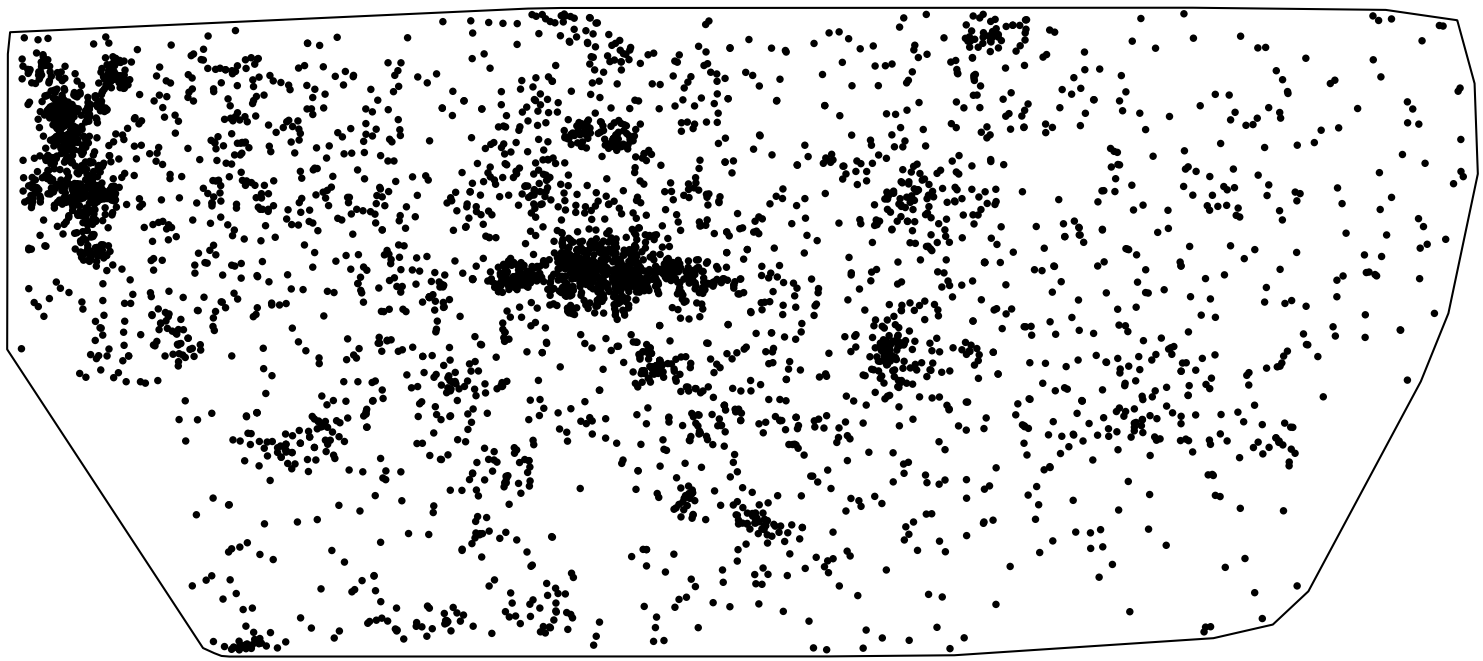
Some examples of point patterns: 2D

Locations of some (Langerin) **proteins** in a living cell.



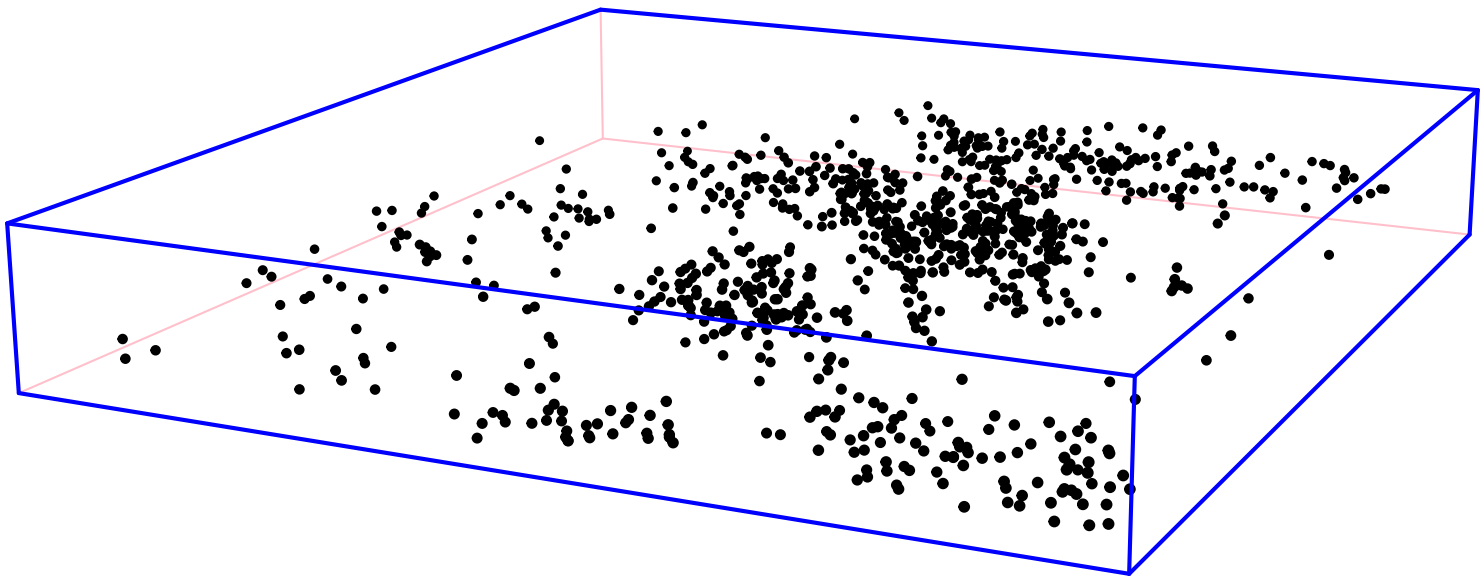
Some examples of point patterns: 2D

Positions of 4215 **galaxies** in the Shapley Supercluster.



Some examples of point patterns: 3D

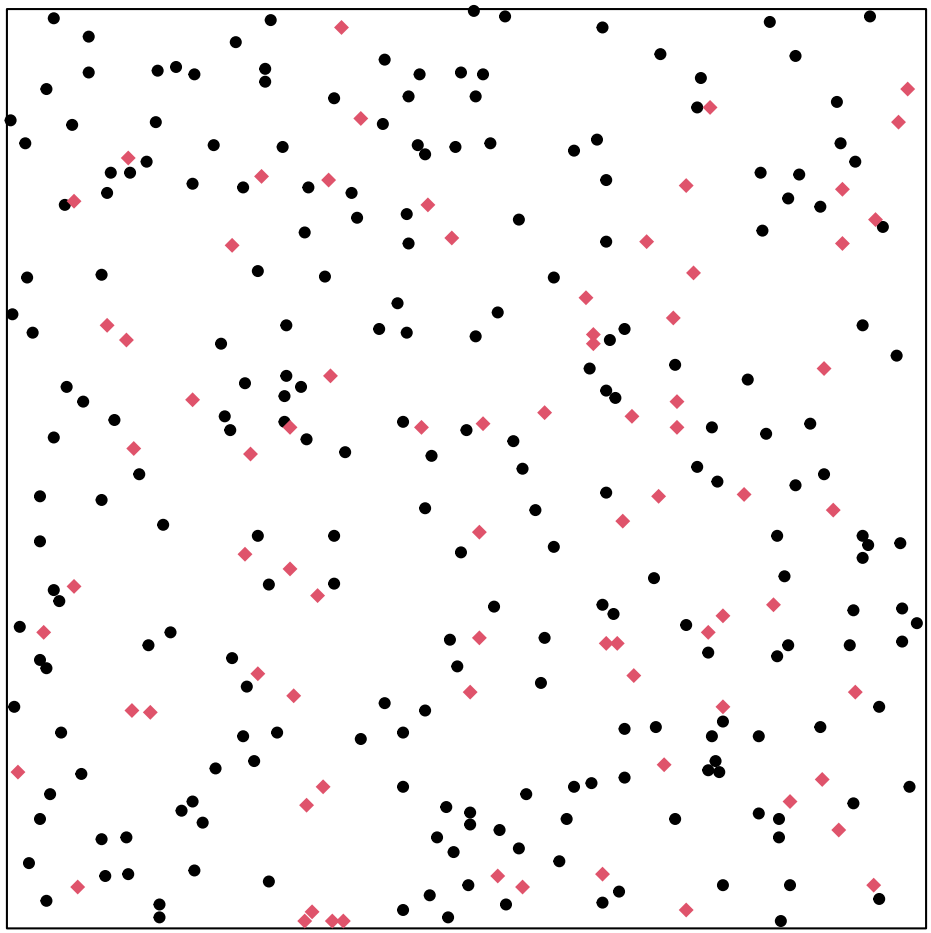
Locations of **nucleosomes** in a cell nucleus of a brain.



Some examples of point patterns: with discrete marks

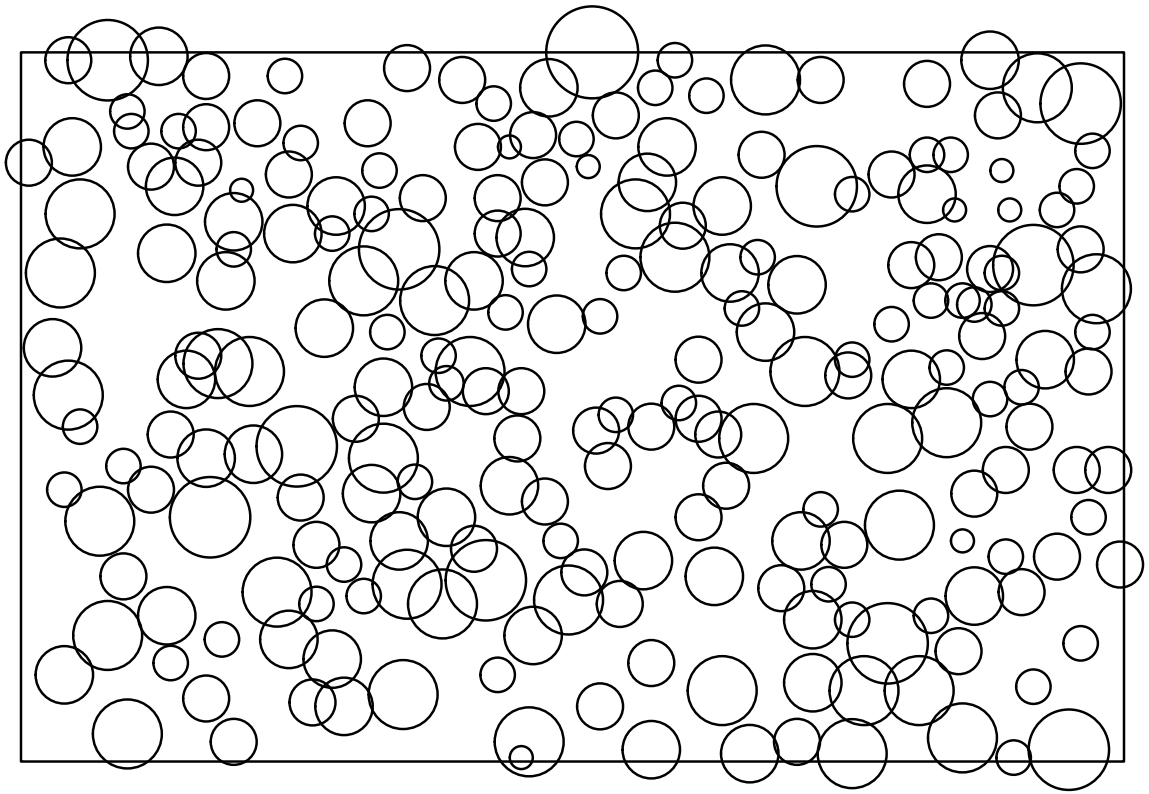
Cell nuclei in hamster kidney, subject to a metastatic lymphoma.

In black: "dividing" cells
In red: "pyknotic" cells, i.e. dying cells



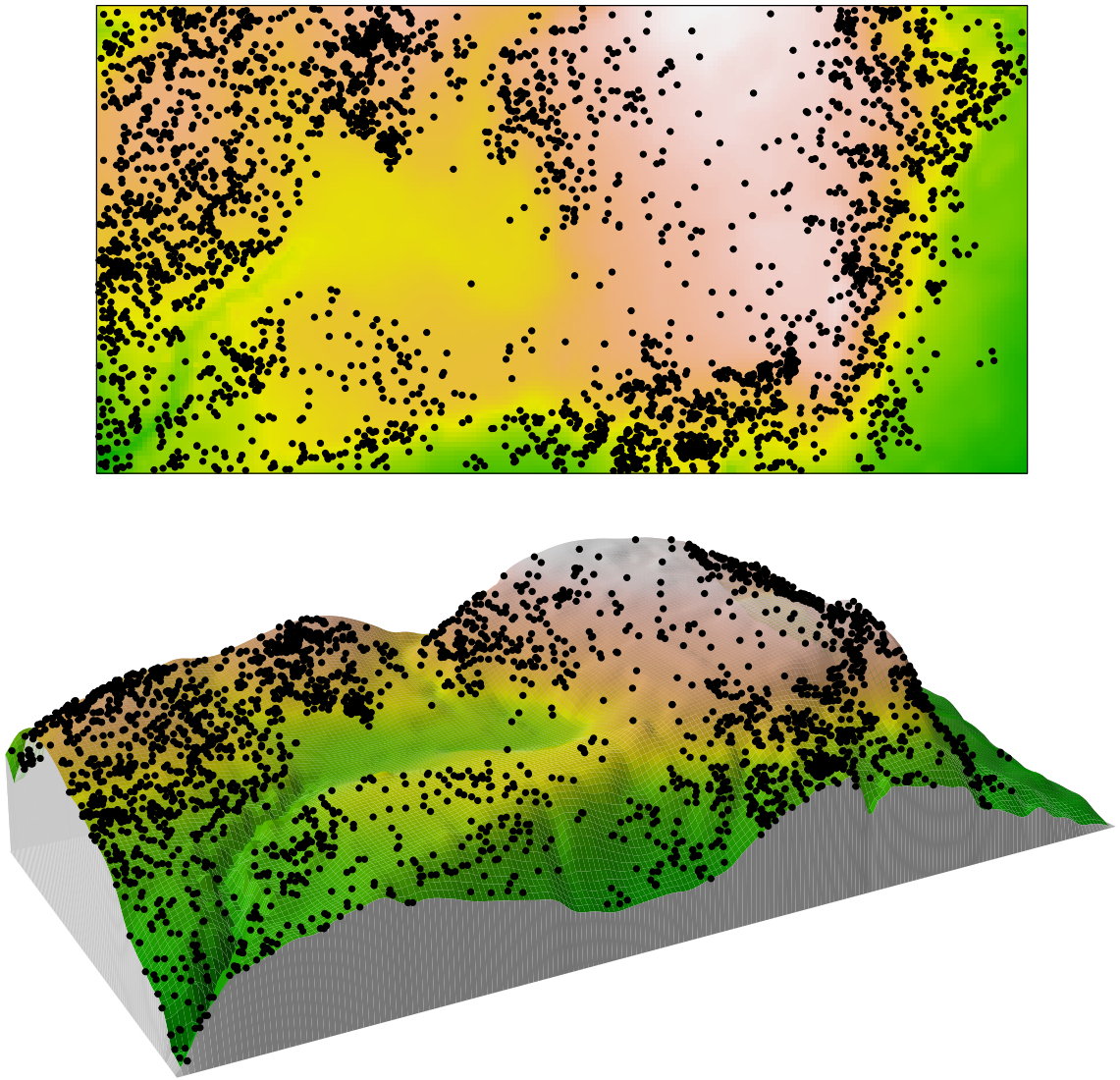
Some examples of point patterns: continuous marks

Locations and diameters of **sea anemones** (at Quiberon)



Some examples of point patterns: with auxiliary information

Locations of **trees** in a tropical rain forest.
Auxiliary information: **elevation** in the study region



The setting

In this presentation, for simplicity:

- we work on a **bounded** set $S \subset \mathbb{R}^d$ (typically $d = 2$)
- the point patterns are **simple** (no duplicated points)
- there is **no mark** (neither discrete, nor continuous)
- there are **no auxiliary information** (no covariate)
- we do not consider temporal, nor space-time, point processes

The main goal is to analyse/model the repartition of points.

What is a spatial point process?

A **point pattern x** in S is a locally finite subset of S.

