Contribution à l'étude des valeurs extrêmes dans un contexte spatial

C. Lantuéjoul
christian.lantuejoul@ensmp.fr

Centre de Géostatistique
Ecole des Mines de Paris

Introduction

Source d'inspiration:

– le rapport "Dépendance des extrêmes" de J.N. Bacro sur les coefficients de dépendance asymptotique;

– les modèles de type "tempête" étudiés par M. Schlather dans "Models for stationary max-stable random fields";

– la théorie des ensembles aléatoires de Matheron.

Thèmes abordés:

– coefficients de dépendance asymptotique des extrêmes;

– caractérisation statistique des modèles "tempête";

– utilisation des outils de la théorie des ensembles aléatoires à l'étude des extrêmes.

Coefficients de dépendance asymptotique des extrêmes

Quelques questions

Soient X et Y deux variables aléatoires positives. Pour étudier la dépendance asymptotique des extrêmes de X et de Y, deux fonctions χ et $\bar{\chi}$ ont été introduites:

$$\chi(z) = P\{Y > z \mid X > z\} \qquad \bar{\chi}(z) = \frac{\ln[P\{X > z\}P\{Y > z\}]}{\ln P\{X > z, Y > z\}}$$

On a ensuite posé $\chi = \lim_{z \longrightarrow \infty} \chi(z)$ ainsi que $\bar{\chi} = \lim_{z \longrightarrow \infty} \bar{\chi}(z)$.

L'introduction de ces coefficients suscite les questions suivantes:

– pourquoi $\bar{\chi}$ est symétrique en X et en Y, et pas χ?

– pourquoi l'introduction des logarithmes?

– quels sont les liens exacts entre ces coefficients?

– d'autres coefficients peuvent ils être envisagés?

– les répartitions spatiales des faibles et des fortes valeurs peuvent elles être étudiées de la même façon?

Cas de variables échangeables

Plutôt que de considérer X et Y, il est plus commode de s'intéresser au minimum $X \wedge Y$ et au maximum $X \vee Y$ de X et de Y.

Dans ce cas, il n'existe que deux lois non triviales reliant ces deux variables:

$$\alpha(z) = P\{X \wedge Y > z \mid X \vee Y > z\} \qquad \beta(z) = P\{X \vee Y < z \mid X \wedge Y < z\}$$

Ces quantités se récrivent

$$\alpha(z) = \frac{P\{X \wedge Y > z\}}{P\{X \vee Y > z\}} \qquad \beta(z) = \frac{P\{X \vee Y < z\}}{P\{X \wedge Y < z\}}$$

Il n'existe pas de relation simple entre $\alpha(z)$ et $\beta(z)$ hormis

$$\alpha_{X,Y}(z^{-1}) = \beta_{X^{-1},Y^{-1}}(z) \qquad \beta_{X,Y}(z^{-1}) = \alpha_{X^{-1},Y^{-1}}(z)$$

Quelques exemples

Exemple 1: X et Y sont complétement indépendantes de même loi F.

$$\alpha(z) = \frac{1 - F(z)}{1 + F(z)} \qquad \beta(z) = \frac{1 - \bar{F}(z)}{1 + \bar{F}(z)}$$

où $\bar{F} = 1 - F$ est la fonction de répartition complémentaire de F. Lorsque z croit, α est décroissante de 1 vers 0 tandis que β est croissante de 0 vers 1. Leur vitesse de croissance ou de décroissance dépendent de F.

Exemple 2: X et Y sont totalement dépendantes $(X = Y)$. On trouve immédiatement $\alpha(z) = \beta(z) = 1$.

