Mélange de distributions des valeurs extrêmes généralisées

Pascal Alain Dkengne Sielenou

Travail en collaboration avec Stéphane Girard (INRIA)

Réunion de synchronisation du 12 janvier 2024

Rappels : Loi des valeurs extremes généralisées

- Soit X une variable aléatoire de fonction de répartion F.
- Soient $X_1, \ldots, X_n$ des copies indépendantes de X.
- Soit $M_n = \max\{X_1, \ldots, X_n\}$, la suite des maximums de X.

Theorem

S'il existent des suites de constantes $a_n > 0$ et $b_n \in \mathbb{R}$ telles que

$$\lim_{n \to +\infty} \mathbb{P}\left\{\frac{M_n - b_n}{a_n} \leq x\right\} = \lim_{n \to +\infty} F^n\left(a_n x + b_n\right) = G(x) \tag{1}$$

pour une fonction non dégénérée G, alors G appartient à la famille des lois des valeurs extrêmes généralisées (GEV)

$$G(x) = G(x; \gamma, \sigma, \mu) = \exp\left\{-\left[1 + \gamma\left(\frac{x - \mu}{\sigma}\right)\right]^{-\frac{1}{\gamma}}\right\}, \tag{2}$$

définies sur $\left\{x \in \mathbb{R} :\ 1 + \gamma\left(\frac{x-\mu}{\sigma}\right) > 0\right\}$, où $\gamma \neq 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.

Si F vérifie (1), on dit qu'il appartient au domaine d'attraction de la function G.

Résultats

Definition (Melange des distributions de probabilités)

Une loi de probabilités est dite **loi de mélange** si sa fonction de répartition est une **moyenne pondérée algébrique**, une **moyenne pondérée géométrique** ou une **moyenne pondérée harmonique** de plusieurs fonctions de répartition.

Example

Considérons une suite de p fonctions de répartition F_j, $j = 1, \cdots, p$ et un vecteur $\omega = (\omega_1, \cdots, \omega_p) \in [0, 1]^p$ tel que $\sum_{j=1}^{p} \omega_j = 1$. Les lois de mélange moyennes pondérées algébriquement, géométriquement et harmoniquement ont respectivement les formes (3) et (4) ci-dessous

$$F_{\mathbb{S}}(x;\omega) = \sum_{j=1}^{p} \omega_j F_j(x), \tag{3}$$

$$F_{\mathbb{P}}(x;\omega) = \prod_{j=1}^{p} F_j^{\omega_j}(x), \tag{4}$$

$$F_{\mathbb{H}}(x;\omega) = \left(\sum_{j=1}^{p} \frac{\omega_j}{F_j(x)} \right)^{-1}. \tag{5}$$

Résultats

Definition (Mélange de distributions GEV)

- Désignons par $\omega = (\omega_1, \cdots, \omega_p) \in [0, 1]^p$ un vecteur tel que $\sum_{j=1}^{p} \omega_j = 1$.
- Désignons par G_j une fonction de répartition de la loi GEV par (2).
- Désignons par $\Theta = (\Theta_j,\, j = 1, \cdots, p)$ où $\Theta_j = (\gamma_j, \sigma_j, \mu_j)$ un vecteur des paramètres de la distribution GEV nommée G_j.

On définit les modèles de mélange $G_{\mathbb{S}}$, $G_{\mathbb{P}}$ et $G_{\mathbb{H}}$ des lois GEV nommées G_j par

$$\begin{aligned} G_{\mathbb{S}}(x;\omega,\Theta) &= \sum_{j=1}^{p} \omega_j G_j(x;\Theta_j), \\ G_{\mathbb{P}}(x;\omega,\Theta) &= \prod_{j=1}^{p} G_j^{\omega_j}(x;\Theta_j), \\ G_{\mathbb{H}}(x;\omega,\Theta) &= \left(\sum_{j=1}^{p} \frac{\omega_j}{G_j(x;\Theta_j)} \right)^{-1}. \end{aligned} \tag{6}$$

Résultats

Definition

Considérons une suite de p fonctions de répartition G_j de la loi GEV définie par

$$G_j(x) = G(x; \gamma_j, \sigma_j, \mu_j) = \exp\left\{-\left[1 + \gamma_j\left(\frac{x-\mu_j}{\sigma_j}\right)\right]^{-\frac{1}{\gamma_j}}\right\}, \tag{7}$$

sur l'ensemble $\left\{x \in \mathbb{R} : \ 1 + \gamma_j\left(\frac{x-\mu_j}{\sigma_j}\right) > 0\right\}$, où $\gamma_j \neq 0$, $\mu_j \in \mathbb{R}$, $\sigma_j > 0$.

