Statistique des événements extrêmes Master 2 Biostat
2010-2011 J.-N. Bacro, M. Ribatet

EXAMEN

Durée : 2 heures.

Les documents manuscrits sont autorisés.

Les exercices sont indépendants.

Exercice 1. Soit $Y \sim GPD_{\xi,\beta}$ (loi de Pareto généralisée de paramètres de forme ξ et d'échelle β).

1. Soit y_1 et y_2 deux réels appartenant au domaine de définition de la loi $GPD_{\xi,\beta}$. Montrer que $P(Y > y_1 + y_2 \mid Y > y_1)$ s'exprime au travers d'une $GPD_{\xi,\beta+\xi y_1}$.
2. Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires de loi F inconnue.
   (a) Rappeler brièvement comment on peut, en pratique, utiliser la loi de Pareto généralisée pour caractériser le domaine d'attraction du maximum de F.
   (b) On décide de considérer successivement différents seuils $u_1 < u_2 < \ldots < u_k$. En vous appuyant sur le résultat de la question 1, proposer alors une approche régression pour estimer ξ permettant de prendre en compte l'information donnée par l'ensemble des dépassements de $u_1 < u_2 < \ldots < u_k$.
   (c) D'un point de vue théorique, en quoi cette méthode vous paraît-elle plus intéressante que l'estimation de ξ selon les dépassements d'un u_i, i fixé ?
   D'un point de vue pratique, en quoi ces avantages doivent-ils être nuancés ?

Exercice 2. Soit $\{Z(x)\}$, $x \in \mathbb{R}$ un processus stochastique donné par

$$Z(x) = c + \max_{i\geq 1} \left\{ \xi_i - \frac{(x - U_i)^2}{2\sigma^2} \right\}, \qquad x \in \mathbb{R}, \quad \sigma > 0, \quad c \in \mathbb{R}, \tag{1}$$

où $\{(\xi_i, U_i)\}_{i\geq 1}$ sont les points d'un processus de Poisson sur $\mathbb{R}\times\mathbb{R}$ d'intensité $\mathrm{d}\Lambda(\xi, u) = e^{-\xi}\mathrm{d}\xi\, \mathrm{d}u$, du étant la mesure de Lebesgue sur $\mathbb{R}$.

1. Donner une interprétation conceptuelle de ce processus max-stable. Vous pourrez vous appuyer sur un schéma.
2. Montrer que pour tout $k \in \mathbb{N}_*$, $x_1, \ldots, x_k \in \mathbb{R}^2$ et $z_1, \ldots, z_k \in \mathbb{R}$,

$$\Pr[Z(x_1) \leq z_1, \ldots, Z(x_k) \leq z_k] = \exp\left[-\int_{\mathbb{R}} \max_{j=1,\ldots,k} e^{-z_j} e^c \exp\left\{ -\frac{(x_j - u)^2}{2\sigma^2} \right\} \mathrm{d}u \right]. \tag{2}$$

3. Quelle valeur faut il donner à c pour que les marges de $\{Z(x)\}$ soient Gumbel unitaires, c.à.d. $\Pr[Z(x) \leq z] = \exp\{-\exp(-z)\}$, $z \in \mathbb{R}$, pour tout $x \in \mathbb{R}$?

4. Soit $U \sim N(0, \sigma^2)$ et notons φ_{σ^2} sa densité de probabilité. Montrer que
$$\Pr\left[e^{-z_1}\varphi_{\sigma^2}(U) > e^{-z_2}\varphi_{\sigma^2}(U + x_2 - x_1)\right] = \Phi\left(\frac{z_1 - z_2}{a} + \frac{a}{2}\right),$$
où $a = |x_1 - x_2|/\sigma$ et Φ est la fonction de répartition d'une loi Normale centrée réduite.
5. En déduire que lorsque $c = -\frac{1}{2}\log(2\pi\sigma^2)$ on a
$$-\log \Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = e^{-z_1}\Phi\left(\frac{z_1 - z_2}{a} + \frac{a}{2}\right) + e^{-z_2}\Phi\left(\frac{z_2 - z_1}{a} + \frac{a}{2}\right). \quad (3)$$
6. i) Déduire de la question précédente la fonction du coefficient extrême $\theta(h)$, $h > 0$.
   ii) Faire un graphe de cette fonction en indiquant où se trouve la dépendance totale et l'indépendance.
   iii) Expliquer comment varie la dépendance en fonction de σ.
7. Faire un développement limité à l'ordre 1 afin de caractériser le comportement de $\theta(h)$ près de l'origine.

Exercice 3. Soit $\{Z(x)\}$, $x \in \mathbb{R}^d$, un processus max-stable simple isotrope dont la fonction du coefficient extrême, notée $\theta_Z(h)$, vérifie $\theta_Z(h) \longrightarrow 2$ lorsque $h \to \infty$.

On considère le processus suivant
$$Z_*(x) = \max\{\alpha_1 Z(x), \alpha_2 \tilde{Z}\}, \qquad 0 \leq \alpha_1, \alpha_2 \leq 1, \qquad x \in \mathbb{R}^d, \quad (4)$$
où $\tilde{Z}$ est une variable aléatoire Fréchet unitaire, c.à.d. $\Pr[\tilde{Z} \leq z] = \exp(-1/z)$, $z > 0$.

1. Quelles contraintes sur α_1 et α_2 faut il imposer pour que $\{Z_*(x)\}$ ait des marges Fréchet unitaires ?
2. Considérons maintenant un cas particulier de (4) pour lequel nous avons
$$Z_*(x) = \max\{\alpha Z(x), (1-\alpha)\tilde{Z}\}, \qquad 0 \leq \alpha \leq 1.$$
Vérifier que $\{Z_*(x)\}$ est un processus max-stable simple.
3. i) Déterminer la fonction du coefficient extrême, appelons la $\theta_{Z_*}(h)$, de $\{Z_*(x)\}$ et tracer son graphe.
   ii) Quel rôle joue α ?
   iii) Expliquer le comportement de Z_* lorsque $\alpha = 0$.
4. **(Question Bonus)** Considérons maintenant le cas où dans (4) nous remplaçons $\tilde{Z}$ par un bruit Fréchet unitaire $\tilde{Z}(x)$ c.à.d. que pour tout $x_1, x_2 \in \mathbb{R}^d$, $\tilde{Z}(x_1)$ est indépendant de $\tilde{Z}(x_2)$ et que $\tilde{Z}(x)$ est une variable Fréchet unitaire pour tout $x \in \mathbb{R}^d$.
   i) Que vaut maintenant la fonction du coefficient extrême $\theta_{Z_*}(h)$? Tracer son graphe.
   ii) Quel rôle joue α ?
   iii) Expliquer le comportement de Z_* lorsque $\alpha = 0$.