Examen Théorie des Valeurs Extrêmes 2011-2012

Seconde session

Master 2 SAF Pro

Sans document - Sans calculatrice

Question de cours (6 points):

1. On rappelle que les lois "max-stable" sont définies par

$$\text{Fréchet } (\alpha > 0): \quad \Phi_\alpha(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \exp\{-x^{-\alpha}\} & \text{si } x > 0. \end{cases}$$

$$\text{Weibull } (\alpha > 0): \quad \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{\alpha}\} & \text{si } x \leq 0, \\ 1 & \text{si } x > 0. \end{cases}$$

$$\text{Gumbel}: \quad \Lambda(x) = \exp\{-e^{-x}\} \qquad x \in \mathbf{R}.$$

Montrer que: $X \sim \Phi_\alpha \Leftrightarrow \ln(X^\alpha) \sim \Lambda \Leftrightarrow -\Psi^{-1} \sim \Phi_\alpha$.

2. Qu'est-ce que le domaine d'attraction d'une loi max-stable? Quelles sont les conditions qui définissent le domaine d'attraction de la loi de Fréchet?

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

Ensae 2018

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

Exercice 1 (8 points):

Supposons que Y a une distribution Pareto Généralisée GPD(σ, ξ) $(\xi \neq 0)$ telle que

$$P(Y < y) = 1 - \left[1 + \xi\left(\frac{y}{\sigma}\right)\right]_+^{-1/\xi}.$$

TD1 exo4

1. Donner le domaine de définition de Y, c'est-à-dire l'ensemble des valeurs de réalisation possibles de la variable aléatoire Y.

2. De quelle loi connue s'agit-il lorsque $\xi = -1$? Est-il possible de définir une loi si $\xi = 0$?

3. Donner la densité de cette loi. Peut-on trouver des expressions analytiques pour l'estimateur du maximum de vraisemblance de (σ, ξ).

4. Montrer que si $U \sim U[0,1]$, alors

$$Y \sim \sigma\left(\frac{U^{-\xi} - 1}{\xi}\right).$$

Donner une procédure pour simuler une loi GPD(σ, ξ).

5. Donner l'expression de la médiane, de la moyenne ($\xi < 1$) et de la variance ($\xi < 1/2$) d'une loi GPD(σ, ξ). Donner un estimateur simple de ξ basé sur les deux premiers moments lorsque $\xi < 1/2$.

6. Montrer que $Y - v|Y > v$ a une distribution GPD$(\sigma + \xi v, \xi)$. A partir du calcul de l'espérance de $Y - v|Y > v$, en déduire une procédure simple pour estimer ξ.

TD2
Q11

7. Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n = \max(X_1, ..., X_n)$. Montrer que si, lorsque $u \to x^F = \sup\{x : F(x) < 1\}$

$$\frac{X-u}{a(u)}|X > u \xrightarrow{L} GPD(1, \xi),$$

alors

$$\lim_{n\to\infty} P\left(\frac{M_n - b_n}{a_n} \le x\right) = \begin{cases} \exp\left(-[1+\xi x]_+^{-1/\xi}\right) & \text{si } \xi \neq 0, \\ \exp(-\exp(-x)) & \text{si } \xi = 0, \end{cases}$$

avec $a_n = a(b_n)$ et $b_n = F^{-1}(1 - n^{-1})$.

TD1
ex2

Exercice 2 (6 points):

Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n = \max(X_1, ..., X_n)$. On note la statistique d'ordre la manière suivante

$$X_{(n)} \le X_{(n-1)} \le \ldots \le X_{(1)}$$

et $B_n(x) = \sum_{i=1}^n \mathbb{I}_{\{X_i > x\}}$.

1. Montrer que $X_{(k)} \le x$ si et seulement si $B_n(x) < k$.

2. Donner la loi de $B_n(x)$. En déduire que

$$P\left(X_{(k)} \le x\right) = \sum_{r=0}^{k-1} C_n^r \left(F(x)\right)^{n-r} \left(1 - F(x)\right)^r.$$

3. Calculer la transformée de Laplace de $B_n(x)$:

$$\varphi_n(t, x) = \mathbb{E}\left(\exp\left(tB_n(x)\right)\right).$$

4. Montrer que s'il existe une suite (u_n) telle que $\lim_{n\to\infty} n \Pr(X_i > u_n) = \tau$, alors

$$\lim_{n\to\infty} \ln \varphi_n(t, u_n) = \tau(e^t - 1).$$

5. Donner la transformée de Laplace d'une loi de Poisson de paramètre τ (Rappel: si N suit une loi de Poisson de paramètre τ, alors $P(N = n) = e^{-\tau}\tau^n/n!$).

6. Déduire des questions précédentes l'expression analytique de la loi limite de $X_{(2)}$ si l'on peut trouver deux suites (a_n) et (b_n) telles que

$$\lim_{n\to\infty} P\left(\frac{X_{(1)} - b_n}{a_n} \le x\right) = \exp\left(-\exp(-x)\right).$$

IV Annale ISFA 2012

Question de cours

1 - • Si $X \sim \Phi_\alpha$ alors $P(X \leq x) = e^{-x^{-\alpha}}$

alors $\mathbb{P}(\ln X^\alpha \leq y) = \mathbb{P}(X \leq e^{\frac{y}{\alpha}}) = e^{-e^{-y}} \sim \Lambda$

alors $\ln(X^\alpha) \sim \Lambda$

• Si $\ln(X^\alpha) \sim \Lambda$ alors $\mathbb{P}(\ln(X^\alpha) \leq x) = e^{-e^{-x}}$

alors
$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(X^\alpha \leq x^\alpha) \\ &= \mathbb{P}(\ln(X^\alpha) \leq \ln(x^\alpha)) \\ &= \exp\{-\exp\{-\ln x^\alpha\}\} \\ &= \exp\{-\exp\{\ln x^{-\alpha}\}\} \\ &= \exp\{-x^{-\alpha}\} \end{aligned}$$

donc $X \sim \Phi_\alpha$. Ainsi (1) $\Leftrightarrow$ (2)

de plus, on peux noter que $-\Psi^{-1} \sim \Phi_\alpha \Leftrightarrow (-\Phi_\alpha)^{-1} \Leftrightarrow \Psi$

• Si $X \sim \Phi_\alpha$ alors $\mathbb{P}(X \leq x) = e^{-x^{-\alpha}}$

$\mathbb{P}(-X^{-1} \leq x) = \mathbb{P}(X^{-1} > -x) = \mathbb{P}(X \leq -\frac{1}{x}) = e^{-(-x)^\alpha} \sim \Psi$ donc $-X^{-1} \sim \Psi$

• Si $-X^{-1} \sim \Psi$,

alors
$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(X^{-1} > x^{-1}) \\ &= \mathbb{P}(-X^{-1} \leq -x^{-1}) \\ &= \exp\{-x^{-\alpha}\} \end{aligned}$$

donc $X \sim \Phi_\alpha$

On a donc montré (2) $\Leftrightarrow$ (1) $\Leftrightarrow$ (3) d'où l'équivalence des 3.

2 - Le domaine d'attraction d'une loi max-stable est l'étude de la distribution des valeurs extrêmes associée éventuellement à une dist.

Le domaine d'attraction de G est $F \in D(G)$ l'ensemble des dist. F tq $X_1 \dots X_n$ iid et qu'il existe (a_n) et (b_n) tq $F^n(a_n x + b_n) \longrightarrow G(x)$.

Les conditions qui définissent le domaine de Frechet sont $\alpha > 0$ et $x > 0$, $F \in D(\text{Frechet})$ ssi $\bar{F}(x) = x^{-\alpha} L(x)$ où L est une f° à variat° lentes, $d_n = 0$ et $c_n = F^{\leftarrow}(1 - n^{-1})$ = quantile de niveau $1 - \frac{1}{n}$.

$F^{\leftarrow}$: l'inverse généralisé.

3) Lien entre GPD et GEV :

La convergence du maximum de v.a. iid correctement normalisé est équivalente à la convergence de la dist des dépassements vers une GPD :

$$\frac{M_n - b_n}{a_n} \xrightarrow[n \infty]{L} GEV(0, 1, \xi) \Leftrightarrow \frac{X - u}{\beta(u)} \mid X > u \xrightarrow[u \to x_F]{} GPD(1, \xi)$$

Exercice 1 _

1 _ 5 _ (cf TD1 exo 4)

6 _ (cf TD2 Q15)

$E(Y - v \mid Y > v)$ graph du ME plot.

ξ est estimer dès que le graph devient linéaire.
pente du graph $= \frac{\xi}{1-\xi}$

7 _ On a $\mathbb{P}\left(\frac{X-u}{a(u)} > x \mid X > u\right) = \mathbb{P}(X > a(u)x + u \mid X > u)$

$$= \frac{\mathbb{P}(X > a(u)x + u)}{\mathbb{P}(X > u)}$$

$$= \frac{1 - F(a(u)x + u)}{1 - F(u)} \longrightarrow GPD(1, \xi)$$

Or, on pose $u = b_n$ et remarquons $b_n = F^{-1}(1 - n^{-1}) \Leftrightarrow 1 - F(b_n) = \frac{1}{n}$

donc $\mathbb{P}\left(\frac{X-u}{a(u)} > x \mid X > u\right) = n\left(1 - F(a(b_n)x + b_n)\right) \longrightarrow GPD(1, \xi)$

alors $\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n \leq a_n x + b_n)$

$$= F^n(a_n x + b_n)$$

$$= \exp\{n \ln(1 - \bar{F}(a_n x + b_n))\}$$

$$\sim \exp\{n \bar{F}(a_n x + b_n)\}$$

$$\sim \exp\{n(1 - F(a(b_n)x + b_n))\}$$

$$\xrightarrow[n\infty]{} \exp\{GPD(1, \xi)\}$$

$$\xrightarrow[n\infty]{} \begin{cases} \exp\{-[1 + \xi x]_+^{-1/\xi}\} & \text{si } \xi \neq 0 \\ \exp\{-\exp(-x)\} & \text{si } \xi = 0 \end{cases}$$

Examen Théorie des Valeurs Extrêmes 2012-2013

Master 2 SAF Pro

Sans document - Sans calculatrice - 2h00

Question de cours (6 points):

Répondre aux questions suivantes de manière la plus littérale possible, c'est-à-dire avec le minimum de mathématique.

1. Qu'est-ce que la propriété de max-stabilité? Est-ce que les lois suivantes sont max-stables:

$$\text{Fréchet :} \quad F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \exp\{-x^{-\alpha}\} & \text{si } x > 0. \end{cases}$$

$$\text{Exponentielle :} \quad F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - \exp\{-x\} & \text{si } x > 0. \end{cases}$$

$$\text{Gumbel :} \quad \Lambda(x) = \exp\{-e^{-x}\} \quad x \in \mathbb{R}$$

2. <u>Domaine d'attraction et max-stabilité:</u>

isfa 2012 (a) Qu'est-ce que le domaine d'attraction d'une loi max-stable?

(b) A quel domaine d'attraction appartient la loi de Pareto?

(c) Est-ce que deux distributions qui ont des comportements de leur fonction de survie identiques pour des grandes valeurs peuvent appartenir à deux domaines d'attraction différents?

(d) Comment utilise-t-on un graphique Quantile-Quantile pour identifier le domaine d'attraction d'une distribution?

3. Les distributions Pareto généralisées GPD(β, ξ) sont définies par

$$G^p_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{si } \xi = 0. \end{cases}$$

Rappeler leur domaine de définition. Quelle est la propriété principale de ces distributions? Quel lien existe-t-il avec les distributions max-stables?

Exercice 1 (4 points):

Soit $X_1, ..., X_n$ une suite de variables aléatoires IID.

1. Montrer que s'il existe deux suites (a_n) et (b_n) telles que

$$\lim_{n\to\infty} \Pr(\max(-X_1, \ldots, -X_n) \leq a_n x + b_n) = H(x)$$

pour une fonction de distribution H non-dégénérée, alors il existe deux suites (c_n) et (d_n) telles que

$$\lim_{n\to\infty} \Pr(\min(X_1, \ldots, X_n) \leq c_n x + d_n) = 1 - H(-x).$$

Quelles relations existe-t-il entre les suites (a_n), (b_n), (c_n) et (d_n)?

2. On rappelle que les distributions des extrêmes généralisées (GEV) sont caractérisées par les fonctions de répartition

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

Caractériser alors les distributions limites possibles pour le minimum de variables aléatoires indépendantes.

3. Montrer qu'elles sont min-stables et donner les coefficients de normalisation.

Exercice 2 (5 points):

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID, $M_n = \max(X_1, ..., X_n)$.

0. Soit $\bar{F} = 1 - F$. Montrer que, pour deux suites (a_n) et (b_n), et une fonction de répartition H, les deux conditions suivantes sont équivalentes:

$$\begin{array}{ll} 1) & \lim_{n\to\infty} \Pr(M_n \leq a_n x + b_n) = H(x), \\ 2) & \lim_{n\to\infty} n\bar{F}(a_n x + b_n) = -\ln H(x). \end{array}$$

1. Soit N une variable aléatoire de Poisson de paramètre λ, indépendante des X_i.

(a) Montrer que

$$\Pr(M_N \leq x) = \exp(-\lambda(1 - F(x))).$$

On rappelle que $\Pr(N = n) = e^{-\lambda}\lambda^n/n!$, $n \in \mathbb{N}$.

(b) Supposons que N_n une variable aléatoire de Poisson de paramètre λ_n. Quelle condition faut-il mettre sur λ_n pour que

$$\frac{M_{N_n} - b_n}{a_n} \xrightarrow{L} H$$

quand $n \to \infty$?

2. Soit N une variable aléatoire Géométrique de paramètre q, indépendante des X_i.

(a) Montrer que

$$\Pr(M_N \leq x) = \frac{qF(x)}{1 - (1-q)F(x)}.$$

On rappelle que $\Pr(N = n) = q(1-q)^{n-1}$, $n \geq 1$.

(b) Supposons que N_n une variable aléatoire Géométrique de paramètre q_n. Est-il possible de trouver q_n telle que

$$\frac{M_{N_n} - b_n}{a_n} \xrightarrow{L} H$$

quand $n \to \infty$?

3. Quelle conclusion tirez-vous de cet exercice?

Exercice 3 (5 points):

Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution de Fréchet de paramètre 1, i.e. $\Pr(X_1 \leq x) = \exp\left(-x^{-1}\right)$.

0. Donner les constantes a_n et b_n telles que

$$\frac{\max(X_1, ..., X_n) - b_n}{a_n} \overset{L}{=} X_1.$$

On définit le processus max-autoregressif de la manière suivante:

$$Y_i = \max\left(\alpha Y_{i-1}, (1-\alpha)X_i\right)$$

avec $0 \leq \alpha < 1$.

1. Montrer que si $Y_i = \max_{j\geq 0} \alpha^j (1-\alpha) X_{i-j}$ alors Y_{i+1} a la même distribution que Y_i. Il s'agit de la distribution stationnaire. Montrer que cette distribution est la distribution de Fréchet de paramètre 1.

2. Montrer que

$$\Pr\left(\max(Y_1, ..., Y_n) \leq x\right) = \Pr\left(Y_1 \leq x, (1-\alpha)X_2 \leq x, \ldots, (1-\alpha)X_n \leq x\right).$$

3. En déduire que

$$\Pr\left(\max(Y_1, ..., Y_n) \leq x\right) = \exp\left(-\left[1 + (1-\alpha)(n-1)/x\right]\right)$$

et que

$$\lim_{n\to\infty} \Pr\left(\frac{\max(Y_1, ..., Y_n) - b_n}{a_n} \leq x\right) = \exp\left(-(1-\alpha)x^{-1}\right).$$

4. Pour quelle valeur de α les lois asymptotiques des maxima des X_i et des Y_i normalisés coïncident?

TVE _ Annale ISFA 2013

Questions de cours

1) Une distribution G est dite max-stable si $\forall n$, $\exists\, \alpha_n > 0, \beta_n$ tq
$G^n(\alpha_n x + \beta_n) = G(z)$
où G^n est la fct de répartition de $M_n = \max(X_1 \dots X_n)$, (X_i) iid ayant chacune pour fonction de répartition G.

