UFR DE MATHEMATIQUES

Master Mathématiques Appliquées
Statistiques

Parcours Ingénierie Statistique et Numérique

Rapport de Travail Encadré de Recherche
(T.E.R)

LES VALEURS EXTREMES

Encadrant
G.CASTELLAN

Auteurs
DIALLO Ibrahima
SANE Sidy

8 mai 2017

Table des matières

1 Introduction

La théorie des valeurs extrêmes est une branche des statistiques qui s'intéresse aux valeurs extrêmes des distributions de probabilité.

La théorie des valeurs extrêmes est appliquées en hydrologie pour prévoir les crues , en démographie pour prévoir la distribution de probabilité de l'âge maximum que l'être humain pourra atteindre , en assurance pour prévoir les grands sinistres , en finance ou encore en météorologie .

2 Exemples de comportement du maximum

2.1 La loi du maximum

Si on a $(X_1 \ ,......,\ X_n)$ une suite de variables aléatoires i.i.d de fonction de repartition F . Alors la fonction de repartition de la suite $M_n = maxX_i$ est donnée par :

$$F_n(x) = F^n(x)$$

Si on a $(X_1 \ ,......,\ X_n)$ une suite de variables aléatoires i.i.d de densité f , alors on a :

$$f_n(x) = nf(x)F^{n-1}(x)$$

Démonstration. On a par definition :

$$F_n(x) = \mathbf{P}(M_n \leq x) = \mathbf{P}(maxX_i \leq x) = \mathbf{P}(X_1 \leq x, .., X_n \leq x) = \bigcap_{i=1}^{n} \mathbf{P}(X_i \leq x)$$

Comme les X_i sont i.i.d , on a

$$F_n(x) = \prod_{i=1}^{n} \mathbf{P}(X_i \leq x) = (P(X_1 \leq x))^n = F^n(x)$$

Pour obtenir la densité de M_n , il suffit de derive sa fonction de repartition . $\square$

2.2 Loi Uniforme

On suppose que la loi de X_1 est la loi uniforme sur $[0\ ,\theta]$, $\theta > 0$. La fonction de repartition de la loi est $F(x) = x/\theta$ pour $x \in [0, \theta]$. La suite $(M_n\ ,\ n \geq 1)$ converge presque surement vers θ

la suite $(n(\frac{M_n}{\theta} - 1)\ ,\ n \geq 1)$ converge en loi vers W de fonction de repartition definie par :

$$P(W \leq x) = e^x\ ,\ x \leq 0$$

. La loi de W est une loi de Weibull . Dans un cas particulier , la loi de $-W$ est la loi exponentielle de parametre 1.

Démonstration. On note F_n la fonction de repartition de $(n(\frac{M_n}{\theta})$. Comme $M_n < \theta$, On a $F_n(x) = 1 \ \ si \ \ x \geq 0$. Considerons le cas $x < 0$:

$$F_n(x) = P(M_n \leq \theta + \theta(\frac{x}{n}) = P(X_1 \leq \theta + \theta(\frac{x}{n})^n = (1 + \frac{x}{n})^n$$

. Il vient $\lim_{x \to +\infty} F_n(x) = e^x$ pour $x \leq 0$. On deduit que $n(\frac{M_n}{\theta} - 1)$ converge en loi vers W de fonction de repertition $x \to min(e^x, 1)$. □

En faisant une simulation sur un échantillon de taille 1000 on observe une convergence presque sure du maximum vers 1 , comme l'illustre la figure suivante .

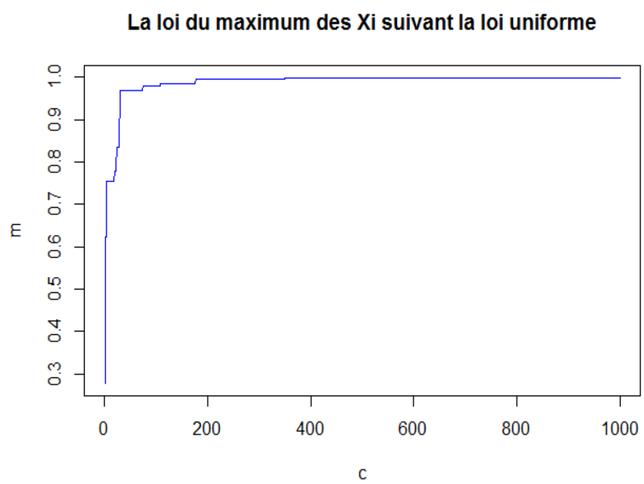

FIGURE 1 –

Le graphe de la fonction densité théorique de M_n est :

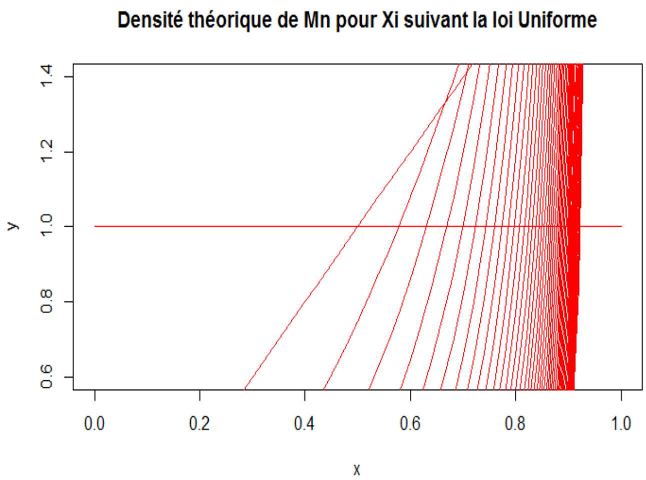

FIGURE 2 –

Le graphe de la fonction densité du maximum nous montre une forte concentration de la masse vers 1 ce qui nous permet de confirmé une convergence presque sur vers 1 du maximum.

2.3 Loi exponetielle

On suppose X_1 de loi exponentielle de parametre $\lambda > 0$. La fonction de repartition de cette loi est $F(x) = 1 - e^{-\lambda x}$ pour $x \geq 0$.comme $F(x) = +\infty$, la suite $(M_n, n \leq 1)$ diverge vers l'infini .

2.4 lemme

La suite $(\lambda M_n - log(n)\ ,\ n \geq 1)$ converge en loi vers G de fonction de repartition definie par

$$P(G \leq x) = e^{-e^{-x}}\ ,\ x \in R$$

. La loi de G est la loi de Gumbel .