Exemple 3: X et Y sont totalement dépendantes avec la probabilité r ou bien complétement indépendantes de même loi F avec la probabilité complémentaire $1 - r$. On a alors

$$\alpha(z) = \frac{1 - (1-r)F(z)}{1 + (1-r)F(z)} \qquad \beta(z) = \frac{1 - (1-r)\bar{F}(z)}{1 + (1-r)\bar{F}(z)}$$

A la limite, on obtient

$$\lim_{z \longrightarrow \infty} \alpha(z) = \lim_{z \longrightarrow 0} \beta(z) = \frac{r}{2-r} \qquad \lim_{z \longrightarrow \infty} \beta(z) = \lim_{z \longrightarrow 0} \alpha(z) = 1$$

Fonctions aléatoires tempêtes

Définition

Ingrédients de base:

– Π processus de Poisson homogène de densité μ sur $I\!R^d \times I\!R_+$;

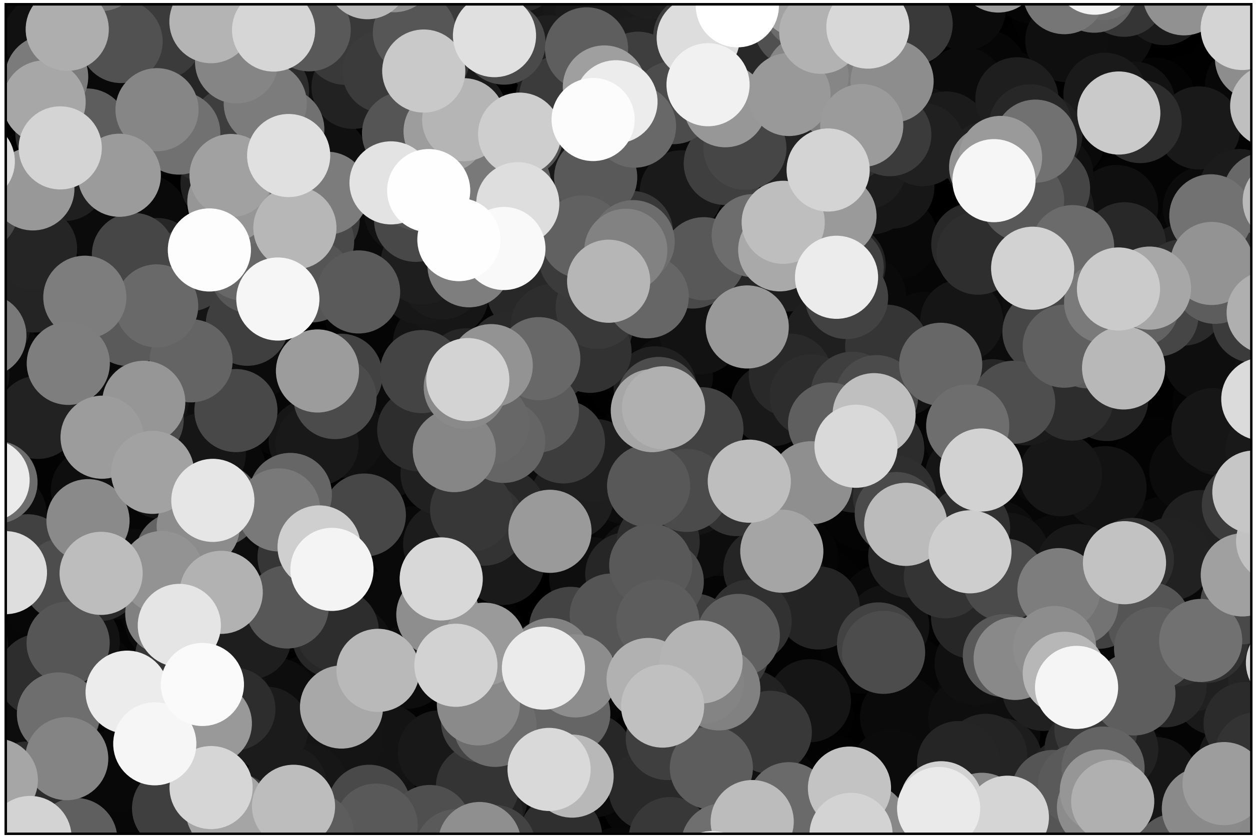
– $\left(Y_{y,t},\ y \in I\!R^d,\ t \in I\!R_+\right)$ copies indépendantes d'une même fonction aléatoire stationnaire Y définie sur $I\!R^d$ et à valeurs positives.