La propriété de max-stable est une propriété satisfaite par des dist. pour lesquelles l'opération de prendre l'échantillon maxima conduit à une dist identique à l'exception d'un changement d'échelle et de position.

Frechet : $n^{-1/\alpha} M_n \overset{d}{=} X_n$ donc max-stable.

Exponentielle & Gumbel : $M_n - \ln(n) \overset{d}{=} X$ donc max-stable.

2) a) cf ISFA 2012

b) La loi de Pareto appartient aux domaines d'attraction de Frechet.

c) Non. • le domaine d'attract° de Weibull contient des lois dont la fct de survie n'a pas de queue de dist.

• Gumbel : fct de survie ↘ exponentielle, i-e, à queue légère.
• Frechet : ———— ↘ polynomiale, i-e, à queue lourde.

d) QQplot compare des dist empirique et théorique :

• Si alignés, on est dans le m̂ domaine d'attraction.
• Si concave : on surestime la queue de dist.
• Si convexe : on sousestime la queue de dist.

⇒ donne une idée sur le ξ.
Si les observations ressemblent à une Pareto alors nos observat° ∈ au domaine de Frechet ⇒ $\xi > 0$.
———— Normale ————
de Gumbel ⇒ $\xi = 0$.

⇒ On peut compléter cette analogie avec le mean excess plot.

3) $D_f = \begin{cases} [0; +\infty[& \text{si } \xi \geq 0 \\ [0; -\frac{\beta}{\xi}] & \text{si } \xi < 0 \end{cases}$

Propriété principal = seuil stable.

Lien max-stable :

Si $\frac{M_n - b_n}{a_n} \longrightarrow GEV(0,1,\xi)$ alors $\left(\frac{X-u}{h(u)} > x \mid X > u\right) \underset{u \to x^F}{\longrightarrow} GPD(1, \xi)$ où $\begin{cases} b_n = F^{-1}(1 - \frac{1}{n}) \\ a_n = h(b_n) \end{cases}$

les seules dist. max-stable sont les GEV.
GPD et GEV ont le m̂ param. de queue ξ.

Maxima ~ GEV alors les excédents ~ GPD.

Exercice 3 -

$\overline{F}(a_n x + b_n) = 1 - \exp(-(a_n x + b_n)^{-1})$

0) $\mathbb{P}(M_n \leq a_n x + b_n) = F^n(a_n x + b_n)$
$$= (1 - \overline{F}(a_n x + b_n))^n$$
$$= \exp(n \ln(1 - \overline{F}(a_n x + b_n)))$$
$$(\sim \exp(-n \overline{F}(a_n x + b_n)))$$
$$\sim \exp\{n(-(a_n x + b_n)^{-1})\} \longrightarrow \exp(-x^{-1})$$

Par identification : $b_n = 0$ et $a_n = n$.

1) $Y_{i+1} = \max(\alpha Y_i, (1-\alpha) X_{i+1})$
$$= \max(\alpha \max_{j \geq 0}(\alpha^j (1-\alpha) X_{i-j}), (1-\alpha) X_{i+1})$$
$$= \max(\max_{j \geq 0} \alpha^{j+1} (1-\alpha) X_{i-j} ; (1-\alpha) X_{i+1})$$
$$= \max(\max_{j \geq 1} \alpha^j (1-\alpha) X_{i-j+1}, \alpha^0 (1-\alpha) X_{i+1-0})$$
$$= \max(\max_{j \geq 0} \alpha^j (1-\alpha) X_{i-j})$$
$$= \max_{j \geq 0} \alpha^j (1-\alpha) X_{i-j}$$
$$= \underline{Y_i}$$

$$\mathbb{P}(Y_i \leq x) = \mathbb{P}(\max_{j \geq 0}(\alpha^j (1-\alpha) X_{i-j}) \leq x)$$
$$= [\mathbb{P}(\alpha^{-1}(1-\alpha) X_{i-1} \leq x)]^i$$
$$= [\mathbb{P}(X_{i-1} \leq \frac{\alpha}{1-\alpha} x)]^i$$
$$= \exp\{-i(\frac{\alpha x}{1-\alpha})^{-1}\}$$
$$= \exp\{-i(\frac{1-\alpha}{\alpha}) x^{-1}\}$$
$$= \exp\{-i(\frac{1}{\alpha} - 1) x^{-1}\}$$

2) $Y_1 = Y_1$

$Y_2 = \max(\alpha Y_1, (1-\alpha) X_2)$

$Y_3 = \max(\alpha Y_2, (1-\alpha) X_3)$

Pour $n = 2$, $\max(Y_1, Y_2) = \max(Y_1, \max(\alpha Y_1, (1-\alpha) X_2))$

Si $Y_1 \leq x$ alors $\{\alpha Y_1 \leq x\} \subset \{Y_1 \leq x\}$ $(\alpha \in [0, 1[)$

alors $\max(Y_1, Y_2) = \max(Y_1, (1-\alpha) X_2)$

et ainsi de suite ...

$\Rightarrow \underline{\mathbb{P}(\max(Y_1 \dots Y_n) \leq x) = \mathbb{P}(Y_1 \leq x, (1-\alpha) X_2 \leq x, \dots, (1-\alpha) X_n \leq x)}$

3) $\mathbb{P}(\max(Y_1 \dots Y_n) \le x) = \mathbb{P}(Y_1 \le x, (1-\alpha)X_2 \le x, \dots, (1-\alpha)X_n \le x)$

$$= \mathbb{P}(Y_1 \le x)\ \mathbb{P}((1-\alpha)X_2 \le x)^{n-1}$$
$$= \mathbb{P}(Y_1 \le x)\ \mathbb{P}\left(X_2 \le \frac{x}{1-\alpha}\right)^{n-1}$$
$$= e^{-x^{-1}}\ e^{-(n-1)\left(\frac{x}{1-\alpha}\right)^{-1}}$$
$$= \underline{\exp\left\{-\left[\frac{1}{x} + (n-1)(1-\alpha)\frac{1}{x}\right]\right\}}$$

$\mathbb{P}\left(\frac{\max(Y_1 \dots Y_n) - b_n}{a_n} \le x\right) = \mathbb{P}(\max(Y_1 \dots Y_n) \le a_n x - b_n)$ où $a_n = n$ et $b_n = 0$ (Frechet)

$$= \exp\left\{-\left[\frac{1}{xn} + \frac{(n-1)}{n}(1-\alpha)\frac{1}{x}\right]\right\}$$
$$\xrightarrow[n\to\infty]{} \underline{\exp\left\{-(1-\alpha)\frac{1}{x}\right\}}$$

4) les lois coïncident lorsque $\exp\{-(1-\alpha)x^{-1}\} = \exp\{-x^{-1}\}$

$\Rightarrow \underline{\alpha = 0}$

Exercice 1

1) On peut remarquer que $\min(X_1 \dots X_n) = -\max(-X_1 \dots -X_n)$

alors $\mathbb{P}(m_n \le c_n x + d_n) = \mathbb{P}(-\max(-X_1 \dots -X_n) \le c_n x + d_n)$

$$= \mathbb{P}(\max(-X_1, \dots -X_n) \ge -c_n x - d_n)$$
$$= 1 - \mathbb{P}(\max(-X_1 \dots -X_n) \le -c_n x - d_n)$$
$$= 1 - \mathbb{P}(\max(-X_1 \dots -X_n) \le c_n(-x) - d_n)$$
$$\xrightarrow[n\to\infty]{} \underline{1 - H(-x)}$$

Par identification, $\underline{a_n = c_n \text{ et } d_n = -b_n}$.

2) On vient de montrer que $\mathbb{P}(m_n \le c_n x + d_n) \xrightarrow[n\to\infty]{} 1 - G_\xi(-x) = \underline{1 - \exp\left(-(1+\xi(-x))_+^{-\frac{1}{\xi}}\right)}$

(en effet $H \sim GEV$)

3) Min stable : $u_n X + v_n \overset{\mathcal{L}}{=} X$

$\mathbb{P}(u_n m_n + v_n \le x) = \mathbb{P}\left(m_n \le \frac{x - v_n}{u_n}\right)$

$$= 1 - \mathbb{P}\left(m_n > \frac{x - v_n}{u_n}\right)$$
$$= 1 - \mathbb{P}\left(X_1 > \frac{x - v_n}{u_n}\right)^n$$
$$= 1 - \left(1 - \mathbb{P}\left(X_1 \le \frac{x - v}{u_n}\right)\right)^n \qquad 1 - H(-x)$$
$$= 1 - \left(H\left(-\frac{x - v_n}{u_n}\right)\right)^n$$
$$= 1 - \exp\left\{-n\left(1 + \xi\left(-\frac{x - v_n}{u_n}\right)\right)_+^{-1/\xi}\right\}$$
$$= 1 - \exp\left\{-\left(n^{\xi} + n^{\xi}\xi\left(-\frac{x - v_n}{u_n}\right)\right)_+^{-1/\xi}\right\}$$

$n^{\xi} + \xi v_n = 1$

donc $\underline{u_n = n^{\xi} \text{ et } v_n = -n^{\xi}\xi^{-1}}$

Exercice 2 -

0) $\mathbb{P}(M_n \leq a_n x + b_n) \to H(x) \iff F^n(a_n x + b_n) \to H(x)$
$\iff n \ln(F(a_n x + b_n)) \to \ln H(x)$
$\iff n \ln(1 - \bar{F}(a_n x + b_n)) \to \ln H(x)$
$\iff -n \bar{F}(a_n x + b_n) \to \ln(H(x))$
$\iff \underline{n \bar{F}(a_n x + b_n) \to -\ln H(x)}$

1) a) $\mathbb{P}(M_N \leq x) = \mathbb{E}(\mathbb{P}(M_n \leq x \mid N = n))$
$= \mathbb{E}(F^n(x) \mid N = n) = \mathbb{E}(F^N(x))$
$= \mathbb{E}(e^{N \ln F(x)})$,

or $\boxed{\mathbb{E}(e^{Nt}) = e^{\lambda(e^t - 1)}}$ (transformée de Laplace)

donc $\underline{\mathbb{P}(M_N \leq x) = e^{\lambda(F(x) - 1)}}$

1) b) $\mathbb{P}\left(\frac{M_{N_n} - b_n}{a_n} \leq x\right) = \mathbb{P}(M_{N_n} \leq a_n x + b_n)$
$= \exp\{-\lambda_n(1 - F(a_n x + b_n))\}$
$= \exp\{-\lambda_n \bar{F}(a_n x + b_n)\}$

donc $\mathbb{P}\left(\frac{M_{N_n} - b_n}{a_n} \leq x\right) \to H(x) \iff -\lambda_n \bar{F}(a_n x + b_n) \to \ln H(x)$
$\iff \underline{\lambda_n = n}$ (d'après hyp 2 question 0)

$\sum_{n=0}^{\infty} q^k = \frac{1}{1-q}$

2) a) $\mathbb{P}(M_N \leq x) = \mathbb{E}\left[e^{N \ln F(x)}\right]$

or $\mathbb{E}\left[e^{tN}\right] = \sum_{n=1}^{\infty} e^{tn} q(1-q)^{n-1} = q e^t \sum_{n=1}^{\infty} \left[e^t(1-q)\right]^{n-1} = \frac{q e^t}{1 - e^t(1-q)}$

donc, $\underline{\mathbb{P}(M_N \leq x) = \frac{q F(x)}{1 - (1-q) F(x)}}$

2) b) $\mathbb{P}\left(\frac{M_{N_n} - b_n}{a_n} \leq x\right) = \mathbb{P}(M_{N_n} \leq a_n x + b_n)$
$= \frac{q F(a_n x + b_n)}{1 - (1-q) F(a_n x + b_n)}$
$= \frac{1 - \bar{F}(a_n x + b_n)}{1 + \frac{1}{q} \bar{F}(a_n x + b_n)}$

$\sim \frac{1}{1 + \frac{\bar{F}(a_n x + b_n)}{q}} \longrightarrow \frac{1}{1 - \ln H(x)} \neq H(x)$

donc $\underline{(N_n) \text{ n'est pas max-stable}}$.

3) Si je randomise, la loi limite n'est pas toujours max-stable.

Examen Théorie des Valeurs Extrêmes 2013-2014

Master 2 SAF Pro

Sans document - Sans calculatrice - 1h30

Question de cours (8 points):

Répondre aux questions suivantes de manière la plus littérale possible, c'est-à-dire avec le minimum de mathématique.

1. Qu'appelle-t-on une distribution à queue épaisse et comment l'identifie-t-on dans la pratique?

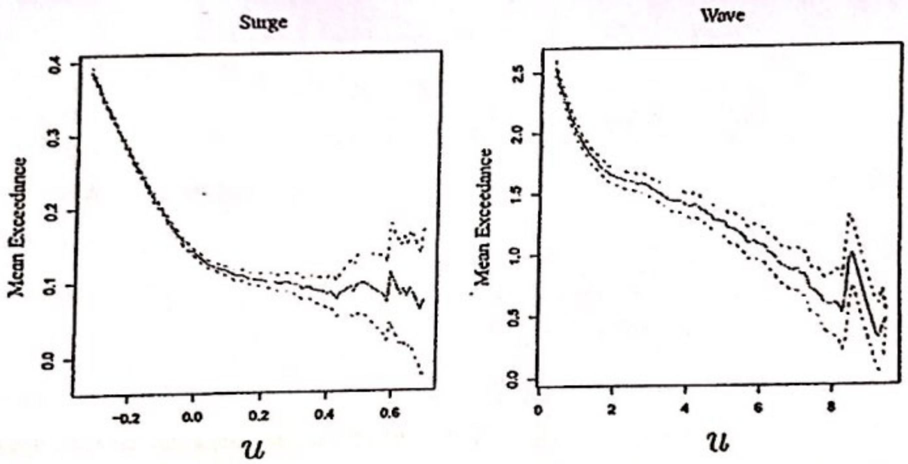
isfa2012

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

ensae 2018

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

4. Quelle utilisation fait-on de la fonction de dépassement moyen empirique? Vous trouverez ci-dessous deux représentations graphiques de la fonction de dépassement moyen empirique pour deux grandeurs physiques (Surge et Wave) en fonction d'un seuil u. Quel comportement de la fonction de dépassement est attendu pour les grandes valeurs de u?



ensae 2018

Exercice 1 (8 points):

On considère la distribution donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp\left(-\frac{x}{1-x}\right) & \text{si } 0 \le x \le 1, \\ 1 & \text{si } x > 1. \end{cases}$$

enoncé 2018

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID de distribution F et $M_n = \max(X_1, ..., X_n)$.

0. Est-ce que M_n peut avoir une distribution non dégénérée avec une suite de seuils non-linéaires?

1. En choisissant

$$a_n = \frac{1}{(1+\ln n)^2} \quad \text{et} \quad b_n = \frac{\ln n}{1+\ln n},$$

trouver la loi limite de $a_n^{-1}(M_n - b_n)$.

Soient $E_1, ..., E_n$ une suite de variables aléatoires IID de distribution exponentielle de paramètre 1 et $M_n^E = \max(E_1, ..., E_n)$.

2. Donner les constantes de normalisation a_n^E et b_n^E telles que $(a_n^E)^{-1}(M_n^E - b_n^E)$ converge en distribution vers la loi de Gumbel.

3. On considère la fonction

$$g(x) = \frac{x}{1+x}, \quad x > 0.$$

a. Montrer que $X_1 \stackrel{d}{=} g(E_1)$. En déduire que

$$\frac{M_n - b_n}{a_n} \stackrel{d}{=} \frac{g(M_n^E) - g(\ln n)}{g'(\ln n)}.$$

b. Montrer qu'il existe ζ_n tel que

$$g(M_n^E) - g(\ln n) = (M_n^E - \ln n)\, g'(\zeta_n)$$

avec $\ln n \leq \zeta_n \leq M_n^E$ si $\ln n \leq M_n^E$ et inversement.

c. Montrer que

$$\frac{M_n^E}{\ln n} \xrightarrow{P} 1,$$

et en déduire que

$$\frac{g'(\zeta_n)}{g'(\ln n)} \xrightarrow{P} 1.$$

Retrouver le résultat de la question 1.

