# Semis de Points

Université Paris-Est Créteil
Serge Lhomme

Maître de conférences en géographie
http://sergelhomme.fr/
serge.lhomme@u-pec.fr

# Analyse de semis de points

## Présentation

Cette partie traite de la répartition d'ensemble de lieux qui correspondent aux différentes localisations d'un phénomène.

Ces lieux peuvent être des habitations, des commerces, des personnes, des clients...

Ces lieux peuvent être traités comme des points à un certain degré de généralisation. On parlera donc de semis de points.

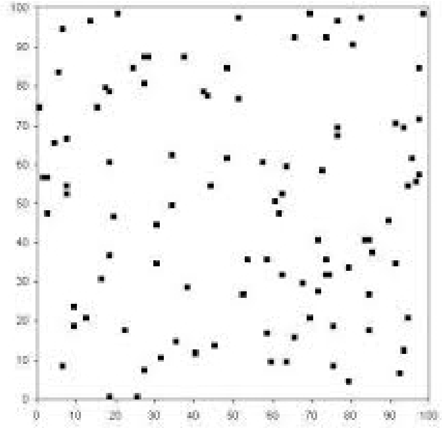
Pour comparer des semis de points ou pour mettre en exergue certaines de leurs spécificités, on va être amené à étudier leur forme.

Dans ces analyses, l'espace est souvent considéré comme homogène.

L'avantage de travailler avec des méthodes adaptées à des données ponctuelles (individuelles), c'est qu'elles ne sont en théorie pas sensibles aux problèmes d'échelle.

# Analyse de semis de points

## Présentation



# Analyse de semis de points

## Identifier le centre d'un semis de points : la position moyenne

Le point moyen :

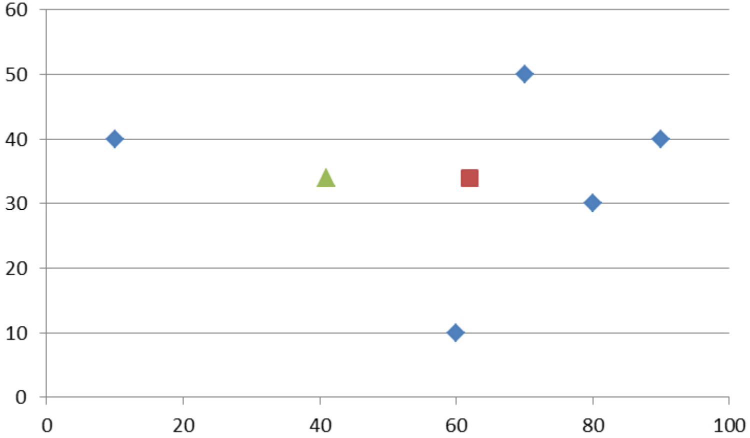
$$\bar{X} = \frac{1}{N} \times \sum_{i=1}^{N} X_i \text{ et } \bar{Y} = \frac{1}{N} \times \sum_{i=1}^{N} Y_i$$

Le point moyen pondéré :

$$\bar{X_p} = \frac{\sum_{i=1}^{N} (P_i \times X_i)}{\sum_{i=1}^{N} P_i} \text{ et } \bar{Y_p} = \frac{\sum_{i=1}^{N} (P_i \times Y_i)}{\sum_{i=1}^{N} P_i}$$

# Analyse de semis de points

## Identifier le centre d'un semis de points : la position moyenne



Ayant déterminé le point moyen, on peut chercher à mesurer la dispersion des lieux autour de ce point central. On parle de distance-type :

$$\sigma_D = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2 + (Y_i - \bar{Y})^2} = \sqrt{\sigma_X^2 + \sigma_Y^2}$$

# Analyse de semis de points

## Mesurer la concentration d'un semis de points

Une distribution est aléatoire si :

1. Tous les emplacements de l'espace ont la même probabilité d'accueillir un point.
2. La position d'un nouveau point est indépendante de la position des points précédents.