Denoting by $n(A)$ the cardinality of any set $A \subseteq S$, this means:

$$\begin{cases} \mathbf{x} \subset S, \\ \text{for all } A \subseteq S, \ n(\mathbf{x} \cap A) < \infty. \end{cases}$$

Hence, it makes sense to write $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$ where $n(\mathbf{x}) = n(\mathbf{x} \cap S)$ and $x_i \in S$ for all $i = 1, \ldots, n(\mathbf{x})$.

What is a spatial point process?

A **point pattern x** in S is a locally finite subset of S.

Denoting by $n(A)$ the cardinality of any set $A \subseteq S$, this means:

$$\begin{cases} \mathbf{x} \subset S, \\ \text{for all } A \subseteq S, \ n(\mathbf{x} \cap A) < \infty. \end{cases}$$

Hence, it makes sense to write $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$ where $n(\mathbf{x}) = n(\mathbf{x} \cap S)$ and $x_i \in S$ for all $i = 1, \ldots, n(\mathbf{x})$.

A **spatial point process X** is a **random point pattern** in S.

- The number of points $n(\mathbf{X})$ is random.
- The locations of the points in S are random.

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

First order intensity ρ

The first order intensity ρ answers two questions:

- How many points can be expected?
- Where are the points more likely to occur?

First order intensity ρ

The first order intensity ρ answers two questions:

- How many points can be expected?
- Where are the points more likely to occur?

By definition, the intensity ρ satisfies:

$$\forall A \subseteq S, \quad \mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\sum_{u \in \mathbf{X}} \mathbf{1}_{u \in A}\right) = \int_A \rho(u) du.$$

Intuitively, for $u \in S$, $\rho(u) \approx \mathbb{P}(\mathbf{X} \text{ has a point at } u)$.

In particular: $\mathbb{E}(n(\mathbf{X})) = \int_S \rho(u) du$.

First order intensity ρ

An important special case : the **homogeneous** case

$$\rho(u) = \rho \quad \text{for any } u \in S.$$

Then,

- for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \rho|A|$
- the points are equally likely to appear anywhere.

First order intensity ρ

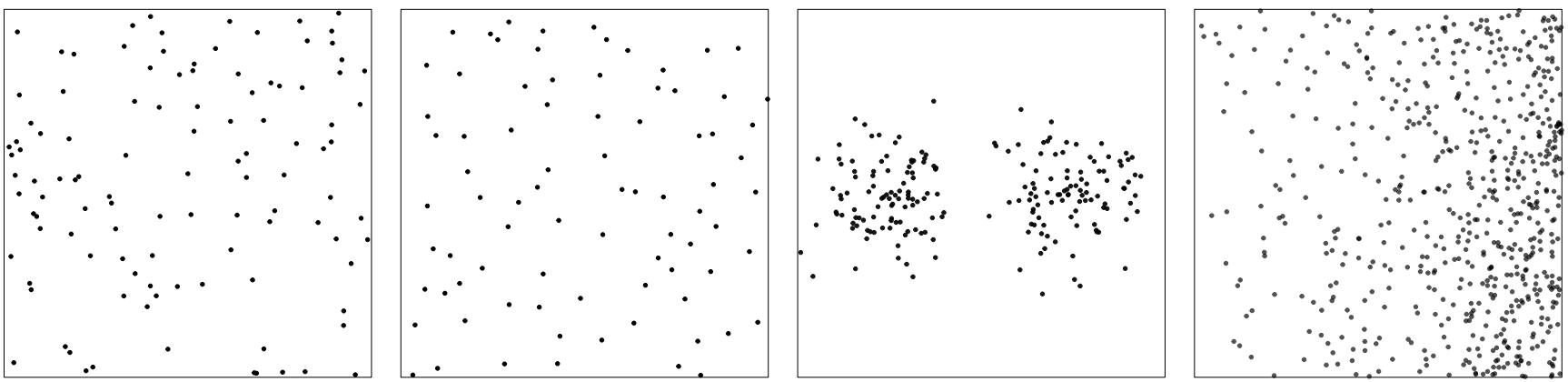
An important special case : the **homogeneous** case

$$\rho(u) = \rho \quad \text{for any } u \in S.$$

Then,

- for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \rho|A|$
- the points are equally likely to appear anywhere.

Otherwise, the point process is **inhomogeneous**.



Homogeneous Homogeneous Inhomogeneous Inhomogeneous

First order intensity ρ: estimation

Assume we observe a realisation $\mathbf{x}$ of $\mathbf{X}$ in S. How can we estimate ρ?

First order intensity ρ: estimation

Assume we observe a realisation $\mathbf{x}$ of $\mathbf{X}$ in S. How can we estimate ρ?

If $\mathbf{X}$ is assumed to be homogeneous : $\hat{\rho} = n(\mathbf{x})/|S|$.

First order intensity ρ: estimation

Assume we observe a realisation $\mathbf{x}$ of $\mathbf{X}$ in S. How can we estimate ρ?

If $\mathbf{X}$ is assumed to be homogeneous : $\hat{\rho} = n(\mathbf{x})/|S|$.

If $\mathbf{X}$ is **not** assumed to be homogeneous:

- non parametric estimation of $u \mapsto \rho(u)$: for a density kernel k (typically the standard Gaussian density), and a bandwidth $h > 0$,

$$\hat{\rho}(u) = \sum_{v \in \mathbf{x}} \frac{1}{h^d} k\left(\frac{v-u}{h}\right) / K_h(v), \quad u \in S,$$

where $K_h(v) = h^{-d} \int_S k\left(\frac{v-u}{h}\right) du$ accounts for edge effects ($K_h(v) \approx 1$ if v is far from the border of S).

Remark: $\int_S \hat{\rho}(u)du = n(x)$ in agreement with $\int_S \rho(u)du = \mathbb{E}(n(x))$.

The choice of h is crucial: adaptive choices are available.

First order intensity ρ: estimation

Assume we observe a realisation $\mathbf{x}$ of $\mathbf{X}$ in S. How can we estimate ρ?

If $\mathbf{X}$ is assumed to be homogeneous : $\hat{\rho} = n(\mathbf{x})/|S|$.

If $\mathbf{X}$ is **not** assumed to be homogeneous:

- non parametric estimation of $u \mapsto \rho(u)$: for a density kernel k (typically the standard Gaussian density), and a bandwidth $h > 0$,

$$\hat{\rho}(u) = \sum_{v \in \mathbf{x}} \frac{1}{h^d} k\left(\frac{v-u}{h}\right) / K_h(v), \quad u \in S,$$

  where $K_h(v) = h^{-d} \int_S k\left(\frac{v-u}{h}\right) du$ accounts for edge effects ($K_h(v) \approx 1$ if v is far from the border of S).

  Remark: $\int_S \hat{\rho}(u)du = n(x)$ in agreement with $\int_S \rho(u)du = \mathbb{E}(n(x))$.

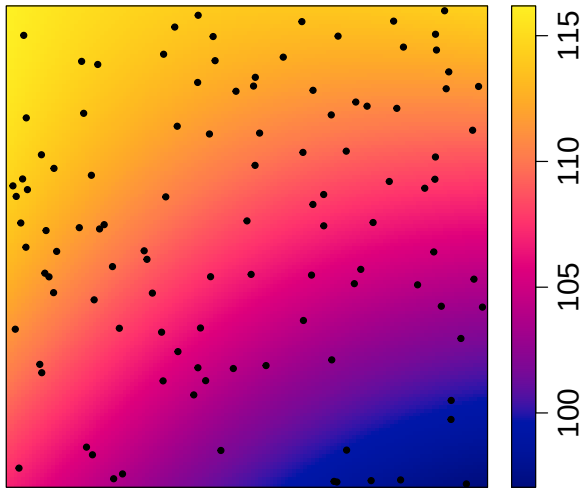
  The choice of h is crucial: adaptive choices are available.

- if that makes sense, we can assume a parametric form for $\rho(u)$, i.e. $\rho(u) = \rho_\theta(u)$, and estimate the parameter θ by specific methods.

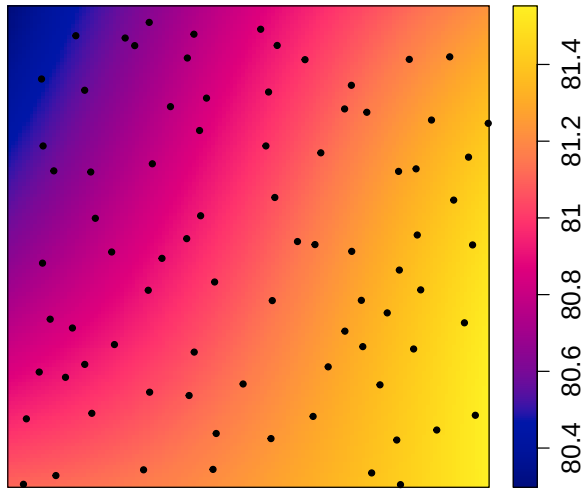
  Ex: $\rho_\theta(u) = \exp(\theta_1 z_1(u) + \theta_2 z_2(u))$ where z_1, z_2 are auxiliary variables.

First order intensity ρ: estimation

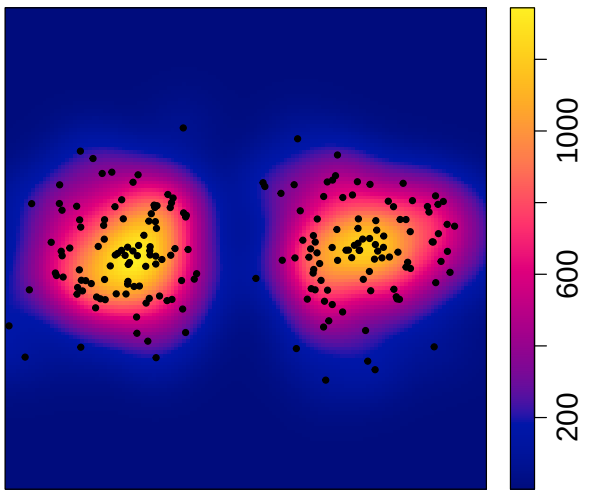
Examples of non-parametric estimation of the intensity



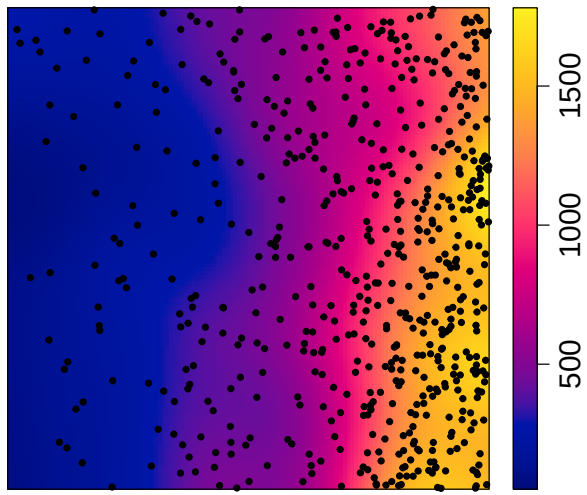
Homogeneous ($\rho = 100$)



Homogeneous ($\rho = 81$)



Inhomogeneous



Inhomogeneous

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Second order moments : the second order intensity

Aim: analyse the **interaction** between the points ($\approx$ spatial covariance)

Second order moments : the second order intensity

Aim: analyse the **interaction** between the points ($\approx$ spatial covariance)

The **second order intensity** $\rho^{(2)}(u, v)$ satisfies

$$\forall A, B \subseteq S, \quad \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \mathbf{1}_{u \in A,\, v \in B}\right) = \int_A \int_B \rho^{(2)}(u, v) du dv.$$

Intuitively, $\rho^{(2)}(u, v) \approx P(\mathbf{X}$ has a point at u and a point at $v)$.

Second order moments : the second order intensity

Aim: analyse the **interaction** between the points ($\approx$ spatial covariance)

The **second order intensity** $\rho^{(2)}(u,v)$ satisfies

$$\forall A, B \subseteq S, \quad \mathbb{E}\left(\sum_{u,v\in\mathbf{X}}^{\neq} \mathbf{1}_{u\in A,\, v\in B}\right) = \int_A \int_B \rho^{(2)}(u,v)dudv.$$

Intuitively, $\rho^{(2)}(u,v) \approx P(\mathbf{X}$ has a point at u and a point at $v)$.

Interpretation:

- If the points are located independently of each other, then
$$\rho^{(2)}(u,v) = \rho(u)\rho(v)$$
- If there is a positive dependence (attraction), then
$$\rho^{(2)}(u,v) > \rho(u)\rho(v)$$
- If there is a negative dependence (inhibition), then
$$\rho^{(2)}(u,v) < \rho(u)\rho(v)$$

Second order moments: the pcf

In spatial statistics, we rather use the **pair correlation function** (pcf) g

$$g(u,v) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)}.$$

- If $g(u,v) = 1$, there is no interaction between u and v,
- If $g(u,v) > 1$, there is attraction,
- If $g(u,v) < 1$, there is inhibition.

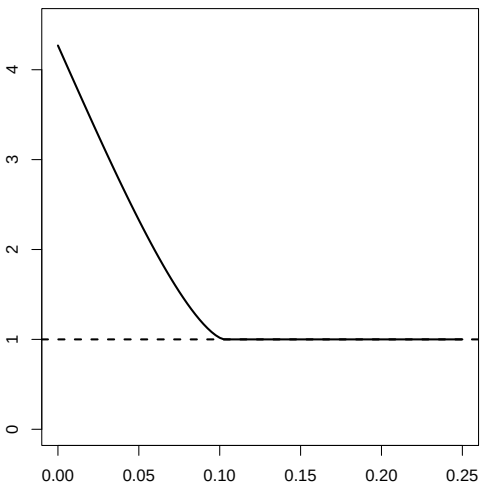
In most applications, it is assumed that $g(u,v)$ only depends on $\|u-v\|$. We then simply define for $r > 0$

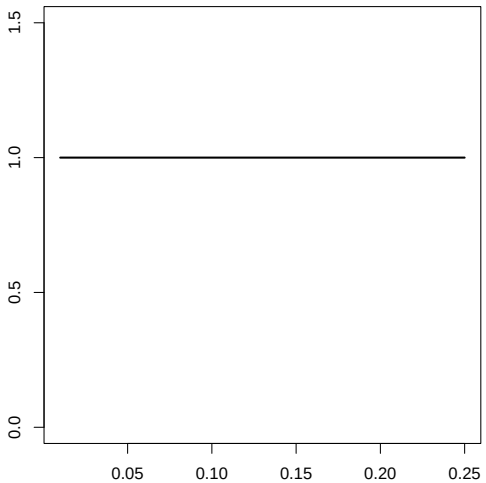
$$g(r) = \frac{\rho^{(2)}(u,v)}{\rho(u)\rho(v)} \quad \text{for } \|u-v\| = r.$$

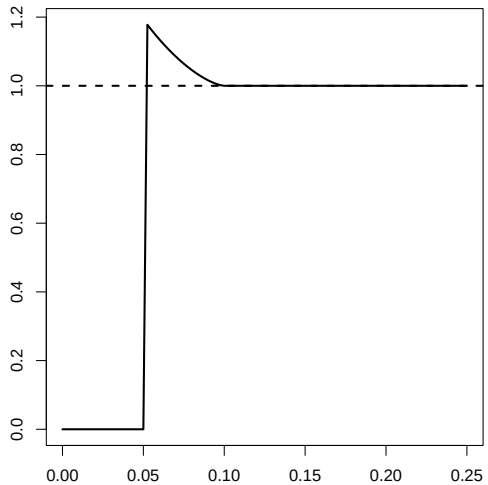
This is the case for a stationary and isotropic point process.

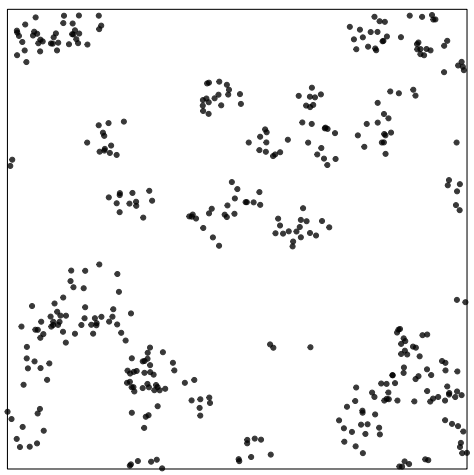
Second order moments: the pcf

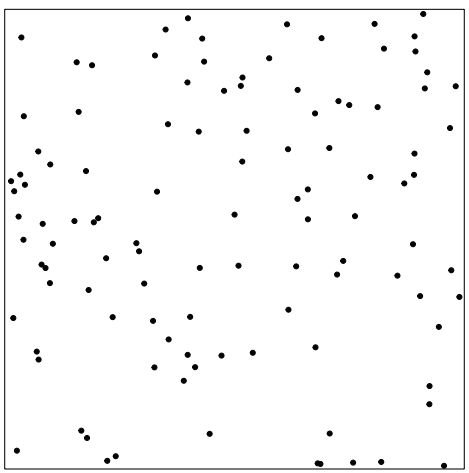
Examples of theoretical pcf g, with one realisation for each.