On définit les modèles de mélange $G_{\min}$, $G_{\max}$, $G_{\inf}$, et $G_{\sup}$ par

$$G_{\min}(x) = \min_{1\leq j\leq p} G_j(x), \tag{8}$$

$$G_{\max}(x) = \max_{1\leq j\leq p} G_j(x), \tag{9}$$

$$G_{\inf}(x) = G\left(x; \max_{1\leq j\leq p}\gamma_j, \max_{1\leq j\leq p}\sigma_j, \max_{1\leq j\leq p}\mu_j\right), \tag{10}$$

$$G_{\sup}(x) = G\left(x; \min_{1\leq j\leq p}\gamma_j, \min_{1\leq j\leq p}\sigma_j, \min_{1\leq j\leq p}\mu_j\right). \tag{11}$$

Propriétés 1

Theorem (Inégalités entre les modèles des distributions GEV)

Les inégalités suivantes sont satisfaites

$$G_{\text{inf}}(x) \leq G_{\text{min}}(x) \leq G_{\mathbb{H}}(x;\omega,\Theta) \leq G_{\mathbb{P}}(x;\omega,\Theta) \leq G_{\mathbb{S}}(x;\omega,\Theta) \leq G_{\text{max}}(x) \leq G_{\text{sup}}(x). \quad (12)$$

La preuve de résultats exploite les relations entre les moyennes arithmétiques, géométriques et harmoniques disponibles sur la page web suivante
https://en.wikipedia.org/wiki/QM-AM-GM-HM_Inequalities

Propriétés 2

Theorem (Stabilité de la famille $\{G_{\mathbb{P}}(\cdot;\omega,\Theta)\}$)

Pour tout entier positif m et pour tout réel x, la propriété suivante est satisfaite

$$G_{\mathbb{P}}^{m}(x;\omega,\Theta) = \prod_{j=1}^{p} \left[G_j(x;\Theta_j(m))\right]^{\omega_j} = G_{\mathbb{P}}(x;\omega,\Theta(m)). \tag{13}$$

Ici, $\Theta(m) = (\Theta_j(m), j = 1, \cdots, p)$ *où* $\Theta_j(m) = (\gamma_j(m), \sigma_j(m), \mu_j(m))$
avec $\gamma_j(m) = \gamma_j, \quad \sigma_j(m) = \sigma_j\, m^{\gamma_j}, \quad \mu_j(m) = \mu_j + \sigma_j \left(\frac{m^{\gamma_j} - 1}{\gamma_j}\right).$

La propriété (13) montre que si la loi d'une v.a. X appartient à la famille des lois de probabilités $\{G_{\mathbb{P}}(\cdot;\omega,\Theta)\}$, alors la loi du maximum de m copies indépendantes de X appartient également à cette même famille des lois de probabilités.

Propriétés 3

Theorem (Equivalence des queues de distributions)

On considère les distributions de probabilités $F_{\mathbb{S}}$ *et* $F_{\mathbb{P}}$ *définies resp. par (3) et (4). S'ils existent des constantes* $\tau_j > 0$ *telles que pour tout* $j = 1, \cdots, p$ *nous avons*

$$\lim_{x \to x_{F_j}} \frac{1 - F_j(x)}{1 - G_j(x; \Theta_j)} = \tau_j, \tag{14}$$

où $x_{F_j} = \sup\{x \in \mathbb{R} : \; F_j(x) < 1\}$ *est la borne supérieure du support de la distribution* F_j*, alors les limites suivantes sont satisfaites*

$$\lim_{x \to x^\star} \frac{1 - F_{\mathbb{H}}(x;\omega)}{1 - G_{\mathbb{P}}(x;\omega,\Theta(\tau))} = 1, \; \lim_{x \to x^\star} \frac{1 - F_{\mathbb{S}}(x;\omega)}{1 - G_{\mathbb{P}}(x;\omega,\Theta(\tau))} = 1, \; \lim_{x \to x^\star} \frac{1 - F_{\mathbb{P}}(x;\omega)}{1 - G_{\mathbb{P}}(x;\omega,\Theta(\tau))} = 1, \tag{15}$$

où $x^\star = \max\{x_{F_j}, j = 1, \cdots, p\}$ *et* $\tau = (\tau_1, \cdots, \tau_p)$.

Ici, $\Theta(\tau) = (\Theta_j(\tau_j), j = 1, \cdots, p)$ *où* $\Theta_j(\tau_j) = (\gamma_j(\tau_j), \sigma_j(\tau_j), \mu_j(\tau_j))$

avec $\gamma_j(\tau_j) = \gamma_j, \quad \sigma_j(\tau_j) = \sigma_j \, \tau_j^{\gamma_j}, \quad \mu_j(\tau_j) = \mu_j + \sigma_j \left(\frac{\tau_j^{\gamma_j} - 1}{\gamma_j} \right).$

Preuve de la Propriété 3 (1/3)