Une distribution aura tendance à être concentrée si :

1. Certains emplacements de l'espace ont plus de chances d'accueillir un point.
2. La localisation d'un premier point favorise l'apparition d'autres points à proximité.

Une distribution aura tendance à être régulière si :

1. Tous les emplacements de l'espace ont la même probabilité d'accueillir un point
2. La localisation d'un premier point défavorise l'apparition d'autres points à proximité.

# Analyse de semis de points

## Mesurer la concentration d'un semis de points

La méthode des quadrats permet de mesurer des concentrations (des densités) dans un semis de points :

1. Soit un semis de N points distribués sur un espace E.
2. On recouvre l'espace E d'un ensemble de K mailles d'une forme régulière (carré, rectangle, cercle).
3. Le nombre moyen de points théorique par maille est égale à $D=N/K$.
4. On associe à chaque maille i le nombre de points qu'elle contient, puis on calcule la variance du nombre de points par maille V(D) et on en déduit un indice de concentration(Ic). $IC=V(D)/D$.

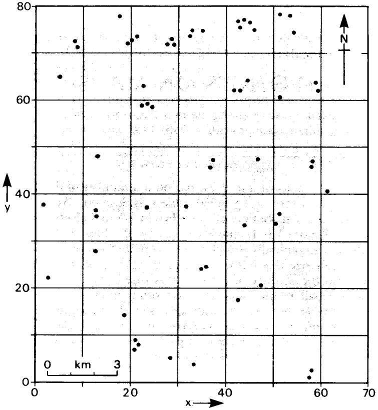
Si $IC = 1$, la distribution est aléatoire.

$IC > 1$, la distribution est plutôt concentrée.

$IC < 1$, la distribution est plutôt régulière.

# Analyse de semis de points

## Mesurer la concentration d'un semis de points



# Analyse de semis de points

## Mesurer la concentration d'un semis de points

| Nombre de points **n** | Nombre de quadrats **K** | Nombre de points **n.K** |
|---|---|---|
| 0 | 25 | 0 |
| 1 | 15 | 15 |
| 2 | 8 | 16 |
| 3 | 5 | 15 |
| 4 | 1 | 4 |
| 5 | 2 | 10 |
| Total | 56 | 60 |

# Analyse de semis de points

## Mesurer la concentration d'un semis de points

| Nombre de points<br>**n** | Nombre de quadrats<br>**K** | Nombre de points<br>**n.K** | Ecarts à la moyenne<br>**(n-D)** | <br>**K(n - D)²** |
|---|---|---|---|---|
| 0 | 25 | 0 | -1.071 | 28.676 |
| 1 | 15 | 15 | -0.071 | 0.076 |
| 2 | 8 | 16 | 0.929 | 6.904 |
| 3 | 5 | 15 | 1.929 | 18.605 |
| 4 | 1 | 4 | 2.929 | 8.579 |
| 5 | 2 | 10 | 3.929 | 30.874 |
| Total | 56 | 60 | | 93.714 |

Densité moyenne **D** = nb. de points / nb. de quadrats = 60/56 **= 1.071**
Variance **V(D)** = 93.714 / 55 = **1.704**
Indice de concentration IC = V(D)/D= **1.590**

# Analyse de semis de points

## Les problèmes d'échelle et de zonage

| | | |
|---|---|---|
| 500 | 100 | 700 |
| 200 | 400 | 100 |
| 300 | 600 | 200 |

population

| | | |
|---|---|---|
| 25 | 40 | 100 |
| 100 | 80 | 6 |
| 30 | 150 | 85 |

population pauvre

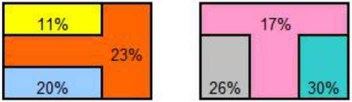
population et population pauvre dans 9 unités géographiques

| | | |
|---|---|---|
| 5% | 40% | 14% |
| 50% | 20% | 6% |
| 10% | 25% | 43% |

| |
|---|
| 13% |
| 27% |
| 24% |

taux de pauvreté avec 2 échelles différentes

**MAUP : Effet d'échelle**




taux de pauvreté avec 2 zonages différents

**MAUP : Effet de zonage**

# Analyse de semis de points

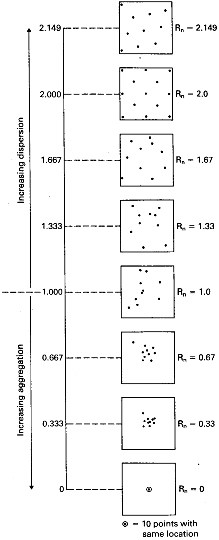
## Mesurer la forme d'un semis de points