Exercice 2 (4 points):

Soient $X_1, ..., X_n$ une suite de variables aléatoires indépendantes de loi Gaussienne centrée et réduite, et $M_n = \max(|X_1|, ..., |X_n|)$. On a dans ce cas là:

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda$$

où Λ a une distribution de Gumbel et

$$\begin{aligned} a_n &= (2\ln(2n))^{-1/2} \\ b_n &= (2\ln(2n))^{1/2} - \frac{1}{2}(2\ln(2n))^{-1/2}[\ln(\ln(2n)) + \log(4\pi)]. \end{aligned}$$

1. Montrer que

$$\frac{M_n}{b_n} \xrightarrow{P} 1.$$

On cherche à caractériser la loi asympotique du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.

2. En utilisant une identité remarquable, montrer que si

$$\frac{\max(|X_1|, ..., |X_n|) - b_n}{a_n} \xrightarrow{d} \Lambda,$$

alors

$$\frac{\max(X_1^2, ..., X_n^2) - b_n^2}{2a_n b_n} \xrightarrow{d} \Lambda.$$

En déduire les constantes de normalisation pour la convergence du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.

TVE _ Annale ISFA 2014

Questions de cours _

1) Une distribution à queue épaisse donne une proba forte aux évènement extrêmes. On peut les identifier en comparant leur graphique de densité avec une loi usuelle à queue de dist. fine ou normale ou en regardant si $h'(x) \xrightarrow[x \to x_F]{} \infty$ ou ξ ⊕ ξ est grand ⊕ la queue est épaisse.

⇒ Frechet : f° de survie à queue lourde.

2) cf ISFA 2012
3) cf ENSAE 2018

4) Fonction de dépassement moyen = outil graphique important permet de comparer les comportements et de déterminer le seuil.

Pour les grandes valeurs de u, peu de données ⇒ plus volatile.

Exercice 2

1) $\mathbb{P}\left(\left| \frac{M_n}{b_n} - 1 \right| > \varepsilon \right) = \mathbb{P}\left(\frac{a_n}{b_n} \left| \frac{M_n - b_n}{a_n} \right| > \varepsilon \right)$

Slutsky : Si $X \xrightarrow{\mathcal{L}} X$, $Y \xrightarrow{\mathcal{L}} c \in \mathbb{R}$ $\}$ alors $XY \xrightarrow{\mathcal{L}} cX$

or $\frac{M_n - b_n}{a_n} \longrightarrow \Lambda$

et $\frac{a_n}{b_n} = \frac{1}{1 - \frac{1}{2}\left[\ln(\ln 2n) + \log(4\pi) \right]} \xrightarrow[n\infty]{} 0$

$\}$ $\frac{a_n}{b_n} \left| \frac{M_n - b_n}{a_n} \right| \xrightarrow[n\infty]{} 0$

donc $\frac{M_n}{b_n} \xrightarrow{\mathbb{P}} 1$

2) $\frac{\max(X_1^2, \dots, X_n^2) - b_n^2}{2 a_n b_n} = \frac{(\max(|X_1| \dots |X_n|) - b_n)(\max(|X_1| \dots |X_n|) + b_n)}{2 a_n b_n}$

$= \frac{(M_n - b_n)}{a_n} \times \frac{M_n + b_n}{2 b_n}$

$= \frac{M_n - b_n}{a_n} \times \frac{\frac{M_n}{b_n} + 1}{2}$

or, $\frac{M_n}{b_n} \longrightarrow 1$ donc $\frac{\frac{M_n}{b_n} + 1}{2} \longrightarrow \frac{1+1}{2} = 1$

$X_i \sim \mathcal{N}(0,1)$
$\Rightarrow X_i^2 \sim \chi_2^1$

donc $\frac{\max(X_1^2 \dots X_n^2) - b_n^2}{2 a_n b_n} \longrightarrow \Lambda$

de plus, $X_i \sim \mathcal{N}(0,1)$ donc $X_i^2 \sim \chi^1$ avec $\begin{cases} a_n^* = 2 a_n b_n \\ b_n^* = b_n^2 \end{cases}$

/

Examen Théorie des Valeurs Extrêmes 2014-2015

Master 2 SAF Pro

Sans document - Sans calculatrice - 1h30

Rabat 2018

Exercices d'applications du cours (10 points):

1. Soit B une variable aléatoire de distribution Binomiale(n,p)

$$P(B=k)=C_n^k p^k (1-p)^{n-k}$$

et N une variable aléatoire de distribution Poisson(λ). On pose $\lambda = np$. Montrer que

$$P(B=k)=P(N=k)\left(\frac{(1-\frac{\lambda}{n})^n}{e^{-\lambda}}\frac{1}{(1-\frac{\lambda}{n})^k}\prod_{i=1}^{k}\left(1-\frac{i-1}{n}\right)\right).$$

En déduire que si $n\to\infty$, $p\to 0$ tel que $\lambda = np$, alors B converge en distribution vers N.

Expliquer comment ce résultat peut être utilisé pour caractériser la loi du maximum et des statistiques d'ordre élevé.

2. Qu'est-ce que la propriété de max-stabilité? Soit X une variable aléatoire de loi exponentielle standard. Est-ce que X ou $-X$ a une loi max-stable? Si oui, donner les coefficients permettant de déduire la max-stabilité.

Rabat 2018

3. Expliquer ce que veut dire "F appartient au domaine d'attraction d'un loi max-stable G ($F\in D(G)$)". On considère la variable aléatoire de distribution H telle que

$$H(x)=F(\alpha x+\beta)$$

avec $\alpha>0$. A quelle domaine d'attraction appartient H. Donner les relations qui existent entre les suites de normalisation pour H et F.

4. On rappelle que la distribution de Pareto généralisée GPD(β,ξ) est définie par

$$G^p_{\xi,\beta}(x)=\begin{cases}1-[1+\xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi\neq 0,\\ 1-e^{-x/\beta} & \text{si } \xi=0,\end{cases}.$$

Expliquer à quel domaine d'attraction appartient la distribution de Pareto généralisée GPD(β,ξ).

5. Soit $X_1,...,X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n=\max(X_1,...,X_n)$. On rappelle que:

i) Pour un $\tau>0$ et une suite de réels (u_n), les conditions suivantes sont équivalentes: quand $n\to\infty$

1) $\Pr(M_n\le u_n)\to e^{-\tau}$,

2) $n\bar F(u_n)=n(1-F(u_n))\to\tau$.

rabat 2018

ii) Il existe une suite (u_n) satisfaisant $n\bar{F}(u_n) \to \tau$ quand $n \to \infty$ si et seulement si

$$\lim_{x \uparrow x^F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1.$$

où $x^F = \sup\{x : F(x) < 1\}$. Peut-on trouver une suite (u_n) telle que $\lim_{n\to\infty} \Pr(M_n \leq u_n) = e^{-\tau}$ si

- les X_i ont une loi Exponentielle de paramètre λ

$$\Pr(X_i > x) = \exp(-\lambda x), \quad x > 0;$$

- les X_i ont une loi Géométrique de paramètre p

$$\Pr(X_i = n) = p(1-p)^n, \quad n \in \mathbb{N};$$

- les X_i ont une loi de Poisson de paramètre λ

$$\Pr(X_i = n) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n \in \mathbb{N}.$$

Exercice 1 (5 points):

On rappelle qu'une distribution F appartient au domaine d'attraction de la loi de Gumbel Λ, si et seulement si il existe une fonction positive g telle que :

$$\lim_{x \to x^F} \frac{\bar{F}(x + tg(x))}{\bar{F}(x)} = \exp(-t), \qquad t \in \mathbb{R},$$

et qu'un choix possible pour la fonction g est donnée par

$$g(x) = \frac{\int_x^{x^F} \bar{F}(t)dt}{\bar{F}(x)} = \mathbb{E}(X - x | X > x).$$

Les coefficients a_n et b_n tels que

$$\Pr(M_n \leq a_n x + b_n) \to \Lambda(x) = \exp(-\exp(-x)), \qquad x \in \mathbb{R},$$

ont alors la forme suivante : $b_n = F^{\leftarrow}(1 - n^{-1})$, $a_n = g(b_n)$.

1. Supposons que X soit une variable aléatoire positive de distribution F telle que $x^F = \infty$.

- Donner quelques éléments d'explication (sans trop de détails mathématiques) pour justifier que l'on a $\lim_{x\to\infty} g(x)/x = 0$.

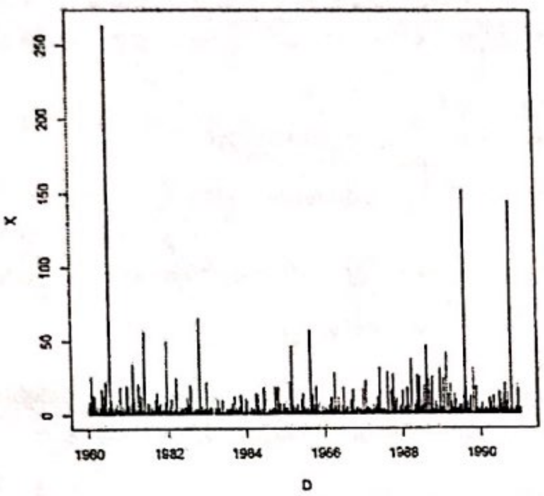
- Montrer que si $F \in D(\Lambda)$ où Λ est la distribution de Gumbel, alors la distribution de $-X^{-1}$ appartient aussi au domaine d'attraction de la distribution de Gumbel. On partira de l'écriture de la condition $F \in D(\Lambda)$ en fonction de M_n, le maximum, pour en déduire la réponse.

2. On considère une transformation croissante T, deux fois dérivable sur $\mathbb{R}$ et on suppose que le point extremal $x^F < \infty$ est fini. Montrer que $F \circ T^{-1}$ appartient au domaine d'attraction de la distribution de Gumbel et donner les coefficients de normalisation.

Indication: on utilisera un développement limité au voisinage de x^F.

rabat 2018

Question méthodologique (5 points):
Un réassureur dispose des données de sinistres suivantes pour tarifier un excédent de sinistres de priorité 100 et de portée illimitée.



Quelle méthodologie lui proposez-vous? Donner une réponse la plus littérale possible en décrivant la formule que vous allez utiliser, les paramètres dont vous avez besoin, les procédures d'estimation que vous allez utiliser.

Examen Théorie des Valeurs Extrêmes 2015-2016

Master 2 SAF Pro

Sans document - Sans calculatrice - 1h30

Questions de cours et exercices d'applications du cours (7 points):

ensae 2018

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

isfa 2018

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

ensae 2018

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.

ensae 2018

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

Exercice 1 (8 points):

isfa 2013

Soit $X_1, \ldots, X_n$ une suite de variables aléatoires IID.

1. Montrer que s'il existe deux suites (a_n), $a_n > 0$, et (b_n) telles que

$$\lim_{n\to\infty} \Pr(\max(-X_1, \ldots, -X_n) \le a_n x + b_n) = H(x)$$

pour une fonction de distribution H non-dégénérée, alors il existe deux suites (c_n), $c_n > 0$, et (d_n) telles que

$$\lim_{n\to\infty} \Pr(\min(X_1, \ldots, X_n) \le c_n x + d_n) = 1 - H(-x).$$

Quelles relations existe-t-il entre les suites (a_n), (b_n), (c_n) et (d_n)?

2. On rappelle que les distributions des extrêmes généralisées (GEV) sont caractérisées par les fonctions de répartition

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

Caractériser alors les distributions limites possibles pour le minimum de variables aléatoires indépendantes.

3. Montrer qu'elles sont min-stables et donner les coefficients de normalisation.

ensae 2018

4. Posons $M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$ et supposons qu'il existe des constantes $a_n > 0$, $\alpha_n > 0$, b_n, β_n telles que

$$P(M_n \leq a_n x + b_n) \to G_1(x) \quad \text{et} \quad P(-m_n \leq \alpha_n x + \beta_n) \to G_2(x).$$

Montrer la convergence

$$P(M_n \leq a_n x + b_n, -m_n \leq \alpha_n y + \beta_n) \to G_1(x) G_2(y).$$

Qu'en concluez-vous?

Exercice 2 (5 points):

istа 2013

Soit $X_1, \dots, X_n$ une suite de variables aléatoires IID de fonction de distribution de Fréchet de paramètre 1, i.e. $\Pr(X_1 \leq x) = \exp(-x^{-1})$.

0. Donner les constantes a_n et b_n telles que

$$\frac{\max(X_1, \dots, X_n) - b_n}{a_n} \overset{L}{=} X_1.$$

On définit le processus max-autoregressif de la manière suivante:

$$Y_i = \max(\alpha Y_{i-1}, (1-\alpha) X_i)$$

avec $0 \leq \alpha < 1$.

1. Montrer que si $Y_i = \max_{j \geq 0} \alpha^j (1-\alpha) X_{i-j}$ alors Y_{i+1} a la même distribution que Y_i. Il s'agit de la distribution stationnaire. Montrer que cette distribution est la distribution de Fréchet de paramètre 1.

2. Montrer que

$$\Pr(\max(Y_1, \dots, Y_n) \leq x) = \Pr(Y_1 \leq x, (1-\alpha) X_2 \leq x, \dots, (1-\alpha) X_n \leq x).$$

3. En déduire que

$$\Pr(\max(Y_1, \dots, Y_n) \leq x) = \exp(-[1 + (1-\alpha)(n-1)]/x)$$

et que

$$\lim_{n\to\infty} \Pr\left(\frac{\max(Y_1, \dots, Y_n) - b_n}{a_n} \leq x\right) = \exp(-(1-\alpha) x^{-1}).$$

4. Pour quelle valeur de α les lois asymptotiques des maxima des X_i et des Y_i normalisés coïncident?

Examen Théorie des Valeurs Extrêmes 2016-2017

Master 2 SAF Pro

Sans document - Sans calculatrice - 1h30

Remarque préliminaire: les questions des exercices peuvent se traiter en utilisant les résultats des questions précédentes même si vous n'avez pas réussi à y répondre.

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

Rabat 2018

1. Expliquer ce que veut dire "F appartient au domaine d'attraction d'un loi max-stable G ($F \in D(G)$)". On considère la variable aléatoire de distribution H telle que

$$H(x) = F(\alpha x + \beta)$$

avec $\alpha > 0$. A quel domaine d'attraction appartient H? Donner les relations qui existent entre les suites de normalisation pour H et F.

2. Supposons que F a une densité positive f et un point extrémal infini, x^F, telle que

$$\lim_{x \to x^F} \frac{(1+\beta x) f(x)}{\bar{F}(x)} = c.$$

A quel domaine d'attraction appartient F?

Rabat 2018

3. On rappelle que la distribution de Pareto généralisée GPD(β, ξ) est définie par

$$G^{p}_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{si } \xi = 0, \end{cases}$$

en 2

Expliquer à quel domaine d'attraction appartient la distribution de Pareto généralisée GPD(β, ξ).

