- La méthode des moments pondérés

égaliser les moments théoriques de GEV avec les moments empiriques

Estimation du paramètre de forme ξ à partir de l'ensemble des données

$\xi \in \mathbb{R}$ $\quad \hat{\xi}^{Pickands}_{k,n} = \left[\ln \dfrac{X_{(k)} - X_{(2k)}}{X_{(2k)} - X_{(4k)}} \right] / \ln 2$

$$\hat{\xi}^{DEdH}_{k,n} = \hat{\xi}^{H(1)}_{k,n} + 1 - \frac{1}{2} \left[1 - \frac{\left(\hat{\xi}^{H(1)}_{(k)}\right)^2}{\hat{\xi}^{H(2)}_{(k)}} \right]^{-1}$$

avec $\hat{\xi}^{H(r)}_{k,n} = \dfrac{1}{k} \sum_{i=1}^{k-1} \left[\ln X_{(i)} - \ln X_{(k)} \right]^r \quad r = 1, 2, \ldots$

$\xi > 0$ $\quad \hat{\xi}^{Hill}_{k,n} = \dfrac{1}{k} \sum_{i=1}^{k} \ln X_{(i)} - \ln X_{(k+1)}$

Les estimateurs sont faiblement convergents

Hypothèses supplémentaires sur k et F :

$$k^{1/2} \left(\hat{\xi}^{P}_{k,n} - \xi \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \xi^2 \frac{(2^{2\xi+1} + 1)}{(2(2^{\xi} - 1) \ln 2)^2} \right)$$

$$k^{1/2} \left(\hat{\xi}^{DEdH}_{k,n} - \xi \right) \xrightarrow{\mathcal{L}} \mathcal{N} (0, 1 + \xi^2), \ \xi \geq 0$$

$$k^{1/2} \left(\hat{\xi}^{H}_{k,n} - \xi \right) \xrightarrow{\mathcal{L}} \mathcal{N} (0, \xi^2)$$

Estimation des paramètres de la loi Pareto généralisée - POT method

GDP utiliseé pour modéliser les dépassements au-delà d'un certain seuil

$Y_{u,1}, \dots, Y_{u,n}$: dépassements au-delà du seuil u $\sim GPD(\sigma_u, \xi)$:

$$\mathbb{P}(Y_u < y \mid Y_u > 0) = 1 - \left[1 + \xi\left(\frac{y}{\sigma_u}\right)\right]_+^{-1/\xi}$$

Choix du seuil u :

- utiliser un seuil peu élevé permet d'augmenter le nb de données mais l'approximation par GPD est mauvaise
- utiliser un seuil trop élevé limite le nb de données et les estimateurs sont moins précis

On peut utiliser les propriétés de la GPD pour déterminer le seuil u.

- Propriétés des lois GPD

  - Stabilité par seuil
    Si $Y_u \mid Y_u > 0 \sim GPD(\sigma_u, \xi)$
    Alors $Y_v \mid Y_v > 0 \sim GPD(\underbrace{\sigma_v}_{\sigma_u + \xi(v-u)}, \xi)$ $\quad v \geq u$

  - Fonction de dépassement moyen
    Si $X = Y_u + u$ où $Y_u \mid Y_u > 0 \sim GPD(\sigma_u, \xi)$
    Alors $\mathbb{E}(X - v \mid X > v) = \dfrac{\sigma_u + \xi(v-u)}{1-\xi}$
    ↳ linéaire en v avec une pente $\xi/(1-\xi)$

- Procédure de sélection du seuil

  - Graphique de dépassement moyen
    si la va suit une GPD de seuil u,
    alors graphique $\sim$ linéaire au delà du seuil

  - Graphique de stabilité du paramètre d'échelle
    $\sigma^* = \sigma_v - \xi v$: ne dépend plus de $v \geq u$ si la va suit une loi de GPD pour le seuil u

$(\hat{\sigma}^*, \hat{\xi})$ pour différents seuils
Si va suit GPD pour seuil u
Alors $(\sigma^*, \xi) \sim$ constants pour $v > u$
$\Rightarrow$ On choisit u comme le seuil minimal pour lequel la propriété de stabilité est satisfaite

2. Analyse des extrêmes dans des cadres dynamiques et multivarié

2.1 Lois limites pour le maximum et dépassements de seuils dans un cadre dynamique

Extrêmes d'une série temporelles $\neq$ suite de va iid
↳ dépendance temporelle peut affecter amplitude et dynamique

$X_1, X_2, \ldots$ strictement stationnaire $\forall h, n$ dist. $(X_{h+1}, \ldots, X_{h+n}) \perp h$
$X_i \sim F$
$\tilde{X}_1, \tilde{X}_2 \ldots$ série associée de va iid de distribution F