La méthode du plus proche voisin permet aussi d'étudier la dispersion, donc la forme d'un semis de points.

1. Soit un semis de N points distribués sur un espace de surface S. On note D la densité moyenne de points par unité de surface (D=N/S).
2. On calcule pour chaque point i la distance Dmin(i) qui le sépare de son voisin le plus proche.
3. On calcule ensuite la moyenne des distances observées au plus proche voisin D0.
4. On détermine la distance théorique moyenne au plus proche voisin DT dans le cas d'une distribution aléatoire (DT=$0.5/\sqrt{D}$).
5. On calcule l'indice de dispersion qui est le rapport : R=D0/DT.

# Analyse de semis de points

## Mesurer la forme d'un semis de points



# Analyse de semis de points

## Mesurer la forme d'un semis de points

| i | Xi | Yi |
|---|---|---|
| 1 | 1,5 | 7 |
| 2 | 1 | 7 |
| 3 | 1,5 | 6,8 |
| 4 | 0,5 | 5,8 |
| 5 | 2,2 | 7,5 |
| 6 | 0,3 | 7 |
| 7 | 0,6 | 4,8 |
| 8 | 1,8 | 4,1 |
| 9 | 2,1 | 5,2 |
| 10 | 4,3 | 5,8 |
| 11 | 1,6 | 7,2 |
| 12 | 3,1 | 6,4 |
| 13 | 0,7 | 2,9 |
| 14 | 0,1 | 2,6 |
| 15 | 1,5 | 4,4 |
| 16 | 3,1 | 5,3 |
| 17 | 5,2 | 6,2 |
| 18 | 5,1 | 7,9 |
| 19 | 1,7 | 1 |
| 20 | 2,4 | 1,8 |
| 21 | 4,2 | 5 |
| 22 | 7 | 6,1 |
| 23 | 6,8 | 3,8 |
| 24 | 7,2 | 0,3 |

# Analyse de semis de points

## Mesurer la forme d'un semis de points

| i | dmin |
|---|---|
| 1 | 0,2 |
| 2 | 0,5 |
| 3 | 0,2 |
| 4 | 1,0 |
| 5 | 0,7 |
| 6 | 0,7 |
| 7 | 1,0 |
| 8 | 0,4 |
| 9 | 1,0 |
| 10 | 0,8 |
| 11 | 0,2 |
| 12 | 1,1 |
| 13 | 0,7 |
| 14 | 0,7 |
| 15 | 0,4 |
| 16 | 1,0 |
| 17 | 1,0 |
| 18 | 1,7 |
| 19 | 1,1 |
| 20 | 1,1 |
| 21 | 0,8 |
| 22 | 1,8 |
| 23 | 2,3 |
| 24 | 3,5 |

$D0 = 0.99$

Comme la surface est égale 64 et l'effectif est égal à 24, on obtient une densité de 0,375 et par conséquent $DT = 0.816$

$R = D0/DT = 1{,}22$

# Analyse de semis de points

## Analyser la configuration d'un semis de points : la fonction K de Ripley

Les valeurs synthétiques précédentes peuvent masquer des phénomènes plus complexes.

C'est pourquoi on préfère généralement utiliser la fonction K de Ripley qui permet des analyses plus qualitatives, sans biais d'échelle et relativement exhaustives.