Démonstration. On note F_n la fonction de repartition de $\lambda M_n - log(n)$. On a pour $x + log(n) > 0$,

$$F_n(x) = \mathbf{P}(\lambda \mathbf{M_n} - \mathbf{log(n)} \leq \mathbf{x}) = P(M_n \leq \frac{(x + log(n))}{\lambda} = \mathbf{P}(\mathbf{X_1} \leq ((\mathbf{x} + \mathbf{log(n)}\lambda)^{\mathbf{n}})) = (1 - \frac{e^{(}-}{n}$$

On a alors
$$\lim_{x\to+\infty} F_n(x) = e^{-e^{-x}}, x \in R$$
. On en deduit que la suite $(\lambda M_n - log(n)\ ,\ n \geq 1)$ converge en loi vers G , de fonction de repartition $x \to e^{-e^{-x}}$. □

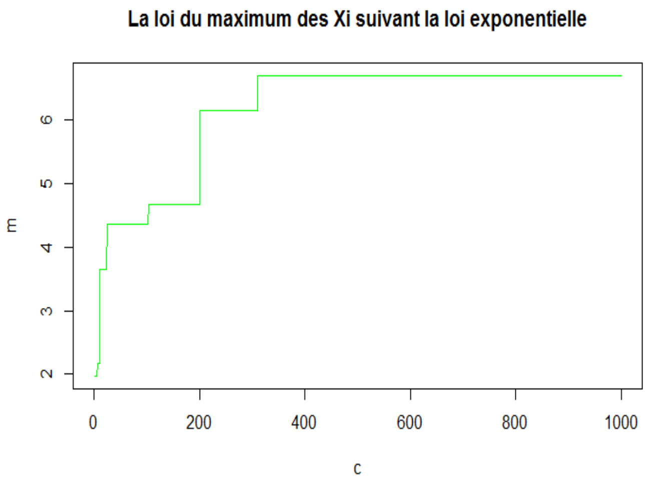

FIGURE 3 –

La figure 3 est la simulation d'un echantillon de taille 1000 suivant la loi exponentielle on constate une divergence du maximum

Le graphe de la fonction densité théorique de M_n est :

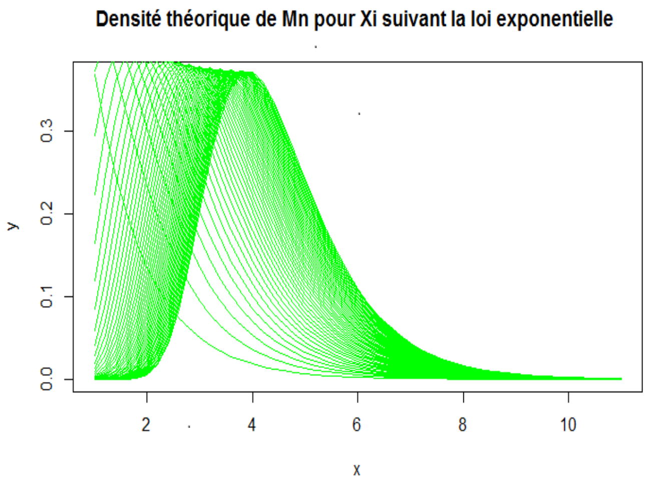

FIGURE 4 –

2.5 Loi de cauchy

On suppose X_1 suit la loi de cauchy (de parametre a $= 1$) . La densité de loi $f(x) = \frac{1}{\pi(1+x^2)}$. Comme le support de la densité est non borné , il est clair que la suite $(M_n, n \geq 1)$ diverge.

2.6 lemme

La suite $(\frac{\pi M_n}{n} \;\;,\;\; n \geq 1)$ converge en loi vers W de fonction de repartition definie par

$$\mathbf{P}(\mathbf{W} \leq \mathbf{x}) = e^{\frac{-1}{x}} \;\;,\;\; x > 0$$

La loi de W appartient à la famille des lois de Frechet.

Démonstration. On note F_n la fonction de repartition de $\frac{\pi M_n}{n}$ on a

$$F_n(x) = \mathbf{P}(\mathbf{M_n} \leq \frac{\mathbf{nx}}{\pi}) = P(x_1 \leq \frac{nx}{\pi})^n = (1 - \int_{\frac{nx}{\pi}}^{\infty} \frac{1}{\pi(1+y^2)} dy)^n$$

Pour $x > 0$, on a

$$\int_{\frac{nx}{\pi}}^{\infty} \frac{1}{\pi(1+y^2)} dy) = \int_{\frac{nx}{\pi}}^{\infty} \frac{1}{\pi(y^2)} dy) + \int_{\frac{nx}{\pi}}^{\infty} [\frac{1}{\pi(1+y^2)} - \frac{1}{\pi y^2}] dy) = \frac{1}{nx} + O((nx)^{-3}) \quad (1)$$

On a alors pour

$$x \geq 0, F_n(x) = (1 - \frac{1}{nx} + \mathbf{O(nx)^{-3})n}$$

. On en deduit que

$$\lim_{x \to +\infty} F_n(x) = e^{\frac{-1}{x}}, x > 0$$

Ainsi la suite $(\frac{\pi M_n}{n}, n \geq 1)$ converge en loi vers W de fonction de repartition

$$\mathbf{P(W \leq x)} = e^{\frac{-1}{x}} \quad , \quad x > 0$$

□

3 Théoréme :Comportement en loi

3.1 Définition

la loi $\mathcal{L}_0$ est dite max-stable si pour tout $n \geq 0$ $(W_1, ..., W_n)$ etant des variables aléatoires independantes de loi $\mathcal{L}_0$, il existe $a_n > 0$ et $b_n \in \mathbf{R}$, tel que $a_n^{-1}(max_{i \in (1,..,n)} W_i - b_i)$ suit la loi $\mathcal{L}_0$.
Remarque
Si les $(X_i, i \geq 0)$ est une suite de variables aléatoire indépendantes et de meme loi , telle que $(a_n^{-1}(max_{i \in (1,..,n)} X_i - b_i), n \geq 1$ converge en loi pour une suite appropriée $a_n > 0$ et $b_n \in \mathcal{R}$ vers une limite non triviale , c'est-à-dire vers une variable aléatoire non constante , alors la limite est une loi max-stable.

