Théorie des Valeurs Extrêmes – Exercices

A – Inverse généralisé et quantile

Exercice [A-E1] – Soit la fonction $\Phi : [0, \infty[\mapsto [0, \infty[$ définie par :

$$\Phi(x) = \begin{cases} x^2 & \text{si } x \in [0, 1[, \\ 1 & \text{si } x \in [1, 2[, \\ \sqrt{3} & \text{si } x \in [2, 3], \\ \sqrt{x} & \text{si } x > 3. \end{cases}$$

Donner l'expression de l'inverse généralisée $\Phi^{\leftarrow}$ de Φ.

Exercice [A-E2] – Donner l'expression de la fonction quantile pour une loi uniforme sur $[a, b]$.

Exercice [A-E3] – Donner l'expression de la fonction quantile $q :]0, 1[\to \mathbb{R}$ définie par $q(\alpha) = F^{\leftarrow}(1 - \alpha)$ avec

$$F(x) = \begin{cases} 0 & \text{si } x < 1, \\ (x-1)/3 & \text{si } x \in [1, 2[, \\ (2/3)^{3/\max(x,3)} & \text{si } x \geq 2. \end{cases}$$

B – Domaines d'attraction

Exercice [B-E1] – Soit F la fonction de répartition d'une loi uniforme sur $[0, 1]$. Pour tout $\alpha \in [0, 1]$, on note $q(\alpha) := F^{\leftarrow}(1 - \alpha)$ le quantile d'ordre $1 - \alpha$ d'une loi uniforme sur $[0, 1]$.

i) Donner l'expression de $q(\alpha)$.

ii) Soient $U_1, \ldots, U_n$ des variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. Montrer que si α_n est une suite de $[0, 1]$ telle que $n\alpha_n \to 0$ alors, il existe une suite v_n telle que

$$v_n \left(\frac{U_{n,n}}{q(\alpha_n)} - 1 \right) \xrightarrow{d} Y,$$

où Y est une variable aléatoire non dégénérée dont vous préciserez la loi.

Exercice [B-E2] – Vérifier que le domaine d'attraction de la loi

i) exponentielle de paramètre $\lambda = 1$ est le domaine d'attraction de Gumbel.

ii) de Pareto de paramètres $\gamma > 0$ et $a > 0$ est le domaine d'attraction de Fréchet.

iii) uniforme sur $[0,1]$ est le domaine d'attraction de Weibull.

Exercice [B-E3] – Soit F la fonction de répartition d'une loi de Gumbel donnée par :

$$F(x) := \exp\left(-\exp(-x)\right), \ x \in \mathbb{R}.$$

Donner les suites (a_n) et (b_n) telles que $F^n(a_n x + b_n) = F(x)$.

Exercice [B-E4] – Quel est le domaine d'attraction de la fonction de répartition F telle que $F(x) = \mathrm{e}^x$ si $x < 0$ et $F(x) = 1$ si $x \geq 0$. Même question pour la fonction de répartition G telle que $G(x) = 1 - \mathrm{e}^{1/x}$ si $x < 0$ et $G(x) = 1$ si $x \geq 0$.

Exercice [B-E5] – Soit la fonction de répartition

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - (1 + x^c)^{-d} & \text{si } x > 0, \end{cases}$$

avec $c > 0$ et $d > 0$.

i) Montrer que $\bar{F} := 1 - F$ est une fonction à variations régulières dont vous préciserez l'indice qui sera notée ξ dans la suite. A quel domaine d'attraction appartient F ? (donner la valeur de γ en fonction de ξ).

ii) Ecrire $\bar{F}$ sous sa représentation de Karamata.

iii) Proposer des suites de normalisation $a_n > 0$ et b_n telles que $F^n(a_n x + b_n) \to H_\gamma(x)$ lorsque $n \to \infty$ en tout point de continuité de H_γ.

Exercice [B-E6] – On rapppelle que la densité d'une loi de Cauchy est donnée par

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}.$$

i) Calculer la fonction de répartition d'une loi de Cauchy.

ii) En utilisant les équivalences suivantes : $\tan(x) \sim x$ et $\arctan(x) \sim x$ lorsque $x \to 0$ et le fait que pour $x > 0$,

$$\arctan\left(\frac{1}{x}\right) + \arctan(x) = \frac{\pi}{2},$$

donner le domaine d'attraction de la loi de Cauchy ainsi qu'un choix possible pour les suites de normalisation.

Exercice [B-E7] – Montrer que la loi normale centrée et réduite appartient au domaine d'attraction de Gumbel avec pour suites de normalisation

$$a_n = (2\log n)^{-1/2} \text{ et } b_n = (2\log n)^{1/2} - 1/2 \frac{\log\log n + \log(4\pi)}{(2\log n)^{1/2}}$$

Montrer également que la loi normale $\mathcal{N}(0,1)$ est une loi à queue de type Weibull d'indice $\theta = 1/2$.