4. Soit $X_1, \dots, X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n = \max(X_1, \dots, X_n)$. On rappelle que:

i) Pour un $\tau > 0$ et une suite de réels (u_n), les conditions suivantes sont équivalentes: quand $n \to \infty$,

$$1) \quad \Pr(M_n \le u_n) \to e^{-\tau},$$
$$2) \quad n\bar{F}(u_n) = n(1 - F(u_n)) \to \tau.$$

ii) Il existe une suite (u_n) satisfaisant $n\bar{F}(u_n) \to \tau$ quand $n \to \infty$ si et seulement si

$$\lim_{x \uparrow x^F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1,$$

où $x^F = \sup\{x : F(x) < 1\}$. Peut-on trouver une suite (u_n) telle que $\lim_{n \to \infty} \Pr(M_n \le u_n) = e^{-\tau}$ si

- les X_i ont une loi Exponentielle de paramètre λ

$$\Pr(X_i > x) = \exp(-\lambda x), \quad x > 0;$$

Rabat 2018

- les X_i ont une loi Géométrique de paramètre p

$$\Pr(X_i = n) = p(1-p)^n, \quad n \in \mathbb{N};$$

- les X_i ont une loi de Poisson de paramètre λ

$$\Pr(X_i = n) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n \in \mathbb{N}.$$

Exercice 1 (8 points):

On considère la distribution donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp\left(-\frac{x}{1-x}\right) & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID de distribution F et $M_n = \max(X_1, ..., X_n)$.

0. Est-ce que M_n peut avoir une distribution non dégénérée avec une suite de seuils non-linéaires?

1. En choisissant

$$a_n = \frac{1}{(1+\ln n)^2} \quad \text{et} \quad b_n = \frac{\ln n}{1+\ln n},$$

trouver la loi limite de $a_n^{-1}(M_n - b_n)$.

ensae 2018

Soient $E_1, ..., E_n$ une suite de variables aléatoires IID de distribution exponentielle de paramètre 1 et $M_n^E = \max(E_1, ..., E_n)$.

2. Donner les constantes de normalisation a_n^E et b_n^E telles que $(a_n^E)^{-1}(M_n^E - b_n^E)$ converge en distribution vers la loi de Gumbel.

3. On considère la fonction

$$g(x) = \frac{x}{1+x}, \quad x > 0.$$

a. Montrer que $X_1 \stackrel{d}{=} g(E_1)$. En déduire que

$$\frac{M_n - b_n}{a_n} \stackrel{d}{=} \frac{g(M_n^E) - g(\ln n)}{g'(\ln n)}.$$

b. Montrer qu'il existe ζ_n tel que

$$g(M_n^E) - g(\ln n) = (M_n^E - \ln n)\, g'(\zeta_n)$$

avec $\ln n \leq \zeta_n \leq M_n^E$ si $\ln n \leq M_n^E$ et inversement.

c. Montrer que

$$\frac{M_n^E}{\ln n} \xrightarrow{P} 1.$$

et en déduire que

$$\frac{g'(\zeta_n)}{g'(\ln n)} \xrightarrow{P} 1.$$

Retrouver le résultat de la question 1.

ensae 2018

Exercice 2 (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 $\bar{F}(x) = x^{-\alpha}L(x)$.

ensae 2016

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 à l'aide d'un graphique 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 $X_1, \ldots, X_n$ n variables positives, indépendantes, identiquement distribuées de distribution caractérisée par $\bar{F}(x) = x^{-\alpha}L(x)$. On note $S_n = X_1 + \ldots + X_n$. 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 $M_n = \max(X_1, \ldots, X_n)$. Utiliser la formule du binôme de Newton.

Interpréter ce résultat en terme de gestion des risques extrêmes.

Annale TVE ISFA 2017

Questions de cours

1) cf Rabat 2018

2) On a $h(x) = \frac{\bar{F}(x)}{f(x)} \xrightarrow[x \to x^F]{} \frac{1+\beta x^F}{c}$

donc $h'(x) \xrightarrow[x \to x^F]{} \frac{\beta}{c}$

Si $\beta = 0$, $D(F) =$ Gumbel
Si $\frac{\beta}{c} > 0$, $D(F) =$ Frechet
Si $\frac{\beta}{c} < 0$, $D(F) =$ Weibull.

Examen Théorie des Valeurs Extrêmes 2016-2017
Seconde session

Master 2 SAF Pro

Sans document - Sans calculatrice

isfa 2013

Question de cours (8 points):

1. On rappelle que les lois "max-stables" sont définies par

$$\text{Fréchet } (\alpha > 0): \quad \Phi_\alpha(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \exp\{-x^{-\alpha}\} & \text{si } x > 0. \end{cases}$$

$$\text{Weibull } (\alpha > 0): \quad \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\} & \text{si } x \leq 0, \\ 1 & \text{si } x > 0. \end{cases}$$

$$\text{Gumbel}: \quad \Lambda(x) = \exp\{-e^{-x}\} \quad x \in \mathbb{R}.$$

isfa 2012 Montrer que: $X \sim \Phi_\alpha \Leftrightarrow \ln(X^\alpha) \sim \Lambda \Leftrightarrow -\Psi^{-1} \sim \Phi_\alpha$.

isfa 2012 2. Qu'est-ce que le domaine d'attraction d'une loi max-stable? Quelles sont les conditions qui définissent le domaine d'attraction de la loi de Fréchet?

isfa 2012 3. Quel lien relie les distributions GEV (Generalised Extreme Value distibutions) et les distributions GPD (Generalised Pareto distributions)?

ensae 2018 4. Quelles sont les procédures statistiques à utiliser pour estimer les paramètres des distributions GEV (discuter suivant la nature des observations)?

TD1 exo2

Exercice 1 (6 points):

isfa 2013

Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution de Fréchet de paramètre 1, i.e. $\Pr(X_1 \leq x) = \exp(-x^{-1})$.

0. Donner les constantes a_n et b_n telles que

$$\frac{\max(X_1, ..., X_n) - b_n}{a_n} \stackrel{L}{=} X_1.$$

On définit le processus max-autoregressif de la manière suivante:

$$Y_t = \max(\alpha Y_{t-1}, (1-\alpha) X_t)$$

avec $0 \leq \alpha < 1$.

ista 2013

1. Montrer que si $Y_i = \max_{j \geq 0} \alpha^j (1-\alpha) X_{i-j}$ alors Y_{i-1} a la même distribution que Y_i. Il s'agit de la distribution stationnaire. Montrer que cette distribution est la distribution de Fréchet de paramètre 1.

2. Montrer que

$$\Pr(\max(Y_1, ..., Y_n) \leq x) = \Pr(Y_1 \leq x, (1-\alpha)X_2 \leq x, ..., (1-\alpha)X_n \leq x).$$

3. En déduire que

$$\Pr(\max(Y_1, ..., Y_n) \leq x) = \exp(-\{1 + (1-\alpha)(n-1)\}/x)$$

et que

$$\lim_{n\to\infty} \Pr\left(\frac{\max(Y_1, ..., Y_n) - b_n}{a_n} \leq x\right) = \exp(-(1-\alpha)x^{-1}).$$

4. Pour quelle valeur de α les lois asymptotiques des maxima des X_i et des Y_i normalisés coïncident?

Exercice 2 (6 points):

TD1 exo2

Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n = \max(X_1, ..., X_n)$. On note la statistique d'ordre la manière suivante

$$X_{(n)} \leq X_{(n-1)} \leq \ldots \leq X_{(1)}$$

et $B_n(x) = \sum_{i=1}^n \mathrm{I}_{\{X_i > x\}}$.

1. Montrer que $X_{(k)} \leq x$ si et seulement si $B_n(x) < k$.
2. Donner la loi de $B_n(x)$. En déduire que

$$P(X_{(k)} \leq x) = \sum_{r=0}^{k-1} C_n^r (F(x))^{n-r} (1-F(x))^r.$$

3. Calculer la transformée de Laplace de $B_n(x)$:

$$\varphi_n(t, x) = \mathrm{E}(\exp(tB_n(x))).$$

4. Montrer que s'il existe une suite (u_n) telle que $\lim_{n\to\infty} n \Pr(X_i > u_n) = \tau$, alors

$$\lim_{n\to\infty} \ln \varphi_n(t, u_n) = \tau(e^t - 1).$$

5. Donner la transformée de Laplace d'une loi de Poisson de paramètre τ (Rappel: si N suit une loi de Poisson de paramètre τ, alors $P(N = n) = e^{-\tau}\tau^n/n!$).

6. Déduire des questions précédentes l'expression analytique de la loi limite de $X_{(2)}$ si l'on peut trouver deux suites (a_n) et (b_n) telles que

$$\lim_{n\to\infty} P\left(\frac{X_{(1)} - b_n}{a_n} \leq x\right) = \exp(-\exp(-x)).$$

Examen Théorie des Valeurs Extrêmes 2013-2014

ENSAE

Sans document - Sans calculatrice - 2h00

Question de cours (8 points):

Répondre aux questions suivantes de manière la plus littérale possible, c'est-à-dire avec le minimum de mathématique.

ista 2014

1. Qu'appelle-t-on une distribution à queue épaisse et comment l'identifie-t-on dans la pratique?

ista 2012

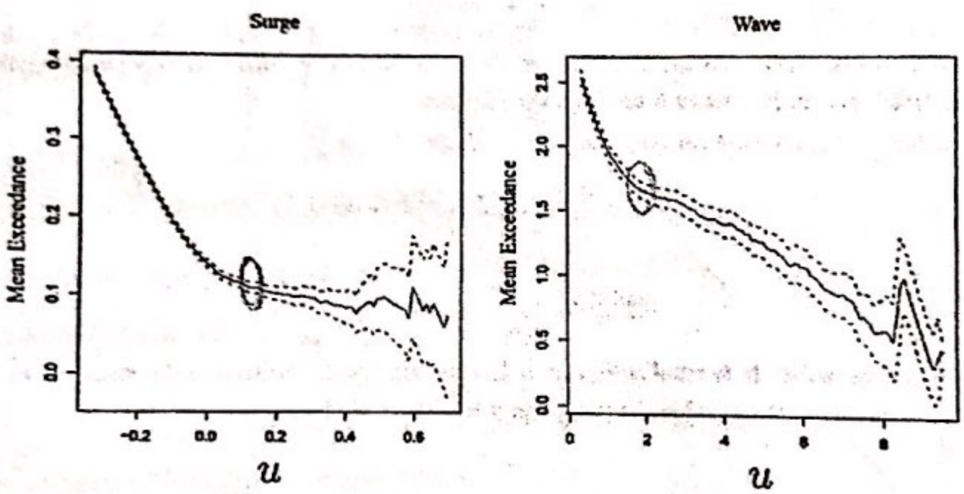
2. Quel lien relie les distributions GEV (Generalised Extreme Value distibutions) et les distributions GPD (Generalised Pareto distributions)?

ensae 2018

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

isfa 2014

4. Quelle utilisation fait-on de la fonction de dépassement moyen empirique? Vous trouverez ci-dessous deux représentations graphiques de la fonction de dépassement moyen empirique pour deux grandeurs physiques (Surge et Wave) en fonction d'un seuil u. Quel comportement de la fonction de dépassement est attendu pour les grandes valeurs de u?



à vérifier → **Exercice 1 (2 points):**

Soient (X_i) une suite de variables i.i.d. de loi absolument continue et $M_n = \max(X_1, \ldots, X_n)$. On définit le temps du premier record de la manière suivante :

$$L = \inf\{j > 1 : M_j > X_1\}.$$

1. Donner la loi de L.
2. Montrer que si $y > x$

$$\Pr(X_L > y | X_1 = x) = \frac{\bar{F}(y)}{\bar{F}(x)}.$$

Indication: On décomposera la probabilité par rapport aux valeurs possibles de L et on utilisera la formule des probabilités totales.

ok **Exercice 2 (3 points):**

Soient $X_1, ..., X_n$ une suite de variables aléatoires indépendantes de loi Gaussienne centrée et réduite, et $M_n = \max(|X_1|, ..., |X_n|)$. On a dans ce cas là:

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda$$

où Λ a une distribution de Gumbel et

$$\begin{aligned} a_n &= (2\ln(2n))^{-1/2} \\ b_n &= (2\ln(2n))^{1/2} - \frac{1}{2}(2\ln(2n))^{-1/2}[\ln(\ln(2n)) + \log(4\pi)]. \end{aligned}$$

isfa 2014

ok 1. Montrer que

$$\frac{M_n}{b_n} \xrightarrow{P} 1.$$

On cherche à caractériser la loi asympotique du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.

2. En utilisant une identité remarquable, montrer que si

$$\frac{\max(|X_1|, ..., |X_n|) - b_n}{a_n} \xrightarrow{d} \Lambda,$$

ok alors

$$\frac{\max(X_1^2, ..., X_n^2) - b_n^2}{2a_nb_n} \xrightarrow{d} \Lambda.$$

En déduire les constantes de normalisation pour la convergence du maximum de variables aléatoires indépendantes avec une distribution du Khi-deux à un degrès de liberté.

?

ok **Exercice 2 (7 points):**

On considère la distribution donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp\left(-\frac{x}{1-x}\right) & \text{si } 0 \le x \le 1, \\ 1 & \text{si } x > 1. \end{cases}$$

ensae 2018

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID de distribution F et $M_n = \max(X_1, ..., X_n)$.

0. Est-ce que M_n peut avoir une distribution non dégénérée avec une suite de seuils non-linéaires?

1. En choisissant

$$a_n = \frac{1}{(1+\ln n)^2} \quad \text{et} \quad b_n = \frac{\ln n}{1+\ln n},$$

trouver la loi limite de $a_n^{-1}(M_n - b_n)$.

Soient $E_1, ..., E_n$ une suite de variables aléatoires IID de distribution exponentielle de paramètre 1 et $M_n^E = \max(E_1, ..., E_n)$.

2. Donner les constantes de normalisation a_n^E et b_n^E telles que $(a_n^E)^{-1}(M_n^E - b_n^E)$ converge en distribution vers la loi de Gumbel.

3. On considère la fonction

$$g(x) = \frac{x}{1+x}, \quad x > 0.$$

a. Montrer que $X_1 \stackrel{d}{=} g(E_1)$. En déduire que

$$\frac{M_n - b_n}{a_n} \stackrel{d}{=} \frac{g(M_n^E) - g(\ln n)}{g'(\ln n)}.$$

b. Montrer qu'il existe ζ_n tel que

$$g(M_n^E) - g(\ln n) = (M_n^E - \ln n)\, g'(\zeta_n)$$

avec $\ln n \leq \zeta_n \leq M_n^E$ si $\ln n \leq M_n^E$ et inversement.

c. Montrer que

$$\frac{M_n^E}{\ln n} \xrightarrow{P} 1,$$

et en déduire que

$$\frac{g'(\zeta_n)}{g'(\ln n)} \xrightarrow{P} 1.$$

Retrouver le résultat de la question 1.