- $\beta_{n,\ell}(\tau) = \sup |\mathbb{P}(X_i \leq u_n(\tau), i \in A \cup B)$
  $- \mathbb{P}(X_i \leq u_n(\tau), i \in A)\mathbb{P}(X_i \leq u_n(\tau), i \in B)|$
  $A \subset \{1, \ldots, k\}$, $B \subset \{k+\ell, \ldots, n\}$

  Condition $D(u_n(\tau))$: deux évènements de type $\{\max_{i \in I_1} X_i \leq u_n(\tau)\}$ et $\{\max_{i \in I_2} X_i \leq u_n(\tau)\}$ deviennent asymptotiquement indépendants lorsque n augmente et que les ensembles d'indices $I_i \subset \{1, \ldots, n\}$ sont séparés par une distance $\ell_n = O(n)$
  ↳ satisfaite si $\exists\, \ell_n = O(n)$ tq $\ell_n \xrightarrow[n \to \infty]{} \infty$
  $\beta_{n,\ell_n}(\tau) \xrightarrow[n \to \infty]{} 0$

- $n\bar{F}(u_n(\tau)) \longrightarrow \tau$ et condition $D(u_n(\tau))$ satisfaite
  Si $\mathbb{P}(M_n \leq u_n(\tau))$ converge
  Alors $\mathbb{P}(M_n \leq u_n(\tau)) \longrightarrow \exp(-\theta\tau)$ $\quad 0 \leq \theta \leq 1$
  θ : indice extrémal de (X_n)
  Distributions limites différentes pour (X_i) et $(\tilde{X}_i)$ sauf si $\theta = 1$
  $\theta > 0$, $u_n(\tau) = a_n x + b_n$, $\tau = (-\ln G(x))$, $G \stackrel{\mathcal{L}}{=} GEV(\mu, \sigma, \xi)$
  $\Rightarrow \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \longrightarrow G^*(x)$, $G \stackrel{\mathcal{L}}{=} GEV(\mu^*, \sigma^*, \xi^*)$
  $\xi = \xi^*$, $\sigma = \sigma^* \theta^{\xi}$, $\mu = \mu^* - \sigma^* \frac{1 - \theta^{\xi}}{\xi}$

Condition $D'(u_n(\tau))$ satisfaite
$\lim_{k\to\infty} \lim_{n\to\infty} \sup n \sum_{j=2}^{[n/k]} \mathbb{P}(X_1 > u_n(\tau), X_j > u_n(\tau)) = 0$

- Conditions $D(u_n(\tau))$ et $D'(u_n(\tau))$ satisfaites
  $\mathbb{P}(M_n \leq u_n(\tau)) \to \exp(-\tau)$
  $\theta = 1$

- Processus ponctuel de dépassement de seuil

$\mathcal{F}_{p,q} = \mathcal{F}_{p,q}(\tau)$: σ-algèbre générée par $\{X_i > u_n(\tau)\}$, $p \leq i < q$
$\alpha_{n,\ell}(\tau) \equiv \sup |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) : A \in \mathcal{F}_{1,t}, B \in \mathcal{F}_{t+\ell,\infty}, t \geq 1|$
Condition $\Delta(u_n(\tau))$: $\exists \ell_n = O(n)$ tq $\ell_n \xrightarrow[n\to\infty]{} \infty$
$\alpha_{n,\ell_n}(\tau) \xrightarrow[n\to\infty]{} 0$

$N_n^{(\tau)}(B) = \sum_{i=1}^{n} \mathbb{1}_{\{i/n \in B, X_i > u_n(\tau)\}}$ $\forall B \subset E := (0,1]$

- Condition $\Delta(u_n(\tau))$ satisfaite pour $(u_n(\tau))$ tq $n\bar{F}(u_n(\tau)) \xrightarrow[n\to\infty]{} \tau$
  Si $N_n^{(\tau)}$ existe c'est un processus de Poisson composé homogène d'intensité $\tau\theta|\cdot|$ et de distribution des clusters π.
  $\pi_n(m; q_n, u_n(\tau)) = \mathbb{P}(N_n^{(\tau)}((0; q_n/n]) = m \mid N_n^{(\tau)}((0; q_n/n]) > 0)$

- θ existe
  Condition nécessaire et suffisante de convergence de $N_n^{(\tau)}$:
  $\pi_n(m; q_n, u_n(\tau)) \xrightarrow[n\to\infty]{} \pi(m)$
  (q_n) tq $\exists (\ell_n)$ tq $\ell_n = O(q_n)$, $q_n = O(n)$ et $n q_n^{-1} \alpha_{n,\ell_n}(\tau) \to 0$

  - si condition $\Delta(u_n(\tau))$ satisfaite $\forall \tau > 0$
    $\to \theta$ et $\pi \perp\!\!\!\perp \tau$
    $\to \theta^{-1} = \sum_{k=1}^{\infty} k\pi(k)$