Définition:

$Z(x)$ est le maximum pris par les fonctions de base au point x, pondérées par leur temps d'arrivée

$$Z(x) = \sup_{(y,t)\in\Pi} \frac{Y_{y,t}(x-y)}{t} \qquad x \in I\!R^d$$

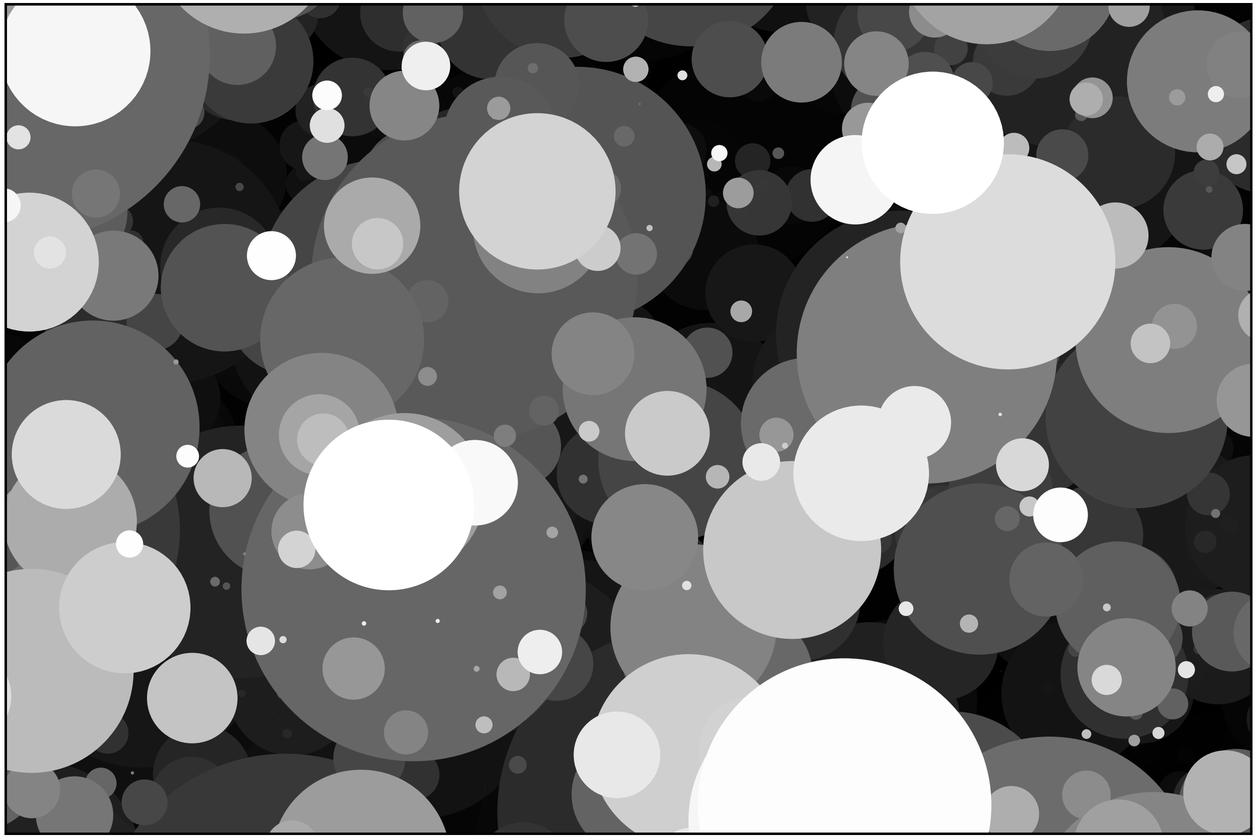
Exemple: disques de rayon constant



Rayon = 10

Champ 300 x 200

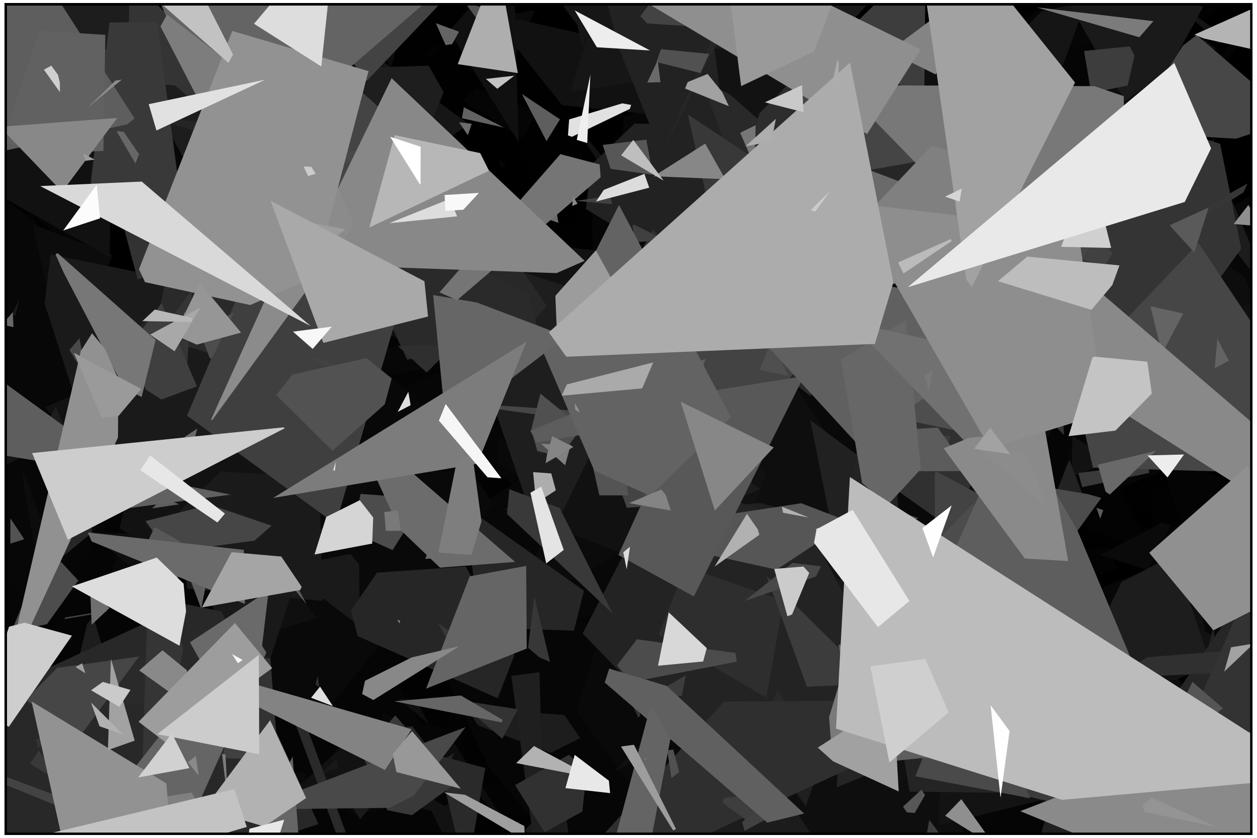
Exemple: disques de rayon exponentiel