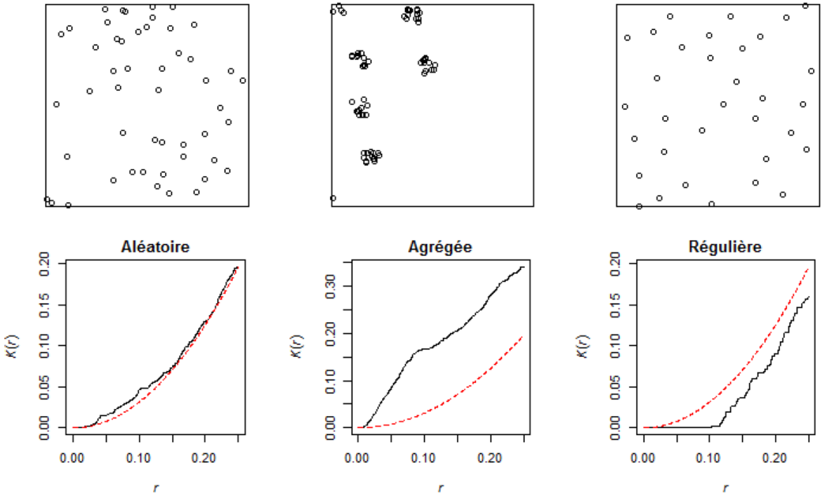
La fonction K(r) de Ripley est une fonction cumulative qui calcule simplement le nombre moyen de voisins de chaque point situés à une distance inférieure à r.

Ce nombre moyen est standardisé par l'intensité du processus (la densité n / W où W est l'aire étudiée).

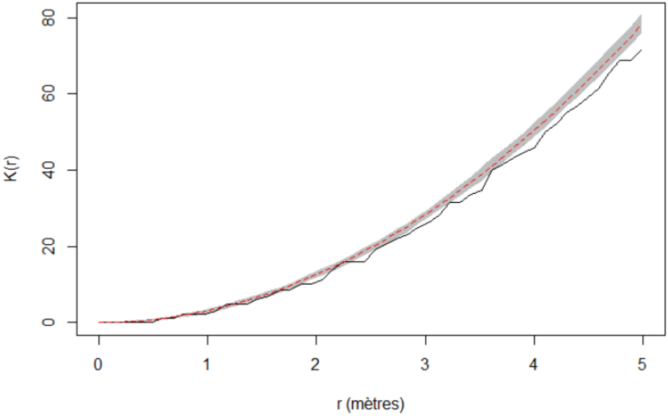
Il convient ensuite de comparer cette fonction standardisée avec celle d'une distribution aléatoire (un processus de Poisson) homogène. Cette fonction est simplement égale à $\pi r^2$.

# Analyse de semis de points

## Analyser la configuration d'un semis de points : la fonction K de Ripley



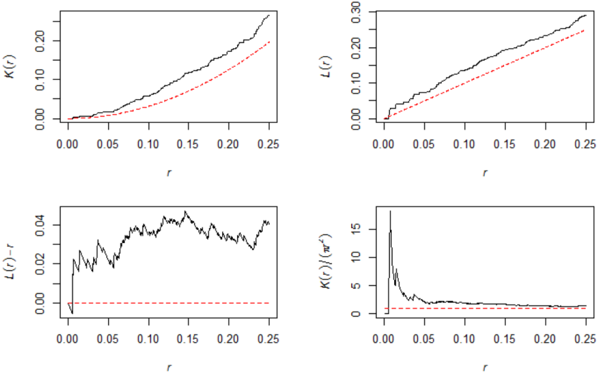
La technique la plus courante pour tester la significativité des valeurs obtenues est le recours à la simulation d'un intervalle de confiance par la méthode de Monte Carlo.



# Analyse de semis de points

## Analyser la configuration d'un semis de points : la fonction K de Ripley

Il existe de nombreuses variantes à la fonction de Ripley comme par exemple la fonction L de Besag (qui permet de comparer les différentes valeurs de la fonction en rapportant notamment K par $\pi$) et la fonction D de Diggle notamment (qui compare la fonction K à des processus non-homogènes).