3.2 theoreme

soit $(X_i \quad , \quad i \geq 1)$ une suite de variables aleatoires independantes et de meme loi . Supposons qu'il existe une suite $((a_n \quad , \quad b_n) \quad , \quad n \geq 0)$ tel que $a_n > 0$ et la suite $(a_n^{-1}(max_{i \in (1,..,n)} X_i - b_i) \quad , \quad n \geq 0)$ converge en loi vers une limite non triviale . Alors à une translation et un changement d'echelle pres la $fonction\ de\ rpartition$ de la limite est de la forme suivante :

loi de Weibull

$$\Psi_\alpha(x) = \begin{cases} exp(-(-x)^\alpha) & , x \leq 0 \ et \ \alpha > 0 \\ \\ 1 & , x > 0 \end{cases} \tag{2}$$

loi de Gumbel
$$\Lambda(x) = exp(-exp(-x)), x \in \mathbf{R} \tag{3}$$

loi de Frechet
$$\Phi_\alpha(x) = \begin{cases} 0 & , x \leq 0 \\ \\ exp(-x^{-\alpha}) & , x > 0 \ et \ \alpha > 0 \end{cases} \tag{4}$$

Démonstration. Supposons qu'il existe une suite $((a_n, b_n) \quad , \quad x \geq 0)$ tel que $a_n > 0$ et la suite de terme $a_n^{-1}(M_n - b_n)$ converge vers une limite de X non constante
Pour une fonction g continue bornée , on a
$$\lim_{n \to +\infty} \mathbf{E}[g(a_n^{-1}(M_n - b_n))] = \mathbf{E}[g(W)]$$
.
Supposons par simplicité que la loi de X_1 posséde une densité $f > 0$. Alors la loi de M_n poséde une densité $nF(x)^{n-1}f(x)$
On a donc
$$I_n = \mathbf{E}[g(a_n^{-1}(M_n - b_n))] = \int_{\mathbf{R}} g(\frac{x - b_n}{a_n})nF(x)^{n-1}f(x)dx$$

Comme $f > 0$, la fonction F est inversible et d'inverse continue . On pose pour $t > 1$
$$U(t) = F^{-1}(1 - \frac{1}{t})$$

En paticulier , on a
$$U(t) = x \Longleftrightarrow 1 - \frac{1}{t} = F(x) \Longleftrightarrow \mathbf{P(X > x)} = \frac{1}{t}$$
On effectue le changement $F(x) = 1 - \frac{v}{n}$, i.e. $x = U(\frac{n}{v})$. on obtient
$$I_n = \int_{\mathbf{R}} g(\frac{U(\frac{n}{v}) - b_n}{a_n})(1 - \frac{v}{n})^{n-1}\mathbf{1}_{]\mathbf{0} \ , \ \mathbf{n}]}(\mathbf{v})dv$$

Remarquons que $(1-\frac{v}{n})^{n-1}\mathbf{1}_{]\mathbf{0}\ ,\ \mathbf{n}]}(\mathbf{v})$ converge en croissant vers $exp(-v)\mathbf{1}_{\mathbf{v}>\mathbf{0}}$. Comme par hypothése I_n converge pour tout g , il est naturel , mais erroné à priori , de penser que pour tout $v > 0$, la suite de terme $J_n(v) = \frac{U(\frac{n}{v})-b_n}{a_n}$ converge . Supposons malgré tout que cette convergence ait lieu . On en deduit en considérant $J_n(\frac{1}{w}) - J_n(1)$ que pour tout $w > 0$, $\frac{U(wn)-U(n)}{a_n}$ converge quand n tend vers l'infini , vers une limite que l'on note $h(w)$. Comme la variable aléatoire W est non triviale , cela implique que la fonction h n'est pas égale à une constante . Comme la fonction U est croissante , la fonction h est également croissante . Supposons que plus généralement , on ait pour tout $w > 0$

$$\frac{U(wx) - U(x)}{a(x)} x \to +\infty \longrightarrow h(w)$$

où $a(x) = a_{[}x]$pour $x \geq 1$, $x_{[}x]$ designant la partie entiére de x . Soit w_1 , $w_2 > 0$ on a

$$\frac{U(xw_1w_2) - U(x)}{a(x)} = \frac{U(xw_1w_2) - U(xw_1)}{a(xw_1)}\frac{a(xw_1)}{a(x)} + \frac{U(xw_1) - U(x)}{a(x)}$$

En faisant tendre x vers l'infini dans l'égalité ci-dessus , on obtient que $\frac{a(xw_1)}{a(x)}$ converge pour tout $w_1 > 0$. On note $l(w_1)$ la limite , et il vient

$$h(w_1w_2) = h(w_1)l(w_1) + h(w_1). \tag{5}$$

La fonction l est mesurable et localement bornée. Comme la fonction h est croissante et non constante , on en deduit que l est strictement positive . De plus en posant $yw' = x$, on a pour $w' > 0$

$$l(w) = \lim_{x\to+\infty} \frac{a(xw)}{a(x)} = \lim_{y\to+\infty} \frac{a(yw'w)}{a(y)}\frac{a(y)}{a(yw')} = \frac{l(w'w)}{l(w')}$$

Ainsi on a pour tous w , $w' > 0$,

l(w'w) = l(w) l(w') (6)

où l est une fonction strictement positive mesurable localement bornée . Verifions que les solutions non nulles de cette equation fonctionelle sont :

$l(w) = w^{\xi}$ où $\xi \in \mathbf{R}$. En intégrant (6) pour w' entre 1 et 2 , il vient en effectuant le changement de variable

$$ww' = u \ , \quad \frac{1}{w}\int_w^{2w} l(u)du = l(w)\int_1^2 l(w')dw'$$

.

On en deduit que l est continu puis dérivable . On obtient alors en dérivant (6) par rapport à w' et en évaluant en $w' = 1$ et que $wl(w') = \xi l(w)$ où $\xi \in \mathbf{R}$

.

Ceci implique que $l(w) = cw^{\xi}$. Comme d'aprés (6) , $l(1) = l(1)^2$ et que l est strictement positive , on en déduit que $l(1) = 1$ et que $l(w) = w^{\xi}$ pour $w > 0$ On retrouve ξ l'indice de la loi de valeur extrémes généralisées . Léquation (5) se récrit $h(w_1w_2) = h(w_2)w_1^{\xi}$ pour tout $w_1 \ , \ w_2 > 0$

Pour $\xi = 0$ on obtient l'équation fonctionnelle $h(w_1w_2) = h(w_1) + h(w_2)$

.

Un raisonnement semblable à celui effectué à partir de l'équation fonctionnelle (6) assure que les solutions mesurables localement bornées sur $]0 \ , \ \infty[$ de cette équation foctionnelle sont $h(w) = clog(w)$, avec $c > 0$.

Pour $\xi \neq 0$, par symétrie , on a

$$h(w_1w_2) = h(w_2)w_1^{\xi} + h(w_1) = h(w_1)w_2^{\xi} + h(w_2)$$

.