Rayon moyen $= 7.072$

Champ 300 x 200

Exemple: polygones poissonniens



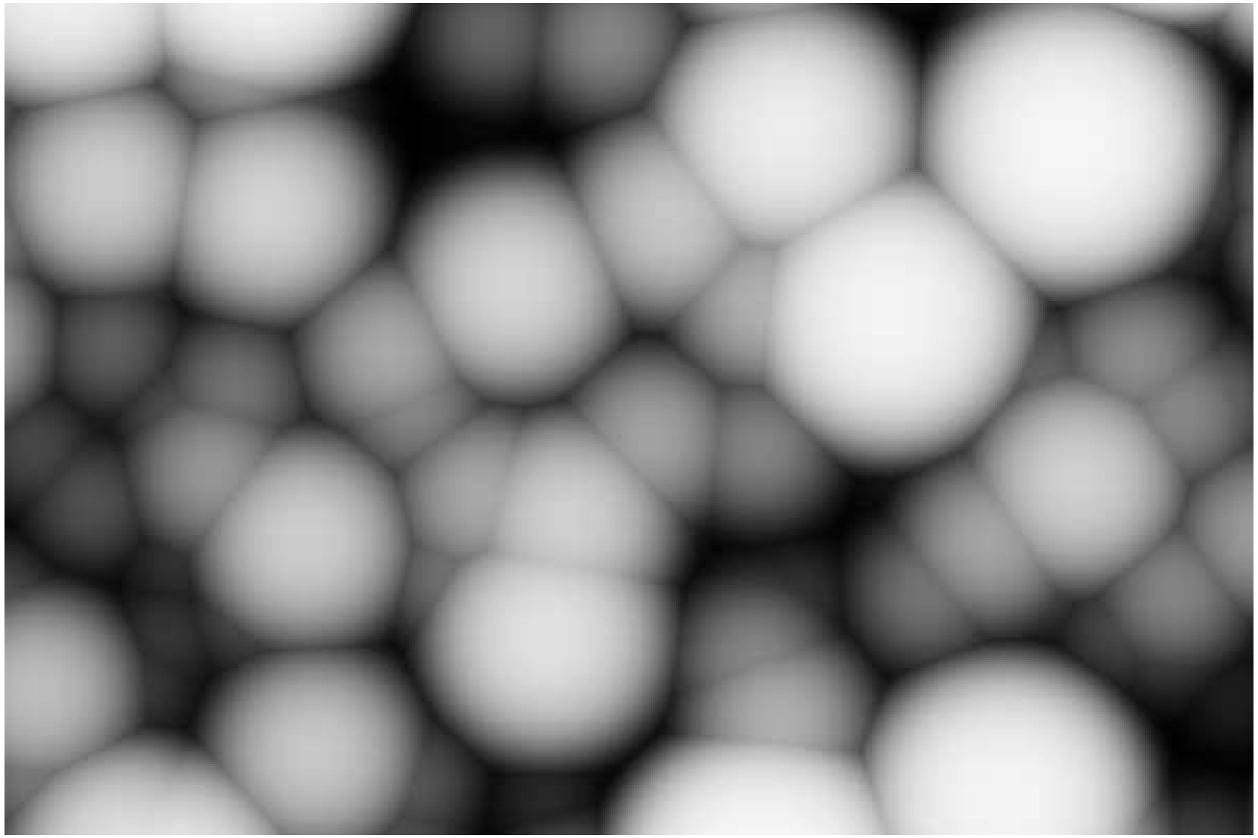
Surface moyenne d'un polygone = 314

Champ 300 x 200

Exemple: chapeaux chinois

$$Y(x) = \left(1 - \frac{|x|}{R}\right) 1_{|x|<R}$$

Exemple: fonctions gaussiennes



$$Y(x) = \exp\left(-\frac{|x|^2}{\sigma^2}\right)$$

Loi ponctuelle

On pose

