

Analyse exploratoire des données spatialisées en utilisant des processus ponctuels

Clément Dugué



Stage de 2ème année encadré par Radu Stoica

8 septembre 2017

1 Introduction

- Présentation de l'entreprise
- Contexte du stage

2 Cahier des charges du projet

- Fonctions à réaliser
- Insertion dans un code existant
- Travail supplémentaire : l'intensité

3 Méthode de travail

- Outils utilisés
- Manière de procéder
- Échanges avec le tuteur

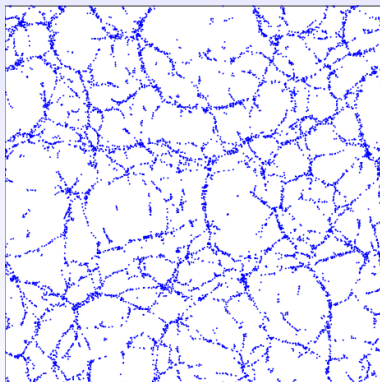
4 Solutions apportées et résultats

- Programmation des fonctions
- Automatisation
- Résultats

5 Bilan du stage



- Recherche en Lorraine
- Institut Elie Cartan
- Equipe Probabilités et Statistiques

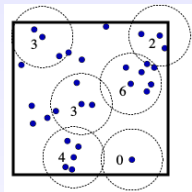


Exemple de configuration de point - Jeu de données fourni par M. Stoica



Crimes sur 2 semaines dans les environs de l'Université de Chicago [Baddeley 2016]

- K : moyenne normalisée du nombre de voisin



- F : fonction d'espace vide

$$F(r) = \mathbb{P}\{d(u, \mathbf{x}) \leq r\} \quad (1)$$

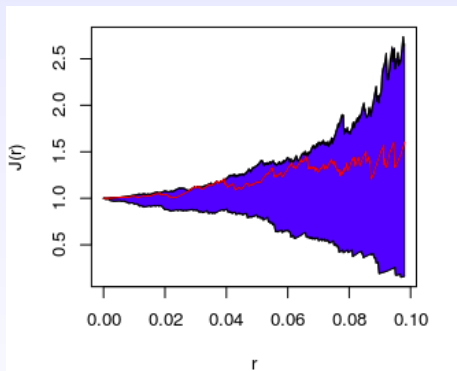
- G : fonction du plus proche voisin

$$G(r) = \mathbb{P}\{d(u, \mathbf{x} \setminus u) \leq r \mid u \text{ est un point de } \mathbf{x}\} \quad (2)$$

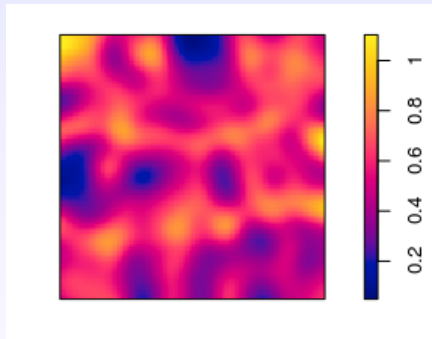
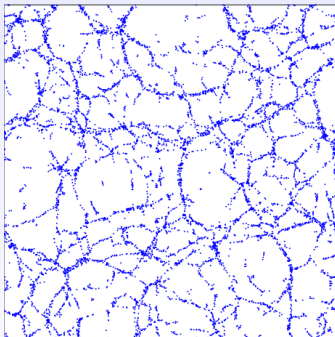
- J :

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \quad (3)$$

- Test d'enveloppe
- "Strauss"
- "Area Interaction"



Exemple d'enveloppe pour la
fonction J avec la méthode
"Strauss"



Exemple d'intensité calculée à partir de la configuration de point à gauche - Créée avec la librairie R-Spatstat



- langage utilisé par mon tuteur
- langage orienté objet



- adapté aux tests statistiques
- affichage des courbes

- ❶ Assimilation de la théorie
 - ▶ Lecture de la documentation
 - ▶ Explications de M. Stoica

- ❷ Discussions avec le tuteur
 - ▶ sur les structures de code à adopter
 - ▶ sur les méthodes à utiliser

- ❸ Implémentation
 - ▶ création des fonctions
 - ▶ vérification des résultats (corrections si besoin)