En particulier , on a

$$h(w_1)(1 - w_2^{\xi}) = h(w_2)(1 - w_1^{\xi})$$

.

Cela implique $h(w) = 0$ si $w = 1$, et sinon $\frac{h(w)}{w^{\xi}-1}$ est constant .

Donc on obtient $h(w) = c(w^{\xi} - 1)$. A un changement d'echelle prés , on peut choisir $c = \frac{1}{\xi}$. A une translation prés , on peut choisir $U(n) = b_n$.

En définitive , il vient

$$\lim_{n\to+\infty} \frac{U(\frac{n}{v}) - b_n}{a_n} = h(\frac{1}{v}) = \begin{cases} \frac{v^{-\xi}-1}{\xi} & si\ \xi \neq 0 \\ \\ -log(v) & si\ \xi = 0 \end{cases}$$

On peut maintenant calculer la limite de $I_n = \mathbf{E}[g(a_n^{-1}(M_n - b_n))]$. Il vient par convergence dominée

$$\lim_{n\to+\infty} I_n = \int g(\frac{v^{-}\xi - 1}{\xi})e^{-v}\mathbf{1}_(\mathbf{v > 0})dv = \int g(y)\mathbf{1}_(\mathbf{1 + \xi y > 0})d(e^{-(1+\xi y)^{-\frac{1}{\xi}}})$$

,
où on a posé $y = \frac{v^{-\xi}-1}{\xi}$ si $\xi \neq 0$ et $y = log(v)$ si $\xi = 0$

$\square$

Remarque
On peut rassembler les trois familles de lois en une seule fammile paramétrique $(H(\xi)\ ,\ \xi \in \mathbf{R})$ dites famille des trois lois de valeurs extrémes généralisées . Elle est paramétrée par une seule variable $\xi \in \mathbf{R}$, mais toujours à un facteur de changement d'echelle et de translation prés . La fonction de répartition est pour $\xi \in \mathbf{R}$

$$H(\xi)(x) = e^{-(1+\xi x)^{-\frac{1}{\xi}}}, si 1 + \xi x > 0 \tag{7}$$

Pour $\xi = 0$ on obtient $H(0)(x) = e^{-e^{-x}}$ en faisant tendre ξ vers 0

4 Domaines d'attraction

Aprés avoir caractérisé les lois limites , il reste à determiner les lois $\mathcal{L}$ pour lesquelles la loi du maximum renormalisé converge vers une loi max-stable donnée $\mathcal{L}_{\prime}$. On dit alors que la loi $\mathcal{L}$ (ou sa fonction de répartition F) appartient au bassin d'attraction de $\mathcal{L}_{\prime}$ (ou de sa fonction de répartition F_0) pour la convergence du maximum renormalisé . On le notera $\mathcal{L} \in D(\mathcal{L}_{\prime})$ (ou $F \in D(F_0)$).

4.1 Caratérisation générales

On définit la fonction U par

$$U(t) = F^{-1}(1 - \frac{1}{t}) \quad , \quad t > 1$$

où F^{-1} est l'inverse généralisé de F . On sait que si $F \in D(H(\xi))$, alors on a , à un changement d'echelle prés ,

$$\lim_{s\to+\infty} \frac{U(sw) - U(s)}{a(s)} = \frac{w^{\xi} - 1}{\xi}. \tag{8}$$

En particulier , si $x \quad , \quad y > 0\ et\ y \neq 0$, on a

$$\lim_{s\to+\infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \lim_{s\to+\infty} \frac{U(sx) - U(s)}{a(s)} \frac{a(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^{\xi}-1}{y^{\xi}-1} & ,\xi \neq 0 \\ \\ \frac{log(x)}{log(y)} & ,\xi = 0 \end{cases} \tag{9}$$

4.2 Proposition

On a $F \in D(H(\xi))$ si et seulement si

$$n\overline{F}(xa_n + b_n) \; n \overrightarrow{\rightarrow} +\infty \; -logH(\xi)(x) \tag{10}$$

pour une certaine suite $((a_n, b_n) \;\; , \;\; n \geq 1)$ où $a_n > 0 \; et \; b_n \in \mathbf{R}$. On a alors la convergence en loi de $(a_n^{-1}(M_n - b_n) \;\; , \;\; n \geq 1)$ vers une variable aléatoire de la fonction de répartition $H(\xi)$

Démonstration. Si $F \in D(H(\xi))$, alors on a $(1 - \overline{F}(xa_n + b_n))^n \; n \overrightarrow{\rightarrow} +\infty$ $H(\xi)(x)$ pour une certaine suite $(a_n^{-1}(M_n - b_n) \;\; , \;\; n \geq 1)$ où $a_n > 0 \; et \; b_n \in \mathbf{R}$.En prenant le logarithme de cette expression il vient

$$n(1 - \overline{F}(xa_n + b_n)) \; n \overrightarrow{\rightarrow} +\infty \; logH(\xi)(x)$$

Ceci implique que pour $1 + \xi x > 0 \;\; , \;\; \overline{F}(xa_n + b_n)$ tend vers vers 0 et

$$n\overline{F}(xa_n + b_n) \; n \overrightarrow{\rightarrow} +\infty \; -logH(\xi)(x)$$

La réciproque est claire : si on a la convergence ci-dessus , alors $F \in D(H(\xi))$ □

4.3 Domaines d'attraction des lois de Fréchet et Weibull

4.3.1 Definition

On dit qu'une fonction L est à variation lente si $L(t) > 0$ pour t assez grand et si pour tout $x > 0$, on a

$$\lim_{t \to \infty} \frac{L(tx)}{L(t)} = 1. \tag{11}$$

4.3.2 Proposition

soit L une fonction à variation lente . Il existe deux fonctions mesurables $c > 0$ et k telles que :

$\lim_{x \to \infty} c(x) = c_0 \in]0, \infty[\;\; et \;\; \lim_{x \to \infty} k(x) = 0,$ (12)et $a \in \mathbf{R}$, pour tout $x > a$,

$$L(x) = c(x)exp(\int_a^x \frac{k(u)}{u})du \tag{13}$$

4.3.3 Théoreme

La fonction de répartition F appartient au domaine d'attraction de la loi de Fréchet de parametre α si et seulement si

$$\overline{F}(x) = x^{-\alpha}L(x)$$

où la fonction L est à variation lente . En particulier $x_F = +\infty$.De plus si $F \in D(\Phi_\alpha)$, alors avec $a_n = U(n) = F^{-1}(1 - \frac{1}{n})$, la suite $(a_n^{-1}M_n, n \geq 1)$ converge en loi vers une variable aléatoire de fonction de répartition Φ_α .