$$m = E\left\{\int_{\mathbb{R}^d} Y(x)\,dx\right\} = \int_{\mathbb{R}^d} E\{Y(x)\}\,dx$$

Si $0 < m < \infty$ (ce qui se produit lorsque Y est à support compact ou bien à décroissance rapide à l'infini), alors $0 < Z < \infty$ p.s.

Dans ce cas, Z est une fonction aléatoire stationnaire à loi marginale Fréchet unité:

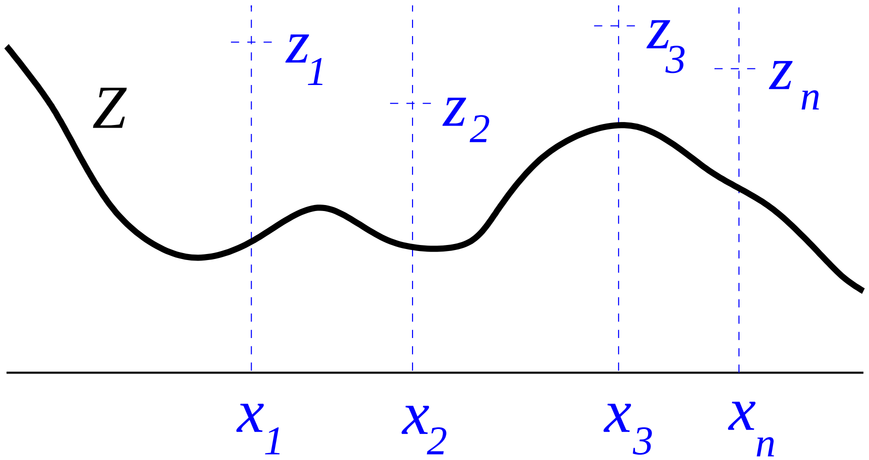
$$P\{Z(x) < z\} = \exp\left(-\frac{\mu m}{z}\right)$$