- Un gros RDV par semaine
 - ▶ Aide à la compréhension de la théorie
 - ▶ Discussions sur les méthodes d'implémentaion
 - ▶ Réponse aux questions
 - ▶ Planification du travail

- Petites rencontres régulières
 - ▶ Confirmations, validation du travail
 - ▶ Petits changements / corrections
 - ▶ Aide

- fonction K :

$$K(r) = \frac{\sum_{i=1}^n \mathbb{1}\{b_i \geq r\} \sum_{\substack{j=1 \\ j \neq i}}^n \mathbb{1}\{d_{ij} \leq r\}}{\lambda \sum_{i=1}^n \mathbb{1}\{b_i \geq r\}} \quad (4)$$

- intensité :

$$\lambda(u) = \sum_{i=1}^n \frac{1}{(\int_W K(x_i - v) dv)} k(u - x_i) \quad (5)$$

- adopter méthodes et stratégies pour implémentations optimales (tris, discrétisation d'intégrales,...)

- Fichiers de sortie



F.txt



G.txt



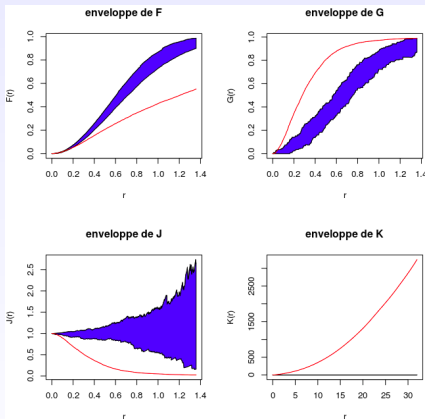
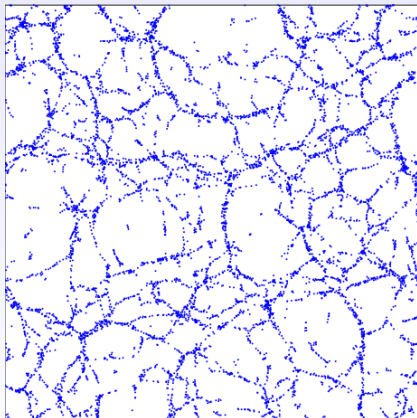
J.txt

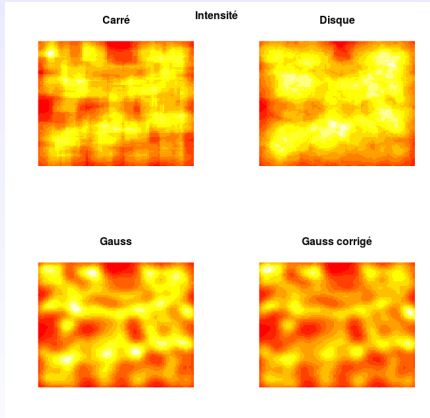
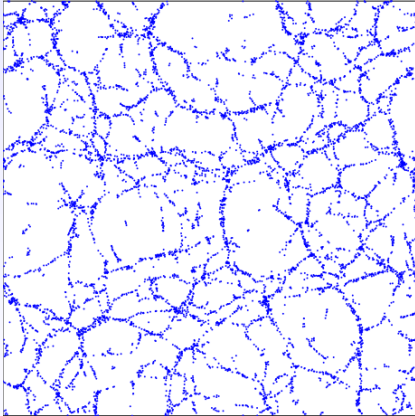


K.txt

- Scripts shell

```
clement@clement-K56CB:~/Documents/Stage2A/codeIntensite$ ./visualisation.sh  
fichier: POINTS/dataset3.txt  
dossier: RESULTS/dataset3  
m: 100  
c: 10  
r: 8  
sigma: 7  
  
xMin : 0.00502  
xMax : 127.986  
yMin : 0.02076  
yMax : 127.999
```





- Projet impliquant un travail avec des personnes d'autres domaines scientifiques
- Découverte ou approfondissement de domaines de connaissances
 - ▶ en mathématique
 - ▶ en informatique
 - ▶ dans le domaine de la recherche
- Objectif atteint avec de nombreuses perspectives

Merci pour votre attention.
Avez-vous des questions ?



Baddeley.

Spatial point patterns - methodology and applications with r.
CRC Press, 2016.