# Analyse de semis de points

## M de Marcon et Puech et fonction intertype

L'indicateur M de Marcon et Puech est un indicateur cumulatif relatif qui va comparer la proportion de points d'intérêt dans un voisinage à celle que l'on observe sur l'ensemble du territoire analysé.

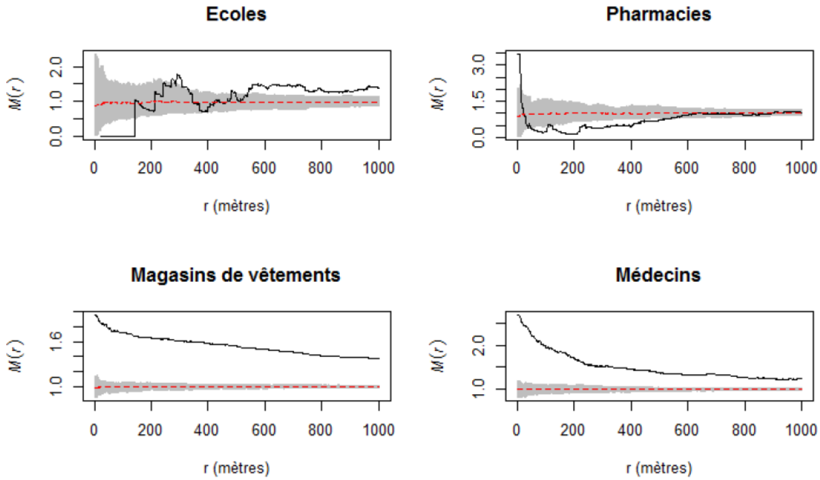
L'intérêt d'une approche relative, c'est qu'elle peut facilement être appliqué à deux types de configuration de points en comparant une proportion locale à une proportion globale mais où le type de points voisins d'intérêt n'est pas le même type que celui des points centre.

Par exemple, si nous suspectons une attraction des points de type T par ceux de type S, nous allons comparer la proportion locale de voisins du type T autour de points du type S à la proportion globale observée sur tout le territoire considéré.

Si l'attraction entre les points de type T autour de type S est réelle, la proportion de points de type T autour de ceux du type S devrait être localement plus importante que celle observée sur toute l'aire d'étude.

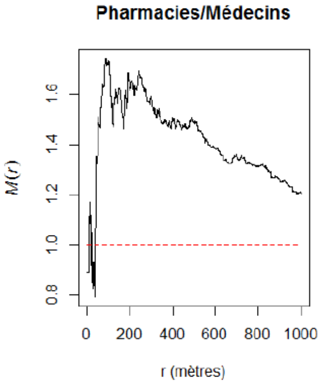
# Analyse de semis de points

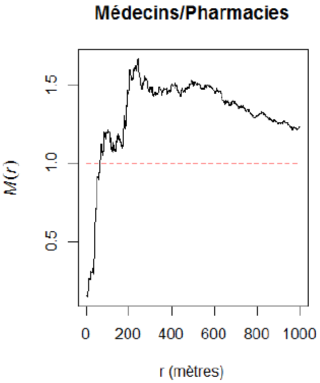
## Analyser la configuration d'un semis de points : la fonction K de Ripley



# Analyse de semis de points

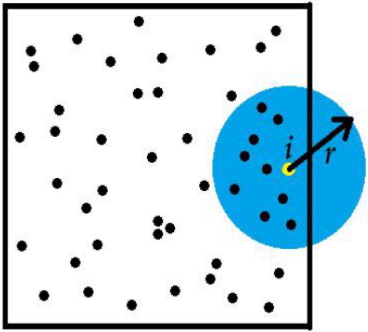
## Analyser la configuration d'un semis de points : la fonction K de Ripley



# Analyse de semis de points

## Attention aux effets de bord




Généralement, quel que soit le domaine d'application, ce biais potentiel est jugé suffisamment sévère pour que l'on recoure à une technique correctrice prenant en compte les "effets de bord". On a pour cela souvent recours à du lissage spatial...