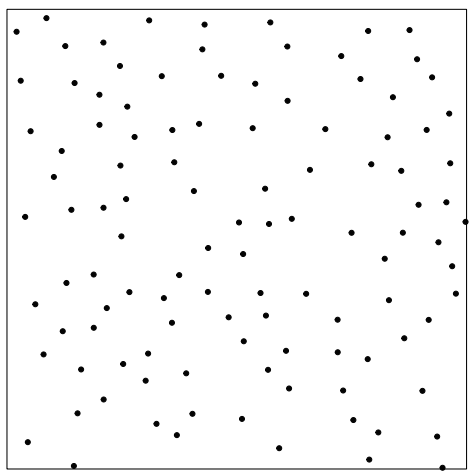












Second order moments: Ripley's K function

Assume that $\mathbf{X}$ is stationary and isotropic, with intensity ρ.

The **Ripley's K function** is defined for any $r > 0$ by

$\rho K(r) =$ expected number of neighbours within distance r of a point $u \in \mathbf{X}$.

Second order moments: Ripley's K function

Assume that $\mathbf{X}$ is stationary and isotropic, with intensity ρ.

The **Ripley's K function** is defined for any $r > 0$ by

$\rho K(r) =$ expected number of neighbours within distance r of a point $u \in \mathbf{X}$.

Mathematically, denoting $B(u, R)$ the ball centred at u with radius R,

$$K(r) = \frac{1}{\rho}\mathbb{E}\Big(n(\mathbf{X} \cap B(u, r)) - 1 \Big| u \in \mathbf{X}\Big),$$
$$= \frac{1}{\rho}\mathbb{E}\Big(n(\mathbf{X} \cap B(0, r)) - 1 \Big| 0 \in \mathbf{X}\Big) \quad \text{by stationarity of } \mathbf{X}.$$

Second order moments: Ripley's K function

Assume that $\mathbf{X}$ is stationary and isotropic, with intensity ρ.

The **Ripley's K function** is defined for any $r > 0$ by

$\rho K(r) =$ expected number of neighbours within distance r of a point $u \in \mathbf{X}$.

Mathematically, denoting $B(u, R)$ the ball centred at u with radius R,

$$\begin{aligned} K(r) &= \frac{1}{\rho}\mathbb{E}\Big(n(\mathbf{X} \cap B(u, r)) - 1 \Big| u \in \mathbf{X}\Big), \\ &= \frac{1}{\rho}\mathbb{E}\Big(n(\mathbf{X} \cap B(0, r)) - 1 \Big| 0 \in \mathbf{X}\Big) \quad \text{by stationarity of } \mathbf{X}. \end{aligned}$$

We can prove that

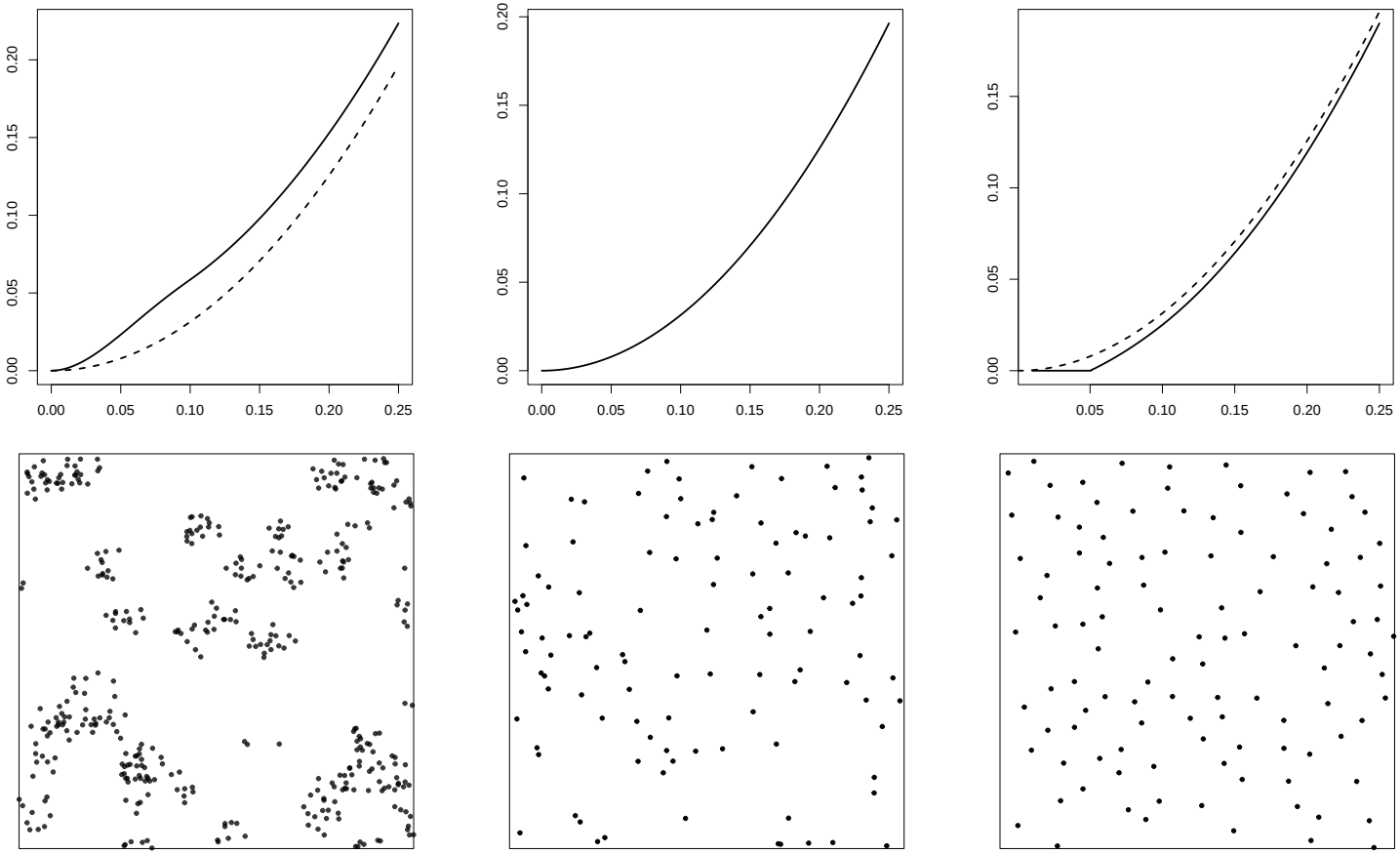
$$K(r) = \int_{B(0,r)} g(\|u\|)du = d\omega_d \int_0^r t^{d-1} g(t)dt,$$

where $\omega_d = |B(0, 1)| = \pi^{d/2}/\Gamma(1 + d/2)$.

Second order moments: Ripley's K function

Remember: $g(r) = 1 \Rightarrow$ no interaction. So:

$$\begin{cases} K(r) = \omega_d r^d \Rightarrow & \text{no interaction,} \\ K(r) > \omega_d r^d \Rightarrow & \text{attraction,} \\ K(r) < \omega_d r^d \Rightarrow & \text{repulsion.} \end{cases}$$



Second order moments: estimation of $K(r)$

Second order moments: estimation of $K(r)$

Since

$\rho K(r) =$ expected number of neighbours within distance r of a point $u \in \mathbf{X}$,

a natural estimator for $K(r)$ based on a realisation $\mathbf{x}$ is

$$\tilde{K}(r) = \frac{1}{\hat{\rho}} \frac{1}{n(\mathbf{x})} \sum_{u \in \mathbf{x}} [n(\mathbf{x} \cap B(u, r)) - 1] = \frac{1}{\hat{\rho}} \frac{1}{n(\mathbf{x})} \sum_{u,v \in \mathbf{x}}^{\neq} \mathbf{1}_{\|u-v\|<r}.$$

Second order moments: estimation of $K(r)$

Since

$\rho K(r) =$ expected number of neighbours within distance r of a point $u \in \mathbf{X}$,

a natural estimator for $K(r)$ based on a realisation $\mathbf{x}$ is

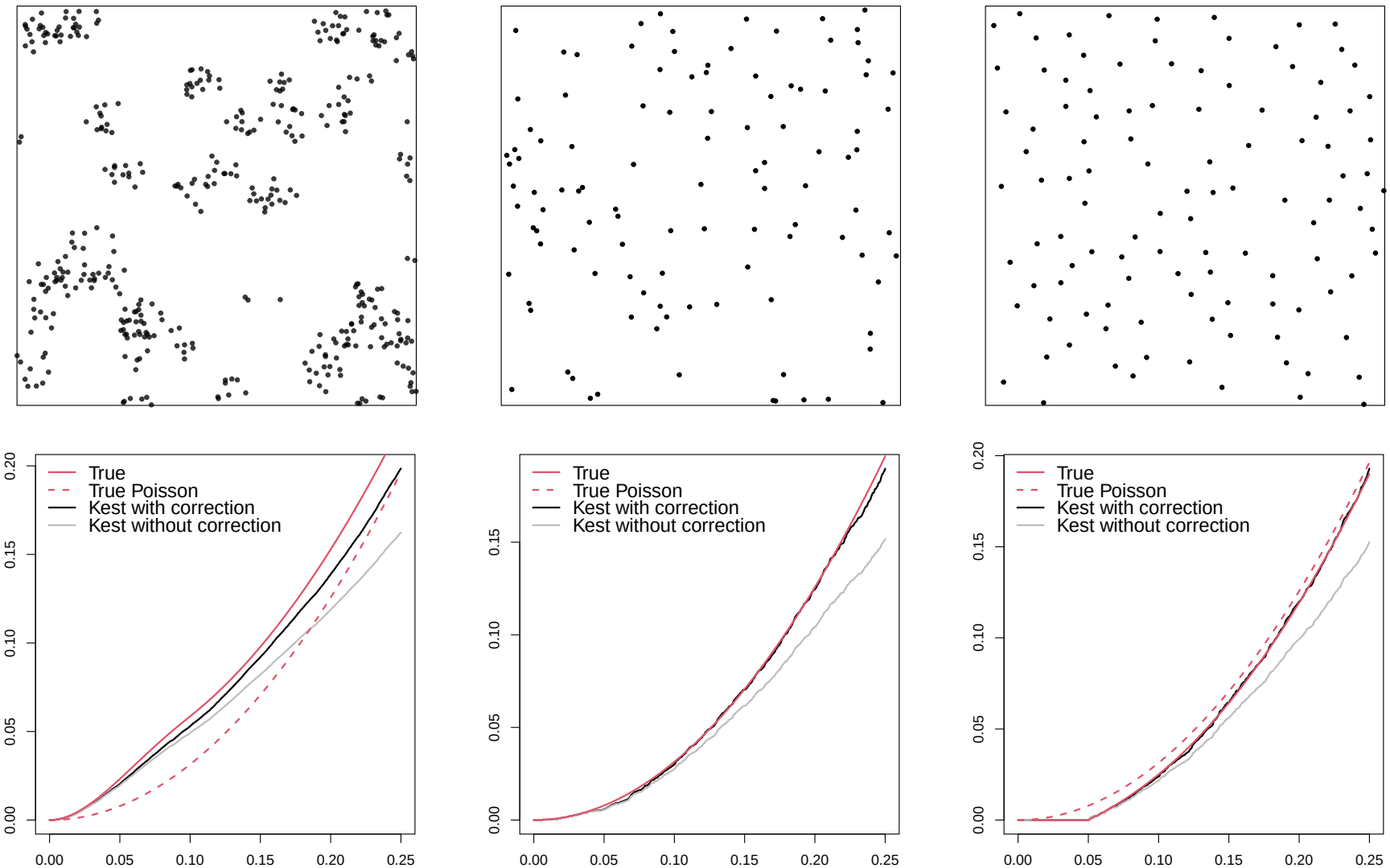
$$\tilde{K}(r) = \frac{1}{\hat{\rho}} \frac{1}{n(\mathbf{x})} \sum_{u \in \mathbf{x}} [n(\mathbf{x} \cap B(u, r)) - 1] = \frac{1}{\hat{\rho}} \frac{1}{n(\mathbf{x})} \sum_{u,v \in \mathbf{x}}^{\neq} \mathbf{1}_{\|u-v\|<r}.$$

But there are edge effects! One solution is to use:

$$\hat{K}(r) = \frac{1}{\hat{\rho}^2} \sum_{u,v \in \mathbf{x}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|},$$

where $S_{u-v} = S + (u - v)$ is the translation of S by $(u - v)$.

Second order moments: estimation of $K(r)$



Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

$$\text{(by def of } \rho^{(2)}) \quad = \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

$$\text{(by def of } \rho^{(2)}) \quad = \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{v \in S} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} g(\|u-v\|) du dv$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

$$\text{(by def of } \rho^{(2)}) \quad = \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{v \in S} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} g(\|u-v\|) du dv$$

$$(w = u - v) \quad = \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{u-w \in S} \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) du dw$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

$$\text{(by def of } \rho^{(2)}) \quad = \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{v \in S} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} g(\|u-v\|) du dv$$

$$(w = u - v) \quad = \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{u-w \in S} \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) du dw$$

$$= \rho^2 \int \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) \left(\int \mathbf{1}_{u \in S \cap S_w} du\right) dw$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

$$\text{(by def of } \rho^{(2)}) \quad = \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{v \in S} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} g(\|u-v\|) du dv$$

$$(w = u - v) \quad = \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{u-w \in S} \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) du dw$$

$$= \rho^2 \int \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) \left(\int \mathbf{1}_{u \in S \cap S_w} du\right) dw$$

$$= \rho^2 \int \mathbf{1}_{\|w\|<r} g(\|w\|) \frac{|S \cap S_w|}{|S \cap S_w|} dw$$

Second order moments: estimation of $K(r)$

Why is it a good correction? Let us prove that $\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \rho^2 K(r)$.

$$\mathbb{E}\left(\hat{\rho}^2 \hat{K}(r)\right) = \mathbb{E}\left(\sum_{u,v \in \mathbf{X}}^{\neq} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|}\right)$$

(by def of $\rho^{(2)}$)
$$= \int_S \int_S \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} \rho^{(2)}(u,v) du dv$$

$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{v \in S} \frac{\mathbf{1}_{\|u-v\|<r}}{|S \cap S_{u-v}|} g(\|u-v\|) du dv$$

($w = u - v$)
$$= \rho^2 \int \int \mathbf{1}_{u \in S} \mathbf{1}_{u-w \in S} \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) du dw$$

$$= \rho^2 \int \frac{\mathbf{1}_{\|w\|<r}}{|S \cap S_w|} g(\|w\|) \left(\int \mathbf{1}_{u \in S \cap S_w} du\right) dw$$

$$= \rho^2 \int \mathbf{1}_{\|w\|<r} g(\|w\|) \frac{|S \cap S_w|}{|S \cap S_w|} dw$$

$$= \rho^2 K(r).$$

Second order moments: estimation of $g(r)$

Second order moments: estimation of $g(r)$

Assume again that $\mathbf{X}$ is stationary and isotropic, with intensity ρ.

Remember that $g(r) = \rho^{(2)}(u, v)/\rho^2$ where $r = \|u - v\|$.