Quand x tend vers $x^\star$, on peut effectuer les équivalents suivants

$$
\begin{aligned}
1 - F_{\mathbb{S}}(x) &= 1 - \sum_{j=1}^{p} \omega_j F_j(x) = \sum_{j=1}^{p} \omega_j \left(1 - F_j(x)\right) \\
&\sim \sum_{j=1}^{p} \omega_j \tau_j \left(1 - G_j(x; \Theta_j)\right), \quad \text{(faisant usage de (14))} \\
&\sim -\sum_{j=1}^{p} \omega_j \log G_j^{\tau_j}(x; \Theta_j) = -\sum_{j=1}^{p} \log G_j^{\omega_j}(x; \Theta_j(\tau_j)) \\
&= -\log \prod_{j=1}^{p} G_j^{\omega_j}(x; \Theta_j(\tau_j)) \\
&\sim 1 - \prod_{j=1}^{p} G_j^{\omega_j}(x; \Theta_j(\tau_j)) \\
&= 1 - G_{\mathbb{P}}(x; \omega, \Theta(\tau)).
\end{aligned}
$$

Preuve de la Propriété 3 (2/3)

Quand x tend vers $x^\star$, on peut effectuer les équivalents suivants

$$\begin{aligned}
1 - F_{\mathbb{P}}(x) &\sim -\log F_{\mathbb{P}}(x) = -\log \prod_{j=1}^{p} F_j^{\omega_j}(x) \\
&= \sum_{j=1}^{p} \omega_j \left(-\log F_j(x)\right) \sim \sum_{j=1}^{p} \omega_j \left(1 - F_j(x)\right) \\
&\sim \sum_{j=1}^{p} \omega_j \tau_j \left(1 - G_j(x;\Theta_j)\right), \quad \text{(faisant usage de (14))} \\
&\sim -\sum_{j=1}^{p} \omega_j \log G_j^{\tau_j}(x;\Theta_j) = -\sum_{j=1}^{p} \log G_j^{\omega_j}(x;\Theta_j(\tau_j)) \\
&= -\log \prod_{j=1}^{p} G_j^{\omega_j}(x;\Theta_j(\tau_j)) \sim 1 - \prod_{j=1}^{p} G_j^{\omega_j}(x;\Theta_j(\tau_j)) \\
&= 1 - G_{\mathbb{P}}(x;\omega,\Theta(\tau)).
\end{aligned}$$

Preuve de la Propriété 3 (3/3)

Quand x tend vers $x^\star$, on peut effectuer les équivalents suivants

$$
\begin{aligned}
1 - F_{\mathbb{H}}(x) &\sim -\log F_{\mathbb{H}}(x) = \log \sum_{j=1}^{p} \frac{\omega_j}{F_j(x)} \sim \sum_{j=1}^{p} \omega_j \left(\frac{1}{F_j(x)} - 1 \right) \\
&\sim \sum_{j=1}^{p} \omega_j \left(-\log F_j(x)\right) \sim \sum_{j=1}^{p} \omega_j \left(1 - F_j(x)\right) \\
&\sim \sum_{j=1}^{p} \omega_j \tau_j \left(1 - G_j(x;\Theta_j)\right), \quad \text{(faisant usage de (14))} \\
&\sim -\sum_{j=1}^{p} \omega_j \log G_j^{\tau_j}(x;\Theta_j) = -\sum_{j=1}^{p} \log G_j^{\omega_j}(x;\Theta_j(\tau_j)) \\
&= -\log \prod_{j=1}^{p} G_j^{\omega_j}(x;\Theta_j(\tau_j)) \sim 1 - \prod_{j=1}^{p} G_j^{\omega_j}(x;\Theta_j(\tau_j)) \\
&= 1 - G_{\mathbb{P}}(x;\omega,\Theta(\tau)).
\end{aligned}
$$

Propriétés 4

Theorem

- *On considère une suite de* p *fonctions de répartition* F_j, $j = 1, \cdots, p$.
- *On considère un vecteur* $\omega = (\omega_1, \cdots, \omega_p) \in [0, 1]^p$ *tel que* $\sum_{j=1}^{p} \omega_j = 1$.
- *On désigne par* $x^\star = \max\left\{x_{F_1}, x_{F_2}, \cdots, x_{F_p}\right\}$, *où* $x_{F_j} \leq +\infty$ *la borne supérieure du support de la distribution* F_j.

Alors, les lois de mélange moyennes pondérées ci-dessous

$$F_{\mathbb{S}}(x;\omega) = \sum_{j=1}^{p} \omega_j F_j(x), \quad F_{\mathbb{P}}(x;\omega) = \prod_{j=1}^{p} F_j^{\omega_j}(x), \quad F_{\mathbb{H}}(x;\omega) = \left(\sum_{j=1}^{p} \frac{\omega_j}{F_j(x)}\right)^{-1}$$

satisfont la propriété suivante :

$$\lim_{x \to x^\star} F_{\mathbb{S}}(x;\omega) = \lim_{x \to x^\star} F_{\mathbb{P}}(x;\omega) = \lim_{x \to x^\star} F_{\mathbb{H}}(x;\omega). \tag{16}$$

