Examen Théorie des Valeurs Extrêmes 2015-2016

ENSAE

Sans document - Sans calculatrice - 2h00

Questions de cours et exercices d'applications du cours (6 points):

1. Que signifie F appartient au domaine d'attraction d'une distribution max-stable G $(F \in D(G))$?

2. Quel lien relie les distributions GEV (Generalised Extreme Value distributions) et les distributions GPD (Generalised Pareto distributions)?

3. On suppose que deux distributions F et G ont le même point extrêmal $(x^F = x^G)$ et

$$\lim_{x \to x^F} \frac{\bar{F}(x)}{\bar{G}(x)} = c \in (0, \infty).$$

Montrer que F et G appartiennent au même domaine d'attraction (disons celui de la GEV$(0,1,\xi)$) et donner le lien entre les constantes de normalisation. On rappelle que la fonction de répartition d'une GEV$(0,1,\xi)$ est

$$\exp\left(-(1+\xi x)_+^{-\frac{1}{\xi}}\right), \quad x \in \mathbb{R}.$$

Réciproquement supposons que F et G ont le même point extrêmal et appartiennent au domaine d'attraction de la Gumbel $(\Lambda(x) = \exp(-(\exp(-x)))$ telles qu'il existe $c_n > 0$ et d_n vérifiant

$$\lim_{n\to\infty} F^n(c_n x + d_n) = \Lambda(x) \quad \text{et} \quad \lim_{n\to\infty} G^n(c_n x + d_n) = \Lambda(x+b).$$

Montrer que

$$\lim_{x \to x^F} \frac{\bar{F}(x)}{\bar{G}(x)} = c \in (0, \infty)$$

et caractériser c.

4. Quelles sont les procédures statistiques à utiliser pour estimer les paramètres des distributions GEV (discuter suivant la nature des observations)?

Exercice 2 (4 points):

Quelles sont les domaines d'attraction des distributions suivantes (justifier):

- Benktander de type I:

$$\bar{F}(x) = (1 + 2(\beta/\alpha)\ln x)\exp\left\{-\left(\beta(\ln x)^2 + (\alpha+1)\ln x\right)\right\}, \quad x \geq 1, \quad \alpha, \beta > 0,$$

- Benktander de type II:

$$\bar{F}(x) = x^{-(1-\beta)}\exp\left\{-\frac{\alpha}{\beta}(x^\beta - 1)\right\}, \quad x \geq 1, \quad \alpha > 0, 0 < \beta < 1,$$

- Personne de type I:

$x-1$

$$\bar{F}(x) = \exp\left(-\frac{x}{1-x}\right), \quad 0 \leq x < 1.$$

- Personne de type II:

$$\bar{F}(x) = \frac{1}{(\ln\ln(x))^{\ln\ln(x)}}, \quad x \geq e.$$

Vous pourrez utiliser la fonction $h(x) = \bar{F}(x)/f(x)$ ou discuter l'épaisseur de la queue de distribution.

Exercice 1 (6 points): On rappelle qu'une fonction L mesurable et positive sur $]0, \infty[$ est *à variations lentes*, si, pour tout $x > 0$, $\lim_{t\to\infty} L(tx)/L(t) = 1$. De plus un distribution F appartient au domaine d'attraction de la loi de Fréchet Φ_α ($\alpha > 0$), si et seulement si $\bar{F}(x) = x^{-\alpha}L(x)$, pour une fonction à variations lentes L.

Soient X_1 et X_2 deux variables positives, indépendantes, identiquement distribuées de distribution caractérisée par $\bar{F}(x) = x^{-\alpha}L(x)$.

1. Montrer que

$$\{X_1 > x\} \cup \{X_2 > x\} \subset \{X_1 + X_2 > x\}.$$

En déduire que

$$P(X_1 + X_2 > x) \geq 2P(X_1 > x)(1 + o(1))$$

quand $x \to \infty$.

2. Montrer que, pour $0 < \delta < 1/2$

$$\{X_1 + X_2 > x\} \subset \{X_1 > (1-\delta)x\} \cup \{X_2 > (1-\delta)x\} \cup \{X_1 > \delta x, X_2 > \delta x\}.$$

En déduire que

$$P(X_1 + X_2 > x) \leq 2P(X_1 > (1-\delta)x)(1 + o(1))$$

quand $x \to \infty$.

3. En jouant sur δ, montrer que

$$P(X_1 + X_2 > x) \sim 2P(X_1 > x).$$

4. Soient $X_1, ..., X_n$ n variables positives, indépendantes, identiquement distribuées de distribution caractérisée par $\bar{F}(x) = x^{-\alpha}L(x)$. On note $S_n = X_1 + ... + X_n$. Montrer que

$$P(S_n > x) \sim nP(X_1 > x)$$

quand $x \to \infty$. En déduire que

$$P(S_n > x) \sim P(M_n > x)$$

quand $x \to \infty$, où $M_n = \max(X_1, ..., X_n)$. Utiliser la formule du binôme de Newton.

Interpréter ce résultat en terme de gestion des risques extrêmes.

Exercice 2 (4 points):

Soit $X_1, X_2, ...$ une suite de variables aléatoires indépendantes et identiquement distribuées de distribution $\Phi_1(x) = \exp(-1/x)$, $x \geq 0$. On note $M_n^X = \max(X_1, ..., X_n)$.

0. Donner a_n et b_n pour que

$$\frac{M_n^X - b_n}{a_n} \sim_d \Phi_1.$$

On définit

$$Y_n = \frac{1}{3} \max(X_n, X_{n+1}, X_{n+3}).$$

1. Donner la distribution de Y_n.

2. Soit $M_n^Y = \max(Y_1, ..., Y_n)$. Quelle est la loi limite de

$$\frac{M_n^Y - b_n}{a_n}$$

quand $n \to \infty$?

Interpréter ce résultat par rapport à celui de la question 0.

3. (Question bonus) Quelle est la dynamique asymptotique de l'arrivée des dépassements de seuils $\{i/n : X_i > a_n x + b_n\}_{i=1,...,n}$ et celle des dépassements de seuils $\{i/n : Y_i > a_n x + b_n\}_{i=1,...,n}$ quand $n \to \infty$?