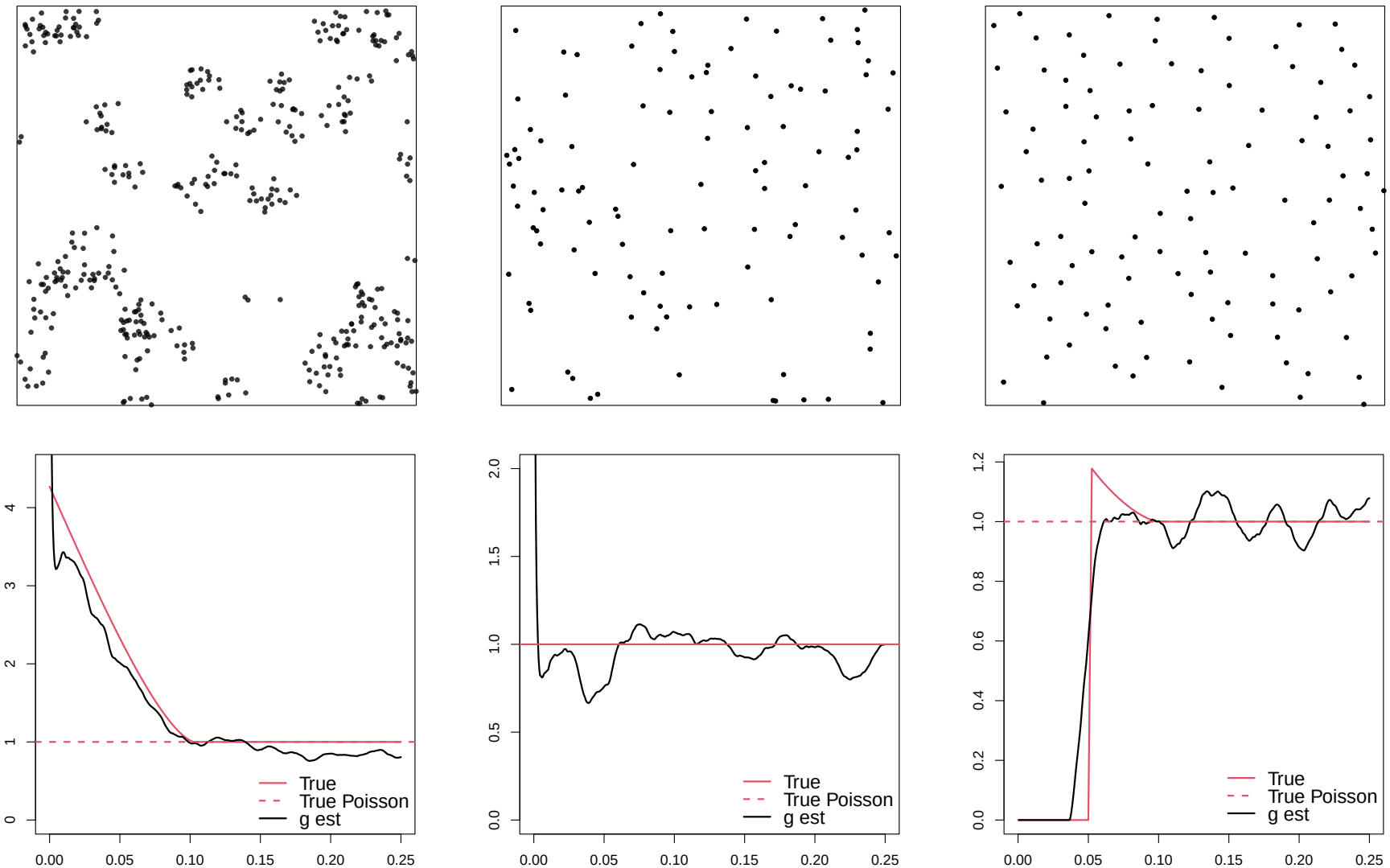
For a kernel k (e.g. a Gaussian density) and a bandwidth $h > 0$,

$$\hat{g}(r) = \frac{1}{\hat{\rho}^2} \frac{1}{d\omega_d r^{d-1}} \sum_{u,v \in \mathbf{x}}^{\neq} \frac{1}{h^d} k\left(\frac{\|v - u\| - r}{h}\right) \frac{1}{|S \cap S_{u-v}|}$$

Red term: due to polar coordinates ($\omega_d = |B(0, 1)|$)
Blue term: same edge correction as before
The sum: we are "counting" the number of pairs (u, v) at distance $\approx r$.

As for all non-parametric method, the choice of h is crucial.

Second order moments: estimation of $g(r)$



1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Other summary statistics

Assume again that $\mathbf{X}$ is stationary and isotropic, with intensity ρ.

- The L-function: $L(r) = (K(r)/\omega_d)^{1/d}$, with $\omega_d = |B(0,1)|$.
- The empty space function

$$F(r) = \mathbb{P}(n(\mathbf{X} \cap B(0,r)) > 0).$$

- The nearest neighbour distribution function

$$G(r) = \mathbb{P}(n(\mathbf{X} \cap B(0,r)) > 1 \,|\, 0 \in \mathbf{X}).$$

- The J-function: $J(r) = (1 - G(r))/(1 - F(r))$.

Interpretation:
No interaction : $L_0(r) = r$, $F_0(r) = G_0(r) = 1 - e^{-\rho\omega_d r^d}$ and $J_0(r) = 1$.
Attraction if $L(r) > r$, $F(r) < F_0(r)$, $G(r) > G_0(r)$, $J(r) < 1$.
Inhibition if $L(r) < r$, $F(r) > F_0(r)$, $G(r) < G_0(r)$, $J(r) > 1$.

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

For each of the considered models, the main questions are:

- What are its main characteristics: attraction? inhibition?
- Do we know its moments (e.g. ρ and g)?
- How can we simulate a realisation?
- How can we fit this model to a point pattern?

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Homogeneous Poisson point process

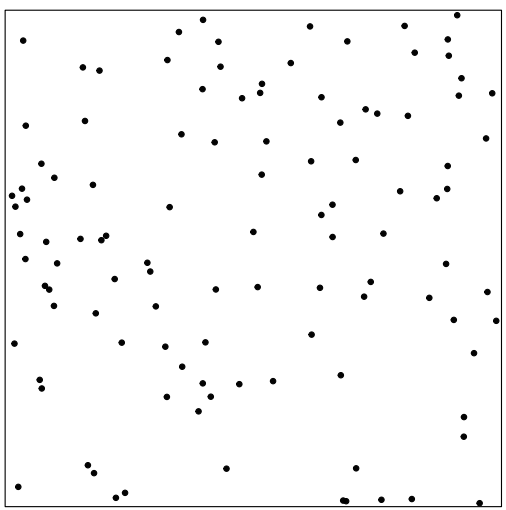
A homogeneous PPP $\mathbf{X}$ with intensity ρ on S satisfies:

- $n(\mathbf{X}) \sim \mathcal{P}(\rho|S|)$, i.e.

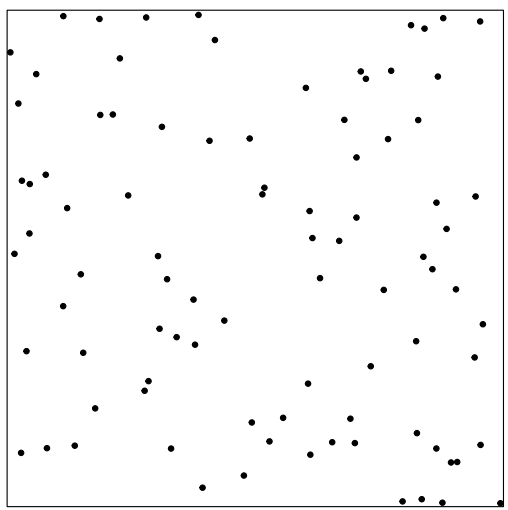
$$P(n(\mathbf{X}) = k) = e^{-\rho|S|} \frac{(\rho|S|)^k}{k!}, \quad \forall\, k = 0, 1, 2, \ldots,$$

- Given $n(\mathbf{X}) = n$, $\{x_1, \ldots, x_n\}$ are **independent** and uniform in S.

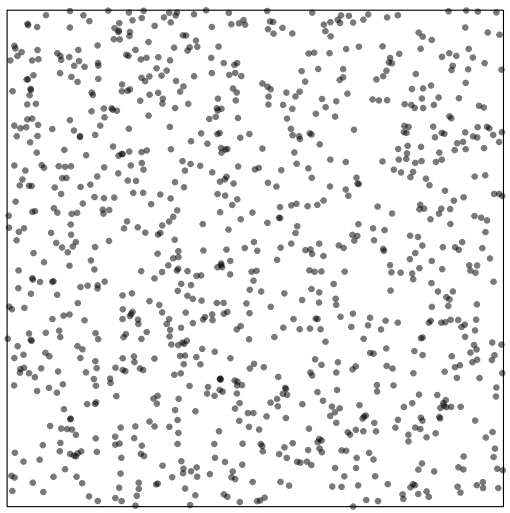
The simulation is straightforward.



$\rho = 100$



$\rho = 100$

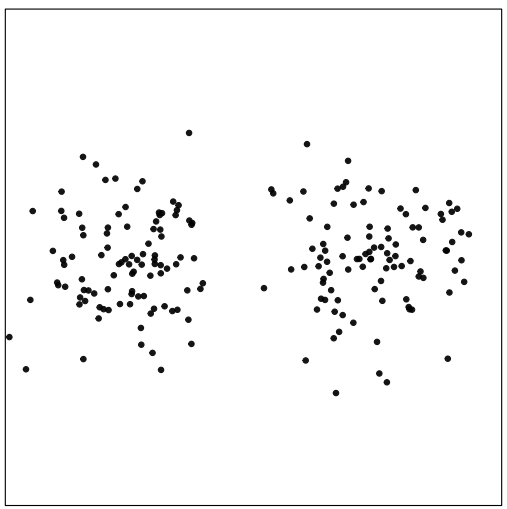


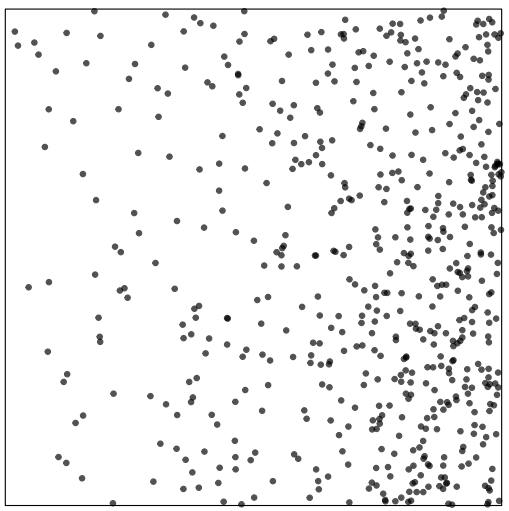
$\rho = 1000$

Inhomogeneous Poisson point process

An inhomogeneous PPP with intensity function $\rho(u)$ on S satisfies:

- $n(\mathbf{X}) \sim \mathcal{P}(\int_S \rho(u)du)$,
- Given $n(\mathbf{X}) = n$, $\{x_1, \ldots, x_n\}$ are **independently** distributed in S according to the density $u \mapsto \rho(u)/(\int_S \rho(u)du)$.





Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

  $$\mathbb{E}(n(\mathbf{X} \cap A)) = \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n)$$

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n)$$

$$= \sum_{n \geq 0} n \int_A \frac{\rho(u)}{c} du\, e^{-c} \frac{c^n}{n!} \quad \left(\text{since } x_i | n(\mathbf{X}) \text{ are iid} \sim \frac{\rho(u)}{c} \right)$$

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n)$$

$$= \sum_{n \geq 0} n \int_A \frac{\rho(u)}{c} du\, e^{-c} \frac{c^n}{n!} \quad \left(\text{since } x_i | n(\mathbf{X}) \text{ are iid} \sim \frac{\rho(u)}{c} \right)$$

$$= \left(\int_A \rho(u)du \right) \sum_{n \geq 1} e^{-c} \frac{c^{n-1}}{(n-1)!} = \int_A \rho(u)du.$$

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

$$\begin{aligned}
\mathbb{E}(n(\mathbf{X} \cap A)) &= \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n) \\
&= \sum_{n \geq 0} n \int_A \frac{\rho(u)}{c} du\, e^{-c} \frac{c^n}{n!} \quad \left(\text{since } x_i | n(\mathbf{X}) \text{ are iid} \sim \frac{\rho(u)}{c} \right) \\
&= \left(\int_A \rho(u)du \right) \sum_{n \geq 1} e^{-c} \frac{c^{n-1}}{(n-1)!} = \int_A \rho(u)du.
\end{aligned}$$

- The second order intensity function is $\rho^{(2)}(u, v) = \rho(u)\rho(v)$.

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

$$\begin{aligned}
\mathbb{E}(n(\mathbf{X} \cap A)) &= \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n) \\
&= \sum_{n \geq 0} n \int_A \frac{\rho(u)}{c} du\, e^{-c} \frac{c^n}{n!} \quad \left(\text{since } x_i | n(\mathbf{X}) \text{ are iid} \sim \frac{\rho(u)}{c} \right) \\
&= \left(\int_A \rho(u)du \right) \sum_{n \geq 1} e^{-c} \frac{c^{n-1}}{(n-1)!} = \int_A \rho(u)du.
\end{aligned}$$

- The second order intensity function is $\rho^{(2)}(u, v) = \rho(u)\rho(v)$.
- Hence $g(r) = 1$ for all $r > 0$ and $K(r) = d\omega_d r^d$.

Poisson point process: moments

For a PPP with intensity function $\rho(u)$:

- The intensity function is indeed $\rho(u)$.

  We need to prove that for any $A \subseteq S$, $\mathbb{E}(n(\mathbf{X} \cap A)) = \int_A \rho(u)du$.
  Denoting $c = \int_S \rho(u)du$ so that $n(\mathbf{X}) \sim \mathcal{P}(c)$,

$$\begin{aligned}
\mathbb{E}(n(\mathbf{X} \cap A)) &= \sum_{n \geq 0} \mathbb{E}(n(\mathbf{X} \cap A) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n) \\
&= \sum_{n \geq 0} n \int_A \frac{\rho(u)}{c} du\, e^{-c} \frac{c^n}{n!} \quad \left(\text{since } x_i | n(\mathbf{X}) \text{ are iid} \sim \frac{\rho(u)}{c} \right) \\
&= \left(\int_A \rho(u)du \right) \sum_{n \geq 1} e^{-c} \frac{c^{n-1}}{(n-1)!} = \int_A \rho(u)du.
\end{aligned}$$

- The second order intensity function is $\rho^{(2)}(u, v) = \rho(u)\rho(v)$.
- Hence $g(r) = 1$ for all $r > 0$ and $K(r) = d\omega_d r^d$.

PPP is the default model for non-interacting points, i.e. Complete Spatial Randomness (CSR).

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) = \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n)$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\begin{aligned}
\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) &= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n) \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) f(u_1, \ldots, u_n) \frac{du_1}{|S|} \cdots \frac{du_n}{|S|}\, e^{-|S|} \frac{|S|^n}{n!}
\end{aligned}$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\begin{aligned}
\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) &= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n) \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) f(u_1, \ldots, u_n) \frac{du_1}{|S|} \cdots \frac{du_n}{|S|}\, e^{-|S|} \frac{|S|^n}{n!} \\
(c = \int_S \rho(u)du) \quad &= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) e^{-c} \prod_{i=1}^{n} \rho(u_i) du_i\, \frac{1}{n!}
\end{aligned}$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\begin{aligned}
\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) &= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n) \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) f(u_1, \ldots, u_n) \frac{du_1}{|S|} \cdots \frac{du_n}{|S|}\, e^{-|S|} \frac{|S|^n}{n!} \\
(c = \int_S \rho(u)du) \quad &= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) e^{-c} \prod_{i=1}^{n} \rho(u_i) du_i\, \frac{1}{n!} \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) \prod_{i=1}^{n} \frac{\rho(u_i)}{c} du_i\, \frac{1}{n!} c^n e^{-c}
\end{aligned}$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\begin{aligned}
\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) &= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n) \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) f(u_1, \ldots, u_n) \frac{du_1}{|S|} \cdots \frac{du_n}{|S|}\, e^{-|S|} \frac{|S|^n}{n!} \\
(c = \int_S \rho(u)du) \quad &= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) e^{-c} \prod_{i=1}^{n} \rho(u_i) du_i\, \frac{1}{n!} \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) \prod_{i=1}^{n} \frac{\rho(u_i)}{c} du_i\, \frac{1}{n!} c^n e^{-c} \\
&= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n)
\end{aligned}$$

Poisson point process: density (or likelihood)

The reference point process $\mathbf{X}_0$ is the PPP with intensity $\rho = 1$.

A PPP $\mathbf{X}$ with intensity function $\rho(u)$ admits the density wrt $\mathbf{X}_0$:

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i), \quad \forall\, \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Proof:

We need to prove that for any test function h, $\mathbb{E}(h(\mathbf{X})) = \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0))$.

$$\begin{aligned}
\mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0)) &= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}_0) f(\mathbf{X}_0) | n(\mathbf{X}_0) = n)\, \mathbb{P}(n(\mathbf{X}_0) = n) \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) f(u_1, \ldots, u_n) \frac{du_1}{|S|} \cdots \frac{du_n}{|S|}\, e^{-|S|} \frac{|S|^n}{n!} \\
(c = \int_S \rho(u)du) \quad &= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) e^{-c} \prod_{i=1}^{n} \rho(u_i) du_i\, \frac{1}{n!} \\
&= \sum_{n \geq 0} \int_{S^n} h(u_1, \ldots, u_n) \prod_{i=1}^{n} \frac{\rho(u_i)}{c} du_i\, \frac{1}{n!} c^n e^{-c} \\
&= \sum_{n \geq 0} \mathbb{E}(h(\mathbf{X}) | n(\mathbf{X}) = n)\, \mathbb{P}(n(\mathbf{X}) = n) = \mathbb{E}(h(\mathbf{X}))
\end{aligned}$$

Poisson point process: inference

The only parameter to fit is the intensity function $\rho(u)$.