Preuve de la Propriété 4 (1/2)

$$
\begin{aligned}
\lim_{x\to x^\star} F_{\mathbb{S}}(x;\omega) &= \lim_{x\to x^\star} \sum_{j=1}^{p} \omega_j F_j(x) = \lim_{x\to x^\star} \exp\left\{\log\left(\sum_{j=1}^{p} \omega_j F_j(x)\right)\right\} \\
&= \lim_{x\to x^\star} \exp\left\{\sum_{j=1}^{p} \omega_j F_j(x) - 1\right\} = \lim_{x\to x^\star} \exp\left\{\sum_{j=1}^{p} \omega_j \left(F_j(x) - 1\right)\right\} \\
&= \lim_{x\to x^\star} \exp\left\{\sum_{j=1}^{p} \omega_j \log F_j(x)\right\} = \lim_{x\to x^\star} \exp\left\{\log\left(\prod_{j=1}^{p} F_j^{\omega_j}(x)\right)\right\} \\
&= \lim_{x\to x^\star} \prod_{j=1}^{p} F_j^{\omega_j}(x) = \lim_{x\to x^\star} F_{\mathbb{P}}(x;\omega).
\end{aligned}
$$

Preuve de la Propriété 4 (2/2)

$$
\begin{aligned}
\lim_{x\to x^\star} F_{\mathbb{H}}(x;\omega) &= \lim_{x\to x^\star}\left(\sum_{j=1}^{p}\frac{\omega_j}{F_j(x)}\right)^{-1} = \lim_{x\to x^\star}\exp\left\{-\log\left(\sum_{j=1}^{p}\frac{\omega_j}{F_j(x)}\right)\right\}\\
&= \lim_{x\to x^\star}\exp\left\{1-\sum_{j=1}^{p}\frac{\omega_j}{F_j(x)}\right\} = \lim_{x\to x^\star}\exp\left\{\sum_{j=1}^{p}\omega_j\left(1-\frac{1}{F_j(x)}\right)\right\}\\
&= \lim_{x\to x^\star}\exp\left\{\sum_{j=1}^{p}\omega_j\log F_j(x)\right\} = \lim_{x\to x^\star}\exp\left\{\log\left(\prod_{j=1}^{p}F_j^{\omega_j}(x)\right)\right\}\\
&= \lim_{x\to x^\star}\prod_{j=1}^{p}F_j^{\omega_j}(x) = \lim_{x\to x^\star}F_{\mathbb{P}}(x;\omega).
\end{aligned}
$$

Propriétés 5

Theorem

On considère les distributions de probabilités $F_{\mathbb{S}}$ *et* $F_{\mathbb{P}}$ *définies resp. par (3) et (4). S'ils existent des suites* $a_{n,j} > 0$ *et* $b_{n,j} \in \mathbb{R}$ *telles que pour tout* $x \in \mathbb{R}$, *nous avons*

$$\lim_{n\to+\infty} n\left[1 - F_j\left(a_{n,j}x + b_{n,j}\right)\right] = -\log G_j(x) \Leftrightarrow \lim_{n\to+\infty} F_j^n\left(a_{n,j}x + b_{n,j}\right) = G_j(x), \tag{17}$$

alors les limites suivantes sont satisfaites pour tout $x \in \mathbb{R}$:

$$\lim_{n\to+\infty} F_{\mathbb{P}}^n\left(v_n(x)\right) \leq G_{\mathbb{P}}(x) \leq \lim_{n\to+\infty} F_{\mathbb{P}}^n\left(u_n(x)\right), \tag{18}$$

$$\lim_{n\to+\infty} F_{\mathbb{S}}^n\left(v_n(x)\right) \leq G_{\mathbb{P}}(x) \leq \lim_{n\to+\infty} F_{\mathbb{S}}^n\left(u_n(x)\right), \quad \lim_{n\to+\infty} F_{\mathbb{H}}^n\left(v_n(x)\right) \leq G_{\mathbb{P}}(x) \leq \lim_{n\to+\infty} F_{\mathbb{H}}^n\left(u_n(x)\right),$$

où les suites $u_n(x)$ *et* $v_n(x)$ *sont définies par*

$$u_n(x) = \max_{1\leq j\leq p}\left\{a_{n,j}x + b_{n,j}\right\}, \quad v_n(x) = \min_{1\leq j\leq p}\left\{a_{n,j}x + b_{n,j}\right\}. \tag{19}$$

De plus, il existe trois suites non linéaires $p_n(x), s_n(x), h_n(x) \in [v_n(x), u_n(x)]$ *qui sont strictement croissantes et satisfont*

$$\lim_{n\to+\infty} F_{\mathbb{P}}^n\left(p_n(x)\right) = G_{\mathbb{P}}(x) = \lim_{n\to+\infty} F_{\mathbb{S}}^n\left(s_n(x)\right) = \lim_{n\to+\infty} F_{\mathbb{H}}^n\left(h_n(x)\right). \tag{20}$$