Démonstration. Supposons que $\overline{F}(x) = x^{-\alpha}L(x)$, où L est à variation lente . Soit $\overline{F}(x) \sim g(x)$ où $g(x) = x^{-\alpha}c_0 exp(\int_a^x \frac{k(u)}{u})du$ est une fonction continue . Posons $a_n = U(n)$. on a $\overline{F}(x) \leq \frac{1}{n} \leq \overline{F}(a_n^-$ et a_n tend vers l'infini avec n.Si F est continue en a_n , alors ona $\overline{F}(a_n) = 1/n$, sinon comme $\overline{F}$ est équvalente en $+\infty$ à une fonction continue , on en déduit que $\overline{F}(a_n) \sim 1/n$ quand n tend vers l'infini. Pour $x > 0$, on a donc

$$\lim_{n\to\infty} n\overline{F}(a_n x) = \lim_{n\to\infty} \frac{\overline{F}(a_n x)}{\overline{F}(a_n)} = x^{-\alpha} \tag{14}$$

□

4.3.4 Théoreme

Pour qu'une fonction de distribution F appartient au domaine d'attraction de la loi $\Phi_\alpha(x)$ il faut et il suffit que l'on ait

$$\lim_{x\to\infty} \frac{1 - F(x)}{1 - F(kx)} = k^\alpha \tag{15}$$

pour toute valeur de $k > 0$

Démonstration. Supposons d'abord que la condition (15) est vérifiée et montrons que la fonction $F(x$ appartient au domain d'attraction de la loi $\Phi_\alpha(x)$. Il est évident , d'apres (15) , que $F(x) < 1$ pour toutes les valeurs de x . Il en resulte que ,pour n suffisamment grand , les valeurs de x donnant lieu à l'inégalité

$$1 - F(x) \leq \frac{1}{n}$$

sont positives .
Définissons a_n comme la plus petite des valeurs de x vèrifiant les inégalités

$$1 - F(x(1+0)) \leq \frac{1}{n} \leq 1 - F(x(1-0)) \tag{16}$$

il résulte de ce qui précede que $a_n \to \infty$ pour $n \to \infty$. D'apres la condition (15) , pour toute valeur de x et pour tout $\epsilon(0 < \epsilon < 1)$ nous avons

$$\frac{1 - F(a_n x)}{1 - F(a_n(1+\epsilon} \to (\frac{1+\epsilon}{x})^{\alpha}$$

et

$$\frac{1 - F(a_n x)}{1 - F(a_n(1-\epsilon} \to (\frac{1-\epsilon}{x})^{\alpha}$$

n tendant vers l'infini . Les premiers membres de ces relation étant fonctions monotones de ϵ et les seconds membres étant fonctions continues de ϵ , la convergence en question est uniforme , ce qui nous permet d'écrire , pour $n \to \infty$,

$$\frac{1 - F(a_n x)}{1 - F(a_n(1+0))} \to \frac{1}{x^{\alpha}}$$

et

$$\frac{1 - F(a_n x)}{1 - F(a_n(1-0))} \to \frac{1}{x^{\alpha}}$$

nous avons

$$\frac{1 - F(a_n x)}{1 - F(a_n(1-0))} \leq n(1 - F(a_n x)) \leq \frac{1 - F(a_n x)}{1 - F(a_n(1+0))}$$

on voit que pour tout $x > 0$ on a

$$n(1 - F(a_n x)) \to x^{-\alpha}$$

pour $n \to \infty$ on sait que pour $n \to \infty$

$$F^n(a_n x) \to \Phi_{\alpha}(x) \qquad (-\infty < x < +\infty)$$

. Supposons maintenant que $F(x)$ appartient au domaine d'attraction de la $\Phi_{\alpha}(x)$, c'est-à-dire supposons que , pour un choix convenable de constantes $a_n > 0$ et b_n , pour toutes les valeurs de $x(x > 0)$ a lieu la relation

$$n(1 - F(a_n x + b_n)) \to x^{-\alpha} \tag{17}$$

pour $n \to \infty$ Pour toute constante $\beta > 1$ nous avons pour $n \to \infty$

$$n\beta(1 - F(a_n x + b_n)) \to \beta x^{-\alpha}.$$

Pour $x > 0$ et $n \to \infty$ on a

$$1 - F(a_n x + b_n) \to 0$$

nous voyons que , pour $n \to \infty$, on doit avoir

$$[n\beta](1 - F(a_n x + b_n)) \to \beta x^{-\alpha}$$

où $[\beta x]$ désigne l'entier de βx .
Par le changement de variable $x = z\beta^{1/\alpha}$ cette relation prend la forme suivane :
pour $n \to \infty$ et pout tout $z > 0$ on a

$$[n\beta](1 - F(a_n \beta^{1/\alpha} z + b_n) \to z^{-\alpha} \tag{18}$$

Il résulte de (39) que pour $n \to \infty$

$$[n\beta](1 - F(a_{[n\beta]} x + b_{[n\beta]}) \to x^{-\alpha} \tag{19}$$

En vertu des Lemmes 5 et 2 et en tenant compte de (40) et (41) ,nous concluons que les relations

$$\frac{a_{[n\beta]}}{a_n} \to \beta^{1/\alpha}, \qquad \frac{b_{[n\beta]} - b_n}{a_{[n\beta]}} \to 0$$

doivent avoir lieu pour $n \to \infty$. En vertu du Lemme 3 la relation (19) ne changera pas si nous posons

$$a_{n\beta} = a_n \beta^{1/\alpha}, b_{n\beta} = b_n \tag{20}$$

posons

$$n_s = [n_{s-1,\beta}], \qquad n_1 = [n\beta]$$

où l'entier n est considérer comme comme fixe.Il résulte de (20) que pour tout nombre naturel s nous avons

$$a_{n_s} = a_n \beta^{s/\alpha}, \qquad b_{n_s} = b_n$$

D'où nous tirons que pour $s \to \infty$

$$\frac{b_{n_s}}{a_{n_s}} \to 0$$

et par conséquent , en vertu du Lemme 3 , que pour $s \to \infty$

$$n_s[1 - F(a_{n_s} x] \to x^{-\alpha} \tag{21}$$

Supposons que $y \to \infty$; pour toute valeur suffisamment grande de y on peut trouver une valeur de s telle que

$$a_{n_s}x \leq y \leq a_{n_{s+1}}x$$

et par suite

$$1 - F(a_{n_{s+1}}x) \leq 1 - F(y) \leq 1 - F(a_{n_s}x)$$

et , pour $k > 0$

$$1 - F(a_{n_{s+1}}kx) \leq 1 - F(ky) \leq 1 - F(a_{n_s}kx)$$

Nous en tirons l'inégalité

$$\frac{1 - F(a_{n_{s+1}}x)}{1 - F(a_{n_s}kx)} \leq \frac{1 - F(y)}{1 - F(ky)} \leq \frac{1 - F(a_{n_s}x)}{1 - F(a_{n_{s+1}}kx)} \tag{22}$$