Loi spatiale

Plus généralement, la loi spatiale de Z est donnée par

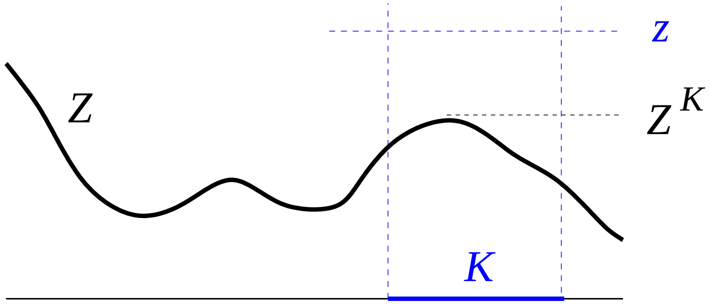
$$P\left\{\bigcap_{i\in I} Z(x_i) < z_i\right\} = \exp\left(-\mu \int_{\mathbb{R}^d} E\left\{\max_{i\in I} \frac{Y(x_i - y)}{z_i}\right\} dy\right)$$

pour toute famille finie de points $(x_i, i \in I)$ et toute famille finie de valeurs $(x_i, i \in I)$.



Loi du maximum sur un compact

Soit $Z^K = \max_{x \in K} Z(x)$ le maximum pris par Z (supposé s.c.s.) sur le compact K de $\mathbb{R}^d$



On peut montrer que ce maximum suit également une loi Fréchet unité:

$$P\{Z^K < z\} = \exp\left(-\frac{\theta(K)}{z}\right) \qquad K \in \mathcal{K}\,,\; z > 0$$

avec

$$\theta(K) = \mu E\left\{\int_{\mathbb{R}^d} Y^{K_y}\, dy\right\} = \mu \int_{\mathbb{R}^d} E\{Y^{K_y}\}\, dy$$

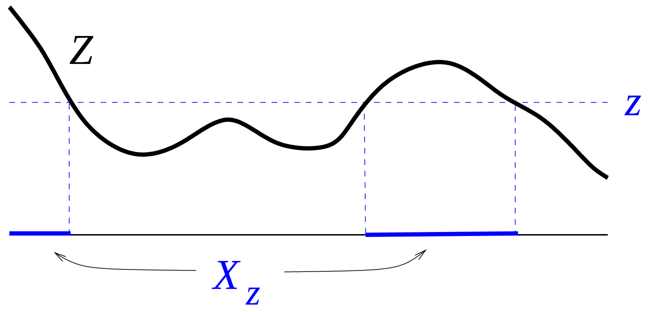
K_y translaté de K selon $\vec{oy}$

Utilisation de la théorie des ensembles aléatoires

Seuils des processus tempêtes

Définition:

Il s'agit des ensembles aléatoires $X_z = \{x \in I\!R^d : Z(x) \geq z\}$ pour tout $z > 0$. Comme Z est s.c.s., ces ensembles sont topologiquement fermés.



Caractérisation statistique:

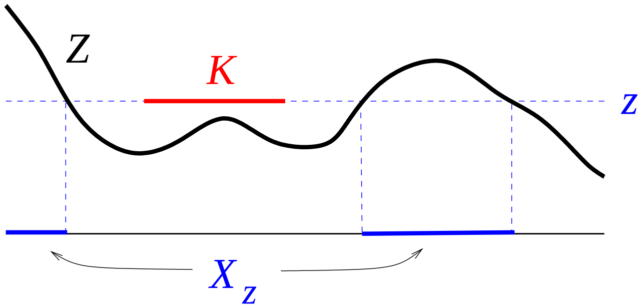
La théorie des ensembles fermés aléatoires (Matheron, 1975) montre les propriétés statistiques de X_z sont données par sa fonctionnelle d'évitement:

$$Q_z(K) = P\{X_z \cap K = \emptyset\} \qquad K \in \mathcal{K}$$

Seuils des processus tempêtes (2)