Exercice [B-E8] – On considère la loi de student à $\mu > 0$ degré de liberté de densité :

$$f(x) = \frac{\Gamma((\mu+1)/2)}{\sqrt{\mu\pi}\Gamma(\mu/2)} \left(1 + \frac{x^2}{\mu}\right)^{-(\mu+1)/2}$$

Montrer que la loi de student appartient au domaine d'attraction de Fréchet en précisant la valeur de l'indice des valeurs extrêmes γ. Donner un choix possible pour les suites de normalisation.

Exercice [B-E9] – Soit $\gamma \in \mathbb{R}$ et F une fonction de répartition appartenant au domaine d'attraction de H_γ. On note x_F le point terminal de F.

i) Montrer que

$$\lim_{t \uparrow x_F} \frac{\lim_{s \uparrow t}(1 - F(s))}{\lim_{s \downarrow t}(1 - F(s))} = 1.$$

ii) Soit X une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda > 0$. Montrer que pour tout $x > \lambda$,

$$\mathbb{P}[X \geq x] \leq \mathrm{e}^{-\lambda} \frac{\mathrm{e}^x \lambda^x}{x^x}.$$

iii) En déduire que la loi de Poisson n'appartient à aucun domaine d'attraction.

C – Exercices complémentaires

Exercice [C-E1] – Soit N une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$. Soient $\{X_i,\ i \in \mathbb{N}^*\}$ des variables aléatoires indépendantes et de même loi de Pareto généralisée de paramètres $\gamma \neq 0$ et $\sigma > 0$ et indépendantes de N. Donner la loi de $M_N := \max(X_1, \ldots, X_N)$.

Exercice [C-E2] – Calculer la fonction de répartition des excès et la fonction excès moyen d'une variable aléatoire X de loi :

i) Exponentielle de paramètre $\lambda > 0$,

ii) Pareto Généralisée de paramètres $\gamma \in \mathbb{R}$ et $\sigma > 0$,

iii) Pareto de paramètres $c > 0$ et $\gamma > 0$ (dont on rappelle que la fonction de répartition est donnée pour $x \geq c$ par $1 - (x/c)^{-1/\gamma}$).

Exercice [C-E3] – Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes et de même fonction de répartition F que l'on supposera continue et strictement croissante. On pose $Z_n := n(1 - F(X_{n,n}))$ où $X_{n,n} = \max(X_1, \ldots, X_n)$. Donner

la loi limite de Z_n.

Exercice [C-E4] – Soient F_1 et F_2 deux fonctions de répartition admettant le même point terminal x^*. Si $F_1 \in \mathcal{DA}(H_{\gamma_1})$ et $F_2 \in \mathcal{DA}(H_{\gamma_2})$ avec $\gamma_1 < \gamma_2$ alors

$$\lim_{x \uparrow x^*} \frac{1 - F_1(x)}{1 - F_2(x)} = 0.$$

Ceci illustre bien le fait que l'indice des valeurs extrêmes contrôle la "lourdeur" de la queue de distribution.

D – Statistiques d'ordre

Exercice [D-E1] – Soit k_n une suite telle que $k_n \to \infty$ et $n/k_n \to \infty$ et soient $U_{1,n} \leq \ldots \leq U_{n,n}$ les statistiques d'ordre d'un échantillon de variables aléatoires de loi uniforme standard. Montrer que $n/k_n U_{k_n+1,n}$ converge presque-sûrement vers 1.

Exercice [D-E2] – On rappelle qu'une loi est une loi à queue de type Weibull si sa fonction de répartition est de la forme :

$$F(x) = 1 - \exp\left\{-V^{\leftarrow}(x)\right\},$$

avec $\theta > 0$ et où V est une fonction à variations régulières d'indice $1/\theta$.

i) Rappeler la définition d'une fonction à variations régulières.

ii) Donner un exemple simple de loi à queue de type Weibull.

On considère à présent un échantillon $(X_1, \ldots, X_n)$ de variables aléatoires indépendantes et de même fonction de répartition :

$$F(x) = 1 - \exp\left(-x^{1/\theta}\right), \ x > 0, \ \theta > 0.$$

iii) Soit (k_n) une suite d'entiers telle que $k_n \to \infty$ et $n/k_n \to 0$ lorsque $n \to \infty$. Montrer que $\{\log X_{n-i+1,n} - \log X_{n-i,n}, \ i = 1, \ldots, k_n\}$ à même loi que

$$\left\{\theta \log \frac{E_{n-i+1,n}}{E_{n-i,n}}, \ i = 1, \ldots, k_n\right\},$$

où $E_1, \ldots, E_n$ sont des variables aléatoires indépendantes et de loi exponentielle de paramètre 1.