Poisson point process: inference

The only parameter to fit is the intensity function $\rho(u)$.

Given $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$, the log-likelihood is

$$\log f(\mathbf{x}) = |S| - \int_S \rho(u)du + \sum_{i=1}^{n(\mathbf{x})} \log \rho(x_i)$$

Poisson point process: inference

The only parameter to fit is the intensity function $\rho(u)$.

Given $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$, the log-likelihood is

$$\log f(\mathbf{x}) = |S| - \int_S \rho(u)du + \sum_{i=1}^{n(\mathbf{x})} \log \rho(x_i)$$

- For a homogeneous PPP ($\rho(u) = \rho$), $\hat{\rho} = n(\mathbf{x})/|S|$ is the MLE.

Poisson point process: inference

The only parameter to fit is the intensity function $\rho(u)$.

Given $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$, the log-likelihood is

$$\log f(\mathbf{x}) = |S| - \int_S \rho(u)du + \sum_{i=1}^{n(\mathbf{x})} \log \rho(x_i)$$

- For a homogeneous PPP ($\rho(u) = \rho$), $\hat{\rho} = n(\mathbf{x})/|S|$ is the MLE.
- For a parametric inhomogeneous PPP ($\rho(u) = \rho_\theta(u)$),

$$\hat{\theta} = \operatorname*{argmax}_\theta \sum_{i=1}^{n(\mathbf{x})} \log \rho_\theta(x_i) - \int_S \rho_\theta(u)du.$$

Poisson point process: inference

The only parameter to fit is the intensity function $\rho(u)$.

Given $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$, the log-likelihood is

$$\log f(\mathbf{x}) = |S| - \int_S \rho(u)du + \sum_{i=1}^{n(\mathbf{x})} \log \rho(x_i)$$

- For a homogeneous PPP ($\rho(u) = \rho$), $\hat{\rho} = n(\mathbf{x})/|S|$ is the MLE.
- For a parametric inhomogeneous PPP ($\rho(u) = \rho_\theta(u)$),

$$\hat{\theta} = \operatorname*{argmax}_\theta \sum_{i=1}^{n(\mathbf{x})} \log \rho_\theta(x_i) - \int_S \rho_\theta(u)du.$$

- For a general inhomogeneous PPP: non-parametric estimator

$$\hat{\rho}(u) = \frac{1}{K_h(u)} \sum_{v \in \mathbf{x}} \frac{1}{h^d} k\left(\frac{v-u}{h}\right), \quad u \in S.$$

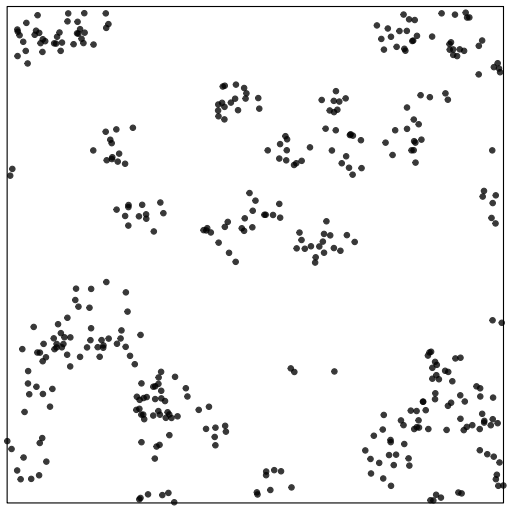
Poisson point process: validation

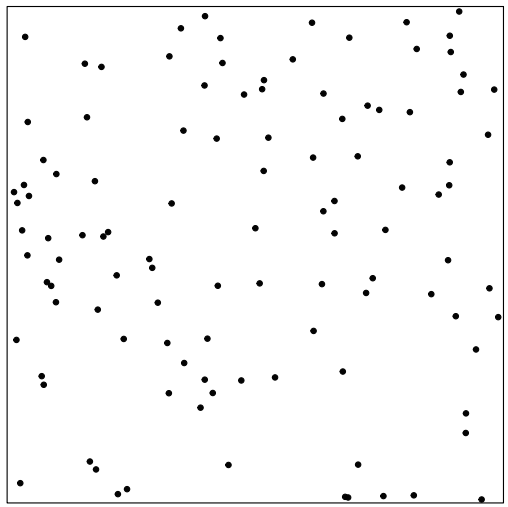
Is a (Poisson) model a good fit to a point pattern?

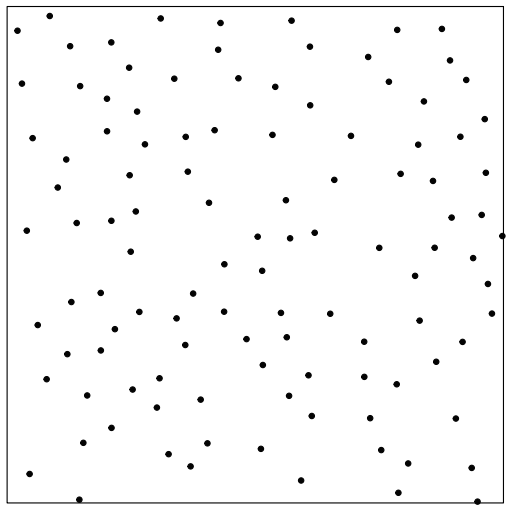
1. Fit the model.
2. Simulate many realisations from the fitted model.
3. Compute a descriptor for each realisation, for instance $\hat{K}(r)$ or $\hat{g}(r)$.
4. Check if the same descriptor for the data is consistent with these simulations.

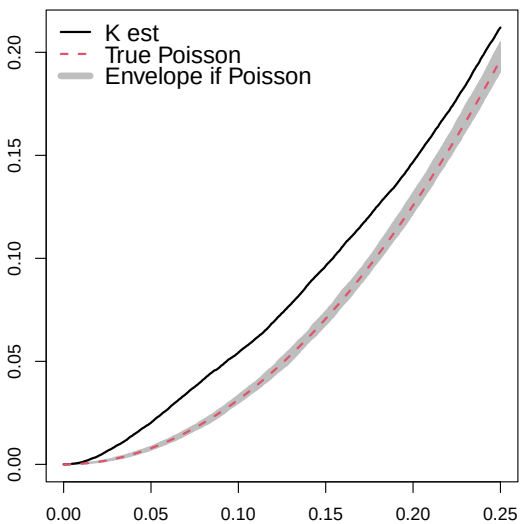
Poisson point process: validation

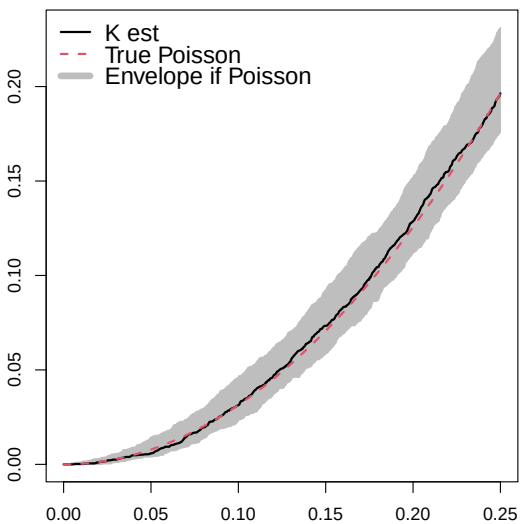
We fit a homogeneous PPP to these point patterns, and check with K

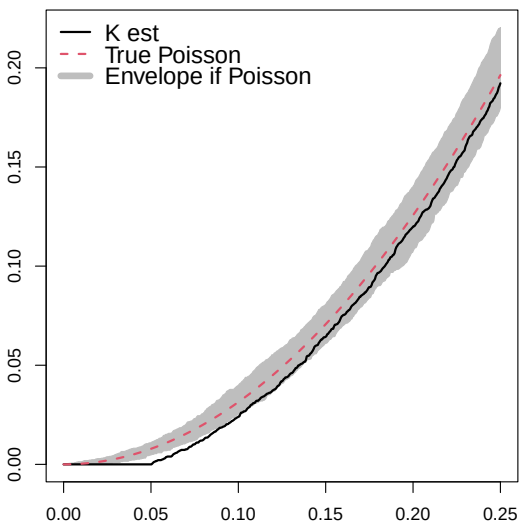






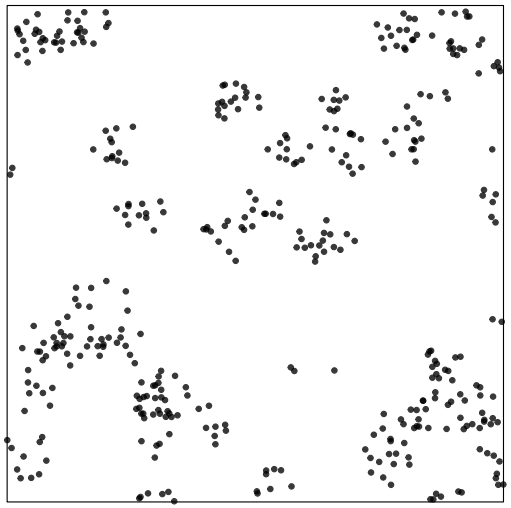


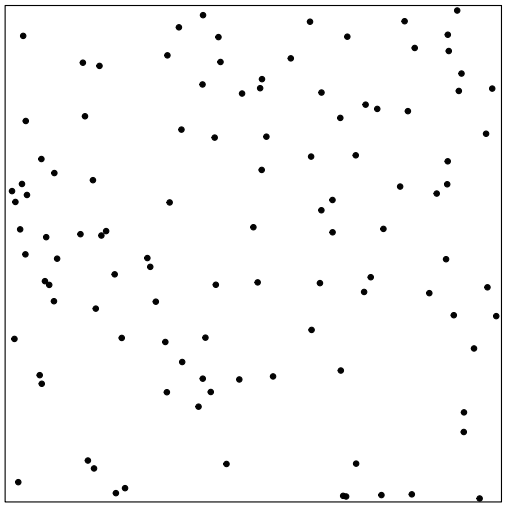


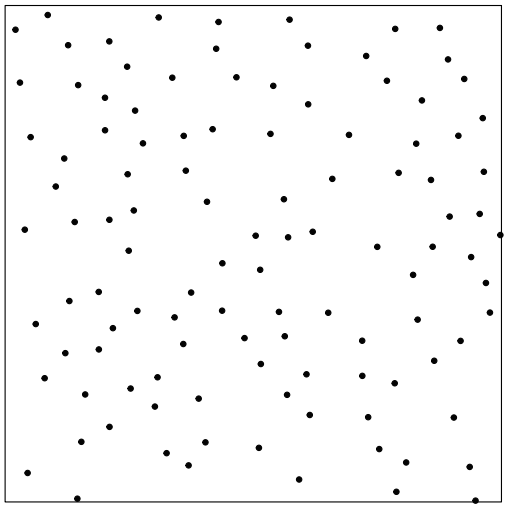


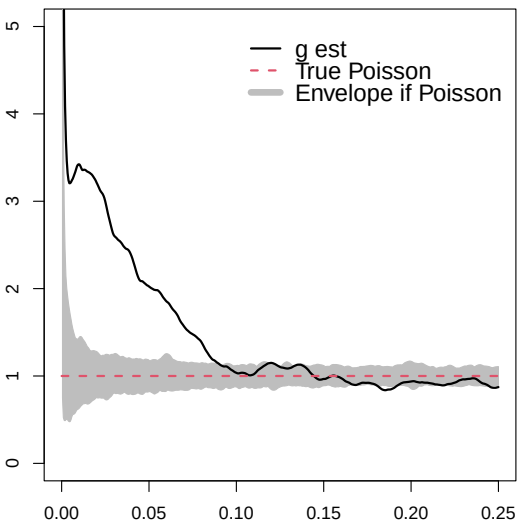
Poisson point process: validation

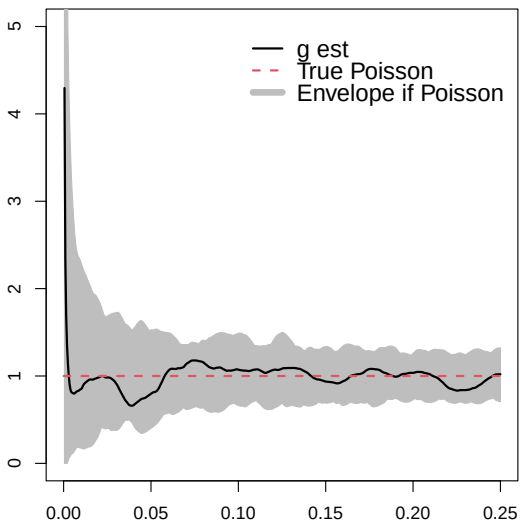
idem with g

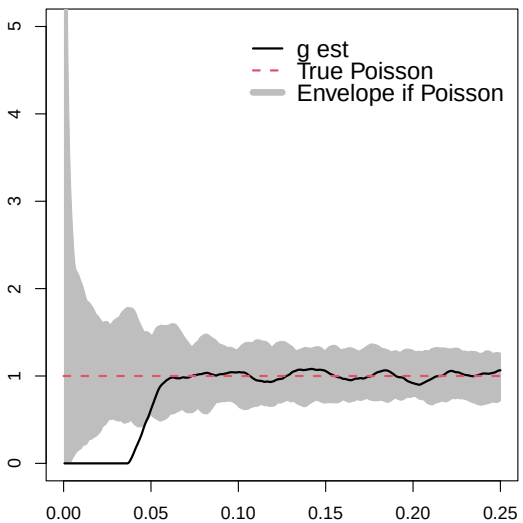












1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Cox processes: definition

Let $\Lambda(u)$, for $u \in S$, be a nonnegative random field on S.

$\mathbf{X}$ is a Cox process if it is a "PPP with random intensity $\Lambda(u)$".

Algorithm of simulation:

1. Generate $\Lambda(u)$.
2. Given $\Lambda(u) = \lambda(u)$, generate a PPP with intensity $\lambda(u)$.

Cox processes: definition

Let $\Lambda(u)$, for $u \in S$, be a nonnegative random field on S.

$\mathbf{X}$ is a Cox process if it is a "PPP with random intensity $\Lambda(u)$".