Insae 2018

$\exists c \in [a,b]$ tq $g'(c)(b-a) = g(b) - g(a)$

Annale TVE - Freire 2014

Exercice 1

1) $\mathbb{P}(L=k) = \mathbb{P}(X_k > X_1, X_{k-1} < X_1, \dots, X_2 < X_1)$
$= \mathbb{P}(X_k > x_1, X_{k-1} < x_1, \dots, X_2 < x_1 \mid X_1 = x_1)\, \mathbb{P}(X_1 = x_1)$
$= \mathbb{P}(X_k > x_1)\, \mathbb{P}(X_{k-1} < x_1) \dots \mathbb{P}(X_2 < x_1)\, \mathbb{P}(X_1 = x_1)$
$= \int_{-\infty}^{+\infty} \bar{F}_{X_k}(x_1) F_{X_{k-1}}(x_1) \dots F_{X_2}(x_1) f_{X_1}(x_1)\, dx_1$
$= \int_{-\infty}^{+\infty} F(x_1)^{k-2} \bar{F}(x_1) f(x_1)\, dx_1$

On pose $u = F(x)$ alors $du = f(x)dx$

Si $x \to -\infty$, $u \to 0$
Si $x \to +\infty$, $u \to 1$

donc, $\mathbb{P}(L=k) = \int_0^1 u^{k-2}(1-u)\,du$
$= \int_0^1 (u^{k-2} - u^{k-1})\,du$
$= \left[\frac{u^{k-1}}{k-1}\right]_0^1 - \left[\frac{u^k}{k}\right]_0^1$
$= \frac{1}{k-1} - \frac{1}{k}$
$= \underline{\frac{1}{k(k-1)}}$

2) $\mathbb{P}(X_L > y \mid X_1 = x) = \mathbb{E}[\mathbb{1}_{\{X_L > y\}} \mid X_1 = x]$
$= \mathbb{E}_x[\mathbb{1}_{\{X_L > y\}}]$
$= \mathbb{E}_x[\sum_{k=2} \mathbb{1}_{\{X_k > y\}} \mathbb{1}_{L=k}]$
$= \mathbb{E}_x[\sum \mathbb{1}_{\{X_k > y\}} \mathbb{1}_{\{X_2 \leq x\}} \dots \mathbb{1}_{\{X_{k-1} \leq x\}}]$
$= \sum_{k=2}^{\infty} \mathbb{E}_x[\mathbb{1}_{X_k > y}]\, \mathbb{E}_x[\mathbb{1}_{\{X_2 \leq x\}}] \dots \mathbb{E}_x[\mathbb{1}_{\{X_{n-1} \leq x\}}]$
$= \sum \bar{F}(y) F(x)^{k-2}$
$= \bar{F}(y) \sum F(x)^{k-1}$
$= \bar{F}(y) F^{-1}(x) \frac{F(x)}{1-F(x)}$
$= \underline{\frac{\bar{F}(y)}{\bar{F}(x)}}$

$\sum_{k=1}^{n} u^k = \frac{u(1-u^n)}{1-u}$

Examen Théorie des Valeurs Extrêmes 2015-2016

ENSAE

Sans document - Sans calculatrice - 2h00

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

ensae2018 1. Que signifie F appartient au domaine d'attraction d'une distribution max-stable G ($F \in D(G)$)?

ista2012 2. Quel lien relie les distributions GEV (Generalised Extreme Value distributions) et les distributions GPD (Generalised Pareto distributions)?

ensae2018 3. On suppose que deux distributions F et G ont le même point extrémal ($x^F = x^G$) 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 extrémal 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

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

Montrer que

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

et caractériser c.

ensae2018 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.

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

- 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 $X_1, \ldots, X_n$ n variables positives, indépendantes, identiquement distribuées de distribution caractérisée par $\bar{F}(x) = x^{-\alpha}L(x)$. On note $S_n = X_1 + \ldots + X_n$. 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 $M_n = \max(X_1, \ldots, X_n)$. Utiliser la formule du binôme de Newton.

Interpréter ce résultat en terme de gestion des risques extrêmes.

Exercice 2 (4 points):

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

0. Donner a_n et b_n pour que

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

On définit

$$Y_n = \frac{1}{3}\max(X_n, X_{n+1}, X_{n+2}).$$

1. Donner la distribution de Y_n.
2. Soit $M_n^Y = \max(Y_1, \ldots, Y_n)$. 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 $\{i/n : X_i > a_n x + b_n\}_{i=1,\ldots,n}$ et celle des dépassements de seuils $\{i/n : Y_i > a_n x + b_n\}_{i=1,\ldots,n}$ quand $n \to \infty$?

Annale TVE Ensae 2016.

Questions de cours

1) cf ensae 2018
2) cf isfa 2012
3) cf ensae 2018
4) cf ensae 2018

Exercice 2 (questions de cours)

– Benktander de type I – $(x^F = +\infty)$

$$\bar{F}(x) = \left(1 + 2\frac{\beta}{n}\ln x\right)\exp\{-(\beta(\ln(x)^2) + (\alpha+1)\ln x)\} \underset{x\to+\infty}{\sim} 2\frac{\beta}{n}\ln(x)e^{-\beta(\ln x)^2}$$

$$f(x) = -\bar{F}'(x) = -\left[\frac{2\beta}{n}x^{-1}e^{-\beta(\ln x)^2} + 2\frac{\beta}{n}\ln(x)\left[-2\beta\frac{\ln(x)}{x}\right]e^{-\beta(\ln x)^2}\right]$$

alors $$h(x) = \frac{\bar{F}(x)}{f(x)} = \frac{2\frac{\beta}{n}\ln(x)}{-\left[\frac{2\beta}{n}x^{-1} + \frac{2\beta}{n}\ln(x)\left[-2\beta\frac{\ln(x)}{x}\right]\right]}$$

$$= \frac{x\ln(x)}{-1 + 2\beta(\ln(x))^2}$$

donc $$h'(x) = \frac{[\ln(x)+1][2\beta(\ln(x))^2 - 1] - x\ln(x)\left[4\beta\frac{\ln(x)}{x}\right]}{(2\beta(\ln(x))^2 - 1)^2}$$

$$= \frac{\ln(x)}{(2\beta\ln(x)^2 - 1)} + \frac{2\beta\ln(x)^2 - 1 - 4\beta\ln(x)^2}{(2\beta\ln(x)^2 - 1)^2}$$

$$= \frac{1}{2\beta\ln(x) - \frac{1}{\ln(x)}} - \frac{1}{(2\beta\ln(x)^2 - 1)^2} - \frac{1}{(2\beta\ln x^2 - 1)\left(1 - \frac{1}{2\beta\ln x^2}\right)}$$

donc $h'(x) \xrightarrow[x\infty]{} 0 = \xi \Rightarrow$ Gumbel

– Benktander de type II – $(x^F = +\infty)$

$$\bar{F}(x) = x^{-(1-\beta)}\exp\left\{-\frac{\alpha}{\beta}(x^\beta - 1)\right\} = \exp\left\{-(1-\beta)\ln(x) - \frac{\alpha}{\beta}(x^\beta - 1)\right\} \underset{x\infty}{\sim} e^{-\frac{\alpha}{\beta}(x^\beta - 1)}$$

$$f(x) = -\bar{F}'(x) = +\frac{\alpha}{\beta}\beta x^{\beta-1}e^{-\frac{\alpha}{\beta}(x^\beta - 1)} = +\alpha x^{\beta-1}\exp\left\{-\frac{\alpha}{\beta}(x^\beta - 1)\right\}$$

donc $$h(x) = \frac{\bar{F}(x)}{f(x)} = \frac{1}{+\alpha x^{\beta-1}}$$

donc $$h'(x) = \frac{-\alpha(\beta-1)x^{\beta-2}}{(\alpha x^{\beta-1})^2} = \frac{1-\beta}{\alpha}x^{\beta-2-2\beta+2} = \frac{1-\beta}{\alpha x^\beta} \xrightarrow[x\infty]{} 0$$

donc $\xi = 0 \Rightarrow$ Gumbel

- **Personne de type I** $(x^F = 1)$

$$\bar{F}(x) = \exp\left\{-\frac{x}{1-x}\right\}$$

$$f(x) = -\bar{F}'(x) = -\left[\frac{-(1-x)-x}{(1-x)^2}\right] e^{-\frac{x}{1-x}} = \frac{1}{(1-x)^2} e^{-\frac{x}{1-x}}$$

donc $h(x) = (1-x)^2$

et $h'(x) = -2(1-x) = 2(x-1) \xrightarrow[x\to 1]{} 0^-$

donc $\xi = 0 \Rightarrow$ Gumbel

- **Personne de type II** $(x^F = +\infty)$

$$\bar{F}(x) = \frac{1}{(\ln \ln(x))^{\ln \ln(x)}}$$

$$= \exp\{-\ln(\ln(x)) \ln(\ln(\ln(x)))\}$$

$$= \frac{\ln(\ln(\ln(x)))}{\ln(x)}$$

$$f(x) = -\bar{F}'(x) = \frac{-\frac{1}{\ln(\ln(x))} \times \frac{1}{\cancel{\ln(x)}} \times \frac{1}{x} \times \cancel{\ln(x)} + \ln[\ln(\ln x)] \times \frac{1}{x}}{\ln(x)^2}$$

alors $h(x) = \dfrac{\bar{F}(x)}{f(x)} = \dfrac{\ln(\ln(\ln x))}{\cancel{\ln(x)}} \times \dfrac{x[\ln(x)]^{\cancel{2}}}{\ln[\ln(\ln(x))] - \frac{1}{\ln(\ln x)}} = \dfrac{\ln(\ln x) \times \ln(\ln(\ln x)) \times x \ln x}{\ln(\ln x) \ln[\ln(\ln x)] - 1}$

donc $h'(x) = \dfrac{[\ln(\ln \ln x) + 1 + \ln(\ln x) \ln(\ln \ln x)(\ln x + 1)][\ln(\ln x) \ln(\ln \ln x) - 1] - \ln(\ln x) \ln(\ln \ln x)[\ln(\ln \ln x) + 1]}{(\ln(\ln x) \ln(\ln \ln x) - 1)^2}$

$$= \frac{\ln(\ln \ln x) + 1 + \ln(\ln x) \ln(\ln \ln x)(\ln x + 1)}{\ln(\ln x) \ln(\ln \ln x) - 1} - \frac{\ln(\ln x) \ln(\ln \ln x)(\ln(\ln \ln x) + 1)}{(\ln(\ln x) \ln(\ln \ln x) - 1)^2}$$

or $\ln(x) = -\ln\left(\frac{1}{x}\right) = -\ln\left(1 + \frac{1}{x} - 1\right) \underset{x\infty}{\sim} 1 - \frac{1}{x}$ $\quad (x \to \infty$ alors $\frac{1}{x} \to 0)$

$\ln(\ln x) \underset{x\infty}{\sim} \ln\left(1 - \frac{1}{x}\right) \sim -\frac{1}{x}$

$\ln(\ln \ln x) \underset{x\infty}{\sim} \ln\left(-\frac{1}{x}\right) = \ln\left(1 - \frac{1}{x} - 1\right) = -\left(1 + \frac{1}{x}\right)$

donc $h'(x) \underset{x\infty}{\sim} \dfrac{-\left(1+\frac{1}{x}\right) + 1 - \left(-\frac{1}{x}\right)\left(1+\frac{1}{x}\right)\left(1 - \frac{1}{x} + 1\right)}{\left(\frac{1}{x}\left(1+\frac{1}{x}\right) - 1\right)} - \dfrac{\frac{1}{x}\left(1+\frac{1}{x}\right)\left(1 - \left(1 + \frac{1}{x}\right)\right)}{\left(\frac{1}{x}\left(1+\frac{1}{x}\right) - 1\right)^2}$

$$\underset{x\infty}{\sim} \frac{-1 + \left(1 + \frac{1}{x}\right)\left(2 - \frac{1}{x}\right)}{1 + \frac{1}{x} - x} + \frac{\left(1 + \frac{1}{x}\right)}{\left(1 + \frac{1}{x} - x^2\right)^2}$$

$$\underset{x\infty}{\sim} \frac{1}{1-x} + \frac{1}{x^4} \xrightarrow[x\infty]{} 0$$

donc $\xi = 0$ Gumbel

Exercice 1

1) Supposons $\omega \in \{X_1 > x\} \cup \{X_2 > x\}$

alors $X_1(\omega) > x$ et $X_2(\omega) > x$

or $X_1(\omega) \geqslant 0$ et $X_2(\omega) \geqslant x$

donc $X_1(\omega) + X_2(\omega) > 2x \geqslant x$

donc $\omega \in \{X_1 + X_2 > x\}$ d'où l'inclusion.

alors $\mathbb{P}(X_1 > x \text{ ou } X_2 > x) \leq \mathbb{P}(X_1 + X_2 > x)$

$\Leftrightarrow 2\mathbb{P}(X_1 > x) - \mathbb{P}(X_1 > x, X_2 > x) \leq \mathbb{P}(X_1 + X_2 > x)$ $(X_1, X_2 \text{ iid})$

$\Leftrightarrow 2\mathbb{P}(X_1 > x) - \mathbb{P}(X_1 > x)^2 \leq \mathbb{P}(X_1 + X_2 > x)$

Pour $x \to \infty$, $\mathbb{P}(X_1 > x) = o(\mathbb{P}(X_1 > x)) = o(1)$

donc $2\mathbb{P}(X_1 > x)(1 - o(1)) \leq \mathbb{P}(X_1 + X_2 > x)$

2) Supposons que $\omega \in \{X_1 + X_2 > x\}$

$0 < \delta < 1/2$, $X_1(\omega) + X_2(\omega) > x$

alors $X_1(\omega) > (1-\delta)x$ ou $X_2(\omega) > (1-\delta)x$ ou $X_1(\omega) > \delta x$ et $X_2(\omega) > \delta x$

$$\mathbb{P}(X_1 + X_2 > x) \leq \mathbb{P}(X_1 > (1-\delta)x) + \mathbb{P}(X_2 > (1-\delta)x) + \mathbb{P}(X_1 > \delta x, X_2 > \delta x)$$

$$\leq 2\mathbb{P}(X_1 > (1-\delta)x) + \mathbb{P}(X_1 > \delta x)^2$$

$$\leq 2\,\mathbb{P}(X_1 > (1-\delta)x)(1 + o(1)).$$

3) D'après (1) et (2)

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

donc si $\delta \to 0$, $\mathbb{P}(X_1 + X_2 > x) \sim 2\mathbb{P}(X_1 > x)$

4) Récurrence :

- Initialisation, d'après (3), $\mathbb{P}(X_1 + X_2 > x) \sim 2\mathbb{P}(X_1 > x)$
- Hérédité : $\mathbb{P}(X_1 + \dots + X_{n+1} > x) = \mathbb{P}(X_1 + \dots + X_n > x)\mathbb{P}(X_{n+1} > x)$

$$\sim n\,\mathbb{P}(X_1 > x) \times \mathbb{P}(X_1 > x)$$

$$\sim (n+1)\,\mathbb{P}(X_1 > n)$$

donc, $\mathbb{P}(S_n > x) \sim n\mathbb{P}(X_1 > x)$

de plus, $\mathbb{P}(M_n > x) = 1 - F^n(x) = 1 - [1 - \bar{F}(x)]^n$

or $[1 - \bar{F}(x)]^n = \sum_{k=0}^{n} C_n^k \, 1^{n-k} (-\bar{F}(x))^k$

$$= \sum_{k=0}^{n} C_n^k (-1)^k (\bar{F}(x))^k$$

$$= 1 - n\bar{F}(x) + \underbrace{\sum_{k=2}^{n} C_n^k \bar{F}^k(x)(-1)^k}_{\text{presque nul quand } n \to \infty}$$

$$\underset{n\infty}{\sim} 1 - n\bar{F}(x)$$

donc $\mathbb{P}(M_n > x) \sim N - (1 - n\bar{F}(x))$
$\sim n\,\mathbb{P}(X_1 > x)$

$\Rightarrow \underline{\mathbb{P}(S_n > x) \sim \mathbb{P}(M_n > x)}$

S_n représente la somme de sinistres sur 1 période. On s'intéresse à la proba que S_n dépasse un montant élevé $(x \to \infty)$, ce qui équivaut à s'intéresser à la proba que le montant du plus grand sinistre dépasse ce seuil x.

Exercice 2

0) $$\mathbb{P}\left(\frac{M_n^X - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n^X \leq a_n x + b_n)$$
$$= F^n(a_n x + b_n)$$
$$= \exp\{-n(a_n x + b_n)^{-1}\} \longrightarrow e^{-\frac{1}{x}}$$

donc $\underline{a_n = n \text{ et } b_n = 0}$

1) $$\mathbb{P}(Y_n \leq x) = \mathbb{P}(\max(X_n, X_{n+1}, X_{n+2}) \leq 3x)$$
$$= \mathbb{P}(X_n \leq 3x)\,\mathbb{P}(X_{n+1} \leq 3x)\,\mathbb{P}(X_{n+2} \leq 3x) \quad \begin{cases} (X_i) \text{ ind.} \\ X_i \text{ id. dist} \end{cases}$$
$$= \Phi_1(3x)^3$$
$$= \exp\left\{-\frac{3}{3x}\right\}$$
$$= \exp\{-1/x\} \qquad \text{donc } \underline{Y_n \sim \Phi_1}$$

2) $$\mathbb{P}\left(\frac{M_n^Y - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n^Y \leq a_n x + b_n)$$
$$= \mathbb{P}(M_n^Y \leq nx)$$
$$= F_Y^n(nx) \qquad (Y \sim \Phi_1)$$
$$= \exp\{-n/nx\}$$
$$= \exp\{-1/x\}$$

donc $\underline{\frac{M_n^Y - b_n}{a_n} \sim \Phi_1}$

$\Rightarrow$ max-stable.

