COURS 6 : statistiques spatiales rapide introduction aux processus spatiaux

Système d'Information Géographique

Arlette Antoni

Université de Bretagne Sud
Université Bretagne Pays de Loire

Année Universitaire 2021 -2022

Types de processus

- **Ponctuel**

  position du point aléatoire

  présence absence d'un événement

  éventuellement une valeur sur les points
- **Latticiel**

  position fixée (en réseau)

  valeur en ces points
- **Continu**

  valeur en tout point de l'espace

  éventuellement non observé mais supposé

Type de support latticiel

- **Cas régulier**
  - points équidistants
  - maillage (grille)

- **Cas irrégulier**
  - épouse les contours d'un objet géométrique
  - ou points non équidistants

Distribution de référence

- On observe N points
- Sur une surface S
- Intensité : N/S
- stationnaire au second ordre :

l'intensité du second ordre n'est pas affectée par la translation
ne dépend que de la différence entre les points :2(x,y) =2(xy).

- isotrope : non affecté la rotation
- La stationnarité au second ordre et l'isotropie sont indispensables pour de nombreux outils de statistique spatiale.

Source de variation

Dans le positionnement ou dans la valeur

- **Variation due à la réceptivité de la localisation à recevoir des points**
  - présence locale favorisant l'événement
  - effet de premier ordre

- **Interdépendance entre les observations**
  - épidémie ou agrégation
  - rejet
  - Ce sont des effets de second ordre

Quelques objectifs

- Le processus générateur des données génère les coordonnées géographiques associées à l'apparition d'une observation.
- On n'observe qu'une fenêtre (spatiale et ou temporelle)
- On veut quantifier l'écart entre la distribution spatiale des observations et une distribution complètement aléatoire dans l'espace.

Distribution de référence

- **Poisson**
  - toutes les positions ont la même probabilité d'accueillir un point
  - la position d'un point nouveau est indépendante des points précédents
- **Concentrée**
  - certaines position ont de plus fortes probabilités d'accueillir des points
  - la localisation d'un premier point favorise l'apparition d'autres points
- **Régulière**
  - toutes les positions ont la même probabilité d'accueillir un point
  - la localisation d'un premier point défavorise l'apparition d'autres points

Propriétés Processus de Poisson

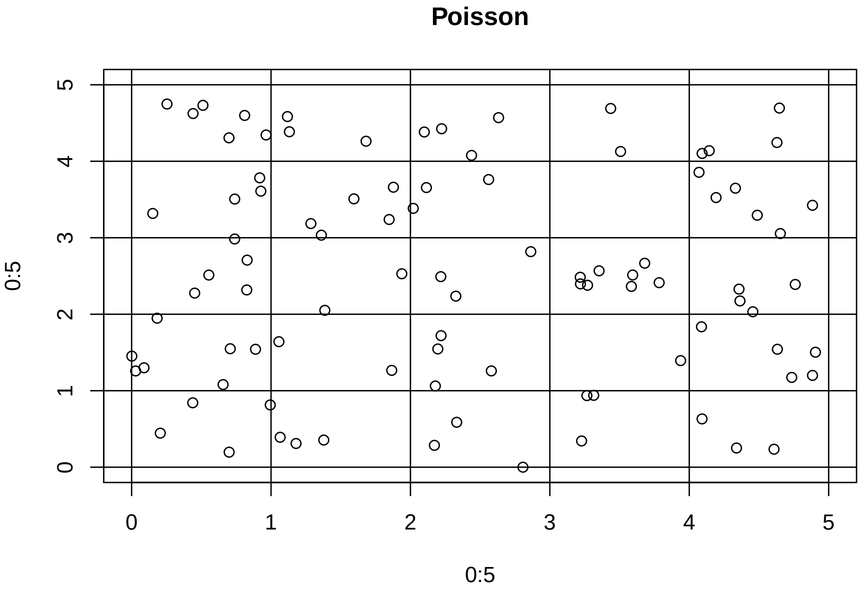
- Les impacts ou semis : suivent une exponentielle de paramètre λ
- Leur comptage dans un support (intervalle ou espace)
- Le processus de comptage $N(s)$ suit la loi de Poisson de paramètre λs.
- La loi conditionnelle des impacts sachant leur comptage suit une loi uniforme sur s