Algorithm of simulation:

1. Generate $\Lambda(u)$.
2. Given $\Lambda(u) = \lambda(u)$, generate a PPP with intensity $\lambda(u)$.

Intuitively, $\mathbf{X}$ will have many points where $\lambda(u)$ takes high values, provided $\lambda(u)$ is smooth enough.

$$\Rightarrow \text{ we expect a clustering behavior}$$

Cox process: example

Example : LGCP (Log Gaussian Cox process)
Let $G(u)$ be a Gaussian random field on S and take

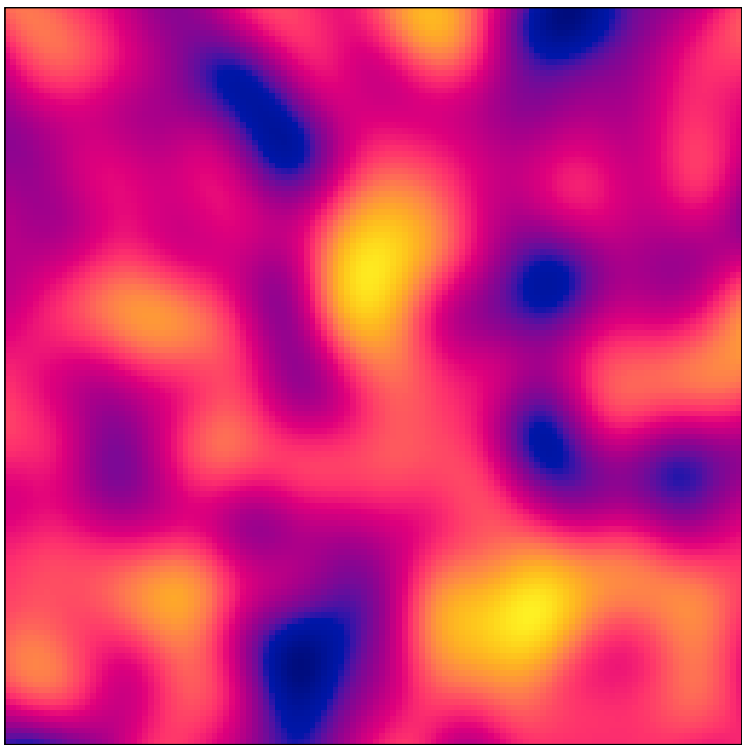
$$\Lambda(u) = \exp(G(u)).$$

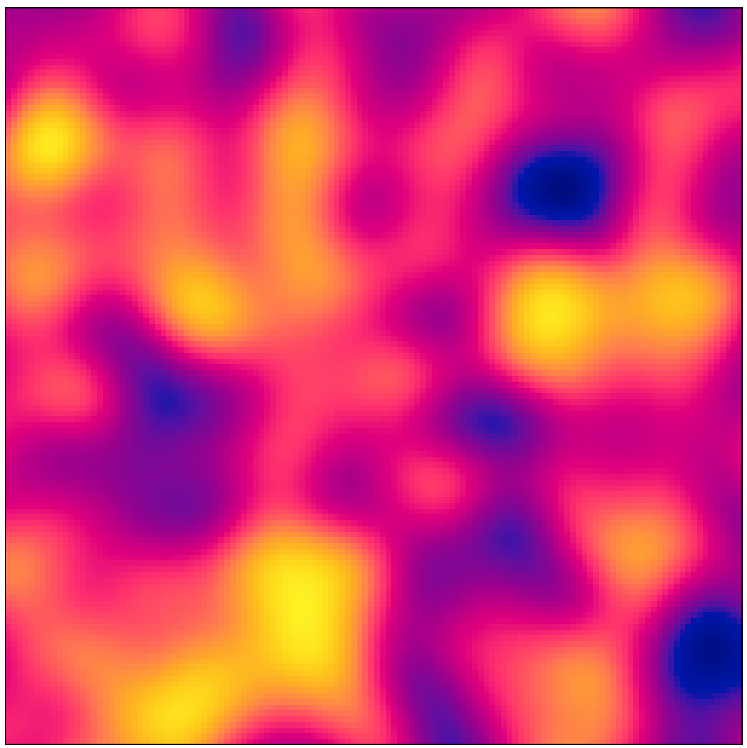
Cox process: example

Example : LGCP (Log Gaussian Cox process)
Let $G(u)$ be a Gaussian random field on S and take

$$\Lambda(u) = \exp(G(u)).$$

1. We simulate the Gaussian field $G(u)$



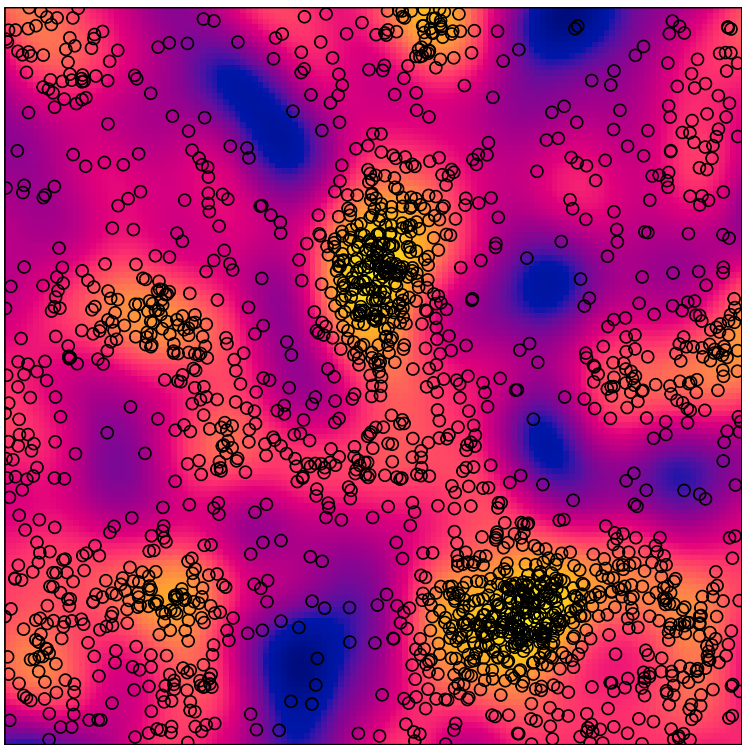


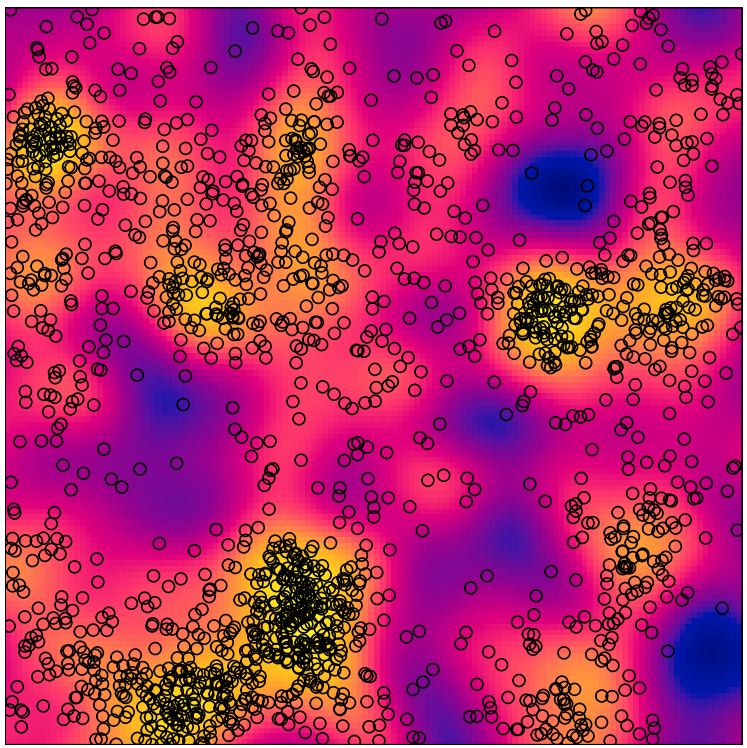
Cox process: example

Example : LGCP (Log Gaussian Cox process)
Let $G(u)$ be a Gaussian random field on S and take

$$\Lambda(u) = \exp(G(u)).$$

1. We simulate the Gaussian field $G(u)$
2. Given $G(u)$, we generate the PPP with intensity $\exp(G(u))$.



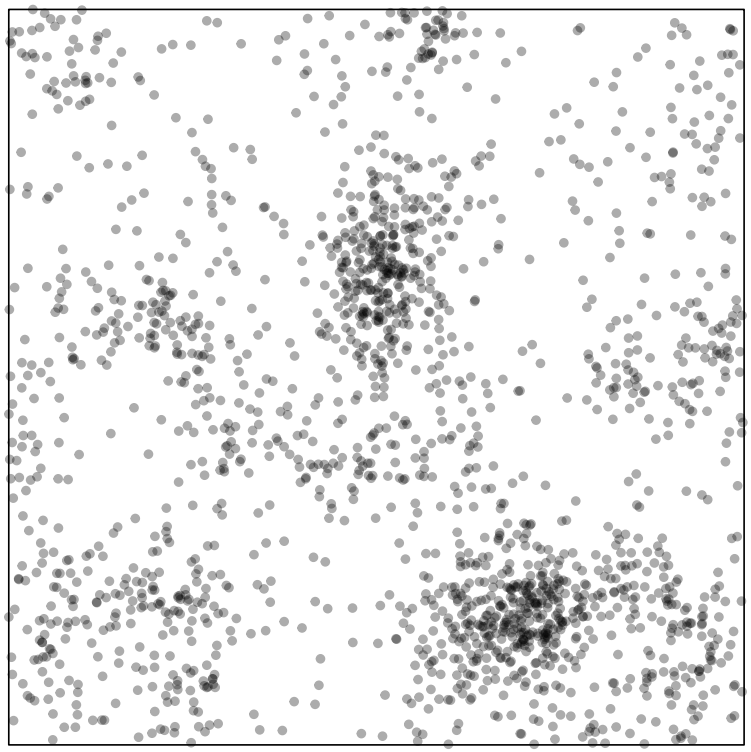


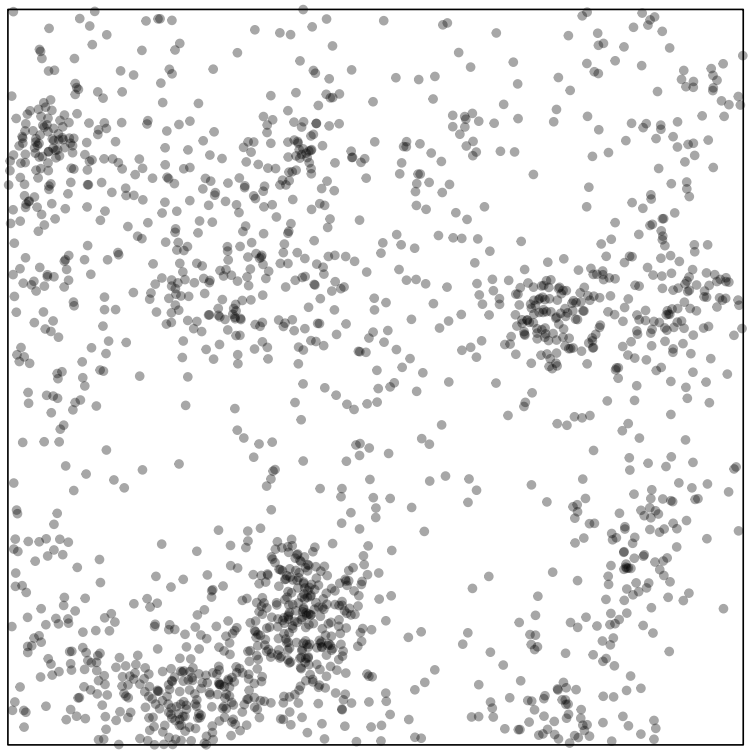
Cox process: example

Example : LGCP (Log Gaussian Cox process)
Let $G(u)$ be a Gaussian random field on S and take

$$\Lambda(u) = \exp(G(u)).$$

1. We simulate the Gaussian field $G(u)$
2. Given $G(u)$, we generate the PPP with intensity $\exp(G(u))$.





Cox process: general properties

Note the difference with an inhomogeneous PPP:

- for a PPP, the clusters are at the same place for every realisations. No interaction between the points.
- for a Cox process, the clusters are located randomly. This creates an **attraction** between the points (see next slide)

Cox process: general properties

- The intensity of a Cox process is $\rho(u) = \mathbb{E}(\Lambda(u))$

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\mathbb{E}(n(\mathbf{X} \cap A)|\Lambda)\right) = \mathbb{E}\left(\int_A \Lambda(u)du\right) = \int_A \mathbb{E}(\Lambda(u))du.$$

Cox process: general properties

- The intensity of a Cox process is $\rho(u) = \mathbb{E}(\Lambda(u))$

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\mathbb{E}(n(\mathbf{X} \cap A)|\Lambda)\right) = \mathbb{E}\left(\int_A \Lambda(u)du\right) = \int_A \mathbb{E}(\Lambda(u))du.$$

- The second order intensity is $\rho^{(2)}(u, v) = \mathbb{E}\left(\Lambda(u)\Lambda(v)\right)$.

Cox process: general properties

- The intensity of a Cox process is $\rho(u) = \mathbb{E}(\Lambda(u))$

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\mathbb{E}(n(\mathbf{X} \cap A)|\Lambda)\right) = \mathbb{E}\left(\int_A \Lambda(u)du\right) = \int_A \mathbb{E}(\Lambda(u))du.$$

- The second order intensity is $\rho^{(2)}(u, v) = \mathbb{E}\left(\Lambda(u)\Lambda(v)\right)$.
- So the pcf is

$$g(u, v) = \frac{\mathbb{E}\left(\Lambda(u)\Lambda(v)\right)}{\mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v))}.$$

If Λ is positively correlated (the typical case), there is attraction:

$$Cov(\Lambda(u), \Lambda(v)) \geq 0 \Leftrightarrow \mathbb{E}\left(\Lambda(u)\Lambda(v)\right) \geq \mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v)) \Leftrightarrow g(u, v) \geq 1.$$

Cox process: general properties

- The intensity of a Cox process is $\rho(u) = \mathbb{E}(\Lambda(u))$

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\mathbb{E}(n(\mathbf{X} \cap A)|\Lambda)\right) = \mathbb{E}\left(\int_A \Lambda(u)du\right) = \int_A \mathbb{E}(\Lambda(u))du.$$

- The second order intensity is $\rho^{(2)}(u, v) = \mathbb{E}\left(\Lambda(u)\Lambda(v)\right)$.
- So the pcf is

$$g(u, v) = \frac{\mathbb{E}\left(\Lambda(u)\Lambda(v)\right)}{\mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v))}.$$

If Λ is positively correlated (the typical case), there is attraction:

$$Cov(\Lambda(u), \Lambda(v)) \geq 0 \Leftrightarrow \mathbb{E}\left(\Lambda(u)\Lambda(v)\right) \geq \mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v)) \Leftrightarrow g(u, v) \geq 1.$$

- The density wrt to $\mathbf{X}_0$ (PPP with intensity 1) is

$$f(\mathbf{x}) = \mathbb{E}\left(e^{|S|-\int_S \Lambda(u)du} \prod_{i=1}^{n(\mathbf{x})} \Lambda(x_i)\right), \quad \forall \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

Cox process: general properties

- The intensity of a Cox process is $\rho(u) = \mathbb{E}(\Lambda(u))$

$$\mathbb{E}(n(\mathbf{X} \cap A)) = \mathbb{E}\left(\mathbb{E}(n(\mathbf{X} \cap A)|\Lambda)\right) = \mathbb{E}\left(\int_A \Lambda(u)du\right) = \int_A \mathbb{E}(\Lambda(u))du.$$

- The second order intensity is $\rho^{(2)}(u, v) = \mathbb{E}\left(\Lambda(u)\Lambda(v)\right)$.
- So the pcf is

$$g(u, v) = \frac{\mathbb{E}\left(\Lambda(u)\Lambda(v)\right)}{\mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v))}.$$

If Λ is positively correlated (the typical case), there is attraction:

$$Cov(\Lambda(u), \Lambda(v)) \geq 0 \Leftrightarrow \mathbb{E}\left(\Lambda(u)\Lambda(v)\right) \geq \mathbb{E}(\Lambda(u))\mathbb{E}(\Lambda(v)) \Leftrightarrow g(u, v) \geq 1.$$

- The density wrt to $\mathbf{X}_0$ (PPP with intensity 1) is

$$f(\mathbf{x}) = \mathbb{E}\left(e^{|S|-\int_S \Lambda(u)du} \prod_{i=1}^{n(\mathbf{x})} \Lambda(x_i)\right), \quad \forall \mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}.$$

None of ρ, $\rho^{(2)}$, g and f is explicit in general for a Cox process!
But some are known for important cases, see next.

Example: Log Gaussian Cox process (LGCP)

Let $G(u)$ be an isotropic Gaussian random field with

- mean m: $\mathbb{E}(G(u)) = m$,
- covariance function $c(r)$: $Cov(G(u), G(u+v)) = c(\|v\|)$

The LGCP driven by $G(u)$ is the Cox process driven by $\Lambda(u) = \exp(G(u))$
We know its moments:

$$\rho = \exp(m + c(0)/2),$$
$$g(r) = \exp(c(r)).$$

But the density does not simplify.

Example: Neymann-Scott process

A Neymann-Scott process is a union of clusters:

$$\mathbf{X} = \bigcup_{c \in C} \mathbf{X}_c$$

- C (centers) is a homogeneous PPP with intensity γ
- Given $c \in C$, $\mathbf{X}_c$ (a cluster) is a PPP with intensity $\alpha k(u - c)$

Interpretation of parameters:

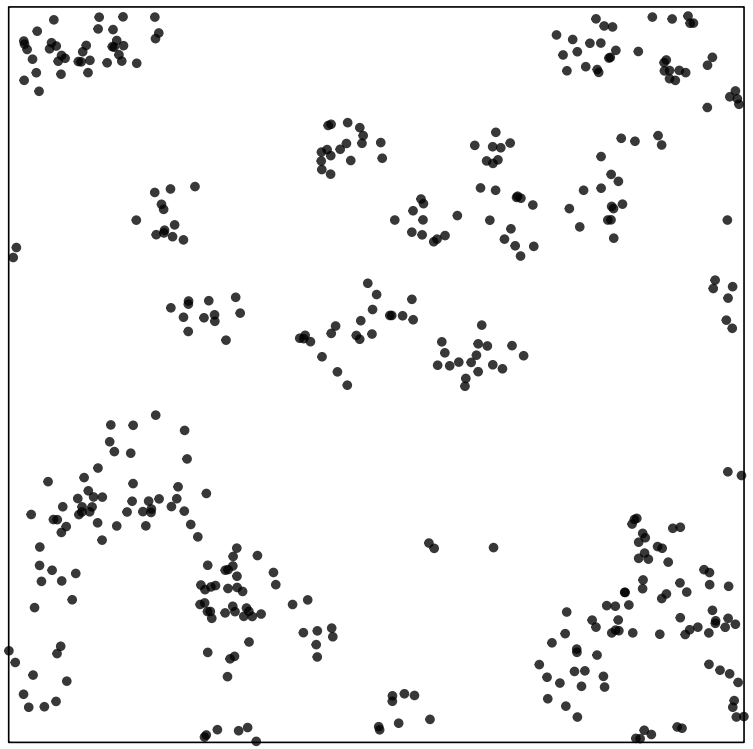
- $\gamma > 0$ is the intensity of clusters,
- $\alpha > 0$ is the mean intensity (or size) of each cluster
- k is a kernel (a density) encoding the spread of each cluster.