Examen Théorie des Valeurs Extrêmes 2016-2017

ENSAE

Sans document - Sans calculatrice - 2h00

Rabat 2018

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 ($F \in D(G)$)?

On rappelle que la distribution de Pareto généralisée GPD(β, ξ) est définie par

$$G_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{si } \xi = 0, \end{cases}$$

Expliquer à quel domaine d'attraction appartient la distribution de Pareto généralisée GPD(β, ξ).

2. Soit B une variable aléatoire de distribution Binomiale(n,p)

$$P(B = k) = C_n^k p^k (1-p)^{n-k}$$

et N une variable aléatoire de distribution Poisson(λ). On pose $\lambda = np$. Montrer que

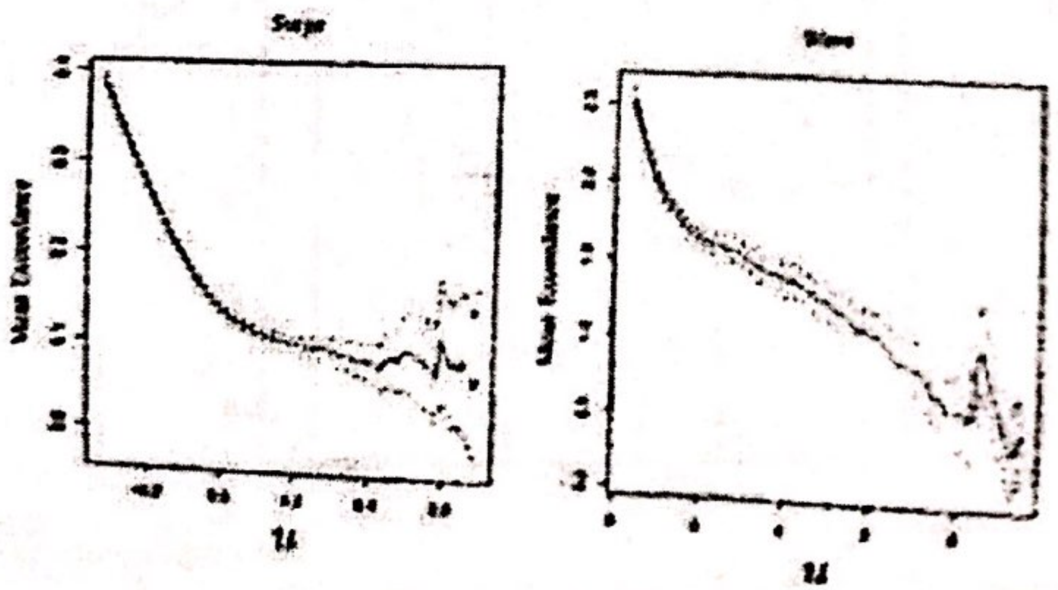
$$P(B = k) = P(N = k)\left(\frac{(1-\frac{\lambda}{n})^n}{e^{-\lambda}} \frac{1}{(1-\frac{\lambda}{n})^k} \prod_{i=1}^{k}\left(1 - \frac{i-1}{n}\right)\right)$$

En déduire que si $n \to \infty$, $p \to 0$ tel que $\lambda = np$, alors B converge en distribution vers N.

Expliquer comment ce résultat peut être utilisé pour caractériser la loi du maximum ou des statistiques d'ordre élevé.

Isfa 2014

3. Quelle utilisation fait-on de la fonction de dépassement moyen empirique? Vous trouverez ci-dessous deux représentations graphiques de la fonction de dépassement moyen empirique pour deux grandeurs physiques (Surge et Wave) en fonction d'un seuil u. Quel comportement de la fonction de dépassement est attendu pour les grandes valeurs de u?



Exercice 1 (6 points):

A. On considère la distribution $F(x) = 1 - e^{1/x}$ pour $x < 0$, et $F(x) = 1$ pour $x \geq 0$.

1. Montrer que la suite

$$u_n = \frac{1}{\ln \tau - \ln n}$$

assure que la condition $n\bar{F}(u_n) \to \tau > 0$ est satisfaite. En déduire la limite de $P(M_n \leq u_n)$ si $M_n = \max(X_1, ..., X_n)$ avec (X_i) i.i.d. de loi F quand $n \to \infty$.

2. En prenant $\tau = e^{-x}$ et en utilisant un développement limité, en déduire qu'il existe $(a_n) > 0$ et (b_n) tel que

$$\lim_{n \to \infty} P(M_n \leq a_n x + b_n) = \exp(-\exp(-x)).$$

B. On considère la distribution $F(x) = K(1 - e^{-x})$ pour $0 < x < x_F$.

1. Donner la valeur de K en fonction de x_F.

2. Trouver $(a_n) > 0$ tel que, pour $x < 0$,

$$\lim_{n \to \infty} P(M_n \leq a_n x + x_F) = \exp(x).$$

Que concluez-vous de cet exercice?

Exercice 2 (5 points):

Soit (X_i) des variables aléatoires i.i.d. de loi F. On suppose qu'il existe deux suites $(u_n^{(1)})$ et $(u_n^{(2)})$ telles que pour $i = 1, 2$

$$n\bar{F}(u_n^{(i)}) \to \tau_i$$

avec $0 < \tau_1 \leq \tau_2$.

1. On définit $N_n^{(i)} = \sum_{i=1}^n \mathbb{1}_{\{X_i \leq u_n^{(i)}\}}$. Il est possible alors de montrer que

$$\lim_{n \to \infty} P\left(N_n^{(1)} = k_1, N_n^{(2)} = k_1 + k_2\right) = \frac{\tau_1^{k_1}}{k_1!} \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!} e^{-\tau_2}.$$

Quel résultat retrouve-t-on si $\tau_1 = \tau_2$.

2. Exprimer l'événement

$$\{X_{(1)} \leq u_n^{(1)}, X_{(2)} \leq u_n^{(2)}\}$$

en fonction de $N_n^{(1)}$ et $N_n^{(2)}$. On rappelle que $X_{(1)}$ et $X_{(2)}$ sont les deux plus grandes statistiques d'ordre de $(X_i)_{i=1,...,n}$.

3. En supposant que F appartient au domaine d'attraction de G pour deux suites $(a_n) > 0$ et (b_n), montrer que pour $x_1 > x_2$

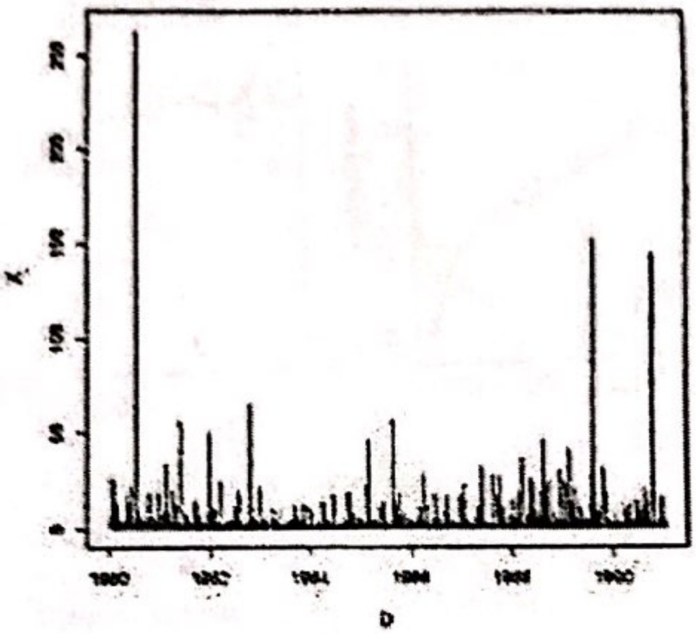
$$\begin{aligned} &\lim_{n \to \infty} P\left(X_{(1)} \leq a_n x_1 + b_n, X_{(2)} \leq a_n x_2 + b_n\right) \\ = \; & G(x_2)\{\ln G(x_1) - \ln G(x_2) + 1\}. \end{aligned}$$

Quel est l'intérêt de cette distribution?

Rabat 2018

Question méthodologique (3 points):

Un réassureur dispose des données de sinistres suivantes pour tarifier un excédent de sinistres de priorité 100 et de portée illimitée.



On rappelle qu'en cas de sinistre X_i, le réassureur rembourse à l'assureur le montant $(X_i - 100)_+$

Quelle méthodologie lui proposez-vous? Donner une réponse littérale.

Annale TVE Ensae 2017

Exercice 1 -

A_1) $n\bar{F}(u_n) = n e^{1/u_n}$

or $u_n = \dfrac{1}{\ln(\tau)-\ln(n)} = \ln\left(\frac{\tau}{n}\right)^{-1}$

donc, $n\bar{F}(u_n) = n e^{\ln(\frac{\tau}{n})} = n\frac{\tau}{n} = \tau \xrightarrow[n\infty]{} \underline{\tau}$

alors $\mathbb{P}(M_n \leq u_n) = F^n(u_n)$
$= \exp\{n \ln(1-\bar{F}(u_n))\}$
$\underset{n\infty}{\sim} \exp\{-n\bar{F}(u_n)\}$
$\xrightarrow[n\infty]{} \underline{e^{-\tau}}$

2) $\tau = e^{-x}$

On sait que $\mathbb{P}(M_n \leq u_n) = \mathbb{P}(M_n \leq \dfrac{1}{-x-\ln n}) \longrightarrow e^{-e^{-x}}$

or au voisinage de 0, on pose $f_n(x) = (x+\ln n)^{-1}$

$f_n(x) \underset{x\to 0}{\sim} f_n(0) + \dfrac{f_n'(0)}{1!}x^1$
$\underset{x\to 0}{\sim} \dfrac{1}{\ln(n)} - \dfrac{1}{\ln n^2}x$

donc $\mathbb{P}(M_n) \leq \dfrac{-1}{\ln n} + \dfrac{1}{\ln n^2}x) \longrightarrow e^{-e^{-x}}$ d'où $\underline{a_n = \dfrac{-1}{\ln n^2} \text{ et } b_n = \dfrac{-1}{\ln n}}$

B_1) On sait que $0 \leq F(x) \leq 1 \Rightarrow 0 \leq K(1-e^{-x}) \leq 1$

alors, $F(x) \geq 0 \Rightarrow K(1-e^{-x}) \geq 0 \Rightarrow \underline{K \geq 0}$
Si $x^F = +\infty$, alors $K \leq 1$
Si $x^F < +\infty$, alors $K \leq (1-e^{-x_F})^{-1}$
$\Rightarrow \underline{K \in \left]0 ; \dfrac{1}{1-e^{-x_F}}\right[}$

2) $\mathbb{P}(M_n \leq a_n x + x_F) = F^n(a_n x + x_F) = [K(1-e^{-(a_n x + x_F)})]^n = e^{n[\ln K + \ln(1-e^{-a_n x - x_F})]}$
$= e^{n[\ln K - \exp\{-a_n x - x_F\}]}$
$\underset{n\infty}{\sim} e^{n[\ln K - [1-a_n x - x_F]]}$
$\underset{n\infty}{\sim} e^{n a_n x}$
$\underset{n\infty}{\sim} e$ (lim du terme de + haut degré d'1 fonction polynomiale)

donc $\underline{a_n = 1/n}$.

$\Rightarrow$ la distribution des max des (X_i) iid de loi F est non dégénérée.

Exercice 2 -

1_ Si $\tau_1 = \tau_2$, $\mathbb{P}(N_n^{(1)} = k_1, N_n^{(2)} = k_1 + k_2) \xrightarrow[n\infty]{} \dfrac{\tau_2^{k_1}}{k_1!}\dfrac{(\tau_2-\tau_2)^{k_2}}{k_2!}e^{-\tau_2}$

alors $P(N_n^{(1)} = k_1, N_n^{(2)} = k_1 + k_2) \xrightarrow[n\to\infty]{} \begin{cases} 0 & \text{si } k_2 \neq 0 \\ \frac{\tau_1^{k_1}}{k_1!} e^{-\tau_1} & \text{si } k_2 = 0 \end{cases}$

On retrouve le fait que la distribution limitée du nombre de dépassement de seuil est une dist. de Poisson de paramètre $\tau_1 > 0$.

2) $\{X_{(1)} \leq u_n^{(1)}, X_{(2)} \leq u_n^{(2)}\}$:=: la 1ère et 2e valeurs des (X_i) sont respectivement $\leq u_n^{(1)}$ et $u_n^{(2)}$.

:=: aucune valeur dépasse $u_n^{(1)}$ et 1 seule valeur dépasse $u_n^{(2)}$

:=: $\{N_n^{(1)} = n, N_n^{(2)} = n-1\}$

3) $\mathbb{P}(X_{(1)} \leq a_n x_1 + b_n, X_{(2)} \leq a_n x_2 + b_n) = \mathbb{P}(X_{(1)} \leq u_n^{(1)}, X_{(2)} \leq u_n^{(2)})$

$= \mathbb{P}(N_n^{(1)} = n, N_n^{(2)} = n-1)$

$= \mathbb{P}(u_n^{(2)} < X_{(1)} \leq u_n^{(1)})$

$= \mathbb{P}(X_{(1)} \leq u_n^{(1)}) - \mathbb{P}(X_{(1)} \leq u_n^{(2)})$

$= \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x_1\right) - \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x_2\right)$

$= F^n(u_n^{(1)}) - F^n(u_n^{(2)})$

exo 2 Q2

Examen Théorie des Valeurs Extrêmes 2017-2018

ENSAE

Sans document - Sans calculatrice - 2h00

Questions de cours et exercices d'applications du cours (10 points):

1. Expliquer ce que veut dire "F appartient au domaine d'attraction d'un loi max-stable G ($F \in D(G)$)".

Quel est le domaine d'attraction de la loi logistique définie par

$$P(Y \leq y) = \frac{1}{1+e^{-y}}, \quad y \in \mathbb{R}.$$

2. Quelles sont les assertions vraies? Justifier vos réponses.

A) Si une distribution F appartient au domaine d'attraction de la $GEV(\mu,\sigma,\xi)$, alors la distribution limite de ses dépassements seuils (correctement normalisés) convergent en loi vers une loi $GPD(\beta,\xi)$.

B) Si une distribution F appartient au domaine d'attraction de la $GPD(\beta,\xi)$, alors la distribution limite de ses dépassements seuils (correctement normalisés) convergent en loi vers une loi $GEV(\mu,\sigma,\xi)$.

C) La loi conditionnelle d'une $GPD(\beta,\xi)$ sachant qu'elle dépasse un niveau x est une $GPD(\beta,\xi)$.

D) La loi Uniforme sur $[0,\beta]$ appartient à la famille des lois $GPD(\beta,\xi)$.

E) Les lois $GPD(\beta,\xi)$ sont utilisées en réassurance pour tarifer des Quote-parts.

3. On suppose que deux distributions F et G ont le même point extrémal ($x^F = x^G$) 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 extrémal 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 1 (7 points):

On considère la distribution donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp\left(-\frac{x}{1-x}\right) & \text{si } 0 \le x \le 1. \\ 1 & \text{si } x > 1. \end{cases}$$

Soient $X_1, ..., X_n$ une suite de variables aléatoires IID de distribution F et $M_n = \max(X_1, ..., X_n)$.

0. Est-ce que M_n peut avoir une distribution non dégénérée avec une suite de seuils non-linéaires?

1. En choisissant

$$a_n = \frac{1}{(1+\ln n)^2} \quad \text{et} \quad b_n = \frac{\ln n}{1+\ln n},$$

trouver la loi limite de $a_n^{-1}(M_n - b_n)$.

Soient $E_1, ..., E_n$ une suite de variables aléatoires IID de distribution exponentielle de paramètre 1 et $M_n^E = \max(E_1, ..., E_n)$.