Remarquons que

$$\frac{n_{s+1}}{n_s} = \frac{n_s\beta - \theta_s}{n_s}$$

où $0 \leq \theta_s < 1$;donc pour $s \to \infty$ on a

$$\frac{n_{s+1}}{n_s} \to \beta$$

Nous en concluons , en vertu de (..) et (..),que

$$\frac{1}{\beta}k^\alpha \leq \lim_{y\to\infty} \frac{1 - F(y)}{1 - F(ky)} \leq \beta k^\alpha$$

or,puisque β peut etre choisi aussi peut différent de l'unité que l'on veut , la condition du théoreme en résulte . Faisons remarquer qu'il résulte de ce qui précede que toute fonction de distribution $F(x)$ appartient au domaine d'attraction de la loi $\Phi_\alpha(x)$ est attirée vers $\Phi_\alpha(x)$ d'une façon plus particuliere , à savoir : pour un choix de constante a_n a lieu l'égalité

$$\lim_{n\to\infty} F^n(a_nx) = \Phi_\alpha(x)$$

□

4.3.5 Théoreme

La fonction de répartition de F appartient au domaine d'attraction de la loi de Weibull de paramétre $\alpha > 0$ si et seulement si $x_F < \infty$ et

$$\overline{F}(x_F - \frac{1}{x}) = x^{-\alpha}L(x)$$

où la fonction L est à variation lente . De plus si $F \in D(\Psi_\alpha)$, alors avec $a_n = x_F - U(n) = x_F - F^{-1}(1 - \frac{1}{n})$, la suite $(a_n^{-1}(M_n - x_F), n \geq 1)$ converge vers la fonction de répartition Ψ_α

4.4 Proposition (Critére de Von Mises)

Soit F la fonction de répartition d'une loi de densité f
1. Si on a

$$\lim_{x\to\infty} \frac{xf(x)}{\overline{F}(x)} = \alpha > 0$$

alors F appartient au domaine d'attraction de la loi de Fréchet de paramétre α .
2. On suppose la loi de densité f trictement positive sur un intervalle (z, x_F) , avec $x_F < \infty$. Si on a

$$\lim_{x\to x_F^-} \frac{(x - x_F)f(x)}{\overline{F}(x)} = \alpha > 0$$

alors F appartient au domaine d'attraction de la loi de Weibull de paramétre α .

5 Statistique d'ordre

Soient $X_1, ..., X_n$ i.i.d de fonction de repartition F

5.1 Définition

La statistique d'ordre de l'echantillon $(X_1, ...X_n)$ est le rearangement croissant de cette echantillon . On la note $(X_{(1,n)}, ..., X_{(n,n)})$. On a

$$X_{(1,n)} \leq ... \leq X_{(n,n)}$$

, et il existe une permutation aleatoire $\sigma_n \in S_n$ l'ensemble des permutations tel que $(X_{(1,n)}, ..., X_{(n,n)}) = (X_{\sigma_{n(1)}}, ..., X_{\sigma_{n(n)}})$ En particulier on a $X_{(1,n)} = minX_i$ et $X_{(n,n)} = MaxX_i$ on note F^{-1} l'inverse generalisé de F .

lemme 1. *Soit $X_1, ..., X_n$ des variables alatoires independantes et de $fonction$ de $reparti-$ tion F.Soit $U_1, ..., U_n$ des variables aleatoires independantes de loi $uniforme$ sur* $[0; 1]$. *Alors* $(F_1(U(1,n), ..., F_1(U(1,n))$ *meme loi que* $(X(1,n), ..., X(n,n))$

6 Estimation du parametres de la loi de valeurs extremes

La vraisemblance d'un modele aléatoire est la densité de sa loi de probabilité . Soit$(X_n, n \geq 1)$ une suite de variables aléatoire independante La vraisemblance du modele generalisé des extrémes s'ecrit

$$l(x; \mu, \sigma, \xi) = (\frac{1}{\sigma})^m \prod_{i=1}^{m} exp(-(1 + \xi(\frac{x_i - \mu}{\sigma})^{\frac{-1}{\xi}})(1 + \xi(\frac{x_i - \mu}{\sigma}))^{-1-\frac{1}{\xi}}) \quad (23)$$

$$L(\mu, \sigma, \xi) = -mlog(\sigma) - (\frac{1}{\sigma} + 1) \sum_{i=1}^{m} log(1 + \xi(\frac{x_i - \mu}{\sigma})) - \sum_{i=1}^{m} (1 + \xi(\frac{x_i - \mu}{\sigma})^{\frac{-1}{\xi}}) \quad (24)$$

on pose

$$\rho = \frac{1}{\sigma}$$

et on introduit λ tel que :

$$log(\lambda) = \frac{\mu}{\sigma}$$

alors la vraisemblance se presente comme suit avec $\xi = 0$

$$L(\mu, \sigma, \xi) = -mlog(\lambda\rho) - \rho(\sum_{i=1}^{m} x_i) - \lambda(\sum_{i=1}^{m} exp(-\rho x_i) \quad (25)$$

6.1 Estimateur de pickand

6.1.1 theoreme

Soit

$$(X_n \ , \ n \geq 1)$$

une suite de variable aleatoire independante de meme fonction de repartition F $\in D(H(\xi))$, ou $\xi \in \mathbf{R}$. Si

$$\lim_{n \to \infty} k(n) = \infty$$

et

$$\lim_{n \to \infty} \frac{k(n)}{n} = 0$$

, alors l'estimateur de Pickand

$$\xi^P_{(k(n),n)} = \frac{1}{log2} log(\frac{X_{(n-k(n)+1,n)} - X_{(n-2k(n)+1,n)}}{X_{(n-2k(n)+1,n)} - X_{(n-4k(n)+1,n)}}) \quad (26)$$

converge en probabilité vers ξ

Démonstration. Pour $\xi \in R$, on a avec le choix t=2s et y=1/2

$$\lim_{t\to\infty} \frac{U(t)-U(t/2)}{U(t/2)-U(t/4)} = 2^{\xi}$$

En fait en utilisant la croissance de U qui se deduit de la croissance de F , on obtient

$$\lim_{t\to\infty} \frac{U(t)-U(t_{c1}(t))}{U(t_{c1}(t))-U(t_{c2}(t))} = 2^{\xi}$$

des que

$$\lim_{t\to\infty} c_1(t) = 1/2 \quad et \lim_{t\to\infty} c_2(t) = 1/4$$

il reste donc à trouvé des etimateurs pour $U(t)$ Soit $k(n)$, $n \geq 1)$ une suite d'entiers telle que

$$1(n) \geq n/4 \quad , \quad \lim_{n\to\infty} k(n) = \infty \quad et \quad \lim_{n\to\infty} \frac{k(n)}{n} = 0.$$