This is in fact a Cox point process driven by the random field

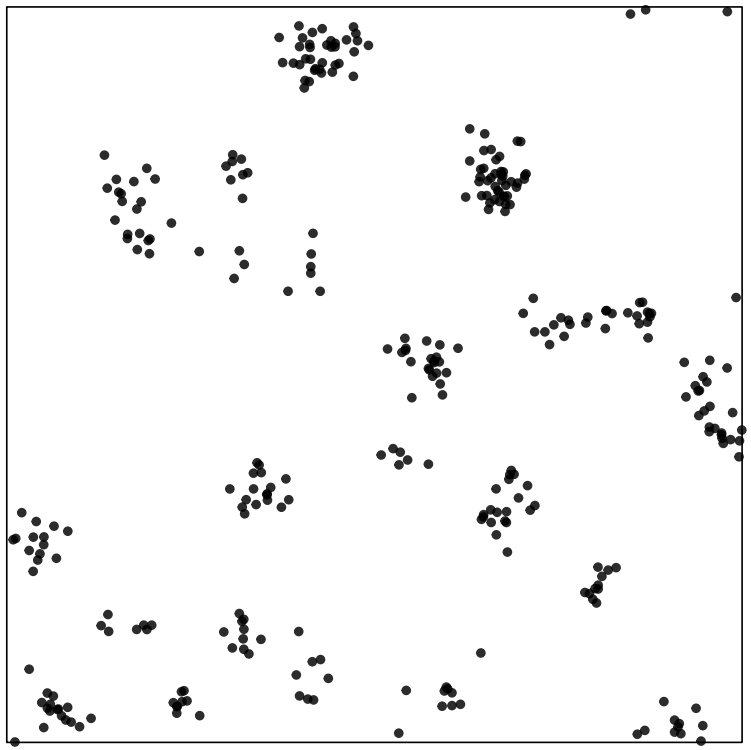
$$\Lambda(u) = \sum_{c \in C} \alpha k(u - c).$$

Example: Neymann-Scott process

Matérn cluster process with radius R: k is the uniform density on $B(0, R)$
Thomas process with variance σ^2: k is the density of a $\mathcal{N}(0, \sigma^2 I_d)$.



Matérn ($R = 0.05$)



Thomas ($\sigma = 0.02$)

Example: Neymann-Scott process

For a Neymann-Scott process with parameters $\gamma > 0$, $\alpha > 0$ and k.

$$\rho = \gamma\alpha,$$
$$g(r) = 1 + \frac{1}{\gamma}\int k(s)k(s+r)ds.$$

(the integral in g is explicit for a Matérn cluster and a Thomas process)

The density does not simplify.

Cox process: inference

Let a parametric Cox process $\mathbf{X}$ with parameter θ.
For instance $\theta = (\gamma, \alpha, \sigma^2)$ for a Thomas process.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

Cox process: inference

Let a parametric Cox process $\mathbf{X}$ with parameter θ.
For instance $\theta = (\gamma, \alpha, \sigma^2)$ for a Thomas process.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

- Minimum contrast estimation based on g or K, that is

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} (\hat{g}(r)^q - g_\theta(r)^q)^2 \, dr$$

$$\text{or} \quad \hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} \left(\hat{K}(r)^q - K_\theta(r)^q\right)^2 dr$$

where typically $q = 1/4$, $r_{\min} = 0$ and $r_{\max}$ is to be chosen.

Cox process: inference

Let a parametric Cox process $\mathbf{X}$ with parameter θ.
For instance $\theta = (\gamma, \alpha, \sigma^2)$ for a Thomas process.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

- Minimum contrast estimation based on g or K, that is

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} (\hat{g}(r)^q - g_\theta(r)^q)^2 \, dr$$

$$\text{or} \quad \hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} \left(\hat{K}(r)^q - K_\theta(r)^q\right)^2 dr$$

  where typically $q = 1/4$, $r_{\min} = 0$ and $r_{\max}$ is to be chosen.

- Maximum likelihood estimation

$$\hat{\theta} = \operatorname*{argmax}_{\theta} \mathbb{E}\left(e^{|S| - \int_S \Lambda_\theta(u) du} \prod_{i=1}^{n(\mathbf{X})} \Lambda_\theta(x_i) \right)$$

  but this requires huge Monte-Carlo simulations to approximate $\mathbb{E}(.)$

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Gibbs point processes

We speak of a **Gibbs model** when we define $\mathbf{X}$ through its density.

Gibbs point processes

We speak of a **Gibbs model** when we define $\mathbf{X}$ through its density.

Examples :

- A PPP with intensity $\rho(u)$ admits the density (wrt to $\mathbf{X}_0$):

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i) = c \, \exp\left(\sum_{u \in \mathbf{x}} \log(\rho(u)) \right)$$

  where c does not depend on $\mathbf{x}$. In the homogeneous case

$$f(\mathbf{x}) = c \, \exp(\log(\rho) \, n(\mathbf{x})).$$

Gibbs point processes

We speak of a **Gibbs model** when we define $\mathbf{X}$ through its density.

Examples :

- A PPP with intensity $\rho(u)$ admits the density (wrt to $\mathbf{X}_0$):

$$f(\mathbf{x}) = e^{|S| - \int_S \rho(u)du} \prod_{i=1}^{n(\mathbf{x})} \rho(x_i) = c \, \exp\left(\sum_{u \in \mathbf{x}} \log(\rho(u)) \right)$$

  where c does not depend on $\mathbf{x}$. In the homogeneous case

$$f(\mathbf{x}) = c \, \exp(\log(\rho) \, n(\mathbf{x})).$$

- We define an homogeneous **pairwise interaction Gibbs model** by

$$f(\mathbf{x}) = c \, \exp\left(\beta n(\mathbf{x}) + \sum_{u,v \in \mathbf{x}}^{\neq} \Phi(\|u - v\|) \right),$$

  where $\beta > 0$ is related (but not equal) to the log-intensity ρ of $\mathbf{x}$, and $\Phi : \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$ is the pairwise interaction function.

Pairwise interaction Gibbs models

Recall : the density wrt to $\mathbf{X}_0$ (i.e., the PPP with $\rho = 1$) is

$$f(\mathbf{x}) = c \exp\left(\beta n(\mathbf{x}) + \sum_{u,v\in\mathbf{x}}^{\neq} \Phi(\|u - v\|)\right).$$

Interpretation:

- If $\Phi = 0$, then we're back to a PPP
- A realisation $\mathbf{x}$ will tend to maximise $\sum_{u,v\in\mathbf{x}}^{\neq} \Phi(\|u - v\|)$

Pairwise interaction Gibbs models

Recall : the density wrt to $\mathbf{X}_0$ (i.e., the PPP with $\rho = 1$) is

$$f(\mathbf{x}) = c \exp\left(\beta n(\mathbf{x}) + \sum_{u,v\in\mathbf{x}}^{\neq} \Phi(\|u-v\|)\right).$$

Interpretation:

- If $\Phi = 0$, then we're back to a PPP
- A realisation $\mathbf{x}$ will tend to maximise $\sum_{u,v\in\mathbf{x}}^{\neq} \Phi(\|u-v\|)$

Existence:
Any Φ is admissible provided the normalizing constant c exists, that is if

$$\mathbb{E}(f(\mathbf{X}_0)) = 1 \Leftrightarrow c^{-1} = \mathbb{E}\left[\exp\left(\beta n(\mathbf{X}_0) + \sum_{u,v\in\mathbf{X_0}}^{\neq} \Phi(\|u-v\|)\right)\right] < \infty.$$

A sufficient assumption is $\Phi \leq 0$.

Example: the Strauss process

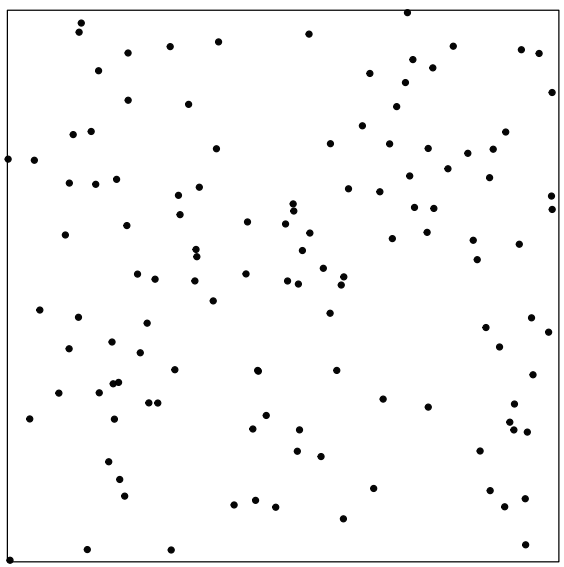
Strauss model with radius $R > 0$ and interaction parameter $\gamma \leq 0$ when

$$\Phi(r) = \gamma \mathbf{1}_{r \leq R}$$

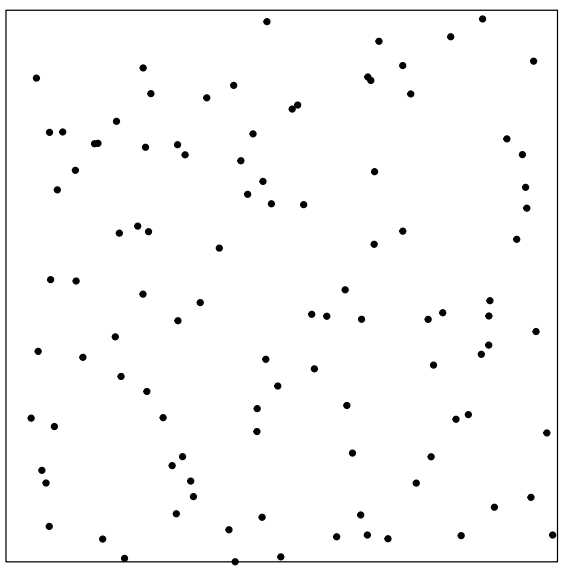
so that

$$f(\mathbf{x}) = c \, \exp\left(\beta n(\mathbf{x}) + \gamma s_R(\mathbf{x})\right),$$

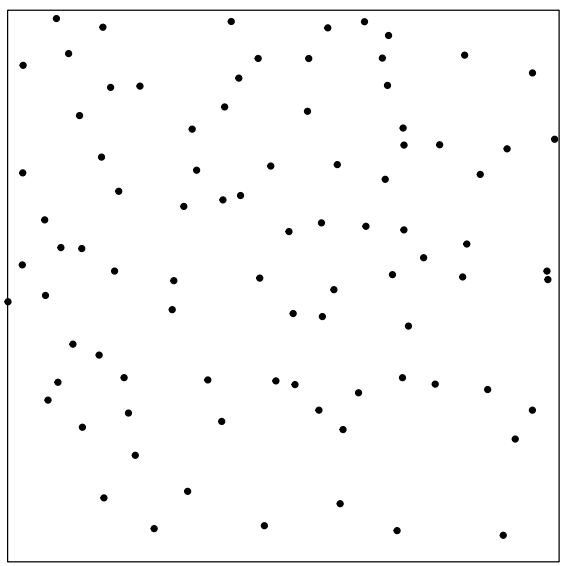
where $s_R(\mathbf{x})$ is the number of R-close pairs of points in $\mathbf{x}$.



$\gamma = \log(0.8)$



$\gamma = \log(0.5)$



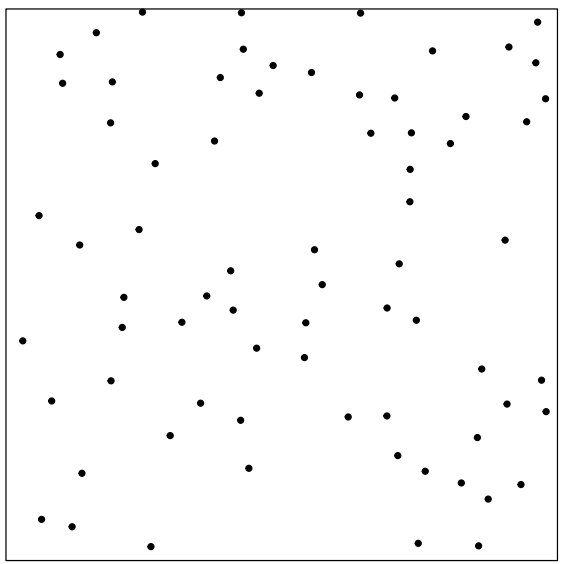
$\gamma = \log(0.2)$

Example: the Hardcore point process

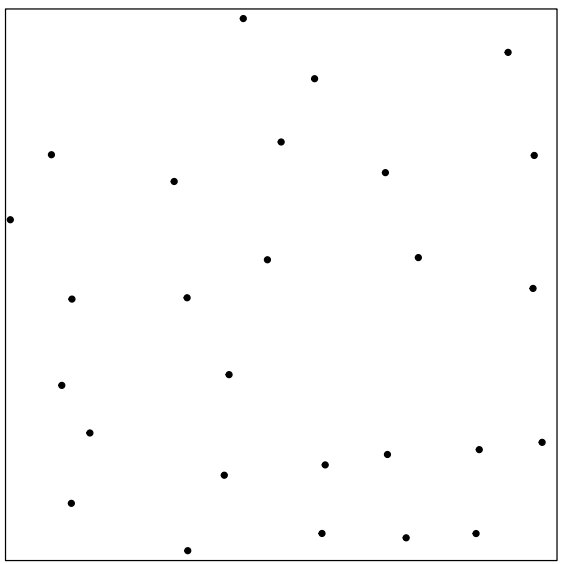
The **Hardcore model** with radius $R > 0$ corresponds to

$$\Phi(r) = \begin{cases} -\infty & \text{if } r \leq R \\ 0 & \text{otherwise} \end{cases}$$

so that no pairs can occur at a distance less than R.



$R = 0.05$



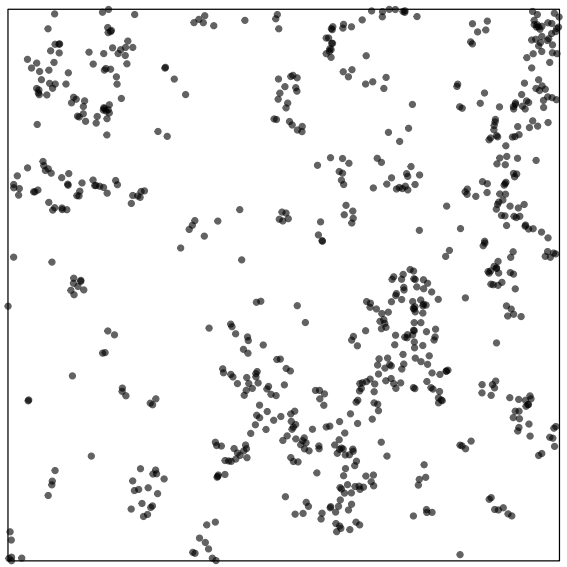
$R = 0.1$

Example: non pairwise interaction

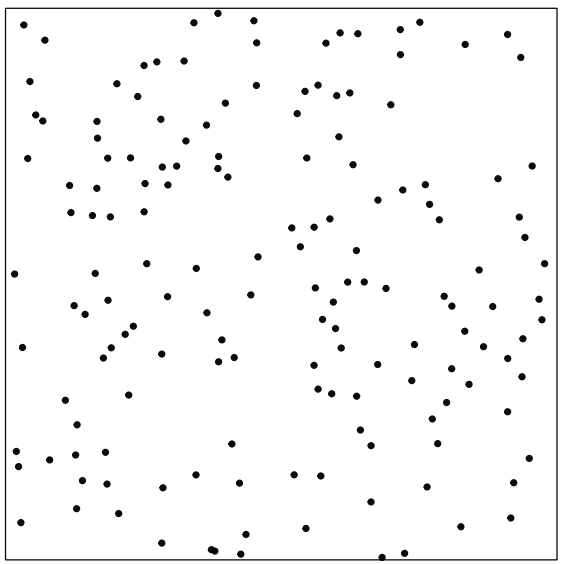
We can consider more complicated density. For instance the **area interaction** process, for $R > 0$ and $\gamma \in \mathbb{R}$,

$$f(\mathbf{x}) = c \exp\left(\beta n(\mathbf{x}) + \gamma \left|\bigcup_{u \in \mathbf{x}} B(u, R)\right|\right),$$

where the interaction depends on the volume of the union of balls.