– La fonctionnelle d'évitement de X_z a pour formule

$$Q_z(K) = P\{X_z \cap K = \emptyset\} = P\{Z^K < z\} = \exp\left(-\frac{\theta(K)}{z}\right)$$



– X_z est infiniment divisible pour la réunion: il est en effet réunion de n copies indépendantes de X_{nz}:

$$Q_z(K) = \exp \quad -\frac{\theta(K)}{z} \quad = \exp \quad -n\frac{\theta(K)}{nz} \quad = \Big[Q_{nz}(K)\Big]^n$$

– X_z est sans point fixe: $Q_z(\{x\}) = P\{Z(x) < z\} = \exp(-1/z) < 1$

Caractérisation des fermés aléatoires infiniment divisibles et sans point fixe

Théorème (Matheron, 1975):

(i) Il existe une mesure positive σ-finie ζ_z sur l'ensemble $\mathcal{F}'$ des fermés non vides de $I\!R^d$ telle que $\zeta_z(\mathcal{F}_K) = -\ln Q_z(K)$ pour tout compact K de $I\!R^d$;

(ii) X_z a la même loi que la réunion des fermés d'un processus de Poisson localement fini sur $\mathcal{F}'$ de densité ζ_z.

$\mathcal{F}_K$ est l'ensemble des fermés rencontrant K

Un processus de Poisson sur $\mathcal{F}'$ est dit localement fini lorsque le nombre de fermés du processus rencontrant tout compact K est fini presque sûrement.

$$\zeta_z(\mathcal{F}_K) = -\ln Q_z(K) = \frac{\theta(K)}{z} \qquad K \in \mathcal{K}$$

En particulier $\zeta_1(\mathcal{F}_K) = \theta(K)$. Pour la suite, on notera ζ au lieu de ζ_1.

Quelques relations de cohérence

$$\theta(K) = \zeta(\mathcal{F}_K) \qquad K \in \mathcal{K}$$

- $\theta(K) \geq 0$ et $\theta(\emptyset) = \zeta(\mathcal{F}_\emptyset) = 0$;
- $K \subset K' \implies \theta(K) = \zeta(\mathcal{F}_K) \leq \zeta(\mathcal{F}_{K'}) = \theta(K')$;
- si $(K_i, i \in I)$ est une famille finie de compacts, alors

$$\begin{aligned}
0 \leq \zeta(\cap_{i\in I}\mathcal{F}_{K_i}) &= \sum_{J\subset I}(-1)^{|J|-1}\zeta(\cup_{j\in J}\mathcal{F}_{K_j}) \\
&= \sum_{J\subset I}(-1)^{|J|-1}\zeta(\mathcal{F}_{\cup_{j\in J}K_j}) \\
&= \sum_{J\subset I}(-1)^{|J|-1}\theta(\cup_{j\in J}K_j)
\end{aligned}$$

Liens avec la représentation de Resnick

Soit $(x_A = (x_a, a \in A)$ une partie finie de $I\!R^d$. D'après Resnick (1987), il existe une mesure H sur le simplexe $S = \{w_A = (w_a, a \in A) : w_a \geq 0 \,\text{et}\, \sum_{a \in A} w_a = 1\}$ telle que

$$\theta(x_B) = \int_S \max_{b \in B}(w_b)\, dH(w)$$

pour toute partie x_B de x_A.

On peut montrer par récurrence sur les points de x_A que l'on a

$$\int_S \min_{b \in B}(w_b)\, dH(w) = \zeta\left(\cap_{b \in B} \mathcal{F}_{x_b}\right)$$

pour toute partie x_b de x_A.

Bibliographie

– Bacro J.N. (2005) - "Dépendance des extrêmes". Rapport non publié.

– Matheron G. (1975) - Random sets and integral geometry. Wiley (New York).

– Resnick S.I. (1987) - Extreme values, iregular variation and point processes. Springer-Verlag (New York).

– Schlather M. (2002) - "Models for stationary max-stable random fields". Extremes, Vol. 5-1, pp. 33-44, 2002.

– Schlather M. et Tawn J.A. (2003) - "A dependence measure for multivariate and spatial extreme values: Properties and inference". Biometrika, Vol. 90.1, pp. 139-156.