Nous ecrivons k pour k(n) . Soit $(V_(1,n),...,V_(n,n))$ la statistique d'ordre d'un echantillon de variables aléatoires independantes de loi de pareto . On note $F_V(x) = 1 - \frac{1}{x}$, pour $x \geq 1$, la fonction de repartittion de la loi de pareto . On a les convergences en probabilites suivantes

$$V_{(n-k+1,n)} \to \infty \tag{27}$$

$$\frac{V_{(n-2k+1,n)}}{V_{(n-k+1,n)}} \to 1/2 \tag{28}$$

$$\frac{V_{(n-4k+1,n)}}{V_{(n-k+1,n)}} \to 1/4 \tag{29}$$

on en deduit donc que la convergence suivante a lieu en probabilité

$$\frac{U(V_{(n-k+1,n)})-U(V_{(n-2k+1)})}{U(V_{(n-2k+1)})-U(V_{(n-4k+1)})} \to 2^{\xi} \tag{30}$$

Il reste maintenant à determiné la loi de $(U(V_(1,n)),...,U(V_(n,n)))$.Remarquons que si $x \geq 1$ alors

$$U(x) = F^{-1}(F_V(x)) \tag{31}$$

on a donc

$$(U(V_{(1,n)}),...,U(V_{(n,n)})) = (F^{-1}(F_V(V_{(1,n)})),...,F^{-1}(F_V(V_{(1,n)}))). \tag{32}$$

ou F_V est la fonctio de repartition de la loi de pareto . On deduit du lemme 1 que le vecteur aleatoire

$$(F^{-1}(F_V(V_{(1,n)})),...,F^{-1}(F_V(V_{(1,n)}))$$

à méme loi que

$$(x_(1,n),...,x_(n,n))$$

la statistique d'ordre d'un echantillon de n variables aleatoires independantes dont la loi à pour fonction de repartition F . Donc la variable aleatoire

$$\frac{U(V_{(n-k+1,n)})-U(V_{(n-2k+1)})}{U(V_{(n-2k+1)})-U(V_{(n-4k+1)})} \tag{33}$$

a meme loi que

$$\frac{X_{(n-k(n)+1,n)}-X_{(n-2k(n)+1,n)}}{X_{(n-2k(n)+1,n)}-X_{(n-4k(n)+1,n)}}) \tag{34}$$

ainsi cette quantité converge en loi vers 2^{ξ} quand n tend vers l'infini . Comme la fonction logarithme est continue sur R_{+}^{*} . On en deduit que l'estimateur de pickand converge en loi vers ξ.Mais comme ξ est une constante , on a egalement la convergence en probabilité $\square$

6.2 Estimateur de Hill

6.2.1 Lemme

soit **L** une fonction a variation lente . Alors on a :pour tout $\rho \geq 0$, $L(x) = o(x^p)$ en $+\infty$ et

$$\int_x^{\infty} t^{-\rho-1}L(t) \sim \frac{1}{\rho}x^{-\rho}L(x) \tag{35}$$

en $+\infty$

6.2.2 Theoreme

Soit $(X_n \ , \ n \geq 1)$ une suite de variable aleatoire independante de meme fonction de repartition $F \in D(H(\xi))$, ou $\xi \geq 0$. Si

$$\lim_{n\to\infty} k(n) = \infty$$

et

$$\lim_{n\to\infty} \frac{k(n)}{n} = 0$$

, alors l'estimateur de Hill définit par

$$\xi^{H}_{(k(n),n)} = \frac{1}{k(n)} \sum_{n-k(n)+1}^{n} logX_{(i,n)} - logX_{(n-k(n)+1,n)} \tag{36}$$

converge en probabilité vers ξ .

7 Applications

A ANNEXE

Définition
Soit la suite de maxima

$$M_1, ..., M_n, ... \tag{37}$$

D'une serie de variables aléatoires mutuellement indépendantes

$$x_1, ...x_n \tag{38}$$

est assujettie à la loi des grands nombres s'il existe des constantes A_n telles que l'on ait

$$\mathbf{P}\left\{|M_n - A_n| \leq \epsilon\right\} \to 1 \tag{39}$$

pour $n \to +\infty$ et tout $\epsilon < 0$ donnée d'avance.
La suite des maximas (37) seras appellée relativement stable si,pour un choix convenable de constantes positives B_n , la relation

$$\lim_{n\to\infty} \mathbf{P}\left\{|\frac{M_n}{B_n} - 1| < \epsilon\right\} = 1 \tag{40}$$

à lieu pour tout $\epsilon > 0$
Si la fonction de distribution $F(x)$ des variables aleatoires de la suite (38) jouit de ctte propriété qu'il existe une valeur x_0 telle que l'on ait

$$F(x_0) \ = \ 1 \ \ et \ \ F(x_0 - \epsilon) \ < \ 1 \tag{41}$$

pour tout $\epsilon > 0$ alors la suite (37) est assujettie à la loi des grands nombres.
Proposition

lemme 2. *Soient F_n et Φ des fonctions de distribution , $\Phi(x)$ n'etant pas unitaire . Si pour certaines suites de nombres réels $a_n \geq 0$, b_n , $\alpha_n \geq 0$, β_n on a*

$$F_n(a_n x + b_n) \to \Phi(x), F_n(\alpha_n x + \beta_n) \to \Phi(x) \tag{42}$$

pour $n \to \infty$

$$\frac{a_n}{\alpha_n} \to 1, \frac{b_n - \beta_n}{a_n} \to 0 \tag{43}$$

lemme 3. *Si $F_n(x)$ est une suite de distribution donnant lieu à la relation*

$$\lim_{n\to\infty} F_n(a_n x + b_n) = \Phi(x) : \tag{44}$$

pour un certains choix de constantes $a_n \geq 0$ et b_n et pour toutes les valeurs de x , alors pour deux suites quelconques de constantes $\alpha_n \geq 0$ et β_n telles que pour $n \to \infty$ $\frac{a_n}{\alpha_n} \to 1$, $\frac{b_n - \beta_n}{\alpha_n} \to 0$ on a $F_n(\alpha_n x + \beta_n) \to \Phi(x)$ pour $n \to \infty$ et toute les valeurs de x