2.2 Lois limites des maxima des composantes d'un vecteur

$(X_i = (X_{i,1}, \dots, X_{i,m}))$ suite de vecteurs aléatoires iid de distribution (multivariée) F

$M_{j,n} = \max(X_{1,j}, \dots, X_{n,j})$: maxima par composante

$$\mathbb{P}(M_{j,n} \leq c_{j,n} x_j + d_{j,n},\ j = 1, \dots, n) = G(x_1, \dots, x_m)$$

$\downarrow$

formes semi-paramétriques

Distributions marginales de G = Fréchet unitaires $\Phi_1(x) = \exp(-x^{-1})$

Représentation de Pickands :

$$G(x_1, \dots, x_m) = \exp\left\{-\int_{S_m} \max_{1 \leq j \leq m} (w_j\, x_j^{-1})\, \mu(dw)\right\}$$

μ mesure finie positive sur $S_m = \{y_j \geq 0,\ j = 1, \dots, m : \sum_{j=1}^{m} y_j = 1\}$
tq $\int_{S_m} w_j\, \mu(dw) = 1$, $j = 1, \dots, m$

Ex $m = 2$

- $\mu(0) = \mu(1) = 1$

$$G(x_1, x_2) = \exp(-x_1^{-1}) \exp(-x_2^{-1}) \rightarrow \text{indépendance}$$

- $\mu(0{,}5) = 2$

$$G(x_1, x_2) = \exp\{-\max(x_1^{-1}, x_2^{-1})\} \rightarrow \text{parfaite dependance}$$

2.3 Caractérisation de la dépendance extrêmale dans le cas bivarié

- proximité dans l'espace
- proximité dans le temps
- covariables communes

Distributions marginales et structure de dépendance

Comparaison de 2 va avec dist. marginales différentes difficile
⇒ transformation des marginales

- Standardisation des distributions marginales et copules

Copule : FdR dist. bivariées ayant des distributions de loi $\mathcal{U}([0,1])$ pour ces dist. marginales

$U = F_X(X) \sim \mathcal{U}([0,1])$
$V = F_Y(Y) \sim \mathcal{U}([0,1])$

$$F_{X,Y}(x,y) = C(u,v) = C(F_X(x), F_Y(y))$$

Standardiser des distributions marginales continues F_X et F_Y en des dist. continues G_X et G_Y :

$$X^* = G_X^{-1}(F_X(X)) \qquad Y^* = G_Y^{-1}(F_Y(Y))$$

Courant de transformer les marginales pour avoir des Φ_1

Comportements limite de dépendance

$(X,Y) \sim F$ avec $F_X \overset{\mathcal{L}}{=} F_Y \overset{\mathcal{L}}{=} \Phi_1$

- X et Y asymptotiquement indépendants :
  $$\mathbb{P}(Y > t \mid X > t) \xrightarrow[t \to \infty]{} 0$$

- X et Y asymptotiquement dépendants :
  $$\mathbb{P}(Y > t \mid X > t) \xrightarrow[t \to \infty]{} c$$

- Coefficient de dépendance de la queue

$$\mathbb{P}(X > t, Y > t) \sim \underbrace{L(t)}_{\text{variations lentes}} \left[\mathbb{P}(X > t)\right]^{1/\eta}$$ ← coef de dependance de la queue $\in [0,1]$

- indépendance asymptotique : $0 < \eta < 1$
- independance : $\eta = 1/2$ et $L(t) = 1$
- parfaite dépendance : $\eta = 1$ et $L(t) = 1$
- dépendance asymptotique : $\eta = 1$ et $L(t) \xrightarrow[t\to\infty]{} c > 0$

(1) - $1/2 < \eta < 1$ et $L(t) \xrightarrow[t\to\infty]{} c > 0$
$\eta = 1$ et $L(t) \xrightarrow[t\to\infty]{} 0$

| obs pour lesquelles des évènements du type $\{X > t\}$ et $\{Y > t\}$ pour des grandes valeurs de t se produise plus souvent que sous l'independance exacte

(2) $\eta = 1/2$: presque indépendants

(3) $0 < \eta < 1/2$: ... moins souvent ...