Simulation d'un Processus de Poisson

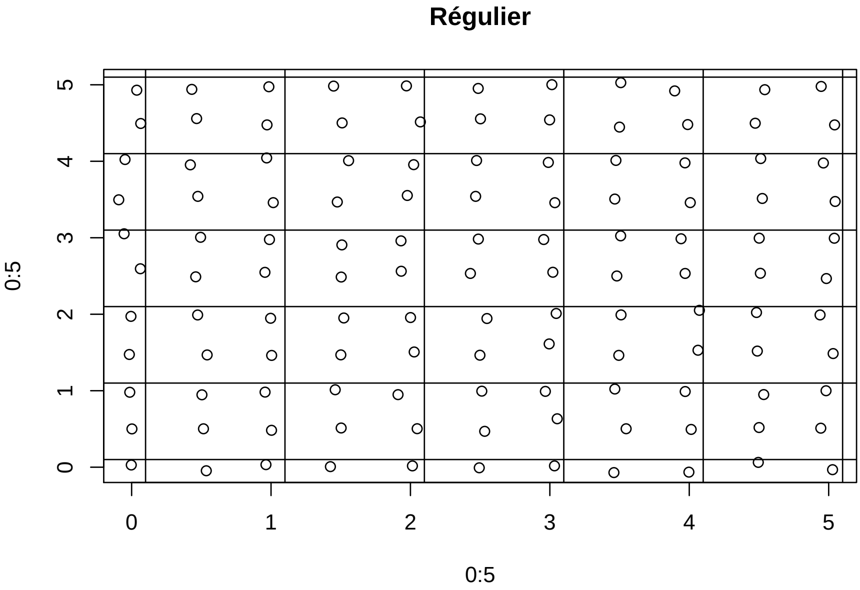
Dans R :

- on se donne une unité de surface S
- on découpe la surface en unités de taille S. par ex 5 lignes 4 colonnes - ce qui fera r=1 à 20 surfaces
- On tire le nombre total n_r de points du semis suivant une loi de Poisson de λS
- On tire successivement et indépendamment n_r points suivant une loi de probabilité uniforme sur S
- on visualise

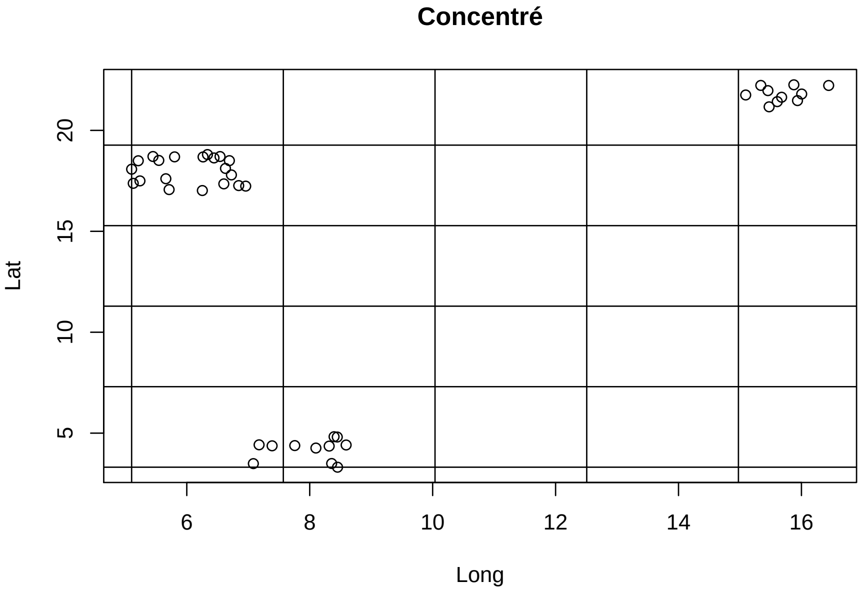
Simulation



Simulation



Simulation



Comptage d'un processus de ponctuel

Méthode des quadrats

- **découpage de l'espace selon un quadrillage**

  Il y a N quadrats
- **Numérotation des quadrats**

  on peut numéroter les quadrats séquentiellement de 1 à N

  ou selon les coordonnées ligne et colonne i et j (notre choix)
- **Comptage du nombre de points K**
  - $K_{i,j} = n_{i,j}$

Comparaison à un processus de Poisson

- **Cas poissonnien**

  $P(K = k) = \frac{\lambda^k exp(-\lambda)}{k!}$

- **Vérification rapide**
  - le coefficient de variation $\frac{\sqrt{var(K)}}{\bar{k}}$
  - si Poisson alors CV = 1
  - si CV > 1 concentré
  - si CV < 1 régulier

Les plus proches voisins

Sur des points répartis sur une surface S (mettons Bounding box)

- **Densité moyenne** : Nombre de points sur la surface S, $D = N/S$
- **Moyenne des plus petite distance**
  pour chaque point r on calcule d_r la distance à son voisin le plus proche
  on en fait la moyenne $\bar{d}$
- Calcul de la distance théorique : $d_t = \frac{0.5}{\sqrt{D}}$
- **Indice de dispersion** $\bar{d}/d_t$

Statistique de Ripley

B.D. Ripley 1976

On calcule le nombre de voisins d'un point dans un disque de rayon r

$$R(k) = \frac{surfTot}{N(N-1)} \Sigma_i \Sigma_{j \neq i} 1\{\|x_i - x_j\| < r\} correctionBord$$

Si la répartition est aléatoire alors

$$R(k) = \pi r^2$$

Selon les types de processus

Ripley.png

On peut aussi normaliser par

$$\sqrt{\frac{K(r)}{\pi}} - r$$