iv) En utilisant le théorème des accroissements finis sur la fonction logarithme, montrer que pour tout $i \in \{1, \ldots, k_n\}$, il existe une variable aléatoire $E^*_{i,n} \in]E_{n-i,n}, E_{n-i+1,n}[$ telle que

$$\log \frac{E_{n-i+1,n}}{E_{n-i,n}} = \frac{E_{n-i+1,n} - E_{n-i,n}}{E^*_{i,n}}.$$

v) En utilisant le représentation de Rényi pour des variables aléatoires exponentielles, montrer que

$$\max_{i=1,\dots,k_n} \frac{E^*_{i,n}}{\log(n/i)} = 1 + o_{\mathbb{P}}(1).$$

vi) En déduire que $\{i(\log X_{n-i+1,n} - \log X_{n-i,n}),\ i = 1, \dots, k_n\}$ à même loi que

$$\left\{\theta \frac{F_i}{\log(n/i)}(1 + o_{\mathbb{P}}(1)),\ i = 1, \dots, k_n\right\},$$

où $F_1, \dots, F_n$ sont des variables aléatoires indépendantes de loi exponentielle de paramètre 1 et où le $o_{\mathbb{P}}(1)$ ne dépend pas de i.

vii) Proposer un estimateur qui converge en probabilité vers θ.

E – Estimation

Exercice [E-E1] – On rappelle que la fonction de répartition d'une loi de Burr de paramètres $\tau > 0$, $\beta > 0$ et $\lambda > 0$ est donnée pour $x > 0$ par

$$F(x) = 1 - \left(1 + \frac{x^\tau}{\beta}\right)^{-\lambda}.$$

i) Montrer que F appartient au domaine d'attraction de Fréchet et préciser la valeur de l'indice γ des valeurs extrêmes.

ii) On rappelle que le quantile d'ordre $\alpha \in]0, 1[$ est $Q(\alpha) = F^{\leftarrow}(1 - \alpha)$. Montrer que pour la loi de Burr, $Q(\alpha) = \alpha^{-\gamma}\ell(\alpha^{-1})$ où ℓ est une fonction à variations lentes que vous préciserez.

iii) Montrer que ℓ est une fonction à variations lentes normalisée *i.e.* il existe x_0 tel que

$$\ell(x) = c \exp\left\{\int_{x_0}^{x} \frac{\Delta(t)}{t} dt\right\},$$

avec $c > 0$ et $\Delta(t) \to 0$ lorsque $t \to \infty$. Montrer également que Δ est une fonction positive, décroissante sur $]0, \infty[$ et à variations régulières d'indice $\rho < 0$ (vous donnerez la valeur de ρ).

Exercice [E-E2] –Soit X une variable aléatoire suivant une loi de Pareto généralisée de paramètres $\gamma \neq 0$ et $\sigma = 1$. On pose $Y = 1 + \gamma X$.

i) Donner la loi de Y.

ii)) Soit $s \in \mathbb{N}^*$. Donner la relation liant γ et s et assurant que $\mathbb{E}(Y^s)$ existe. Calculer alors cette valeur.

iii) Soient $X_1, \ldots, X_n$ des variables aléatoires indépendants et de même loi que X. A l'aide de la question précédente, proposer un estimateur de γ.

iv) Donner la loi limite de l'estimateur proposé à la question précédente (attention aux conditions sur γ).

Exercice [E-E3] – On considère une fonction de répartition F appartenant au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $\gamma > 0$.

i) Donner l'expression de F en fonction de γ et d'une fonction L dont on rappellera la principale caractéristique.

ii) Pour $u > 0$, donner l'expression de $G_u(t) := \mathbb{P}(\log(X/u) \leq t | X > u)$, où X admet F pour fonction de répartition.

On rappelle que si Y est une variable aléatoire positive,

$$\mathbb{E}(Y) = \int_0^\infty \mathbb{P}(Y > t) dt.$$

On rappelle également le résultat suivant : soit v une fonction à variations régulières d'indice $\alpha < -1$. Alors la fonction

$$V(x) := \int_x^\infty v(t) dt,$$

vérifie $xv(x) \sim -(\alpha + 1)V(x)$ lorsque $x \to \infty$.

iii) Montrer que $\mathbb{E}(\log(X/u)|X > u) \to \gamma$ lorsque $u \to \infty$.

iv) En déduire un estimateur de γ.

Exercice [E-E4] – On considère la fonction

$$f(x) = \beta(\beta + 1)x(1 - x)^{\beta - 1} \mathbb{I}_{\{x \in [0,1]\}}.$$

i) Montrer que f est une densité.

ii) Calculer la fonction de survie $\bar{F}$ associée et en déduire le domaine d'attraction de cette loi (vous donnerez la valeur de l'indice des valeurs extrêmes γ correspondant).

iii) Soit $(X_1, \ldots, X_n)$ un échantillon de variables aléatoires indépendantes de même densité f. En utilisant la méthode des moments, proposer un estimateur de l'indice des valeurs extrêmes.