2. Donner les constantes de normalisation a_n^E et b_n^E telles que $(a_n^E)^{-1}(M_n^E - b_n^E)$ converge en distribution vers la loi de Gumbel.

3. On considère la fonction

$$g(x) = \frac{x}{1+x}, \quad x > 0.$$

a. Montrer que $X_1 \stackrel{d}{=} g(E_1)$. En déduire que

$$\frac{M_n - b_n}{a_n} \stackrel{d}{=} \frac{g(M_n^E) - g(\ln n)}{g'(\ln n)}.$$

b. Montrer qu'il existe ζ_n tel que

$$g(M_n^E) - g(\ln n) = (M_n^E - \ln n)\, g'(\zeta_n)$$

avec $\ln n \le \zeta_n \le M_n^E$ si $\ln n \le M_n^E$ et inversement.

c. Montrer que

$$\frac{M_n^E}{\ln n} \xrightarrow{P} 1,$$

et en déduire que

$$\frac{g'(\zeta_n)}{g'(\ln n)} \xrightarrow{P} 1.$$

Retrouver le résultat de la question 1.

Exercice 2 (3 points):

Supposons que les variables (X_i) sont des variables iid de distributions F. Posons $M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$ et supposons qu'il existe des constantes $a_n > 0$, $\alpha_n > 0$, b_n, β_n telles que

$$P(M_n \leq a_n x + b_n) \to G_1(x) \quad \text{et} \quad P(-m_n \leq \alpha_n x + \beta_n) \to G_2(x).$$

1. Montrer la convergence

$$P(M_n \leq a_n x + b_n, -m_n \leq \alpha_n y + \beta_n) \to G_1(x) G_2(y).$$

Qu'en concluez-vous?

2. On considère le cas où les (X_i) sont des variables aléatoires Gaussiennes centrées et réduites. On rappelle que dans ce cas

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda$$

où Λ a une distribution de Gumbel et

$$\begin{aligned} a_n &= (2\ln(2n))^{-1/2} \\ b_n &= (2\ln(2n))^{1/2} - \frac{1}{2}(2\ln(2n))^{-1/2}[\ln(\ln(2n)) + \ln(4\pi)]. \end{aligned}$$

Montrer que $(M_n + m_n)/a_n$ converge en distribution et caractériser sa loi.

TVE _ Annale ENSAE 2018.

Questions de cours

1) On dit que F appartient au domaine d'attraction de G ($F \in \mathcal{D}(G)$) s'il existe 2 suites $(a_n)>0$ et (b_n) telle que la convergence suivante ait lieu :
$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = [F(a_n x + b_n)]^n \longrightarrow G(x)$.

Domaine d'attraction : $h(x) = \frac{1-F(x)}{f(x)} = \frac{1-F(x)}{F'(x)}$ alors $h'(y) \underset{y \to x^F}{\longrightarrow} \xi$
où x^F point extrémal de D_F.

On a $\mathbb{P}(Y \leq y) = (1+e^{-y})^{-1}$, $y \in \mathbb{R}$ $\Rightarrow$ $x^F = +\infty$ et $F'(y) = -1 \times -e^{-y}(1+e^{-y})^{-2}$

$\Rightarrow h(y) = \frac{1-F(y)}{F'(y)} = \frac{1 - \frac{1}{1+e^{-x}}}{\frac{e^{-x}}{(1+e^{-x})^2}} = (1+e^{-x}) \frac{1+e^{-x}-1}{e^{-x}} = 1+e^{-x}$.

$\Rightarrow h'(y) \longrightarrow 0$

donc $\xi = 0$ donc domaine d'attraction de Gumbel $= \mathbb{R}$.

2) A _ Vraie (th cours 4 p5)
B _ Faux _ distribution limite des dépassements seuils $\rightarrow GPD(\beta, \xi)$
C _ Faux _ Si $X \sim GPD(\beta, \xi)$ alors $X - x \mid X > x \sim GPD(\beta + \xi x, \xi)$
$(X \mid X > x \sim GPD(\beta + (\xi + 1)x, \xi))$

D _ GPD :
- $\xi > 0$: Pareto (type III)
- $\xi = 0$: $X \sim \mathcal{E}(1/\beta)$
- $\xi < 0$ (distribution à supports bornés)
- $\xi = -1$: $X \sim \mathcal{U}[0, \beta]$

Faux _ La loi Uniforme sur $[0, \beta] \in GPD(\beta, -1)$.
E _ Vraie _ En général on utilise les loss ratios pour calculer des quote parts. Mais on peut utiliser la GPD pour simuler les sinistres larges et simuler les 2 autres (attritionnels, CAT) et aggréger les 3 types _

3) • $F \in \mathcal{D}(H)$, $H \sim GEV(0, 1, \xi)$ et $H(x) = \exp\left(-(1+\xi x)_+^{-1/\xi}\right)$

$\Leftrightarrow \exists (a_n) > 0$ et (b_n) tq $\mathbb{P}\left(\frac{M_n^F - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n^F \leq a_n x + b_n) \longrightarrow H(x)$
$= [F_F(x)]^n$

$\Leftrightarrow n \ln(1-\bar{F}) \longrightarrow \ln(H(x))$
$\Leftrightarrow -n\bar{F} \longrightarrow \ln(H(x))$ $\quad (\ln(1+x) \sim x)$
$\Leftrightarrow n\bar{F} \longrightarrow -\ln(H(x))$

or par hyp, $\frac{\bar{F}}{\bar{G}} \rightarrow c$ donc $\bar{F} \rightarrow c\bar{G}$ or $n\bar{F} \rightarrow -\ln(H(x))$
donc $n\bar{G} \longrightarrow -\frac{1}{c}\ln(H(x))$ (1)

• de plus, $G \in \mathcal{D}(H)$ donc $G^n(\tilde{a}_n y + \tilde{b}_n) \longrightarrow H(y)$
$n \ln(G(\tilde{a}_n y + \tilde{b}_n)) \longrightarrow \ln(H(y))$
$n\bar{G}(\tilde{a}_n y + \tilde{b}_n) \longrightarrow -\ln(H(y))$ (2)

D'après (1) et (2), $\frac{1}{c}\ln(H(x)) = \ln(H(y))$ $\Leftrightarrow$ $(1+\xi y)_+^{-1/\xi} = \left[\frac{1}{c}(1+\xi x)^{-1/\xi}\right]_+$
$\Leftrightarrow x = \frac{1}{\xi}\left[c^{-\xi}(1+\xi y) - 1\right]$

donc $a_n x + b_n = \frac{a_n}{\xi c^{\xi}}(1+\xi y) - \frac{a_n}{\xi} + b_n = \underbrace{\frac{a_n}{c^{\xi}}}_{\tilde{a}_n} y + \underbrace{\frac{a_n}{\xi}(c^{-\xi}-1) + b_n}_{\tilde{b}_n}$

Montrer que F et G ∈ aux m̂ domaines de def $\iff \frac{h_F}{h_G} = 1$

$$h_F(x) = \frac{\bar F(x)}{-\bar F'(x)} \sim \frac{c\bar G(x)}{-c\bar G'(x)} = h_G(x) \qquad \left(\text{car } \frac{\bar F}{\bar G} \to c\right)$$

donc F et G ∈ aux m̂ domaines de def.

Réciproquement,
on sait que $F^n \to \Lambda(x)$
$\iff n\ln(F) \to \ln(\Lambda(x))$
$\iff n\ln(1-\bar F) \to \ln(\Lambda(x))$
$\sim \quad n\bar F \to -\ln(\Lambda(x)) \qquad (\ln(1+x) \sim x)$

de même $n\bar G \to -\ln(\Lambda(x+b))$

donc $\frac{\bar F}{\bar G} \to \frac{\ln(\Lambda(x))}{\ln(\Lambda(x+b))} = \frac{e^{-x}}{e^{-(x+b)}} = e^{-x}e^{x+b} = e^{b} = \text{cste}$ □.

4) Estimation des paramètres des distributions GEV:

- QQplot : donne une idée sur les modèles possibles et valeur de ξ
- Estimation par maximum de vraisemblance que si $\xi > -1/2$.
  Si $-1 < \xi < -1/2$: estimateur super-efficace (donc vraisemblance ne marche pas)
  Si $\xi < -1$, l'estimateur du point extrêmal est donné par la + grande des observ^{n}.

Estimation des paramètres des dist. GPD.

- QQplot
- Graphique de dépassement moyen (mean excess plot)

On trace $u \mapsto \hat e(u)$, si $X \sim$ GPD de seuil u alors graph ~ linéaire au delà de ce seuil.
Pente $= \xi/(1-\xi)$

- Si Pente > 0, alors $0 < \xi < 1$: si linéaire, Pareto, si concave | Lognormal Weibull $\tau < 1$
- Si Pente < 0, alors $\xi > 1$ ou $\xi < 0$
- Si cste, alors $\xi = 0$ exponentielle.

- Graphique de stabilité du param. d'échelle (stability plot)
  ⇒ trouver la valeur de ξ.

Exercice 1.

0)

1) $P(M_n \le a_n x + b_n) = F^n(a_n x + b_n)$.

or $F(x) = 1 - \exp\{-x/(1-x)\}$

$$F^n(a_n x + b_n) = \left(1 - \exp\left\{-(a_n x + b_n)/(1-(a_n x + b_n))\right\}\right)^n$$
$$= \left[1 - \exp\left(-\frac{\frac{1}{(1+\ln(n))^2}x - \frac{\ln n}{1+\ln n}}{1 - \frac{x}{(1+\ln(n))^2} - \frac{\ln n}{1+\ln n}}\right)\right]^n$$

$$F^n(a_n x + b_n) = \left[1 - \exp\left(\frac{-\frac{x}{1+\ln(n)} - \ln(n)}{-\frac{x}{1+\ln(n)} + (\ln n + 1 - \ln n)}\right)\right]^n \quad \text{simplifié} \left(\text{par } \frac{1+\ln n}{1+\ln n}\right)$$

$$= \left[1 - \exp\left(-\left(\frac{x}{1+\ln n} + \ln n\right) \times \frac{1}{\frac{-x}{1+\ln n} + 1}\right)\right]^n$$

$$\sim \left[1 - \exp\left(-\left(\frac{x}{1+\ln n} + \ln n\right)\left(1 + \frac{x}{1+\ln n}\right)\right)\right]^n \quad \frac{1}{1-u} \sim 1+u$$

$$\sim \left[1 - \exp\left(-\frac{x}{1+\ln n} - \left(\frac{x}{1+\ln n}\right)^2 - \ln n - \frac{x \ln(n)}{1+\ln n}\right)\right]^n$$

$$\sim \left[1 - \exp\left(-\frac{x}{1+\ln n} - \left(\frac{x}{1+\ln n}\right)^2 - \ln n - \left(\frac{x}{\frac{1}{\ln n}+1}\right)\right)\right]^n$$

$$\underset{n\to\infty}{\longrightarrow} \left[1 - \exp(-\ln n - x)\right]^n$$

$$\underset{n\to\infty}{\longrightarrow} \exp\{n \ln(1 - e(-\ln n - x)\} \sim \exp\{-n e^{-\ln n - x}\} \quad \ln(1-u) \sim -u$$

$$\rightarrow \exp\{-e^{-x}\} = \Lambda \quad \text{(Gumbel)}$$

2) Pour une loi de Gumbel, on sait que $F(x) = 1 - e^{-x}$ donc $f(x) = e^{-x}$

donc $h(x) = \frac{\bar{F}(x)}{f(x)} = 1$

$h'(x) = 0 \Rightarrow \xi = 0$

Par théorème, pour une $GEV(0, 1, \xi)$:

$$\begin{cases} 1 - F(b_n^E) = \frac{1}{n} \\ a_n^E = h(b_n^E) \end{cases} \Leftrightarrow \begin{cases} e^{-b_n^E} = 1/n \\ a_n^E = \frac{\bar{F}(b_n^E)}{f(b_n^E)} = 1 \end{cases} \Leftrightarrow \begin{cases} b_n^E = \ln(n) \\ a_n^E = 1 \end{cases}$$

3) a) $\mathbb{P}(g(E_1) \leq x) = \mathbb{P}\left(\frac{E_1}{1+E_1} \leq x\right)$

$= \mathbb{P}\left(E_1 \leq \frac{x}{1-x}\right)$

$= 1 - e^{-x/(1-x)}$

$= F(x)$ donc $g(E_1) \overset{d}{=} X_1$

De plus, $M_n = \max(X_1 \dots X_n)$

$\overset{d}{=} \max(g(E_1) \dots g(E_n))$

$\overset{g \nearrow}{=} g(\max(E_1 \dots E_n))$

$= g(M_n^E)$

et $g(\ln(n)) = \frac{\ln(n)}{1+\ln(n)} = b_n$

$g'(\ln(n)) = \frac{(1+\ln(n)) - \ln(n)}{(1+\ln(n))^2} = \frac{1}{(1+\ln n)^2} = a_n$

3) b) Th. d'accroissement fini : si g est continue et dérivable sur $[a,b]$ alors $\exists c \in [a,b]$ tq $g(b) - g(a) = g'(c)(b-a)$

g étant continue et dérivable

$\exists\, \xi_n$ tq $g(M_n^E) - g(\ln(n)) = (M_n^E - \ln(n))\, g'(\xi_n)$.

3) c) $\mathbb{P}\left(\left|\frac{M_n^E}{\ln n} - 1\right| > \varepsilon\right) = \mathbb{P}\left(\left|\frac{M_n^E - \ln n}{\ln(n)}\right| > \varepsilon\right)$

$= \mathbb{P}\left(\underbrace{\frac{a_n^E}{b_n^E}}_{\downarrow n\infty \; 0}\ \left|\underbrace{\frac{M_n^E - b_n^E}{a_n^E}}_{\downarrow n\infty \; 1}\right| > \varepsilon\right)$ (car $b_n^E = \ln(n)$ et $a_n^E = 1$)

$\xrightarrow[n\infty]{} 0$ donc $\frac{M_n^E}{\ln(n)} \xrightarrow[n\infty]{} 1$.

de plus, d'après 3b, on sait que $g(M_n^E) - g(\ln n) = (M_n^E - \ln n)\, g'(\xi_n)$

$\Leftrightarrow \frac{g(M_n^E) - g(\ln(n))}{g'(\ln(n))} = \frac{M_n^E - \ln(n)}{1} \times \frac{g'(\xi_n)}{g'(\ln n)}$

$\Leftrightarrow \frac{M_n - b_n}{a_n} = \underbrace{\frac{M_n^E - b_n^E}{a_n^E}}_{\downarrow\ 1} \times \underbrace{\frac{g'(\xi_n)}{g'(\ln n)}}_{\downarrow\ 1}$

donc $\frac{M_n - b_n}{a_n} \xrightarrow[n\infty]{} 1$.

en effet, $\frac{g'(\xi_n)}{g'(\ln n)} = \left(\frac{1 + \ln(n)}{1 + \xi_n}\right)^2 = \left(\frac{1 + \frac{1}{\ln n}}{\frac{1}{\ln n} + \frac{\xi_n}{\ln n}}\right)^2$

or on sait que $\ln(n) \leq \xi_n \leq M_n^E \quad (\Rightarrow)\quad 1 \leq \frac{\xi_n}{\ln(n)} \leq \frac{M_n^E}{\ln(n)} \Rightarrow \frac{\xi_n}{\ln n} \to 1$

donc $\frac{g'(\xi_n)}{g'(\ln n)} \xrightarrow{\mathbb{P}} 1$

Exercice 2.

1) les (X_i) sont iid, donc en particulier, $\overset{X_{(n)}}{M_n}$ indépendant de $\overset{X_{(1)}}{m_n}$

$\Rightarrow M_n$ ind de $-m_n$

donc $\mathbb{P}(M_n \leq a_n x + b_n,\ -m_n \leq \alpha_n y + b_n) = \mathbb{P}(M_n \leq a_n x + b_n)\,\mathbb{P}(-m_n \leq \alpha_n y + b_n)$

$\to G_1(x)\, G_2(y)$

Qu'en concluez vous ?