Preuve de la Propriété 5

$$\lim_{n\to+\infty} F^n_{\mathbb{P}}(v_n(x);\omega) = \lim_{n\to+\infty}\left(\prod_{j=1}^{p} F_j^{\omega_j}(v_n(x))\right)^n \leq \lim_{n\to+\infty}\prod_{j=1}^{p}\left[F_j^n\left(a_{n,j}x+b_{n,j}\right)\right]^{\omega_j}$$
$$= \prod_{j=1}^{p} G_j^{\omega_j}(x), \quad \text{(faisant usage de (17)).} \tag{21}$$

$$\lim_{n\to+\infty} F^n_{\mathbb{P}}(u_n(x);\omega) = \lim_{n\to+\infty}\left(\prod_{j=1}^{p} F_j^{\omega_j}(u_n(x))\right)^n \geq \lim_{n\to+\infty}\prod_{j=1}^{p}\left[F_j^n\left(a_{n,j}x+b_{n,j}\right)\right]^{\omega_j}$$
$$= \prod_{j=1}^{p} G_j^{\omega_j}(x), \quad \text{(faisant usage de (17)).} \tag{22}$$

Faisant usage des égalités (16) et des inégalités (21-22), on peut écrire

$$\lim_{n\to+\infty} F^n_{\mathbb{S}}\left(v_n(x);\omega\right) = \lim_{n\to+\infty} F^n_{\mathbb{H}}\left(v_n(x);\omega\right) = \lim_{n\to+\infty} F^n_{\mathbb{P}}\left(v_n(x);\omega\right) \leq G_{\mathbb{P}}(x;\omega),$$

$$\lim_{n\to+\infty} F^n_{\mathbb{S}}\left(u_n(x);\omega\right) = \lim_{n\to+\infty} F^n_{\mathbb{H}}\left(u_n(x);\omega\right) = \lim_{n\to+\infty} F^n_{\mathbb{P}}\left(u_n(x);\omega\right) \geq G_{\mathbb{P}}(x;\omega).$$

Propriétés 6

Theorem

- *On considère une suite de p fonctions de répartition F_j, $j = 1, \cdots, p$.*
- *On considère un vecteur $\omega = (\omega_1, \cdots, \omega_p) \in [0, 1]^p$ tel que $\sum_{j=1}^{p} \omega_j = 1$.*

La variable aléatoire X associée à la fonction de répartition $F_{\mathbb{P}}(x) = \prod_{j=1}^{p} F_j^{\omega_j}(x)$ est définie par

$$X = \max_{1 \leq j \leq p} \left\{ F_j^{-1} \left(\exp\{-Y_j\} \right) \right\}, \tag{23}$$

où Y_j est une variable aléatoire qui suit la loi exponentielle de paramètre ω_j.

Simulation

La formule (23) suggère les deux étapes suivantes pour simuler une observation x de la distribution $F_{\mathbb{P}}$.

1. Générer les valeurs indépendantes y_j, $j = 1, 2, \cdots, p$ à l'aide de la loi exponentielle de paramètre ω_j.
2. Prendre comme valeur simulée de la distribution $F_{\mathbb{P}}$ la quantité x définie par

$$x = \max_{1 \leq j \leq p} \left\{ F_j^{-1} \left(\exp\{-y_j\} \right) \right\}. \tag{24}$$

Preuve de la Propriété 6

Soit X_j la variable aléatoire associée à la distribution $F_j^{\omega_j}$.

La variable aléatoire $U_j = F_j^{\omega_j}(X_j)$ suit la loi uniforme sur l'intervalle $[0, 1]$.

Ainsi $X_j = F_j^{-1}\left(U_j^{1/\omega_j}\right) = F_j^{-1}\left(\exp\left\{\frac{1}{\omega_j}\log U_j\right\}\right)$.

La variable aléatoire $Y_j = -\frac{1}{\omega_j}\log U_j$ suit la loi exponentielle de paramètre ω_j.

En effet pour tout réel $y > 0$, on a :

$$\begin{aligned}
\mathbb{P}\{Y_j \leq y\} &= \mathbb{P}\left\{-\frac{1}{\omega_j}\log U_j \leq y\right\} \\
&= \mathbb{P}\{U_j > \exp\{-\omega_j y\}\} \\
&= 1 - \exp\{-\omega_j y\}.
\end{aligned}$$

Pour tout $x \in \mathbb{R}$, on a :

$$\begin{aligned}
\mathbb{P}\{X \leq x\} &= \prod_{j=1}^{p} \mathbb{P}\{X_j \leq x\} \\
&= \mathbb{P}\{X_1 \leq x, \cdots, X_p \leq x\} \\
&= \mathbb{P}\{\max\{X_1, \cdots, X_p\} \leq x\}.
\end{aligned}$$

Propriétés 7

Theorem

Soit $\omega = (\omega_1, \cdots, \omega_p) \in [0,1]^p$ *un vecteur de poids, c'est-à-dire* $\sum_{j=1}^p \omega_j = 1$.
On considère les distributions de probabilités $F_{\mathbb{S}}$, $F_{\mathbb{H}}$ *et* $F_{\mathbb{P}}$ *définies par*

$$F_{\mathbb{S}}(x;\omega) = \sum_{j=1}^p \omega_j F_j(x), \quad F_{\mathbb{H}}(x;\omega) = \left(\sum_{j=1}^p \frac{\omega_j}{F_j(x)}\right)^{-1}, \quad F_{\mathbb{P}}(x;\omega) = \prod_{j=1}^p F_j^{\omega_j}(x).$$