Démonstration. :
Soient x_1, x, x_2 $(x_1 \leq x \leq x_2)$ des point de continuité de la fonction $\Phi(x)$. En vertu de notre hypothése nous avons pour n assez grand

$$x_1 \leq x\frac{\alpha_n}{a_n} + \frac{\beta_n - b_n}{a_n} \leq x_2 \tag{45}$$

Puisque

$$\alpha_n x + \beta_n = a_n(\frac{x}{\alpha_n}a_n + \frac{\beta_n - b_n}{a_n}) \tag{46}$$

on a pour n assez grand

$$a_n x_1 + b_n \leq \alpha_n x + \beta_n \leq a_n x_2 + b_n \tag{47}$$

et par suite

$$F_n(a_n x_1 + b_n) \leq F_n(a_n x + b_n) \leq F_n(a_n x_2 + b_n) \tag{48}$$

tenant compte de l'equation (2) ceci montre que

$$\Phi(x_1) \leq \liminf_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \Phi(x_2) \tag{49}$$

en faisant tendre x_1 , x_2 vers x on auras

$$\Phi(x_1) \to \Phi(x) \quad , \quad \Phi(x_2) \to \Phi(x) \tag{50}$$

x etant supposé point de continuité de $\Phi(x)$. On a donc

$$\Phi(x) \leq \liminf_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \Phi(x) \tag{51}$$

$\square$

lemme 4. *Si F est une fonction de distribution et si pour un choix de constantes $a_n \geq 0$ et b_n on a , pour $n \to \infty$ et toutes les valeurs de x*

$$F^n(a_n x + b_n) \to \Phi(x) \tag{52}$$

$\Phi(x)$ etant une fonction de distribution propre , alors ona pour $n \to \infty$,

$$\frac{a_n}{a_{n+1}} \to 1, \frac{b_n - b_{n+1}}{a_n} \to 0 \tag{53}$$

lemme 5. *Pour que l'on ait*

$$F^n(a_n x + b_n) \to \Phi(x) \tag{54}$$

pour toutes les valeurs de x et pour $n \to \infty$, il faut et il suffit d'avoir pour $n \to \infty$

$$n[1 - F_n(a_n x + b_n)] \to -log\Phi(x) \tag{55}$$

pour toutes les valeurs de x telles que $\Phi(x) \neq 0$

Démonstration. Supposons que la relation (48) ait lieu , alors pour toute valeur de a telle que $\Phi(x) \neq 0$ nous avons

$$\lim_{n\to\infty} F(a_n x + b_n) = 1 \tag{56}$$

il est evident que pour ces valeurs de x la condition (48) euivaut à la suivante

$$nlogF(a_n x + b_n) \to \Phi(x) \quad (\Phi(x) \neq 0) \quad pour \quad n \quad \to \infty \tag{57}$$

Or en vertu de (50) , nous avons

$$logF(a_nx+b_n) = -(1-F(a_nx+b_n)) \; - \; \frac{1}{2}(1-logF(a_nx+b_n))^2 \; = ... = \; -(1-F(a_nx+b_n)) \tag{58}$$

D'ou nous voyons que dés que (13) a lieu , la condition (49) est nécessairement remplie . Inversement si c'est la condition (14) qui à lieu , alors (50) a lieu aussi , donc , en vertu de (52)

$$n[1 - F_n(a_n x + b_n)] \quad = \quad nlogF(a_n x + b_n) \tag{59}$$

D'ou , et en vertu de (49) , alors , resulte (51) et ,par consequent , (48) $\square$

Code

simulation de la loi Uniforme

```
n=1000
y<-runif(n,min=0,max=1)

m=matrix(nrow = n, ncol=1, byrow = T)
c=matrix(nrow = n, ncol=1, byrow = T)
for (i in 1 :n)
m[i]=max(y[1 :i])
c[i]=i

plot(c,m,type="l", col="blue",main = "La loi du maximum des Xi suivant
la loi uniforme")
```

la fonction densité théorique de Mn

```
x<-seq(0,1,by=0.020)
y<-vector(length(x))

for (j in 1 :50)
if(j==1)
```

y=dunif(x)
plot(x,y,type = "l",col="red",main = "Densité théorique de Mn pour Xi suivant la loi Uniforme")

else

y=j*dunif(x)*(punif(x))$^{(}j - 1)$
$lines(x, y, type = "l", col = "red")$

simulation de la loi exponentielle
n=1000
x<-rexp(n,1)

m=matrix(nrow = n, ncol=1, byrow = T)
c=matrix(nrow = n, ncol=1, byrow = T)
for (i in 1 :n)
m[i]=max(x[1 :i])
c[i]=i

plot(c,m,type="l", col="green",main = "La loi du maximum des Xi suivant la loi exponentielle")

la fonction densité théorique de Mn
x<-seq(1,11,by=0.20)
y<-vector(length(x))

for (j in 1 :50)
if(j==1)
y=dexp(x)
plot(x,y,type = "l",col="red",main = "Densité théorique de Mn pour Xi suivant la loi exponentielle ")
else

y=j*dexp(x)*(pexp(x))$^{(}j - 1)$
$lines(x, y, type = "l", col = "green")$

simulation de la loi normale

n=1000
y<-rnorm(n)
m=matrix(nrow = n, ncol=1, byrow = T)

c=matrix(nrow = n, ncol=1, byrow = T)
for (i in 1 :n)
m[i]=max(y[1 :i])
c[i]=i

plot(c,m,type="l", col="red",main = "La loi du maximum des Xi suivant la loi normale")

la fonction densité théorique de Mn

x<-seq(1,11,by=0.20)
y<-vector(length(x))

for (j in 1 :50)
if(j==1)
y=dnorm(x)
plot(x,y,type = "l",col="red",main = "Densité théorique de Mn pour Xi suivant la loi normale")

else
y=j*dnorm(x)*(pnorm(x))$^{(}j-1)$
$lines(x, y, type = "l", col = "blue")$

Références

B.Gnedenko , Annals of mathematics, vol 44,pp.423-453

[1] [1] F.Delmas et B. Jourdain. Modèles aléatoires. Applications aux sciences de l'ingénieur et du vivant. Chapitre 11.