$\gamma < 0$



$\gamma > 0$

Gibbs point process: general properties

Advantages:

- Flexible models, easy to interpret
- Mainly used to model inhibition between the points
- But some Gibbs models can yield attraction between points

Drawbacks:

- We do not know the moments (neither ρ, nor g, nor K)
- Not easy to simulate (MCMC methods are needed)
- In the density, we do not know explicitly c: this is problematic when it comes to estimate a parameter θ by MLE (since c depends on θ)
- Very hard to deal with in theory

Gibbs point process: Papangelou conditional intensity

The **Papangelou conditional intensity** is defined for any u and $\mathbf{x}$ by

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup u)}{f(\mathbf{x})}.$$

Intuitively, this is the probability to have a point at u given that the point configuration otherwise is $\mathbf{x}$.

The benefit of λ over f is that $\lambda(u, \mathbf{x})$ does not depend on c.

Gibbs point process: Papangelou conditional intensity

The **Papangelou conditional intensity** is defined for any u and $\mathbf{x}$ by

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup u)}{f(\mathbf{x})}.$$

Intuitively, this is the probability to have a point at u given that the point configuration otherwise is $\mathbf{x}$.

The benefit of λ over f is that $\lambda(u, \mathbf{x})$ does not depend on c.

We can take advantage of that for:

- parametric inference, through the **pseudo-likelihood** estimator:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n(\mathbf{x})} \log \lambda_\theta(x_i, \mathbf{x} \setminus x_i) - \int_S \lambda_\theta(u, \mathbf{x}) du.$$

  (not the similarity with the log-likelihood of a PPP)

Gibbs point process: Papangelou conditional intensity

The **Papangelou conditional intensity** is defined for any u and $\mathbf{x}$ by

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup u)}{f(\mathbf{x})}.$$

Intuitively, this is the probability to have a point at u given that the point configuration otherwise is $\mathbf{x}$.

The benefit of λ over f is that $\lambda(u, \mathbf{x})$ does not depend on c.

We can take advantage of that for:

- parametric inference, through the **pseudo-likelihood** estimator:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n(\mathbf{x})} \log \lambda_\theta(x_i, \mathbf{x} \setminus x_i) - \int_S \lambda_\theta(u, \mathbf{x}) du.$$

  (not the similarity with the log-likelihood of a PPP)

- simulating $\mathbf{X}$ by a birth-death Metropolis-Hastings algorithm, where
  - a birth proposal at u depends (positively) on $\lambda(u, \mathbf{x})$
  - a death proposal of $x_i \in \mathbf{x}$ depends (negatively) on $\lambda(x_i, \mathbf{x} \setminus x_i)$.

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

Determinantal point processes (DPPs)

The n-th order intensity $\rho^{(n)}$ of a point process satisfies

$$\mathbb{E}\left(\sum_{u_1,\ldots,u_n\in\mathbf{X}}^{\neq} \mathbf{1}_{u_1\in A_1,\ldots,u_n\in A_n}\right) = \int_{A_1}\cdots\int_{A_n} \rho^{(n)}(u_1,\ldots,u_n)du_1\cdots du_n,$$

$\forall\, A_1,\ldots,A_n \subseteq S$. Intuitively, $\rho^{(n)}(u_1,\ldots,u_n) \approx P(\{u_1,\ldots,u_n\}\subset \mathbf{X})$.

A DPP is defined through the intensities $\rho^{(n)}$.

Determinantal point processes (DPPs)

The n-th order intensity $\rho^{(n)}$ of a point process satisfies

$$\mathbb{E}\left(\sum_{u_1,\ldots,u_n\in\mathbf{X}}^{\neq} \mathbf{1}_{u_1\in A_1,\ldots,u_n\in A_n}\right) = \int_{A_1}\cdots\int_{A_n} \rho^{(n)}(u_1,\ldots,u_n)du_1\cdots du_n,$$

$\forall\, A_1,\ldots,A_n \subseteq S$. Intuitively, $\rho^{(n)}(u_1,\ldots,u_n) \approx P(\{u_1,\ldots,u_n\}\subset \mathbf{X})$.

A DPP is defined through the intensities $\rho^{(n)}$.

Let $C(u,v)$ be a covariance function. $\mathbf{X}$ is a **DPP with kernel** C if

$$\forall n \geq 1, \quad \rho^{(n)}(u_1,\ldots,u_n) = \det\left[C(u_i,u_j)\right]_{1\leq i,j\leq n},$$

where $[C(u_i,u_j)]_{1\leq i,j\leq n}$ denotes the $n\times n$ matrix with entries $C(u_i,u_j)$.

Stationary DPPs

Recall: $\mathbf{X}$ is a DPP with kernel C if

$$\forall n \geq 1, \quad \rho^{(n)}(u_1, \ldots, u_n) = \det\left[C(u_i, u_j)\right]_{1 \leq i,j \leq n}.$$

If $C(u, v) = C(\|u - v\|)$, then $\mathbf{X}$ is stationary, isotropic and

- its intensity is $\rho^{(1)}(u) = \rho = C(0)$
- its second order intensity $\rho^{(2)}$ is

$$\rho^{(2)}(u, v) = \det \begin{pmatrix} C(0) & C(\|u - v\|) \\ C(\|u - v\|) & C(0) \end{pmatrix} = C(0)^2 - C(\|u - v\|)^2.$$

- Hence its pcf is, for $r = \|u - v\|$, is

$$g(r) = \frac{\rho^{(2)}(u, v)}{\rho(u)\rho(v)} = 1 - \frac{C(r)^2}{C(0)^2}.$$

Since $g \leq 1$, **DPPs are models for inhibitive point processes**.

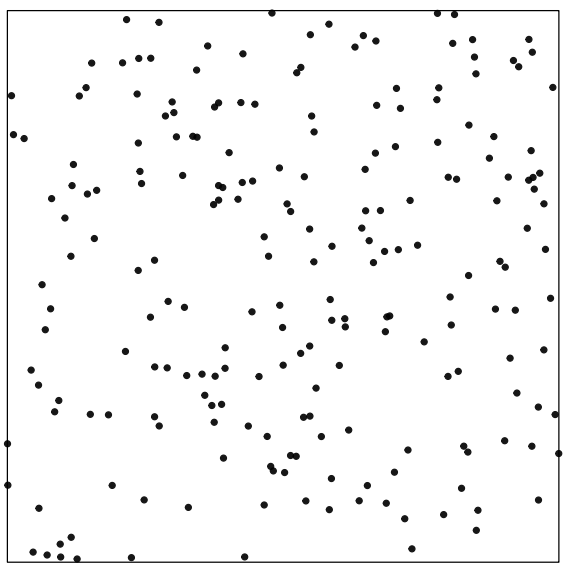
Stationary DPPs: example

Existence: $C(r)$ must be a covariance function such that $\mathcal{F}(C) \leq 1$.

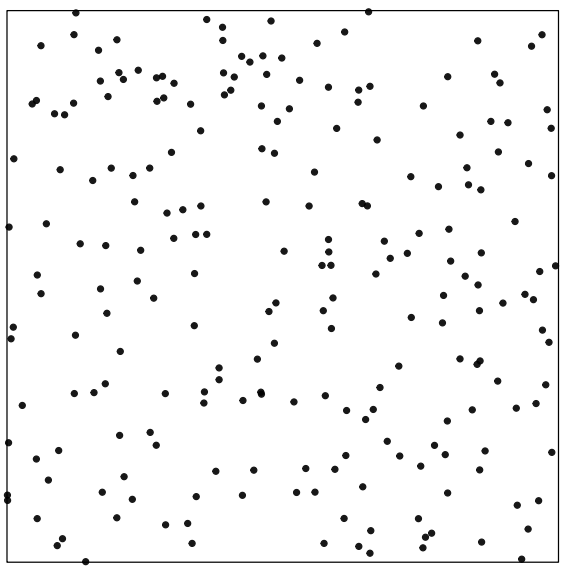
Stationary DPPs: example

Existence: $C(r)$ must be a covariance function such that $\mathcal{F}(C) \leq 1$.

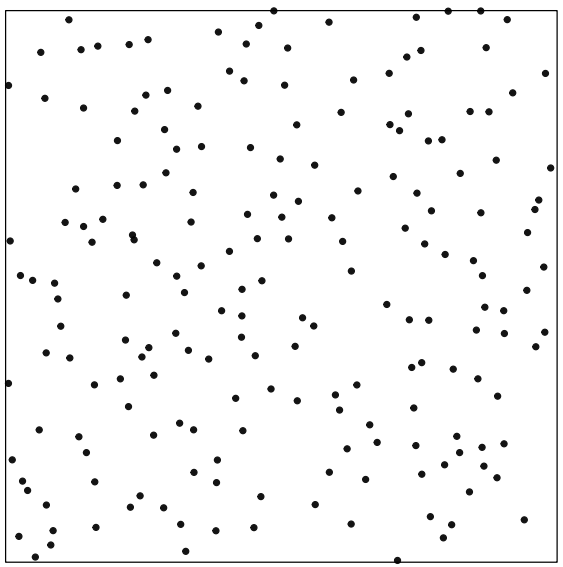
Example: The Gaussian-type DPP
The Gaussian covariance function $C(r) = \rho e^{-r^2/\alpha^2}$ with $\pi\rho\alpha^2 \leq 1$ defines the DPP with intensity ρ and pcf $g(r) = 1 - e^{-2r^2/\alpha^2}$.



small α



medium α



big α

Density of a DPP

In theory, the kernel C admits the eigen-decomposition on S

$$C(u, v) = \sum_{k \geq 1} \lambda_k \Phi_k(u) \Phi_k(v), \quad u, v \in S, \tag{1}$$

where $0 \leq \lambda_k \leq 1$ and $(\Phi_k)_k$ is an orthonormal basis of $L^2(S)$.

Density of a DPP

In theory, the kernel C admits the eigen-decomposition on S

$$C(u,v) = \sum_{k \geq 1} \lambda_k \Phi_k(u) \Phi_k(v), \quad u, v \in S, \tag{1}$$

where $0 \leq \lambda_k \leq 1$ and $(\Phi_k)_k$ is an orthonormal basis of $L^2(S)$.

From this decomposition we can:

- implement a perfect simulation algorithm ;

Density of a DPP

In theory, the kernel C admits the eigen-decomposition on S

$$C(u,v) = \sum_{k\geq 1} \lambda_k \Phi_k(u)\Phi_k(v), \quad u,v \in S, \tag{1}$$

where $0 \leq \lambda_k \leq 1$ and $(\Phi_k)_k$ is an orthonormal basis of $L^2(S)$.

From this decomposition we can:

- implement a perfect simulation algorithm ;
- deduce the density of the DPP when $\lambda_k < 1$, that is

$$f(\mathbf{x}) = c \det\left[L(x_i,x_j)\right]_{1\leq i,j\leq n},$$

where $c = e^{|S|}\prod_{k\geq 1}(1-\lambda_k)$ and L is another kernel defined by

$$L(u,v) = \sum_{k\geq 1} \frac{\lambda_k}{1-\lambda_k}\Phi_k(u)\Phi_k(v), \quad u,v \in S.$$

Density of a DPP

In theory, the kernel C admits the eigen-decomposition on S

$$C(u,v) = \sum_{k\geq 1} \lambda_k \Phi_k(u)\Phi_k(v), \quad u,v \in S, \tag{1}$$

where $0 \leq \lambda_k \leq 1$ and $(\Phi_k)_k$ is an orthonormal basis of $L^2(S)$.

From this decomposition we can:

- implement a perfect simulation algorithm ;
- deduce the density of the DPP when $\lambda_k < 1$, that is

$$f(\mathbf{x}) = c \det\left[L(x_i,x_j)\right]_{1\leq i,j\leq n},$$

where $c = e^{|S|}\prod_{k\geq 1}(1-\lambda_k)$ and L is another kernel defined by

$$L(u,v) = \sum_{k\geq 1} \frac{\lambda_k}{1-\lambda_k}\Phi_k(u)\Phi_k(v), \quad u,v \in S.$$

But (1) is rarely known, unless C is defined through it.
Approximations exist in the stationary case $C(u,v) = C(\|u-v\|)$.

DPP: inference

Let a parametric DPP $\mathbf{X}$ with parameter θ.
For instance $\theta = (\rho, \alpha)$ for a Gaussian-type DPP.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

DPP: inference

Let a parametric DPP $\mathbf{X}$ with parameter θ.
For instance $\theta = (\rho, \alpha)$ for a Gaussian-type DPP.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

- Minimum contrast estimation based on g or K, that is

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} (\hat{g}(r)^q - g_\theta(r)^q)^2 \, dr$$

$$\text{or} \quad \hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} \left(\hat{K}(r)^q - K_\theta(r)^q\right)^2 dr$$

where typically $q = 1/2$, $r_{\min} = 0$ and $r_{\max}$ is to be chosen.

DPP: inference

Let a parametric DPP $\mathbf{X}$ with parameter θ.
For instance $\theta = (\rho, \alpha)$ for a Gaussian-type DPP.

We may estimate θ from a realisation $\mathbf{x}$ of $\mathbf{X}$ by:

- Minimum contrast estimation based on g or K, that is

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} (\hat{g}(r)^q - g_\theta(r)^q)^2 \, dr$$

$$\text{or} \quad \hat{\theta} = \operatorname*{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} \left(\hat{K}(r)^q - K_\theta(r)^q\right)^2 dr$$

  where typically $q = 1/2$, $r_{\min} = 0$ and $r_{\max}$ is to be chosen.

- MLE, if we know the eigen-decomposition of C:

$$\hat{\theta} = \operatorname*{argmax}_{\theta} f_\theta(x).$$

1. Examples of points patterns

2. Summary/descriptive statistics
   - First order moment
   - Second order moments
   - Other summary statistics

3. Point process models
   - Poisson point process (PPP)
   - Cox point processes
   - Gibbs point processes
   - Determinantal point processes
   - Short summary

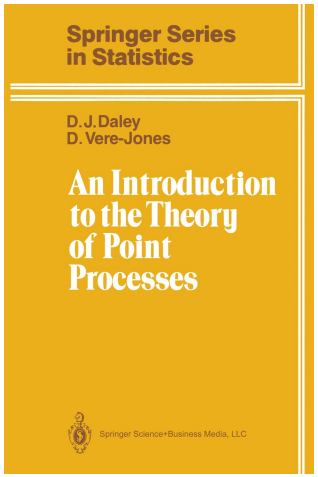
Summary

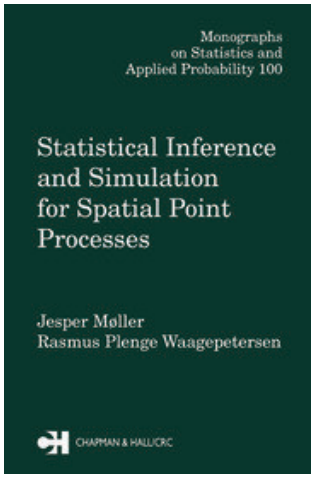
Models	Main features	ρ, g	density
Poisson	Only for CSR.	✓	✓
Cox (LGCP, Matérn, Thomas)	Attraction, clusters.	✓	✗
Gibbs	Very flexible. Mainly inhibition but attraction possible.	✗	✓[1]
DPP	Inhibition. Less flexible than Gibbs.	✓	✓[2]

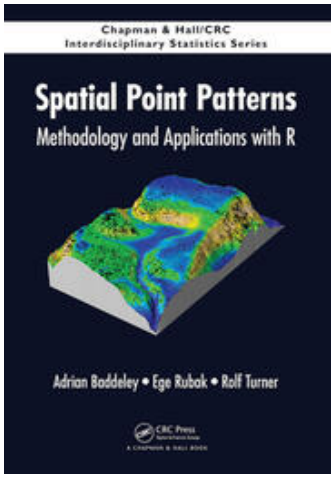
[1] up to the constant; Papangelou conditional density is rather used for inference.

[2] if we know the eigen-decomposition of the kernel.

References (bis)







- D.J Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes.* Springer Series in Statistics, 1998.
- J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton, 2004.
- A. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns : Methodology and Applications with R.* CRC Press, 2015.

All illustrations have been implemented with the **spatstat library** in **R**. Almost all presented datasets also come from this library.