Si $x_j(\alpha)$ *désigne le quantile d'ordre* $\alpha \in [0,1]$ *de la distribution* F_j, *i.e.* $F_j(x_j(\alpha)) = 1-\alpha$, *alors les quantités* $x_1(\alpha) = \min_{1\le j\le p}\{x_j(\alpha)\}$ $x_2(\alpha) = \max_{1\le j\le p}\{x_j(\alpha)\}$ *satisfont pour tout vecteur de poids* ω *les inégalités suivantes :*

- $F_{\mathbb{S}}(x_1(\alpha);\omega) \le 1-\alpha \le F_{\mathbb{S}}(x_2(\alpha);\omega)$,
- $F_{\mathbb{H}}(x_1(\alpha);\omega) \le 1-\alpha \le F_{\mathbb{H}}(x_2(\alpha);\omega)$,
- $F_{\mathbb{P}}(x_1(\alpha);\omega) \le 1-\alpha \le F_{\mathbb{P}}(x_2(\alpha);\omega)$.

Preuve de la Propriété 7 (1/2)

Par hypothèse, pour tout $j = 1, \cdots, p$ on a les deux égalités suivantes vraies

$$F_j^{\omega_j}(x_j(\alpha)) = (1-\alpha)^{\omega_j}, \quad \omega_j F_j(x_j(\alpha)) = \omega_j(1-\alpha), \quad \frac{\omega_j}{F_j(x_j(\alpha))} = \frac{\omega_j}{1-\alpha}.$$

Par définition, pour tout $j = 1, \cdots, p$ on a les inégalités suivantes :

$$F_j\left(\max_{1\leq k\leq p}\{x_k(\alpha)\}\right) \geq F_j(x_j(\alpha)), \quad F_j\left(\min_{1\leq k\leq p}\{x_k(\alpha)\}\right) \leq F_j(x_j(\alpha)).$$

Ce qui implique

$$F_j^{\omega_j}(x_2(\alpha)) \geq (1-\alpha)^{\omega_j}, \quad \omega_j F_j(x_2(\alpha)) \geq \omega_j(1-\alpha), \quad \frac{\omega_j}{F_j(x_2(\alpha))} \leq \frac{\omega_j}{1-\alpha}.$$

$$F_j^{\omega_j}(x_1(\alpha)) \leq (1-\alpha)^{\omega_j}, \quad \omega_j F_j(x_1(\alpha)) \leq \omega_j(1-\alpha), \quad \frac{\omega_j}{F_j(x_1(\alpha))} \geq \frac{\omega_j}{1-\alpha}.$$

Preuve de la Propriété 7 (2/2)

$$F_{\mathbb{P}}(x_2(\alpha)) = \prod_{j=1}^{p} F_j^{\omega_j}(x_2(\alpha)) \geq \prod_{j=1}^{p}(1-\alpha)^{\omega_j} = (1-\alpha)^{\sum_{j=1}^{p}\omega_j} = 1-\alpha.$$

$$F_{\mathbb{P}}(x_1(\alpha)) = \prod_{j=1}^{p} F_j^{\omega_j}(x_1(\alpha)) \leq \prod_{j=1}^{p}(1-\alpha)^{\omega_j} = (1-\alpha)^{\sum_{j=1}^{p}\omega_j} = 1-\alpha.$$

$$F_{\mathbb{S}}(x_2(\alpha)) = \sum_{j=1}^{p} \omega_j F_j(x_2(\alpha)) \geq (1-\alpha)\sum_{j=1}^{p}\omega_j = 1-\alpha.$$

$$F_{\mathbb{S}}(x_1(\alpha)) = \sum_{j=1}^{p} \omega_j F_j(x_2(\alpha)) \leq (1-\alpha)\sum_{j=1}^{p}\omega_j = 1-\alpha.$$

$$\frac{1}{F_{\mathbb{H}}(x_2(\alpha))} = \sum_{j=1}^{p} \frac{\omega_j}{F_j(x_2(\alpha))} \leq \sum_{j=1}^{p} \frac{\omega_j}{1-\alpha} = \frac{1}{1-\alpha}.$$

$$\frac{1}{F_{\mathbb{H}}(x_1(\alpha))} = \sum_{j=1}^{p} \frac{\omega_j}{F_j(x_1(\alpha))} \geq \sum_{j=1}^{p} \frac{\omega_j}{1-\alpha} = \frac{1}{1-\alpha}.$$

Distribution des extrêmes et mélange des lois GEV

- Soit X une v.a. de fonction de répartition F et de borne supérieure x_F.
- Soient $b_1, \cdots, b_p$ une suite de p entiers positifs suffisamment grands.
- On suppose que pour tout $j = 1, \cdots, p$ et pour toute grande valeur $x \in \mathbb{R}$, l'approximation suivante est satisfaite

$$F(x) = \mathbb{P}\{X \leq x\} \sim G_j^{1/b_j}(x; \Theta_j) = G_j(x; \Theta_j(b_j)), \tag{25}$$

où G_j est une distribution GEV de paramètre $\Theta_j = (\gamma_j, \sigma_j, \mu_j) \in \mathbb{R}^3$.