Cours : Q4_c)

éorie des Valeurs Extrêmes 2017-2018

Ecole d'actuariat UIR

Sans document - Sans calculatrice - 2h00

Exercices d'applications du cours (8 points):

1. Soit B une variable aléatoire de distribution Binomiale(n, p)

$$P(B = k) = C_n^k p^k (1-p)^{n-k}$$

et N une variable aléatoire de distribution Poisson(λ). On pose $\lambda = np$. Montrer que

$$P(B = k) = P(N = k)\left(\frac{\left(1-\frac{\lambda}{n}\right)^n}{e^{-\lambda}} \frac{1}{\left(1-\frac{\lambda}{n}\right)^k} \prod_{i=1}^{k}\left(1-\frac{i-1}{n}\right)\right).$$

En déduire que si $n \to \infty$, $p \to 0$ tel que $\lambda = np$, alors B converge en distribution vers N.

Expliquer comment ce résultat peut être utilisé pour caractériser la loi du maximum ou des plus grandes statistiques d'ordre.

2. Expliquer ce que veut dire "F appartient au domaine d'attraction d'un loi max-stable G ($F \in D(G)$)". On considère la variable aléatoire de distribution H telle que

$$H(x) = F(\alpha x + \beta)$$

avec $\alpha > 0$. A quelle domaine d'attraction appartient H. Donner les relations qui existent entre les suites de normalisation pour H et F.

3. On rappelle que la distribution de Pareto généralisée GPD(β, ξ) est définie par

$$G^p_{\xi,\beta}(x) = \begin{cases} 1 - [1 + \xi(x/\beta)]_+^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - e^{-x/\beta} & \text{si } \xi = 0, \end{cases}$$

Expliquer à quel domaine d'attraction appartient la distribution de Pareto généralisée GPD(β, ξ).

4. Soit $X_1, ..., X_n$ une suite de variables aléatoires IID de fonction de distribution F et soit $M_n = \max(X_1, ..., X_n)$. On rappelle que:

i) Pour un $\tau > 0$ et une suite de réels (u_n), les conditions suivantes sont équivalentes: quand $n \to \infty$

$$\begin{aligned} &1) \quad \Pr(M_n \leq u_n) \to e^{-\tau}, \\ &2) \quad n\bar{F}(u_n) = n(1 - F(u_n)) \to \tau. \end{aligned}$$

ii) Il existe une suite (u_n) satisfaisant $n\bar{F}(u_n) \to \tau$ quand $n \to \infty$ si et seulement si

$$\lim_{x \uparrow x^F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1.$$

où $x^F = \sup\{x : F(x) < 1\}$. Peut-on trouver une suite (u_n) telle que $\lim_{n\to\infty} \Pr(M_n \leq u_n) = e^{-\tau}$ si

- les X_i ont une loi Exponentielle de paramètre λ

$$\Pr(X_i > x) = \exp(-\lambda x), \quad x > 0;$$

- les X_i ont une loi Géométrique de paramètre p

$$\Pr(X_i = n) = p(1-p)^n, \quad n \in \mathbb{N};$$

- les X_i ont une loi de Poisson de paramètre λ

$$\Pr(X_i = n) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n \in \mathbb{N}.$$

Exercice 1 (4 points):

On rappelle qu'une distribution F appartient au domaine d'attraction de la loi de Gumbel Λ, si et seulement si il existe une fonction positive g telle que :

$$\lim_{x \to x^F} \frac{\bar{F}(x + tg(x))}{\bar{F}(x)} = \exp(-t), \qquad t \in \mathbb{R},$$

et qu'un choix possible pour la fonction g est donnée par

$$g(x) = \frac{\int_x^{x^F} \bar{F}(t)dt}{\bar{F}(x)} = \mathbb{E}(X - x | X > x).$$

Les coefficients a_n et b_n tels que $P(M_n \leq a_n x + b_n) \to \Lambda(x)$ ont alors la forme suivante : $b_n = F^{\leftarrow}(1 - n^{-1})$, $a_n = g(b_n)$.

1. Supposons que X soit une variable aléatoire positive de distribution F telle que $x^F = \infty$.

Donner quelques éléments d'explication (sans trop de détails mathématiques) pour justifier que l'on a $\lim_{x\to\infty} g(x)/x = 0$.

Montrer que si $F \in D(\Lambda)$ où Λ est la distribution de Gumbel

$$\Lambda(x) = \exp(-\exp(-x)), \qquad x \in \mathbb{R},$$

alors la distribution de $-X^{-1}$ appartient aussi au domaine d'attraction de la distribution de Gumbel.

2. On considère une transformation croissante T, deux fois dérivable sur $\mathbb{R}$ et on suppose que le point extremal $x^F < \infty$ est fini. Montrer que $F \circ T^{-1}$ appartient au domaine d'attraction de la distribution de Gumbel et donner les coefficients de normalisation.

Exercice 2 (4 points):

exam 2018

Supposons que les variables (X_i) sont des variables iid de distributions F. Posons $M_n = \max(X_1, ..., X_n)$ et $m_n = \min(X_1, ..., X_n)$ et supposons qu'il existe des constantes $a_n > 0$, $\alpha_n > 0$, b_n, β_n telles que

$$P(M_n \leq a_n x + b_n) \to G_1(x) \qquad \text{et} \qquad P(-m_n \leq \alpha_n x + \beta_n) \to G_2(x).$$

1. Montrer la convergence

$$P(M_n \leq a_n x + b_n, -m_n \leq \alpha_n y + \beta_n) \to G_1(x)G_2(y).$$

Qu'en concluez-vous?

2. On considère le cas où les (X_i) sont des variables aléatoires Gaussiennes centrées et réduites. On rappelle que dans ce cas

ensae 2018

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda$$

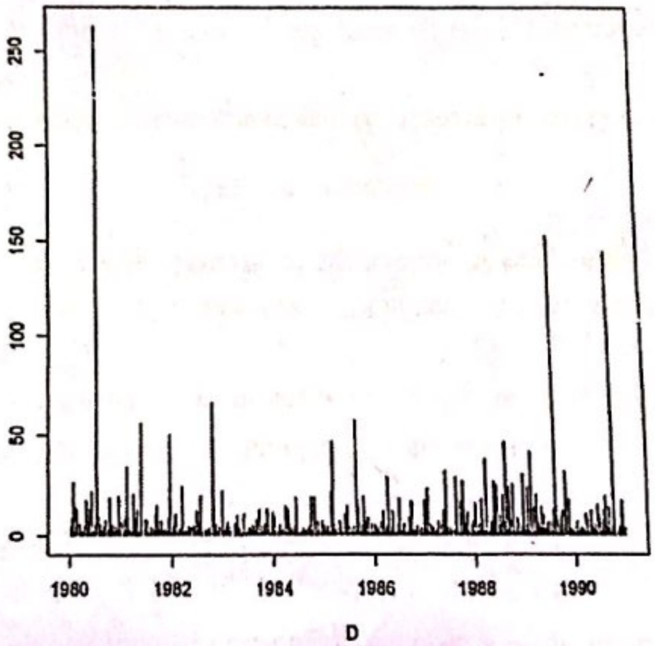
où Λ a une distribution de Gumbel et

$$\begin{aligned} a_n &= (2\ln(2n))^{-1/2} \\ b_n &= (2\ln(2n))^{1/2} - \frac{1}{2}(2\ln(2n))^{-1/2}[\ln(\ln(2n)) + \ln(4\pi)]. \end{aligned}$$

Montrer que $(M_n + m_n)/a_n$ converge en distribution et caractériser sa loi.

Question méthodologique (4 points):

Un réassureur dispose des données de sinistres suivantes pour tarifier un excédent de sinistres de priorité 100 et de portée illimitée.



Quelle méthodologie lui proposez-vous? Donner une réponse littérale.

Annale TVE - Rabat 2018

Question méthodologique

- Présence de montant extrême. On trace le mean excess plot pour déterminer le seuil.
On prend tous les sinistres au dessus de ce seuil et on:
- sépare les attri des larges et on modélise les larges par une GPD
ou - modélise l'ensemble de ces sinistres par une GPD.

Si seuil = 0, on modélise l'ensemble des sinistres avec une GPD.

⇒ On propose au réassureur un traité en XS où on a modélisé la survenance d'1 sinistre et le coût grâce à la GPD.

(NB : on n'utilise pas, dans ce cas la GEV puisqu'en faisant des max par bloc, au max 3 blocks auraient 1 sinistre > 100 (priorité)
⇒ résultat faussé par les autres blocks.)

Exercice d'application du cours

1) On pose $\lambda = np$,

$$\mathbb{P}(N=k)\ \frac{(1-p)^n}{e^{-np}} \times \frac{1}{(1-p)^k} \prod_{i=1}^{k}\left(1-\frac{i-1}{n}\right) = \cancel{e^{-np}} \frac{(np)^k}{k!} \frac{(1-p)^{n-k}}{\cancel{e^{-np}}} \prod_{i=1}^{k}\left(1-\frac{i-1}{n}\right)$$

$$= p^k (1-p)^{n-k} \frac{n^k}{k!} \prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)$$

$$= p^k (1-p)^{n-k} \frac{\cancel{n^k}}{k!} \times \frac{1}{\cancel{n^k}} \frac{n!}{(n-k)!}$$

$$= C_n^k p^k (1-p)^{n-k}$$

$$= \underline{\mathbb{P}(B=k)}$$

de plus,

$$\left(1-\frac{\lambda}{n}\right)^n = \exp\left\{n \ln\left(1-\frac{\lambda}{n}\right)\right\} \underset{n\infty}{\sim} \exp\left\{-n\frac{\lambda}{n}\right\} = e^{-\lambda} = e^{-np}$$

$$\left(1-\frac{\lambda}{n}\right)^k \underset{n\infty}{\sim} \exp\left(-\frac{\lambda k}{n}\right)$$

donc $$\frac{(1-p)^n}{(e^{-np})} \times \frac{1}{(1-p)^k} \prod_{i=1}^{k}\left(1-\frac{i-1}{n}\right) \underset{n\infty}{\sim} \frac{\cancel{e^{-np}}}{\cancel{e^{-np}}} \times e^{\frac{\lambda k}{n}} \underset{n\infty}{\sim} 1$$

donc $$\underline{\mathbb{P}(B=k) \xrightarrow[n\infty]{} \mathbb{P}(N=k)}$$

On utilise ce résultat à l'aide de la suite (u_n) tq $n\bar{F}(u_n) \xrightarrow[n\infty]{} \tau$, alors $N_n = \sum_{i=1}^{n} \mathbb{1}_{\{X_i > u_n\}}$ converge vers une loi de poisson $\mathcal{P}(\tau)$.

N_n est la somme de Bernouilli qui est donc binomiale.
On a : $\tau = np = \lambda$ et $p = \bar{F}(u_n)$.

2) F appartient au domaine d'attraction d'une loi max-stable G ssi il existe 2 suites (a_n) et (b_n) tq

$$\mathbb{P}\left[\frac{M_n-b_n}{a_n} \leq x\right] = F(a_n x+b_n)^n \longrightarrow G(x).$$

$$\cdot\ \mathbb{P}\left(\frac{M_n^H-b_n}{a_n} \leq x\right) = H^n(a_n x+b_n) = F^n(\alpha a_n x + \alpha b_n+\beta)$$

En prenant comme suite de normalisation pour H : $\begin{cases} \tilde{a}_n = \alpha a_n \\ \tilde{b}_n = \alpha b_n+\beta \end{cases}$, on obtient que H appartient au même domaine d'attraction que F.

3) L'indice de queue ξ d'une GPD est le même que celui d'une GEV(μ,σ,ξ).

donc :
- Si $\xi=0$, $D(GPD(\beta,\xi)) = D(\Lambda)$ Gumbel
- Si $\xi<0$, ______ $= D(\Psi_\alpha)$ Weibull
- Si $\xi>0$, ______ $= D(\Phi_\alpha)$ Frechet

4) $\cdot\ \mathbb{P}(M_n \leq u_n) \longrightarrow e^{-\tau} \Leftrightarrow n\overline{F}(u_n) \longrightarrow \tau$

or $n\overline{F}(u_n) = n\exp\{-\lambda u_n\} \xrightarrow[n\to\infty]{} \tau$

alors $n\to\infty$, $n\exp\{-\lambda u_n\} = \tau \Leftrightarrow \boxed{u_n = -\frac{1}{\lambda}\exp\left\{\frac{\tau}{n}\right\}}$

$\cdot$ On sait que $\overline{F}(n) = (1-p)^n$ (proba de survivre jusqu'à u_n au moins)

alors $\dfrac{\overline{F}(n)}{\overline{F}(n-1)} = \dfrac{(1-p)^n}{(1-p)^{n-1}} = 1-p \longrightarrow 1-p \neq 1$

donc, on ne peut pas trouver de suite (u_n).

$\cdot$ On sait que $F(n) = \mathbb{P}(X_1 \leq n) = \sum_{k=1}^{n} \mathbb{P}(X_1=k) = \sum_{k=1}^{n} e^{-\lambda}\frac{\lambda^k}{k!}$

alors $\dfrac{\overline{F}(n)}{\overline{F}(n-1)} = \dfrac{1-\sum_{k=1}^{n} e^{-\lambda}\frac{\lambda^k}{k!}}{1-\sum_{k=1}^{n-1} e^{-\lambda}\frac{\lambda^k}{k!}}$

Supposons $\dfrac{\overline{F}(n)}{\overline{F}(n-1)} \to 1$, alors $\sum_{k=1}^{n} \cancel{e^{-\lambda}}\frac{\lambda^k}{k!} = \sum_{k=1}^{n-1} \cancel{e^{-\lambda}}\frac{\lambda^k}{k!} \Leftrightarrow \frac{\lambda^n}{n!} = 0 \Leftrightarrow \lambda=0$

or $\lambda\neq 0$

donc $\dfrac{\overline{F}(n)}{\overline{F}(n-1)} \not\to 1$

donc on ne peut pas trouver de suite (u_n).

Exercice 1

1) $g(x) = \mathbb{E}(X-x \mid X > x)$ représente la fonction mean excess.

Pour une Gumbel, le ME plot est cst donc $g(x) = \text{cste} \Rightarrow \lim_{x\to\infty} \frac{g(x)}{x} = 0$

• $\mathbb{P}(-X^{-1} \leq x) = \mathbb{P}(X^{-1} > -x) = \mathbb{P}(X \leq -\frac{1}{x}) = F(\frac{-1}{x}) = \exp\{-e^{\frac{1}{x}}\} \sim \Lambda$

$f_{-X^{-1}}(x) = F_{-X^{-1}}'(x) = \frac{-1}{x^2} \times (-e^{-\frac{1}{x}}) \exp\{-e^{\frac{1}{x}}\} = \frac{e^{-1/x}}{x^2} F_{-X^{-1}}(x)$

alors $h(x) = \frac{1 - F_{-X^{-1}}(x)}{f_{-X^{-1}}(x)} = x^2 e^{\frac{1}{x}} \left[\frac{1 - F_{-X^{-1}}(x)}{F_{-X^{-1}}(x)} \right]$

$$h'(x) = e^{\frac{1}{x}} \left[2x \frac{1 - F_s(x)}{F_s(x)} - \frac{x^2}{x^2} \left[\frac{1-F_s(x)}{F_s(x)}\right] + x^2 \frac{-F_s'(x) F_s(x) - (1-F_s) F_s'}{(F_s(x))^2} \right]$$

$$= e^{\frac{1}{x}} \left[\frac{1 - F_s(x)}{F_s(x)} (2x-1) - x^2 \frac{f_s(x)}{[F_s(x)]^2} \right]$$

$$= e^{\frac{1}{x}} \left[(e^{e^{\frac{1}{x}}} - 1)(2x-1) - \underbrace{e^{-1/x} e^{e^{1/x}}}_{1} \right]$$