<u>Ex</u> distribution logistique bivariée, marges de type Fréchet unitaire (Φ_1)

$$\mathbb{P}(X \leq x, Y \leq y) = \exp\left\{-\left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha}\right\} \qquad 0 < \alpha < 1$$

Pour identifier η, développement limité de la fonction de survie jointe :

$$\mathbb{P}(X > t, Y > t) = 1 - \left[\mathbb{P}(X < t) + \mathbb{P}(Y < t) - \mathbb{P}(X \leq t, Y \leq t)\right]$$

$$\mathbb{P}(X \leq x) = \lim_{y \to \infty} \mathbb{P}(X \leq x, Y \leq y) = \exp\left\{-\left(x^{-1/\alpha}\right)^{\alpha}\right\} = \exp\left\{-x^{-1}\right\} = \Phi_1(x)$$

$$\mathbb{P}(X > t, Y > t) = 1 - \left[2\exp\{-t^{-1}\} - \exp\{\underbrace{-\left(2t^{-1/\alpha}\right)^{\alpha}}_{-2^{\alpha}t^{-1}}\}\right]$$

$e^{u} \xrightarrow[u \to 0]{} 1 + u$ $\quad \underbrace{1-(1-t^{-1}}$

$$= 1 - \left[2(1 - t^{-1}) - (1 - 2^{\alpha}t^{-1})\right] = 1 - \left[2 - 2t^{-1} - 1 + 2^{\alpha}t^{-1}\right] = (2 - 2^{\alpha})t^{-1}$$

$$= (2 - 2^{\alpha})\left[1 - \exp\{-t^{-1}\}\right] = (2 - 2^{\alpha})\,\mathbb{P}(X > t)$$

$\Rightarrow \eta = 1$ et $L(t) = (2 - 2^{\alpha}) \Rightarrow$ dépendance asymptotique

- Les mesures de dépendance χ et $\bar{\chi}$

$$\mathbb{P}(V>u \mid U>u) = 2 - \frac{1-\mathbb{P}(U<u,V<u)}{1-\mathbb{P}(U<u)} \underset{u\to 1}{\sim} 2 - \frac{\log C(u,u)}{\log u}$$

$$\mathbb{P}(V>u \mid U>u) = \frac{\mathbb{P}(V>u,U>u)}{\mathbb{P}(U>u)} = \frac{1-[\mathbb{P}(U<u)+\mathbb{P}(V<u)-\mathbb{P}(U<u,V<u)]}{1-\mathbb{P}(U<u)}$$

$$= \frac{2-1-2\mathbb{P}(U<u)+\mathbb{P}(U<u,V<u)}{1-\mathbb{P}(U<u)}$$

$$= 2 - \frac{1-\mathbb{P}(U<u,V<u)}{1-\mathbb{P}(U<u)}$$

$$= 2 - \frac{1-C(u,u)}{1-\mathbb{P}(U<u)} \xrightarrow[u\to 1]{} 2 - \frac{\log C(u,u)}{\log u}$$

car $C(u,u) \underset{u\to 1}{\longrightarrow} 1$ et $\log u \underset{u\to 1}{\sim} 1-u$

$\mathbb{P}(U<u) \underset{u\to 1}{\longrightarrow} 1 = u$

$$\chi(u) = 2 - \frac{\log C(u,u)}{\log u} \qquad 0 \leq u \leq 1$$

- signe $(\chi(u))$ détermine si les variables sont positivement ou négativement associées pour le quantile de niveau u
- $\underbrace{2-\log(2u-1)/\log(u)}_{\substack{-\infty \text{ si } u \leq 1/2 \\ 0 \text{ si } u=1}} \leq \chi(u) < 1$

$\chi = \lim_{u\to 1} \chi(u)$: probabilité que l'une des variables est extrême sachant que l'autre l'est

$\chi = 0$: variables asymptotiquement indépendantes

$$\bar{C}(u,v) = 1 - u - v + C(u,v)$$

$$\bar{\chi}(u) = \frac{2 \log \mathbb{P}(U>u)}{\log \mathbb{P}(U>u, V>u)} - 1 = \frac{2 \log (1-u)}{\log \bar{C}(u,u)} - 1 \quad 0 \leq u \leq 1$$

$$-1 \leq \bar{\chi}(u) \leq 1 \quad \forall\, 0 \leq u \leq 1$$

$$\bar{\chi} = \lim_{u \to 1} \bar{\chi}(u) \qquad -1 < \bar{\chi} < 1$$

$$\bar{\chi} = 2\eta - 1$$

$$\chi = \begin{cases} c & \bar{\chi} = 1 \quad L(t) \xrightarrow[t \to +\infty]{} c > 0 \\ 0 & \bar{\chi} = 1 \quad L(t) \xrightarrow[t \to \infty]{} 0 \\ 0 & \bar{\chi} < 1 \end{cases}$$

→ $\chi \in [0,1]$: $(0,1]$ → dependance asymptotique
→ $\bar{\chi} \in [-1,1]$: $[-1,1)$ → independance asymptotique

→ $(\chi > 0 \,;\, \bar{\chi} = 1)$: dépendance asymptotique d'intensité mesurée par χ
→ $(\chi = 0 \,;\, \bar{\chi} < 1)$: indépendance asymptotique

En pratique, on regarde $\bar{\chi}$:
→ $\bar{\chi} < 1$ → indépendance asymptotique
→ $\bar{\chi} = 1$ ⇒ dependance asymptotique et on regarde χ.