Alors, quel que soit le vecteur $\omega = (\omega_1, \cdots, \omega_p) \in [0, 1]^p$ tel que $\sum_{j=1}^p \omega_j = 1$ et pour toute grande valeur $x \in \mathbb{R}$, on peut faire les approximations suivantes

$$F(x) = \mathbb{P}\{X \leq x\}^{\sum_{j=1}^p \omega_j} = \prod_{j=1}^p \left(\mathbb{P}\{X \leq x\}\right)^{\omega_j} \sim \prod_{j=1}^p \left[G_j\left(x; \Theta_j(b_j)\right)\right]^{\omega_j} = G_{\mathbb{P}}\left(x; \omega, \Theta(b)\right).$$

$$F(x) = \mathbb{P}\{X \leq x\} \sum_{j=1}^p \omega_j = \sum_{j=1}^p \omega_j \mathbb{P}\{X \leq x\} \sim \sum_{j=1}^p \omega_j G_j\left(x; \Theta_j(b_j)\right) = G_{\mathbb{S}}\left(x; \omega, \Theta(b)\right).$$

$$\frac{1}{F(x)} = \frac{\sum_{j=1}^p \omega_j}{\mathbb{P}\{X \leq x\}} = \sum_{j=1}^p \frac{\omega_j}{\mathbb{P}\{X \leq x\}} \sim \sum_{j=1}^p \frac{\omega_j}{G_j\left(x; \Theta_j(b_j)\right)} = \frac{1}{G_{\mathbb{H}}\left(x; \omega, \Theta(b)\right)}.$$

Estimation du paramètre Θ : Cas IID

Soit $\mathcal{X} = (x_1, \ldots, x_n)$ un échantillon d'une une variable aléatoire X ayant une distribution de probabilité **inconnue**.

- Soit $b = \{b_j \in \mathbb{N}^\star, j = 1, \cdots, p\}$ un ensemble de p tailles de blocs assez grandes.
- Soit $(\gamma_j, \sigma_j, \mu_j)$ le vecteur des paramètres de la loi GEV nommée $G_j(\cdot)$ caractérisant la distribution des maximums de b_j obs. consécutives de la v.a. X.

Nous avons adoptés les notations ci-dessous pour désigner l'ensemble des paramètres qui figurent dans les modèles de mélange qui précèdent

- $b = (b_1, \cdots, b_p)$,
- $\Theta(b) = (\Theta_j(b_j), j = 1, \cdots, p)$,
- $\Theta_j(b_j) = (\gamma_j(b_j), \sigma_j(b_j), \mu_j(b_j))$,
- $\gamma_j(b_j) = \gamma_j, \quad \sigma_j(b_j) = \sigma_j\, b_j^{-\gamma_j}, \quad \mu_j(b_j) = \mu_j + \sigma_j \left(\frac{b_j^{-\gamma_j} - 1}{\gamma_j} \right).$

Estimation du paramètre Θ : Cas Stationnaire

Soit X_t, $t = 1, 2, \cdots$ une série temporelle **stationnaire** de processus aléatoire **inconnu**.

- Soit $b = \{b_j \in \mathbb{N}^\star, j = 1, \cdots, p\}$ un ensemble de p tailles de blocs assez grandes.
- Soit $(\gamma_j, \sigma_j, \mu_j)$ le vecteur des paramètres de la loi GEV nommée $G_j(\cdot)$ caractérisant la distribution des maximums de b_j obs. consécutives de la serie X_t.
- Soit $\nu_j \in [0, 1]$ l'**indice extremal** associé au seuil défini par le quantile empirique $x_{n,1/b_j}$ d'ordre $1/b_j$ estimé sur une séquence $\mathcal{X}_n = \{x_1, \cdots, x_n\}$ de la série X_t.

Indice extremal

Un indice extremal ν quantifie le degré de dépendance entre l'occurrence des valeurs extrêmes consécutives. Cette dépendance est **forte** lorsque ν tend vers 0 et **faible** lorsque ν tend vers 1.

Nous avons adoptés les notations ci-dessous pour désigner l'ensemble des paramètres qui figurent dans les modèles de mélange qui précèdent

- $b = (b_1, \cdots, b_p)$, $\Theta(b) = (\Theta_j(b_j), j = 1, \cdots, p)$, $\Theta_j(b_j) = (\gamma_j(b_j), \sigma_j(b_j), \mu_j(b_j))$,
- $\gamma_j(b_j) = \gamma_j, \quad \sigma_j(b_j) = \sigma_j\,(\nu_j b_j)^{-\gamma_j}, \quad \mu_j(b_j) = \mu_j + \sigma_j\left(\frac{(\nu_j b_j)^{-\gamma_j} - 1}{\gamma_j}\right).$

Estimation du paramètre Θ : Cas Non-Stationnaire (1/2)

Soit $X_t,\ t = 1, 2, \cdots$ une série temporelle **non-stationnaire** de processus **inconnu**.
Soit $Y_t = (Y_{1,t}, \cdots, Y_{q,t})$ une série temporelle de q covariables pour la série X_t.
Soit $x_1, \cdots, x_n$ une séquence de n obs. de la série X_t.
On suppose que chaque obs. x_ℓ est associée à un vecteur de q cov. $y_\ell = (y_{1,\ell}, \cdots, y_{q,\ell})$.

- Soit $b = \{b_j \in \mathbb{N}^\star,\ j = 1, \cdots, p\}$ un ensemble de p tailles de blocs assez grandes.
- Soit $(\gamma_j(y_t),\ \sigma_j(y_t),\ \mu_j(y_t))$ le vecteur des paramètres de la loi GEV nommée $G_j(\cdot \mid Y_t = y_t)$ caractérisant la distribution conditionelle des maximums de b_j obs. consécutives de la série X_t.
- Soit $\nu_j \in [0,\ 1]$ l'**indice extremal** associé au seuil défini par le quantile empirique $x_{n,1/b_j}$ d'ordre $1/b_j$ estimé sur une séquence $\mathcal{X}_n = \{x_1, \cdots, x_n\}$ de la série X_t.

Structure des paramètres

- $\mu_j(y_t) = \mu_{0,j} + \mu_{1,j}\, f_1(y_t) + \cdots + \mu_{q,j}\, f_q(y_t),$
- $\sigma_j(y_t) = \exp\left\{\phi_{0,j} + \phi_{1,j}\, g_1(y_t) + \cdots + \phi_{q,j}\, g_q(y_t)\right\},$
- $\gamma_j(y_t) = \gamma_{0,j} + \gamma_{1,j}\, h_1(y_t) + \cdots + \gamma_{q,j}\, h_q(y_t),$

où $f_\ell,\ g_\ell,\ h_\ell$ sont des fonctions continues de supports dans $\mathbb{R}^q$ et à valeurs dans $\mathbb{R}$.

Estimation du paramètre Θ : Cas Non-Stationnaire (2/2)

Nous avons adoptés les notations ci-dessous pour désigner l'ensemble des paramètres qui figurent dans les modèles de mélange qui précèdent

- $b = (b_1, \cdots, b_p)$,
- $\Theta(b) = (\Theta_j(b_j, y_t), j = 1, \cdots, p, t = 1, \cdots, n)$,
- $\Theta_j(b_j, y_t) = (\gamma_j(b_j, y_t), \sigma_j(b_j, y_t), \mu_j(b_j, y_t))$,
- $\gamma_j(b_j, y_t) = \gamma_j(y_t)$,
- $\sigma_j(b_j, y_t) = \sigma_j(y_t) \cdot (\nu_j b_j)^{-\gamma_j(y_t)}$,
- $\mu_j(b_j, y_t) = \mu_j(y_t) + \sigma_j(y_t) \cdot \left(\frac{(\nu_j b_j)^{-\gamma_j(y_t)} - 1}{\gamma_j(y_t)} \right)$.

Estimation du paramètre ω

Estimation des poids

1. On désigne par G_j le modèle GEV ayant pour paramètres $(\gamma_j(b_j), \sigma_j(b_j), \mu_j(b_j))$ dans le cas stationnaire et $(\gamma_j(b_j, y_t), \sigma_j(b_j, y_t), \mu_j(b_j, y_t))$ dans le cas non-stationnaire.
2. On se donne une probabilité a priori $\mathbb{P}\{G_j\}$ pour qu'une observation quelconque x_i soit issue du modèle G_j. On peut prendre par exemple $\mathbb{P}\{G_j\} = 1/p$.
3. La probablité a posteriori pour qu'une (**grande**) observation $x_i > x_0$ soit issue du modèle G_j est (d'après le théorème de Bayes)

$$\mathbb{P}\{G_j \,|\, x_i\} = \frac{\mathbb{P}\{G_j\}\,\mathbb{P}\{x_i \,|\, G_j\}}{\mathbb{P}\{x_i\}} = \frac{\mathbb{P}\{G_j\}\, g_j(x_i)}{\sum_{k=1}^{p} \mathbb{P}\{G_k\}\, g_k(x_i)},$$

   où g_j est la densité de probabilité associée au modèle G_j.
4. On estime le nombre total d'observations n_j issues du modèle G_j par

$$n_j = \sum_{x_i > x_0} \mathbf{1}\left(\mathbb{P}\{G_j \,|\, x_i\} \geq \max_{1 \leq k \leq p} \mathbb{P}\{G_k \,|\, x_i\}\right).$$

5. On estime le poids ω_j du modèle G_j par $\omega_j = \frac{n_j}{\sum_{k=1}^{p} n_k}$.